

# TESIS DOCTORAL

2020

## OPTIMIZACIÓN DE PREDICTORES DE DISRUPCIONES EN ESPACIOS BIDIMENSIONALES

Francisco Javier Hernández Martín

PROGRAMA DE DOCTORADO EN INGENIERÍA DE SISTEMAS  
Y DE CONTROL

Dr. D. JESÚS ANTONIO VEGA SÁNCHEZ

Dr. D. SEBASTIÁN DORMIDO CANTO



## Agradecimientos.

A mis directores de Tesis Doctoral, Dr. D. JESÚS ANTONIO VEGA SÁNCHEZ como Jefe de la Unidad de Análisis Avanzado de Datos del Laboratorio Nacional de Fusión del CIEMAT, y Dr. D. SEBASTIÁN DORMIDO CANTO como Director del Departamento de Informática y Automática de la UNED. Sin ellos, el trabajo que aquí se presenta no habría sido posible. Con sus consejos y dilatada experiencia, me han guiado sabiamente para avanzar en este trabajo de investigación.



## Dedicatoria.

*A mis padres Adelio y Felicísima, que me educaron en la cultura del esfuerzo y de los valores; algo que parece estar cada día más alejado en la sociedad actual. Mención especial a mi padre, del que he obtenido sabias lecciones de vida. Sus consejos, en cada una de las etapas por las que he pasado, han sido aplicados; surtiendo el efecto deseado. También mención especial a mi madre, que ha sabido defender como una leona todo lo que tenía que ver con los suyos (su familia).*

*A mi mujer, compañera de viaje durante más de veinticinco años, siempre cercana, presente en momentos dulces y en otros amargos; pero juntos hemos podido sumar cuando hemos vivido situaciones adversas. Ha sido ella fundamentalmente la que me ha soportado en situaciones de estrés ante exámenes o prácticas que debía entregar a la universidad. Me ha permitido disponer de todo el tiempo necesario para mí, a pesar de no poder dedicárselo ella. Muchas gracias, Rosa.*

*A mi hijo Alejandro, para que encuentre los beneficios que se pueden conseguir mediante la disciplina y el esfuerzo. Y que, a través de ellos, pueda crecer profesionalmente, alcanzando sus metas.*

*A mi hermana Ana, por haber sabido tratar como nadie los asuntos realmente importantes. Manteniéndose en su lugar, para que todo siguiera su curso con la mayor discreción, pero con la fortaleza y determinación que exigen ciertas situaciones adversas.*



# ÍNDICE

---

1. Introducción y Motivación de la Tesis.	1
1.1. El problema de la energía.	1
1.2. COP25.	14
2. Objetivo de la Tesis.	19
2.1. Objetivo de la tesis.	19
2.2. El fenómeno de las disrupciones.	21
2.3. Metodología de centroides en predicción de disrupciones.	23
3. Energía de Fusión y Marco de la Aplicación.	25
3.1. Introducción.	25
3.2. JET (Joint European Torus).	35
3.3. JT-60SA.	39
3.4. DIII-D.	40
4. Estado del arte de la predicción de disrupciones.	43
4.1. Introducción.	43
4.2. Evolución histórica de los métodos predictivos de disrupciones.	43
4.2.1. Utilización de redes neuronales.	43
4.2.2. Máquinas de Vectores Soporte.	49
4.2.3. Predictor para JET simulando entornos en tiempo real.	58
4.2.4. Predictor adaptativo para ASDEX Upgrade.	72
4.2.5. Predictor de disrupciones en tiempo real para JET.	89
4.2.6. Predictor en tiempo real desde cero para JET.	95
4.2.7. Predictor APODIS con entorno MARTE.	108
4.2.8. Predictores adaptativos probabilísticos.	114
4.2.9. Predictor basado en detección de anomalías (SPAD).	129
4.3. Métodos de predicción empleados en la Tesis.	140
5. Predicción basada en el método de los centroides.	143
5.1. Metodología matemática general del método de los centroides.	143
5.2. Metodología matemática de predicción en JET.	145
5.2.1. Fase de entrenamiento para descargas no disruptivas.	145
5.2.2. Fase de entrenamiento para descargas disruptivas.	146
5.2.3. Fase de evaluación.	148
5.3. Metodología matemática de predicción en JT-60U.	152
5.3.1. Fase de entrenamiento para descargas no disruptivas.	152
5.3.2. Fase de entrenamiento para descargas disruptivas.	153
5.3.3. Fase de evaluación.	153
5.4. Predicción de disrupciones en DIII-D.	155
5.4.1. Fase de entrenamiento para descargas no disruptivas.	156
5.4.2. Fase de entrenamiento para descargas disruptivas.	156

5.4.3. Fase de evaluación.	159
6. Aplicación del método de los centroides.	163
6.1. Justificación en el uso de los centroides.	163
6.2. Base de datos utilizada para JET.	166
6.3. Método de los centroides aplicado a JET.	168
6.3.1. Determinación de los centroides no disruptivos en JET.	175
6.3.2. Determinación de los centroides disruptivos en JET.	176
6.3.3. Fase de evaluación en JET aplicando el método de los centroides.	177
6.4. Resultados en JET.	179
6.5. Predicción del tiempo a la disrupción en cada alarma.	193
6.6. Análisis preliminar de datos para JT-60U.	200
6.7. Resultados para JT-60U.	202
6.8. Análisis de datos para DIII-D.	209
6.9. Resultados para DIII-D.	212
7. Conclusiones y líneas futuras de investigación.	221
7.1. Conclusiones.	221
7.2. Líneas futuras de investigación.	224
Bibliografía	227



# LISTA DE FIGURAS

<b>Figura 1.1.</b> Evolución temporal para alcanzar objetivo de cero emisiones en la Unión Europea.	2
<b>Figura 1.2.</b> Demanda energética mundial. Fuente: World Energy Outlook 2018, de la Agencia Internacional de la Energía.	4
<b>Figura 1.3.</b> Plazos e hitos del Programa Europeo de Fusión (ITER). Fuente: Roadmap, de Eurofusion.	6
<b>Figura 1.4.</b> Imagen de la mecánica de fusión dentro del núcleo (Foto © ITER.org).	11
<b>Figura 1.5.</b> Evolución de las instalaciones que albergarán el proyecto ITER.	11
<b>Figura 1.6.</b> Imagen que muestra las dimensiones relativas de la base del Tokamak.	12
<b>Figura 1.7.</b> Ilustración 3D de un reactor de fusión Tokamak.	12
<b>Figura 1.8.</b> Ilustración sobre las reacciones solares de fusión.	13
<b>Figura 1.9.</b> Logo de la COP25.	14
<b>Figura 2.1.</b> Imagen que presenta la inyección del gas ante una alerta previa a la disrupción.	21
<b>Figura 2.2.</b> Imagen de los daños ocasionados en el dispositivo, debidos a consecuencias térmicas.	22
<b>Figura 2.3.</b> Imagen de los daños ocasionados en el dispositivo, debidos a grandes corrientes de electrones muy energéticos.	22
<b>Figura 2.4.</b> Esquema general del método de los centroides.	23
<b>Figura 3.1.</b> Reacción que tiene lugar en el proceso de fusión.	26
<b>Figura 3.2.</b> TOKAMAK.	28
<b>Figura 3.3.</b> Componentes del dispositivo TOKAMAK de ITER (en cifras).	29
<b>Figura 3.4.</b> Esquema general desde la producción de energía hasta su distribución.	30
<b>Figura 3.5.</b> Esquema de funcionamiento de las bobinas magnéticas en el dispositivo TOKAMAK.	30
<b>Figura 3.6.</b> El único isótopo fisionable del uranio en la naturaleza es el uranio 235.	32
<b>Figura 3.7.</b> Reacción típica de fusión nuclear.	32
<b>Figura 3.8.</b> Socios participantes en el proyecto ITER.	33
<b>Figura 3.9.</b> Imagen aérea del CCFE en Oxfordshire (© CCFE).	35
<b>Figura 3.10.</b> Vista interna del recipiente, con una pared metálica completa de berilio y tungsteno (© EUROfusion).	36
<b>Figura 3.11.</b> Imagen dividida que muestra una vista interior del recipiente de vacío JET, con una imagen superpuesta de un plasma JET real tomada con una cámara de luz visible. Solo se pueden ver los bordes fríos del plasma, ya que el centro está tan caliente que irradia solo en la parte ultravioleta del espectro (© EUROfusion).	37
<b>Figura 3.12.</b> Cámara de vacío donde se destacan sus dimensiones (© CCFE).	38
<b>Figura 3.13.</b> Instalación del solenoide central de 4 módulos y 100 toneladas de peso.	39
<b>Figura 3.14.</b> Análisis y control del estado del plasma.	41
<b>Figura 3.15.</b> Sistema de control del estado del plasma.	41
<b>Figura 4.1.</b> (a) Predicción de red para una disrupción del límite de densidad; Las condiciones de alarma se alcanzan en $t = 3.955$ s, 65 ms antes de la disrupción. (b) Predicción de red por una disrupción después de introducir un evento de impureza; las condiciones de alarma se alcanzan en $t = 2.320$ s, 60 ms antes de la disrupción.	46
<b>Figura 4.2.</b> Falsa alarma en ASDEX Upgrade.	46
<b>Figura 4.3.</b> Comparación entre el indicador de desprendimiento (utilizado para controlar una válvula de gas) y la predicción de la red.	47
<b>Figura 4.4.</b> Hiperplanos de separación en espacio bidimensional de un conjunto de muestras separables en dos clases: (a) ejemplo de hiperplano de separación (b) otros ejemplos de hiperplanos de separación, de entre los infinitos posibles.	51
<b>Figura 4.5.</b> Margen de un hiperplano de separación: (a) hiperplano de separación no-óptimo y su margen asociado (no máximo) (b) hiperplano de separación óptimo y su margen asociado (máximo).	51
<b>Figura 4.6.</b> La frontera de decisión debe estar tan lejos de los datos de ambas clases como sea posible.	52
<b>Figura 4.7.</b> Función de decisión para un pulso de prueba. Las regiones grises representan las bandas discriminatorias entre muestras novedosas y no novedosas.	53
<b>Figura 4.8.</b> Arquitectura de la ruta del predictor / novedad; ruta de prueba.	56
<b>Figura 4.9.</b> Esquema del análisis en tiempo real de una descarga, utilizando una serie de clasificadores que analizan en ventanas consecutivas de tiempo de forma paralela.	62
<b>Figura 4.10.</b> Esquema general de entrenamiento de la DF. <b>1.</b> Se almacenan los datos relevantes. <b>2.</b> Se entrena una nueva DF que reemplaza a la anterior.	64
<b>Figura 4.11.</b> Resultados de entrenamiento de tres secuencias. El proceso de entrenamiento se detiene después de cinco iteraciones sin ninguna mejora.	66
<b>Figura 4.12.</b> Tasas de acierto y falsas alarmas de los modelos previamente entrenados en campañas posteriores. Han sido analizados de forma independiente tres períodos diferentes. Primero (izquierda), el periodo que incluye el entrenamiento y validación de la campaña C7 y las anteriores. En segundo lugar (en la zona intermedia), el período de prueba entre las campañas C8 a C14. Y tercero (derecha), después de los cambios estructurales del	

	dispositivo, Período comprendido entre C15 y C19.	68
<b>Figura 4.13.</b>	Tiempos de anticipación (tiempo de disrupción menos tiempo de alarma) para todas las descargas disruptivas desde la campaña C1 hasta la campaña C19. Parte izquierda figura: los porcentajes acumulativos de disrupciones no intencionadas y todas las detectadas. Se comparan las disrupciones para el método SVM con los resultados de JPS. Figura derecha: los porcentajes acumulativos de disrupciones intencionadas y todas las disrupciones. Se comparan las disrupciones detectadas para el método SVM con los resultados de JPS.	70
<b>Figura 4.14.</b>	Histograma acumulativo del intervalo de tiempo ( $t_D - t_{LM}$ ) entre el tiempo LM y la ocurrencia de la disrupción.	75
<b>Figura 4.15.</b>	Algunos parámetros de plasma y función de alarma correspondiente a descarga # 16513.	77
<b>Figura 4.16.</b>	Algunos parámetros de plasma que intervienen en la disrupción de la descarga # 18906, donde la disrupción es provocada por un evento de desplazamiento vertical (VDE).	80
<b>Figura 4.17.</b>	Algunos parámetros de plasma para la descarga disruptiva # 17752, donde el VDE es provocado por la disrupción.	80
<b>Figura 4.18.</b>	Visualización de un SOM para una descarga disruptiva a partir de un espacio de parámetros de seis dimensiones en un espacio bidimensional.	82
<b>Figura 4.19.</b>	SOM 2D que muestra el espacio de parámetros: en azul las muestras no disruptivas de DB 1; en verde las muestras disruptivas de DB 1; en rojo las muestras en los últimos 160 ms de las MA de DB 2.	84
<b>Figura 4.20.</b>	Intervalo de tiempo ( $t_D - t_{LM}$ ) entre la ocurrencia de la disrupción y el activador LM para las descargas en DB 2, frente a los valores de $ q_{95} $ .	86
<b>Figura 4.21.</b>	Intervalo de tiempo ( $t_D - t_{LM}$ ) entre la ocurrencia de la disrupción y el activador NNP para las descargas en DB 2, frente a los valores de $ q_{95} $ .	86
<b>Figura 4.22.</b>	Histograma apilado del rendimiento del sistema adaptativo en relación con los sucesivos procedimientos de reentrenamiento, para las descargas de DB 2.	88
<b>Figura 4.23.</b>	El conjunto de datos de entrenamiento está formado por círculos y cuadrados. La función de decisión (línea punteada roja) separa ambas clases. Con clasificadores SVM, la función de decisión en el espacio de características divide ambas clases por medio de un hiperplano (línea roja), donde el hiperplano se define por las muestras más cercanas de cada clase. Estas muestras aparecen en las líneas discontinuas y se denominan vectores de soporte.	90
<b>Figura 4.24.</b>	Arquitectura de varios niveles de APODIS. La función de decisión está determinada combinando los modelos M1, M2 y M3. La predicción final identifica un comportamiento no disruptivo (-1) o un comportamiento disruptivo (1).	91
<b>Figura 4.25.</b>	Comparación de los tiempos de predicción entre APODIS y el sistema activador de bloqueo del modo de JET.	94
<b>Figura 4.26.</b>	Ejemplo simplificado.	97
<b>Figura 4.27.</b>	Rendimiento de predicción de la tasa de aciertos frente a la tasa de falsas alarmas.	101
<b>Figura 4.28.</b>	Arquitectura APODIS.	102
<b>Figura 4.29.</b>	Resultados de predicción de disrupción desde cero con conjuntos de datos no balanceados.	103
<b>Figura 4.30.</b>	Predicción de disrupciones desde cero teniendo en cuenta datos balanceados.	105
<b>Figura 4.31.</b>	Resultados de predicción de disrupción desde cero con un enfoque híbrido.	106
<b>Figura 4.32.</b>	Predicción de disrupciones desde cero con un enfoque híbrido. Los reentrenamientos se llevan a cabo después de perderse una disrupción, que sucede en las disrupciones 58, 85, 93 y así sucesivamente.	107
<b>Figura 4.33.</b>	Arquitectura APODIS, resaltando las 7 señales de entrada a los clasificadores SVM.	109
<b>Figura 4.34.</b>	Efecto en la lectura de la señal cuando no hay muestra disponible en la RTDN.	111
<b>Figura 4.35.</b>	Arquitectura APODIS utilizando el marco MARTE.	112
<b>Figura 4.36.</b>	Disrupciones acumulativas detectadas de APODIS versus tiempo de disrupción para descargas 82429 a 82905 en la campaña C28.	113
<b>Figura 4.37.</b>	Pasos que implican el proceso de clasificación: inducción y deducción.	115
<b>Figura 4.38.</b>	Si el tiempo de reacción es mayor que el tiempo de anticipación, la acción de mitigación es tardía.	117
<b>Figura 4.39.</b>	En los predictores conformales todas las muestras de entrenamiento se usan con cada nueva predicción.	119
<b>Figura 4.40.</b>	Las cuatro muestras más a la izquierda de este ejemplo tienen etiqueta "Cuadrado" y las cuatro más a la derecha tienen la etiqueta "círculo". $C_s$ y $C_c$ son respectivamente los centroides de las clases "cuadrado" y "círculo".	120
<b>Figura 4.41.</b>	Ventanas temporales de 32 ms formadas desde el instante en que la corriente alcanza el umbral de 750 kA hasta que abandona dicho umbral.	122
<b>Figura 4.42.</b>	Cada fila corresponde a un vector de n características durante la descarga.	122
<b>Figura 4.43.</b>	Vector característico promedio en descargas no disruptivas.	123
<b>Figura 4.44.</b>	Resultados del predictor Venn con las características 2, 3, 4 y 5: evolución de la tasa de aciertos, tasa de falsas alarma y probabilidad de predicción según el número de disrupciones.	127
<b>Figura 4.45.</b>	Resultados de la versión APODIS desde cero.	128
<b>Figura 4.46.</b>	Tiempos de anticipación del predictor Venn con características 2, 3, 4 y 5.	129
<b>Figura 4.47.</b>	Representación de los dos coeficientes de aproximación de la transformada wavelet de Haar	

para (a) descargas no disruptivas y (b) disruptivas como un vector de características en el espacio bidimensional.	130
<b>Figura 4.48.</b> Esquema de la aplicación típica de MARTE y de su integración con RTDN de JET y con el sistema de persistencia.	132
<b>Figura 4.49.</b> Diagrama de implementación de SPAD en MARTE.	132
<b>Figura 4.50.</b> Explicación de las muestras de señales en ventanas deslizantes.	134
<b>Figura 4.51.</b> Proceso de obtención de coeficientes de la transformada wavelet de Haar sobre varias iteraciones de la transformada.	135
<b>Figura 4.52.</b> Representación de la fracción acumulativa de detección de disrupciones con respecto a las disrupciones totales durante todas las campañas ILW de JET.	138
<b>Figura 4.53.</b> Representación del comportamiento del factor atípico SPAD en la descarga 82960.	138
<b>Figura 4.54.</b> Tiempo promedio de ejecución del ciclo para los GAM de SPAD con un nivel de confianza del 97%.	139
<b>Figura 5.1.</b> Procedimiento para la obtención del centroide no disruptivo en la fase de entrenamiento de JET.	145
<b>Figura 5.2.</b> Procedimiento para la obtención del centroide disruptivo en la fase de entrenamiento de JET.	146
<b>Figura 5.3.</b> Procedimiento que determina la naturaleza de una muestra en fase de evaluación de JET.	148
<b>Figura 5.4.</b> Procedimiento del desplazamiento relativo de los centroides en la fase de evaluación de JET.	149
<b>Figura 5.5.</b> Detalle del desplazamiento relativo del centroide disruptivo en la fase de evaluación de JET.	150
<b>Figura 5.6.</b> Desplazamiento relativo del centroide disruptivo de JET con 81 posiciones.	151
<b>Figura 5.7.</b> Desplazamiento relativo de los dos centroides de JET con 81 posiciones.	151
<b>Figura 5.8.</b> Proceso de obtención del centroide no disruptivo en fase entrenamiento de JT-60U.	152
<b>Figura 5.9.</b> Proceso de obtención del centroide disruptivo en fase de entrenamiento de JT-60U.	153
<b>Figura 5.10.</b> Obtención de la distancia de una muestra que está siendo evaluada en tiempo real con respecto a los centroides procedentes de JT-60U.	154
<b>Figura 5.11.</b> Distintos comportamientos de la señal Amplitud para JT-60U.	155
<b>Figura 5.12.</b> Centroides DIII-D con centroide disruptivo considerando 100 ms previos a la disrupción.	157
<b>Figura 5.13.</b> Centroides DIII-D con centroide disruptivo considerando el valor máximo en los 200 ms previos a la disrupción.	157
<b>Figura 5.14.</b> Centroides DIII-D con centroide disruptivo considerando el valor medio de los dos valores máximos en los 100 ms previos a la disrupción.	158
<b>Figura 5.15.</b> Centroides DIII-D con centroide disruptivo considerando el valor medio de los valores máximo y mínimo en los 100 ms previos a la disrupción.	158
<b>Figura 5.16.</b> Centroides DIII-D con centroide disruptivo considerando el valor medio de los diez valores máximos en los 100 ms previos a la disrupción.	158
<b>Figura 5.17.</b> Centroide no disruptivo y contribución al mismo de cada una de las descargas.	159
<b>Figura 5.18.</b> Centroide disruptivo y contribución de sus descargas.	159
<b>Figura 5.19.</b> Descarga con todas sus muestras no disruptivas.	160
<b>Figura 5.20.</b> Descarga con alguna de sus muestras disruptivas.	160
<b>Figura 6.1.</b> Técnica de predicción basada en el centroide más próximo.	164
<b>Figura 6.2.</b> Fundamentos del predictor SIDTIP.	165
<b>Figura 6.3.</b> Vista conceptual de las distribuciones de tiempo de anticipación.	165
<b>Figura 6.4.</b> Representación de la señal ML en el dominio del tiempo.	169
<b>Figura 6.5.</b> Amplitudes de ML en JET.	170
<b>Figura 6.6.</b> Comportamiento de las muestras en el espacio de parámetros.	171
<b>Figura 6.7.</b> Representación de los centroides en JET atendiendo a los datos de la Tabla 6.3.	173
<b>Figura 6.8.</b> Posiciones del $C_D$ debidas a desplazamientos relativos según las barras de error.	174
<b>Figura 6.9.</b> Procedimiento de obtención del centroide global no disruptivo en JET.	175
<b>Figura 6.10.</b> Ejemplos de descargas disruptivas en JET y su contribución al centroide disruptivo.	176
<b>Figura 6.11.</b> Ejemplo de falsas alarmas en JET.	177
<b>Figura 6.12.</b> Ejemplo de alarma tardía en JET.	178
<b>Figura 6.13.</b> Ejemplo de alarma perdida en JET.	178
<b>Figura 6.14.</b> Curva ROC para ML/ $I_p$ .	179
<b>Figura 6.15.</b> Distribución de los tiempos de anticipación.	180
<b>Figura 6.16.</b> Diagramas de dispersión.	181
<b>Figura 6.17.</b> Desplazamiento relativo de los centroides obtenidos de C15 – C20.	185
<b>Figura 6.18.</b> Impacto del desplazamiento de los centroides en la recta que define la frontera.	186
<b>Figura 6.19.</b> Posiciones relativas de los centroides para las nuevas pruebas.	187
<b>Figura 6.20.</b> Gráficas ROC que muestran de entre todos los resultados, el resultado óptimo.	188
<b>Figura 6.21.</b> Distribución de las muestras en torno a la diagonal.	189
<b>Figura 6.22.</b> Disposición de las bandas en torno a la diagonal.	189
<b>Figura 6.23.</b> Emplazamiento de muestras disruptivas en torno a la diagonal para descarga 86268.	190
<b>Figura 6.24.</b> Efectos de las bandas en torno a la diagonal para la descarga 86129.	192
<b>Figura 6.25.</b> Efectos de las bandas con respecto a los tiempos de anticipación.	192
<b>Figura 6.26.</b> Los círculos son los pares (distancia a la diagonal, tiempo de anticipación), que se obtienen	

al reconocer comportamientos disruptivos con ecuación (11) en el conjunto de datos de 113 descargas disruptivas, que se muestran en la Tabla 6.3. La línea roja continua se ajusta a un modelo de Weibull.	197
<b>Figura 6.27.</b> Variación de los parámetros $\alpha$ y $\beta$ del ajuste de Weibull, con límites de confianza del 95% para los casos $k = 1, \dots, 6$ . El eje horizontal muestra el límite derecho de los intervalos.	197
<b>Figura 6.28.</b> Tasas de prueba en las clases respectivas, para cada uno de los 14 modelos. El eje x muestra el límite derecho de los respectivos intervalos de tiempo.	199
<b>Figura 6.29.</b> Representación de señales para una descarga en JT-60U.	200
<b>Figura 6.30.</b> Representación de señales para una descarga con distinto comportamiento en JT-60U.	200
<b>Figura 6.31.</b> Representación de un comportamiento típico no disruptivo en JT-60U.	201
<b>Figura 6.32.</b> Representación de varios comportamientos disruptivos en JT-60U.	201
<b>Figura 6.33.</b> Contribución de las descargas a los centroides en la fase de entrenamiento.	203
<b>Figura 6.34.</b> Espacio de parámetros para distintos comportamientos en el JT-60U.	204
<b>Figura 6.35.</b> Evolución de las señales para una descarga disruptiva en el JT-60U.	205
<b>Figura 6.36.</b> Evolución de las señales para una descarga no disruptiva en el JT-60U.	206
<b>Figura 6.37.</b> Evolución de las señales para descargas con presencia de falsas alarmas en el JT-60U.	207
<b>Figura 6.38.</b> Tiempos de anticipación para la fila 1 de la Tabla 6.15. en el JT-60U.	207
<b>Figura 6.39.</b> Señales proporcionadas para el dispositivo DIII-D.	209
<b>Figura 6.40.</b> Detección con presencia de “ruido” en dispositivo DIII-D.	210
<b>Figura 6.41.</b> Espacio de representación para DIII-D.	211
<b>Figura 6.42.</b> Descarga no disruptiva en DIII-D, con todas sus muestras situadas en la región no disruptiva.	213
<b>Figura 6.43.</b> Descarga con la presencia de una falsa alarma en DIII-D.	213
<b>Figura 6.44.</b> Descarga que presenta comportamiento disruptivo en DIII-D.	214
<b>Figura 6.45.</b> Ampliación de las señales que intervienen para la descarga 162423 de DIII-D.	214
<b>Figura 6.46.</b> Descarga que presenta comportamiento disruptivo, con presencia de alarma prematura.	215
<b>Figura 6.47.</b> Ampliación de las señales que intervienen para la descarga 162457 de DIII-D.	215
<b>Figura 6.48.</b> Descarga que presenta comportamiento disruptivo, con presencia de alarma tardía en DIII-D.	216
<b>Figura 6.49.</b> Ampliación de las señales que intervienen para la descarga 162456 de DIII-D.	216
<b>Figura 6.50.</b> Resultados para cuatro versiones diferentes en la obtención de los centroides en DIII-D.	217
<b>Figura 6.51.</b> Regresión logística aplicada a DIII-D.	219

# LISTA DE TABLAS

<b>Tabla 3.1.</b> Datos relativos a JT-60SA en relación con ITER.	40
<b>Tabla 4.1.</b> Peso y composición de los gránulos utilizados en ASDEX Upgrade.	48
<b>Tabla 4.2.</b> Señales de diagnóstico.	54
<b>Tabla 4.3.</b> Rendimiento del predictor SVM en términos de porcentaje de falsas alarmas (PFA), porcentaje de alarmas perdidas (PMA), porcentaje de alarmas prematuras (PPA) y tasa de éxito en la predicción (PSR).	57
<b>Tabla 4.4.</b> SVM Predictor y detector predictor / novedad. Rendimiento en términos de porcentaje de falsas alarmas (PFA), porcentaje de alarmas perdidas (PMA), porcentaje de alarmas prematuras (PPA) y tasa de acierto en la predicción (PSR).	57
<b>Tabla 4.5.</b> Lista de la base de datos según campañas y número total de descargas utilizadas para el análisis.	59
<b>Tabla 4.6.</b> Lista de las señales analizadas en cada descarga.	60
<b>Tabla 4.7.</b> Ejemplos de los seis casos posibles de clasificaciones correctas e incorrectas que pueden ocurrir en la práctica.	64
<b>Tabla 4.8.</b> Resultados generales de la prueba.	67
<b>Tabla 4.9.</b> Señales en la base de datos.	78
<b>Tabla 4.10.</b> Prestaciones de NNP en el conjunto de prueba de DB 1.	83
<b>Tabla 4.11.</b> Rendimiento de NNP en DB 2.	84
<b>Tabla 4.12.</b> Rendimientos de predicción en descargas disruptivas de DB 2 para LM y para el predictor NN.	85
<b>Tabla 4.13.</b> Rendimiento del predictor neuronal adaptativo en DB 2.	89
<b>Tabla 4.14.</b> Lista de señales que determinan el estado del plasma en JET.	90
<b>Tabla 4.15.</b> Evaluación APODIS: tasa de aciertos, alarmas perdidas y falsas alarmas en las campañas JET C28–C30.	93
<b>Tabla 4.16.</b> Modelos de entrenamiento con conjuntos de datos desbalanceados para el ejemplo simplificado de la Figura 4.26.	98
<b>Tabla 4.17.</b> Modelos de entrenamiento con conjuntos de datos balanceados para el ejemplo simplificado de la Figura 4.26.	99
<b>Tabla 4.18.</b> Conjuntos de prueba para el ejemplo simplificado de la Figura 4.26.	100
<b>Tabla 4.19.</b> Base de datos de campañas JET C28 – C30.	102
<b>Tabla 4.20.</b> Lista de señales.	103
<b>Tabla 4.21.</b> Tasas promedio de éxito y falsas alarmas con la desviación estándar para conjuntos de datos no balanceados para los conjuntos de señales expandido y reducido. Los datos se calculan a partir del modelo 24.	104
<b>Tabla 4.22.</b> Tasas promedio de éxito y falsas alarmas con la desviación estándar para conjuntos de datos balanceados para los conjuntos de señales expandido y reducido. Los datos se calculan a partir del modelo 42 en adelante.	105
<b>Tabla 4.23.</b> Tasas promedio de éxito y falsas alarmas con la desviación estándar para un enfoque híbrido con señal expandida y reducida. Los datos se calculan a partir del modelo 42 en adelante.	107
<b>Tabla 4.24.</b> Rango de señales para selección de descargas.	109
<b>Tabla 4.25.</b> Descargas seleccionadas.	110
<b>Tabla 4.26.</b> Lista de señales para caracterizar el comportamiento disruptivo/no disruptivo en JET.	121
<b>Tabla 4.27.</b> Identificación de características.	125
<b>Tabla 4.28.</b> Resultados de análisis de DPFS con NCT.	126
<b>Tabla 4.29.</b> Resultados basados en distintas combinaciones de características.	126
<b>Tabla 4.30.</b> Tiempos promedio de anticipación AWT con predictores Venn.	128
<b>Tabla 4.31.</b> Comparación de las Tasas de Detección con SPAD.	137
<b>Tabla 4.32.</b> Comparación de Tasas de Detección entre SPAD, APODIS y LMPT.	137
<b>Tabla 6.1.</b> Información de campañas seleccionadas con descargas disruptivas.	167
<b>Tabla 6.2.</b> Información de campañas seleccionadas con descargas no disruptivas.	167
<b>Tabla 6.3.</b> Conjuntos de datos de descargas disruptivas (D) y no disruptivas (ND) para probar un predictor basado en centroides en JET.	172
<b>Tabla 6.4.</b> Resultados del predictor seleccionado por la curva ROC.	180
<b>Tabla 6.5.</b> Resultados según SPAD y según el método de los centroides C36 - C37.	182
<b>Tabla 6.6.</b> Resultados según el método de los centroides para C34.	183
<b>Tabla 6.7.</b> Resultados según centroides con campaña C34 y con pared de carbono.	183
<b>Tabla 6.8.</b> Resultados según el método de los centroides para C15 - C37.	184
<b>Tabla 6.9.</b> Resultados según el método de los centroides con barras de error.	185
<b>Tabla 6.10.</b> Base de datos utilizada usando campañas ILW.	186
<b>Tabla 6.11.</b> Cálculo de los puntos donde se desplaza el centroide disruptivo.	187
<b>Tabla 6.12.</b> Resultados para el conjunto de datos definido en la Tabla 6.10.	187

<b>Tabla 6.13.</b> Aplicación de las bandas a los resultados de la Tabla 6.12.	190
<b>Tabla 6.14.</b> Aplicación de bandas más anchas a los resultados de la Tabla 6.12.	191
<b>Tabla 6.15.</b> Predicciones de JT-60U teniendo en cuenta 11 juegos de prueba aleatorios.	204
<b>Tabla 6.16.</b> Resultados con predicción adaptativa desde cero para JT-60U.	208
<b>Tabla 6.17.</b> Presentación de la doble polaridad en datos DIII-D.	210

# LISTADO DE SIGLAS Y ACRÓNIMOS

---

## A

<b>APODIS</b>	Advance Predictor Of DISruptions.
<b>ASDEX</b>	Axially Symmetric Divertor EXperiment.
<b>AT</b>	Alarm Threshold.

## C

<b>CCFE</b>	Culham Center for Fusion Energy.
<b>CFC</b>	Carbon Fiber Composite.
<b>CIEMAT</b>	Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas.
<b>CMNUCC</b>	Convención Marco de Naciones Unidas para el Cambio Climático.
<b>COP25</b>	25 <sup>th</sup> Conference Of Parts.
<b>CSN</b>	Consejo de Seguridad Nuclear.

## D

<b>DF</b>	Decision Function.
<b>DMS</b>	Disruption Mitigation System.
<b>DPFS</b>	Disruption Predictor From Scratch.

## F

<b><i>F<sub>4E</sub></i></b>	Fusion for Energy.
<b>FCI</b>	Fusión nuclear por Confinamiento Inercial.
<b>FCM</b>	Fusión nuclear por Confinamiento Magnético.
<b>FPGA</b>	Field Programmable Gate Array.

## H

<b>HPC</b>	High Performance Computation.
------------	-------------------------------

**I**

<b>IDR</b>	Initial Decision Rule.
<b>ILW</b>	ITER- Like Wall.
<b>ITER</b>	International Thermonuclear Experimental Reactor.

**J**

<b>JET</b>	Joint European Torus.
<b>JPS</b>	JET Protection System.

**L**

<b>LEP</b>	Linear Equation Predictor.
<b>LM</b>	Locked Mode.

**M**

<b>MARTe</b>	Multithreaded Application Real-Time executor.
<b>MHD</b>	Magneto-Hydrodynamic.

**N**

<b>ND</b>	Novelty Detection.
-----------	--------------------

**O**

<b>OCR</b>	Optical Character Recognition.
<b>OIEA</b>	Organismo Internacional de Energía Atómica.

**R**

<b>ROC</b>	Receiver operating characteristic.
<b>RTPS</b>	Real Time Protection Sequencer.



## **S**

<b>SIDTIP</b>	Signal Increments for Disruption Time Prediction.
<b>SOM</b>	Self-Organizing Map.
<b>SPAD</b>	Single Signal Predictor based on Anomaly Detection.
<b>SVM</b>	Support Vector Machine.

## **T**

<b>TOKAMAK</b>	TOroidalbnaya KAmera MAgnitnim Katushkami (cámara toroidal de bobinas magnéticas).
<b>TTD</b>	Time To the Disruption.

## **U**

<b>UKAEA</b>	United Kingdom Atomic Energy Authority.
--------------	---

## **V**

<b>VDE</b>	Vertical Displacement Event.
------------	------------------------------



# INTRODUCCIÓN Y MOTIVACIÓN DE LA TESIS

1

## 1.1. El problema de la energía.

Aunque Europa es responsable de “tan solo” el 10 % de las emisiones de gases de efecto invernadero, los estados miembros de la Comunidad Europea se han planteado como objetivo para 2050, abandonar completamente los combustibles fósiles. Si los 27 países europeos certifican el acuerdo de tal objetivo, se sentarán las bases para la lucha contra el calentamiento global. Los responsables europeos consideran que la clave para frenar ese problema medio ambiental, es la erradicación de los combustibles fósiles [\[URL, 20\]](#).

Lograr ese objetivo representa una mejora en todos los ámbitos, también en el económico; ya que además de incrementar el PIB en un 2 %, se ahorrará el gran coste económico debido al cambio climático. Si fracasa el Acuerdo de París [\[URL, 30\]](#), se estima ese coste en 240.000 millones de euros [\[URL, 20\]](#).

Las renovables tendrán hasta 2050 un periodo de transición, y en ese periodo aproximadamente un 15 % de la energía procederá de la energía nuclear; que ocasiona problemas por los residuos generados, aunque no expulse  $CO_2$  [\[URL, 20\]](#).

1

La Figura 1.1. presenta los objetivos a cumplir para la obtención de cero emisiones en el ámbito de la Unión Europea.

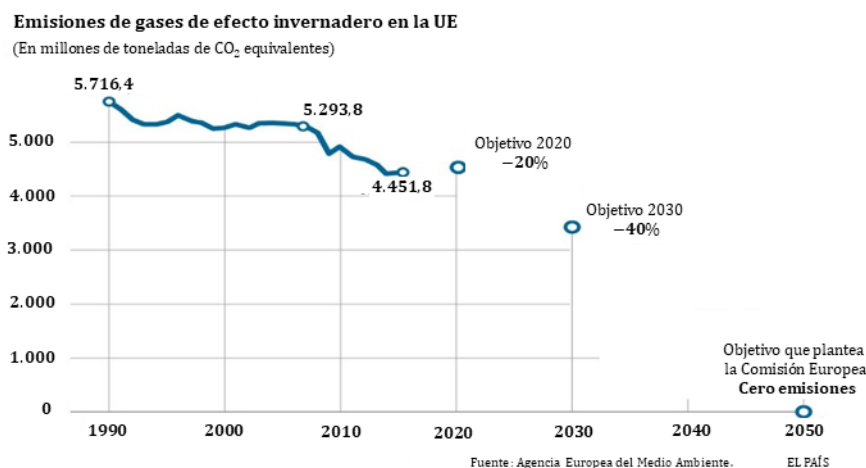


Figura 1.1. Evolución temporal para alcanzar objetivo de cero emisiones en la Unión Europea.

España cuenta por primera vez con un ministerio específico, que pretende el cambio gradual de las energías fósiles a energías renovables. Se trata del Ministerio de Transición Ecológica.

Las fuentes de energía no renovable son las relacionadas con la energía nuclear (fusión y fisión), y las que tienen su origen en los combustibles fósiles (carbón, petróleo y gas natural) [URL, 23].

La energía fósil puede ser biogénica, como el carbón (lignito) y los hidrocarburos líquidos (petróleo crudo) o gaseosos (gas natural); tratándose de una energía no renovable. Esa energía fósil también puede ser mineral, donde los combustibles son sustancias que pueden producir energía nuclear; ya sea mediante la fisión o la fusión nuclear [URL, 27].

Los combustibles fósiles son sustancias procedentes de restos de seres vivos, depositados en distintas cuencas sedimentarias. El carbón, petróleo y gas natural constituyen los principales combustibles de este tipo [URL, 21].

El carbón se forma gracias a la descomposición de vegetales terrestres (hojas, maderas, cortezas), que se van acumulando en zonas pantanosas de poca profundidad. Esta fuente de energía, entre otras aplicaciones, se dedica a la producción de energía eléctrica [URL, 23].

El petróleo se forma gracias a grandes depósitos de materia orgánica. Esos depósitos fueron cubiertos por otras capas de sedimentos hace miles de años. Su combustión produce el polémico CO<sub>2</sub> [URL, 23].

El gas natural se encuentra generalmente junto a los yacimientos de carbón o petróleo. Dentro de los combustibles fósiles, es el más limpio; ya que genera menos emisiones que el carbón o el petróleo [URL, 23].

La energía nuclear se obtiene a partir de las reacciones nucleares, siendo el principal inconveniente el peligro que corre la vida humana ante un mal funcionamiento [URL, 23].

En resumen, las energías no renovables, al emitir  $CO_2$ , contribuyen al cambio climático. De ahí, la firme intención de sustituirlas por otras energías que, además de renovables (ilimitadas), sean más limpias [URL, 23].

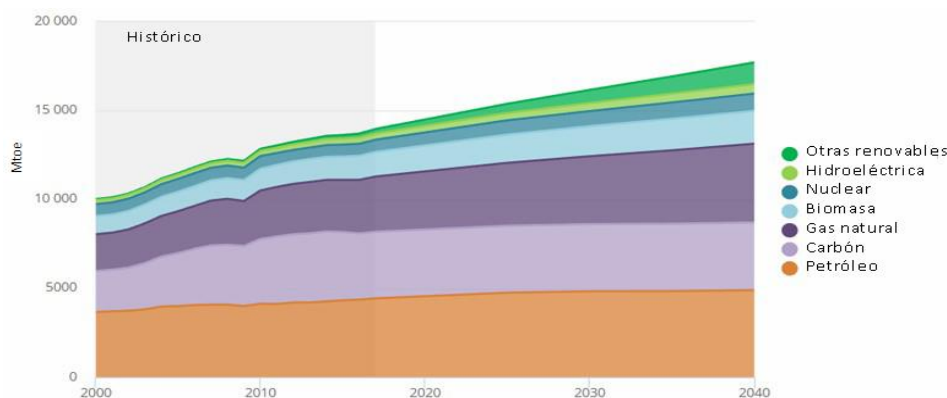
Progresivamente se ha ido sustituyendo el carbón por el petróleo, siendo éste el combustible fósil más utilizado incluso como materia prima para la mayoría de los productos petroquímicos. Esas fuentes de energía fósiles, tanto biogénicas como minerales, siguen siendo los pilares de nuestro actual sistema de suministro de energía. Pero, debido a las crecientes demandas de energía a nivel mundial, las reservas de energía fósil se han reducido. Para compensar dicha reducción de reservas, se está aumentando el uso de energías renovables, para reducir nuestra dependencia de los combustibles fósiles [URL, 27].

La sociedad actual precisa de unos recursos energéticos fiables y abundantes, pero nuestras principales fuentes de combustible, como el petróleo y el gas, son cada vez más escasas, más caras y, en cualquier caso, generan importantes emisiones de gases de efecto invernadero [URL, 1]. El aumento permanente de concentraciones de  $CO_2$  se considera una de las principales causas que agravan el cambio climático [URL, 27]. Aunque se incluyan las tecnologías renovables para la producción de energía, no será suficiente; necesitando desarrollar nuevas fuentes que produzcan energía a gran escala y sin dañar el medio ambiente. La solución pasa por la energía de fusión [URL, 17].

Además del crecimiento de la población mundial, otro factor que condiciona la demanda energética es el aumento del nivel de desarrollo. Las energías fósiles (carbón, petróleo y gas) son limitadas, contaminantes y suelen estar localizadas en zonas con una alta inestabilidad geopolítica. La fisión nuclear es considerada peligrosa por parte de la población, la energía eólica y solar dependen de factores meteorológicos. [URL, 6].

Los expertos estiman que de aquí a 2040, la demanda energética habrá aumentado en un 25%, y de cara a finales de siglo, dicha demanda se habrá duplicado con respecto a la actual (véase Figura 1.2.). Para satisfacer esas previsiones, según el “World Energy Outlook 2018” de la Agencia Internacional de la Energía, se precisan inversiones de más

de 2 millones de millones de dólares al año en nuevos suministros de energía. Y, en caso de no hallar una solución, se podría desencadenar una crisis energética de gran magnitud [URL, 2].



**Figura 1.2.** Demanda energética mundial. Fuente: World Energy Outlook 2018, de la Agencia Internacional de la Energía.

En 1990 el consumo de energía por habitante y año en los países industrializados fue de 5,1 toneladas equivalentes de petróleo; siendo el consumo en los países en vías de desarrollo de la décima parte que en los países industrializados. La evolución de la demanda energética para 2050 podría duplicarse con respecto a 1990 [URL, 5].

El consumo per cápita mundial en 1991 era de aproximadamente 2.2 kW. Actualmente las principales fuentes de energía primaria suponen aproximadamente: petróleo (38 %), carbón (28 %), gas natural (21 %), energía hidroeléctrica (7 %), y energía nuclear (6 %). Ello supone que el 87 % de la energía total mundial, se produce mediante combustibles fósiles (petróleo, carbón y gas natural); que tienen como inconveniente ser recursos no renovables [URL, 28]. Y, además, la contaminación puede ser la causa de graves problemas de salud pública [URL, 29].

Tratar de sustituir los combustibles fósiles por tecnologías energéticas limpias (que no generen emisiones de  $CO_2$ ), no resulta fácil. Alternativas tales como energía nuclear (fusión y fisión), hidroeléctrica, solar, eólica, geotérmica, son incapaces de proporcionar un porcentaje notable entre el 30 y el 40 %. La fisión nuclear en algunos países supone más del 50 % de su energía. Sin embargo, algunos accidentes han mostrado los altos riesgos de seguridad para esta opción energética [URL, 28].

En Europa, Francia es el país que más apuesta por la energía nuclear; debido a que el 73 % de la electricidad que produce es de origen nuclear. A nivel mundial, aproximadamente el 10 % de la energía que utiliza el ser humano, proviene de la fisión

nuclear. A pesar de los residuos nucleares altamente tóxicos generados por la fisión, éstos son almacenados en lugares lo más seguros posibles.

En 2018, un estudio de la revista *European Heart Journal*, reveló que solo en Europa 800.000 personas mueren prematuramente debido a la contaminación del aire [\[URL, 31\]](#). Por otra parte, otro estudio de 2013, elaborado por el Instituto de Estudios Espaciales NASA Goddard y el Instituto de la Tierra de la Universidad de Columbia, reveló que la energía nuclear global ha prevenido una media de 1,84 millones de muertes relacionadas con la contaminación del aire; además de evitar la emisión de 64 gigatoneladas de gases de efecto invernadero, que hubieran resultado de la quema de combustibles fósiles [\[Karecha, 2013\]](#).

La fusión nuclear podría ser la solución para cambiar de forma definitiva los combustibles fósiles a las energías renovables. Para hacer esto posible comercialmente, se necesita tanto el apoyo de organizaciones gubernamentales, como el apoyo de empresas privadas [\[URL, 24\]](#).

Según Michael Delage, Director de Tecnología de General Fusion (2º mayor laboratorio privado de fusión de América del Norte), la capacidad de investigación debe ser considerada como una inversión ya que, dotando de las habilidades necesarias a los graduados, se podrá contribuir al campo de la fusión [\[URL, 24\]](#).

La fusión podría ser considerada la opción más ventajosa dentro de las energías renovables, ya que es la que mayor energía produce y la que menor huella de carbono deja entre todas las fuentes disponibles actualmente [\[URL, 24\]](#).

Para producir energía nuclear de fisión se utilizan sustancias radiactivas como el  $U_{235}$ , aunque también se usa el  $U_{233}$ , y el  $Pu_{239}$ . Este tipo de energía tiene como ventaja que, a partir de pequeñas cantidades de combustible, se produce una gran cantidad de energía; siendo abundantes las reservas de estos materiales nucleares. Como inconvenientes, las reacciones de fisión generan residuos de difícil eliminación; exigiendo medidas muy costosas en cuanto a control y seguridad [\[URL, 21\]](#).

Según Carlos Hidalgo, jefe de la División de Física Experimental del Laboratorio Nacional de Fusión del CIEMAT, “la búsqueda de fuentes de energía se remonta al amanecer del ser humano” [\[URL, 2\]](#).

Una gran parte de la generación de energía procede de la quema masiva de combustibles fósiles, siendo mucho mayor su consumo en los países desarrollados. Con un consumo per cápita como el actual y mediante la energía de fusión, la humanidad podría obtener toda la energía que necesita empleando un gramo de Litio por persona al

año. El empeño por dar solución a la gran demanda energética mundial pasa por el proyecto ITER [URL, 4].

En la actualidad existe un compromiso internacional para la consecución de la fusión nuclear como objetivo, siendo el principal referente el proyecto ITER (véase Figura 1.3.). Las investigaciones actuales en lo relativo al programa de fusión, se llevarán a cabo en tres fases, donde se considera [URL, 2]:

- Corto plazo (< 2020).
- Medio plazo (2030 - 2040).
- Largo plazo (> 2040).

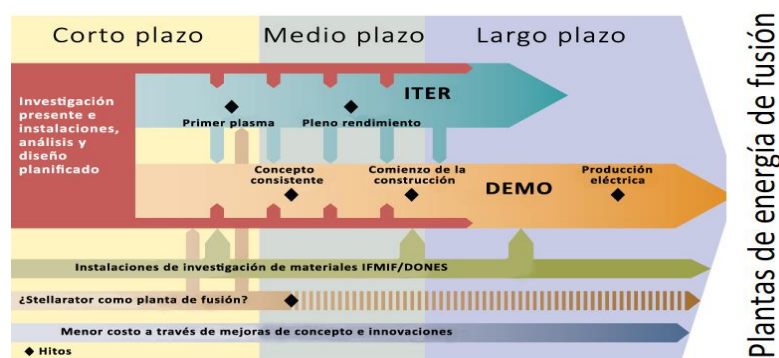


Figura 1.3. Plazos e hitos del Programa Europeo de Fusión (ITER). Fuente: Roadmap, de Eurofusion.

Como desafíos de futuro, se trata de diseñar centrales para la producción de energía de fusión que sean seguras, fiables, sostenibles, que respeten el medio ambiente y que sean económicamente viables [URL, 2].

La fusión nuclear podría solucionar el problema del cambio climático, ya que se trata de una fuente de energía que además de limpia, es inagotable. El combustible requerido por el reactor, es el hidrógeno extraído del agua. Los factores meteorológicos que condicionan la producción de energía para las fuentes renovables actuales, no supondrían obstáculo alguno cuando se trata de la fusión nuclear [URL, 3].

Como ventajas, los elementos que dan lugar a la reacción de fusión son muy abundantes en la naturaleza, ya que se trata del hidrógeno y de sus isótopos; y además no produce residuos radiactivos. Como inconveniente, se precisan temperaturas de 100 millones de grados centígrados para iniciar la reacción [URL, 21]. Por analogía y como dato curioso, decir que se necesitan temperaturas diez veces superiores a la que hay en el núcleo del Sol.



La fusión nuclear, como alternativa a la producción masiva de energía, puede ser muy importante durante la segunda mitad de siglo. Su respeto por el medio ambiente, así como los combustibles que utiliza suponen dos grandes ventajas [\[URL, 4\]](#).

A pesar de que las energías renovables atienden cada vez más demanda energética, se hacen necesarias nuevas alternativas que respeten el medio ambiente, sean seguras y económicamente viables. Aunque todavía hay que superar la complejidad tecnológica de los dispositivos para la fusión termonuclear, ésta sería una de esas alternativas económicas, limpias y seguras [\[URL, 5\]](#).

La fusión nuclear es más fácil de controlar que la reacción de fisión, ya que no se trata de una reacción en cadena. Y además reduce el problema de los residuos nucleares. La fusión nuclear no produce emisiones de carbono, y sería capaz de generar energía según la demanda [\[URL, 6\]](#).

En la reacción de fusión se unen dos núcleos ligeros para formar un núcleo más pesado. La masa del núcleo resultante puede ser ligeramente inferior a la suma de las masas de los núcleos ligeros. Esa diferencia de masa se libera en forma de energía, de acuerdo a la ecuación  $E = mc^2$ . De esta forma, la fusión nuclear resulta muy eficiente para generar energía; liberando mucha mayor energía que un proceso químico. Una reacción de fusión puede liberar aproximadamente 10 millones de veces la energía liberada por un proceso químico, como la quema de madera o petróleo [\[URL, 7\]](#).

La fuerza que mantiene unidos a los protones y neutrones en el núcleo de un átomo es lo que conocemos como energía nuclear. Dicha energía es la que será liberada durante un proceso de fusión o de fisión nuclear. La energía liberada viene dada por la famosa ecuación, desarrollada por Albert Einstein,  $E = mc^2$ . Es decir, la energía es igual a la masa por la velocidad de la luz al cuadrado [\[URL, 29\]](#).

A la energía de fusión se le ha llamado “el santo grial” de la energía; por tratarse de una energía limpia, inagotable y más barata. En la fusión, el residuo que se produce es helio, un gas inofensivo y con valor económico. Este tipo de residuo contrasta con el que se produce mediante la fisión, que implica la desintegración del material radiactivo; arrojando residuos altamente contaminantes, de larga vida y difícil eliminación [\[URL, 8\]](#).

Las reacciones de fusión nuclear pueden liberar o absorber energía. Si los núcleos que se fusionan tienen menor masa que el hierro, se liberará energía. Sin embargo, si los núcleos a fusionar son más pesados que el hierro, se absorberá energía [\[URL, 11\]](#).

Las reacciones de fusión nuclear también son conocidas como reacciones termonucleares, por las altas temperaturas que experimentan. Para que se produzcan reacciones de este tipo, han de cumplirse los siguientes requisitos técnicos [\[URL, 11\]](#):

- Conseguir temperaturas muy elevadas, que permitan vencer las fuerzas de repulsión electrostáticas; permitiendo la aproximación de los núcleos a fusionar. Una vez separados los electrones del núcleo, esos electrones libres y los átomos ionizados forman una masa gaseosa denominada plasma.
- Confinar el plasma durante un periodo mínimo de tiempo a temperaturas elevadas. Existen dos tipos de confinamiento:
  - o FCI (Fusión por Confinamiento Inercial): Crea un medio tan denso que las partículas no tienen apenas posibilidad de escapar sin chocar entre sí.
  - o FCM (Fusión por Confinamiento Magnético): Las partículas eléctricamente cargadas, procedentes del plasma, son atrapadas en un espacio reducido por la acción de campos magnéticos.
- Obtener una densidad del plasma que permita acercar a los núcleos, para así poder generar reacciones de fusión nuclear.

Un ejemplo de reacciones de fusión tiene lugar en el Sol, uniéndose núcleos de hidrógeno para formar helio, y liberar además una gran cantidad de energía que percibimos como luz y calor [\[URL, 12\]](#).

Con la ley de creación del CSN en 1980 se formaron dos organismos [\[URL, 13\]](#):

- El CSN, que se encarga de las funciones de regulación en materia de seguridad nuclear y protección radiológica.
- El CIEMAT, encargado de los proyectos de investigación y desarrollo, entre los cuales se encuentran los relativos a fusión nuclear.

Actualmente, más de 50 estados miembros del OIEA están llevando a cabo investigaciones sobre fusión nuclear controlada y física del plasma; siendo el reto poder demostrar la viabilidad científica de la fusión como fuente de energía. El OIEA promueve la colaboración y coordinación internacional con el fin de subsanar los aspectos físicos, tecnológicos y reguladores; con el fin de avanzar en el desarrollo del uso pacífico de la energía de fusión [\[URL, 18\]](#).

La energía nuclear de fusión es energía limpia, sin emisiones de  $CO_2$ , y que produce una gran cantidad de energía a bajo coste. Siendo los combustibles [\[URL, 19\]](#):

- Deuterio, isótopo estable del hidrógeno formado por un protón y un neutrón, cuyo contenido energético en un litro de agua de mar es equivalente al contenido energético proporcionado por 250 litros de petróleo.
- Tritio, isótopo inestable del hidrógeno, compuesto por un protón y dos neutrones, que se desintegra con bastante rapidez. Aunque el tritio escasea en la naturaleza, puede generarse a partir del litio (elemento abundante en la corteza terrestre).

La energía nuclear, tanto generada por fusión o fisión, no producen emisiones de gases de efecto invernadero. Sin embargo, las reacciones producidas en la fusión son más seguras. Las reacciones de fisión producen reacciones en cadena, pudiendo llegar a originar accidentes. La fusión puede generar de 3 a 4 veces más energía que la fisión [\[URL, 19\]](#).

En el ITER se utilizará tritio de las reservas existentes, producidas en reactores de fisión que operan con agua pesada (agua que contiene deuterio) [\[URL, 14\]](#).

La fusión nuclear aún no está suficientemente desarrollada como para poder aplicarla a la producción real de energía [\[URL, 22\]](#).

Para la producción de energía utilizable a partir del proceso de fusión nuclear se requiere la intervención de diversas áreas científicas como: la física, la ciencia de los materiales, la energía de alta precisión, la robótica, la computación, simulación y la inteligencia artificial [\[URL, 25\]](#).

Uno de los campos donde ya se aplican las reacciones de fusión nuclear es la medicina. Gracias a imanes superconductores, diseñados para controlar las reacciones de fusión nuclear, son empleados para la obtención de imágenes por resonancia magnética [\[URL, 25\]](#).

La fusión genera muchos menos residuos radiactivos que la fisión. Y en el supuesto caso de una reacción fuera de control o una fuga, el plasma se expandiría y se enfriaría con el tiempo. Con lo cual, no provocaría una peligrosa explosión [\[URL, 29\]](#).

A diferencia de la fisión, la fusión presenta claras ventajas en lo relativo a la seguridad y a los residuos nucleares. Un reactor de fusión no puede quedar fuera de control, ya que la reacción se detiene sin riesgo alguno. En caso de producirse el accidente más dramático, como pueda ser la rotura del reactor, en el caso de un reactor de fusión

por confinamiento magnético se liberarían aproximadamente 200 mg de tritio; que es el combustible requerido para 10 segundos de operación. Si a un reactor de fusión se le introduce accidentalmente más combustible del que puede recibir, éste se apagará; debido a que el reactor no podrá calentarlo. Un punto más a favor de la fusión, es que las reservas de combustible disponibles (deuterio y tritio) son, con gran diferencia, mayores que las de cualquier fuente de energía de las que se conocen en la actualidad [URL, 28], y se encuentran disponibles a lo largo de todo el planeta.

En definitiva, y a partir del conocimiento científico-tecnológico disponible en la actualidad, la fusión es la única opción capaz de resolver el problema energético a escala mundial [URL, 28].

ITER es uno de los proyectos energéticos mundiales más ambiciosos. Colaboran en la construcción del Tokamak más grande del mundo 35 países, habiendo firmado un acuerdo para compartir el coste de la construcción, operación y desmantelamiento del proyecto; así como los resultados obtenidos y cualquier propiedad intelectual. Se trata de un dispositivo de fusión magnética diseñado para demostrar la viabilidad económica y técnica de la fusión como fuente de energía a gran escala. Es un proyecto experimental, que tratará de preparar las futuras centrales comerciales de fusión [URL, 9].

El Tokamak de ITER está diseñado para [URL, 9]:

- Desarrollar una potencia de 500 MW. El récord mundial de energía de fusión actual lo tiene el Tokamak europeo JET, que en 1997 produjo 16 MW.
- Demostrar el funcionamiento integrado de las distintas tecnologías que intervienen en una central nuclear de fusión.
- Conseguir una reacción prolongada de deuterio y tritio mediante calentamiento interno.
- Probar la producción de tritio dentro de la vasija de vacío. El suministro mundial de tritio no es suficiente para abastecer las necesidades de las futuras centrales nucleares de fusión.
- Demostrar la seguridad de un dispositivo de fusión. En 2014 ITER obtuvo licencia como operador nuclear en Francia. Siendo uno de los principales objetivos, demostrar que se puede controlar el plasma y las reacciones de fusión sin consecuencias medioambientales.

En la Figura 1.4. se muestra la mecánica de fusión dentro de uno de los dispositivos de fusión.



**Figura 1.4.** Imagen de la mecánica de fusión dentro del núcleo (Foto © ITER.org).

El término “Tokamak” es un acrónimo ruso, que significa “cámara toroidal con bobinas magnéticas”. Se trata de un dispositivo experimental diseñado para aprovechar la energía de fusión. La central de fusión utiliza el calor, que se adhiere a las paredes de la vasija, para producir vapor. A partir de ese vapor se generará electricidad mediante turbinas y generadores. Para poder confinar y controlar el plasma, en el Tokamak se utilizan campos magnéticos muy potentes [\[URL, 9\]](#).

Un reactor de fusión es una instalación en la que se producen las reacciones nucleares de fusión, a partir de deuterio y tritio; liberándose energía en forma de calor, para posteriormente transformar ese calor en energía eléctrica. Tal y como se ha indicado anteriormente, los reactores de fusión del futuro pueden ser de dos tipos: los que utilicen confinamiento inercial, o los que empleen confinamiento magnético [\[URL, 10\]](#).

La instalación de ITER se halla completada al 65 %, habiendo puesto fecha para el <<primer plasma>>, diciembre de 2025. Para observar el estado de las instalaciones, véase Figura 1.5. Para producir energía a nivel comercial, como muy pronto, se estima para 2035 [\[URL, 15\]](#).



**Figura 1.5.** Evolución de las instalaciones que albergarán el proyecto ITER [\[URL, 15\]](#).

## 1.1. El problema de la energía.

Con fecha de 24 de julio de 2019 fue instalada la base y el cilindro principal del Tokamak, un gran logro de la ingeniería teniendo en cuenta sus dimensiones y sus complejas especificaciones [URL, 15]. En la Figura 1.6. se pueden apreciar sus dimensiones.



Figura 1.6. Imagen que muestra las dimensiones relativas de la base del Tokamak [URL, 15].

Este componente, de acero inoxidable, tiene unas dimensiones de  $16.000 m^3$ . Su altura y diámetro son de casi 30 metros, y tiene un peso de 3.850 toneladas. Contendrá los imanes superconductores más grandes del mundo, que se precisan para la generación de campos magnéticos lo suficientemente potentes, como para contener un plasma que alcanzará temperaturas de 150 millones de grados centígrados [URL, 15].

La Figura 1.7. muestra el interior de un dispositivo de fusión Tokamak.



Figura 1.7. Ilustración 3D de un reactor de fusión Tokamak [URL, 16].

Para que el proceso de fusión tenga lugar es necesario que el combustible que interviene se caliente a millones de grados, hasta alcanzar el estado de plasma (cuarto estado de la materia). Los diseños principales para los reactores nucleares de fusión (Tokamaks y Stellarators) prometen todas las ventajas de la fisión nuclear, pero evitando el problema de los residuos nucleares y problemas de seguridad; debido a que el plasma se extingue por sí mismo en caso de accidente. Un Stellarator es un dispositivo utilizado para confinar plasmas calientes mediante campos magnéticos, con el objetivo de mantener reacciones de fusión nuclear de forma controlada [URL, 16].

El proyecto ITER (“camino”, en latín), pretende ser la llave que abra el acceso a la futura energía de fusión. Tratándose de uno de los proyectos más caros de la humanidad, por detrás: del proyecto de la Estación Espacial Internacional [URL, 16], el proyecto Manhattan (la construcción de la primera bomba atómica) y el proyecto Apollo, entre otros.

El premio Nobel de Física Pierre Gilles de Gennes, expuso en una ocasión acerca de la fusión nuclear [URL, 16]:

*“Decimos que pondremos el sol en una caja. La idea es atractiva, pero el problema está en que no sabemos cómo ni de qué hacer la caja”.*

La Figura 1.8. ofrece una visión de las reacciones de fusión solares.



**Figura 1.8.** Ilustración sobre las reacciones solares de fusión [URL, 26].

En resumen, entre las ventajas de la fusión se encuentran [URL, 26]:

- Suministro de combustible ilimitado, ya que básicamente se utiliza deuterio (un isótopo de hidrógeno muy presente en nuestro planeta).
- Energía segura ya que, frente a los riesgos de la fisión, la fusión no es una reacción en cadena; por lo que puede controlarse y cancelarse simplemente retirando el combustible.

- Energía relativamente <<limpia>>, al no producir gases que contribuyan al calentamiento global. El helio que se genera a partir de la reacción de fusión no es nocivo, pero sí produce tritio radiactivo; volviéndose radiactivas las paredes del reactor debido al plasma.

Los avances técnicos en 2014 por el centro de investigación estadounidense Lawrence Livermore National Laboratory [URL, 32], logrando por primera vez que una reacción de fusión liberase más energía de la que absorbe; se considera un paso fundamental para lograr algún día la producción de energía ilimitada, limpia y económicamente viable. Debería ser prioritario a nivel mundial, siempre que los intereses del lobby energético lo permitan, ir desechando a medio plazo las centrales nucleares de fisión; así como las que utilizan combustibles fósiles [URL, 26].

## 1.2. COP25.

Del 2 al 13 de diciembre de 2019, el IFEMA de Madrid ha sido el escenario de la sede de la Cumbre Mundial del Clima COP25; donde más de 25.000 representantes de unos 200 países se han dado cita [URL, 36]. En la Figura 1.9. se muestra el logo de dicha cumbre.

Esta COP25 tenía prevista su celebración en Santiago (Chile), pero debido a la situación del país, fue trasladada a Madrid. Tiene como tarea asegurarse de la implementación del Acuerdo de París de 2015, según los compromisos asumidos por cada uno de los países que ratificaron dicho Acuerdo [URL, 37].



Figura 1.9. Logo de la COP25 [URL, 37].



La CMNUCC es la cumbre más importante del mundo en cuanto a cambio climático. La primera Conferencia de Partes (COP) tuvo lugar en 1995 en la ciudad de Berlín, habiéndose realizado 24 ediciones previas a la anterior; y siendo la ciudad polaca de Katowice en 2018 la sede de la edición número 24. En esta vigésimo quinta cumbre, la organización y presidencia la ostentaba Chile. Las partes han tenido como objetivo llegar a un consenso sobre los mercados de carbono [\[URL, 35\]](#).

La conferencia ha sido diseñada para establecer los pasos sucesivos en el proceso de negociaciones sobre el cambio climático de las Naciones Unidas. Siendo el principal objetivo la revisión de asuntos pendientes para la puesta en funcionamiento del Acuerdo de París sobre el Cambio Climático [\[URL, 33\]](#).

El cambio climático constituye un problema global, que exige la colaboración de todos los países. En la COP1, celebrada en Berlín en 1995, los países participantes consideraron necesaria la reducción de las emisiones de los gases de efecto invernadero (GEI). Desde entonces se celebra anualmente la COP, para avanzar en las negociaciones. De las COP han salido dos acuerdos importantes: el Acuerdo de Kioto y el Acuerdo de París [\[URL, 34\]](#).

El Protocolo de Kioto, que entró en vigor en febrero de 2005, estableció por primera vez objetivos de reducción de emisiones de GEI. Las emisiones de GEI en los países industrializados deberían reducirse en el periodo 2008-2012 en un 5 % como mínimo con respecto a los niveles del año 1990 [\[URL, 34\]](#).

El 12 de diciembre de 2015, en la COP21, se adoptó el Acuerdo de París; cuyo objetivo es evitar que la temperatura media del planeta supere los 2° C con respecto a los niveles preindustriales. Además, se promueven esfuerzos adicionales para conseguir que esa temperatura no supere los 1,5°C. El Acuerdo de París entró en vigor el 4 de noviembre de 2016, donde ratificaron el Acuerdo los países que representan más del 55 % de las emisiones globales. España ratificó el Acuerdo de París con fecha de 12 de enero de 2017 [\[URL, 34\]](#).

Las emisiones alcanzadas por los GEI dan una idea de que, si continua esta tendencia en el aumento de dichas emisiones, las futuras generaciones se encontrarían con impactos más severos del cambio climático. Para limitar el aumento de las temperaturas a 1,5°C con respecto a los niveles preindustriales, se necesita reducir las emisiones en un 7,6 % anual durante el periodo comprendido entre 2020 y 2030. Actualmente no se está haciendo lo suficiente para cumplir los tres objetivos climáticos [\[URL, 37\]](#):

- Reducir las emisiones en un 45 % para 2030.
- Emisiones netas cero de carbono para 2050.
- Estabilizar el aumento de la temperatura global a 1,5°C para finales de siglo.

Los estados deben demostrar un compromiso claro con respecto a su ambición climática, puesto que con los compromisos actuales se prevé un aumento de 3°C de las temperaturas globales hacia finales de siglo. Será clave, para conseguir dicho objetivo, que los países más vulnerables a los impactos del cambio climático conciencien a los países más industrializados y con más recursos. De forma que estos países con más recursos puedan garantizar la financiación, para hacer frente a las pérdidas y daños derivados del cambio climático. La siguiente COP (COP26), se celebrará en Glasgow a finales de 2020 [\[URL, 38\]](#).

Los elogios de la COP25 se reservan para la organización española del evento. Durante dos semanas, representantes de más de 190 países se han reunido para concretar los ‘flecós’ del Acuerdo de París; siendo el asunto más espinoso de la negociación las reglas de funcionamiento de los mercados de carbono. Estados Unidos, China, India, Japón y Brasil han eludido anunciar contribuciones adicionales a las presentadas en Francia en 2005. Aunque 73 países si han expresado su intención para reforzar los planes de acción climática a partir del año 2020, éstos sólo representan el 10 % de las emisiones del planeta. Eludiendo este esfuerzo China (29 % de emisiones), Estados Unidos (16 %), India (7 %) o Japón (4 %) [\[URL, 39\]](#).

En resumen, mientras que unos países piden más ambición, basándose en lo que dice la ciencia, otros países se ajustan al Acuerdo de París; evitando que la actualización al alza de los planes de acción se demore hasta el año 2023 [\[URL, 39\]](#).

Las discrepancias se centraron en las reglas de funcionamiento de los futuros mercados de carbono, ya que los países podrían comprar derechos de emisión sobrantes a otros países. De esta forma mejorarían las economías de los países menos contaminantes, ya que podrían vender sus derechos por la cuota de gases no emitida [\[URL, 39\]](#).

La tesis se ha estructurado de la siguiente manera: en el capítulo 2 se describe el objetivo del trabajo de investigación realizado y se nombran las tareas que se han implementado para lograr dicho objetivo. En el capítulo 3 se muestra una visión general de la energía de fusión como marco de aplicación de la tesis. En el capítulo 4 se realiza

un estudio exhaustivo del estado del arte de la predicción de disrupciones, que se ha desarrollado con anterioridad a la metodología propuesta en la tesis. El capítulo 5 presenta el método basado en los centroides, como metodología general utilizada para los nuevos predictores de disrupciones y se particulariza el procedimiento a cada uno de los dispositivos experimentales de fusión termonuclear que se han analizado. Seguidamente, en el capítulo 6 se presentan y analizan los resultados obtenidos mediante el método de los centroides, para la predicción de disrupciones en los tres dispositivos estudiados. El capítulo 6 termina mostrando, como primera aproximación, cómo se puede utilizar el método basado en los centroides para predecir el tiempo que queda para que ocurra la disrupción. Finalmente, en el capítulo 7, se presentan las conclusiones y las futuras líneas de la investigación realizadas.

Cabe mencionar que, en la tesis se ha intentado traducir al castellano todos los términos utilizados en la redacción. Únicamente, en aquellos casos donde la traducción pudiera llevar a la confusión, se ha optado por mantener el término anglosajón.



# OBJETIVO DE LA TESIS

2

## 2.1. Objetivo de la tesis.

La tesis que aquí se presenta, está enmarcada dentro del campo de la fusión nuclear. El interés por este campo está motivado porque un día podría sustituir a las actuales fuentes de energía.

El objetivo general de la tesis es encontrar un procedimiento automático que, a través de una ecuación física lo más simple posible, permita encontrar predictores de interrupciones que sean efectivos. Esa efectividad se consigue cuando se obtiene una tasa de aciertos por encima del 95%, una tasa de falsas alarmas por debajo del 5%, y si las alarmas se predicen con la suficiente antelación. Si la alarma se predice con mucho tiempo de anticipación, se pueden realizar tareas de elusión; llevando el plasma a una zona de parámetros seguros, para que no se produzca la interrupción y así pueda continuar el procesamiento de la descarga. Si la alarma no se produce con la antelación suficiente, la interrupción resultará inevitable; pudiendo realizar acciones de mitigación, pero dando por perdida la descarga. Una vez que se produzca la interrupción, la descarga resultará irre recuperable.

La tasa de aciertos se basa en el número de alarmas detectadas antes del instante de la interrupción, para descargas que tienen comportamiento disruptivo.

La tasa de falsas alarmas ofrece el número de alarmas detectadas, durante el procesamiento de las descargas, en aquellas que realmente son no disruptivas.

Dependiendo de los dispositivos de fusión, esos procesos automáticos pueden variar notablemente, ya que los datos originales de los que se dispone, son diferentes según el dispositivo del que se trate. Los datos originales son diferentes para los tres dispositivos analizados (JET, JT-60U y DIII-D).

La solución pasa por la determinación de un punto o la ecuación de una recta, que separe el espacio disruptivo del no disruptivo. Tal y como se apuntaba anteriormente, los datos originales disponibles condicionarán en gran medida los procedimientos para la búsqueda de soluciones.

Para conseguir establecer esa frontera entre ambas regiones, con las descargas se realizará una primera fase (entrenamiento), para determinar los puntos medios donde se concentran las descargas disruptivas y las no disruptivas. El procedimiento será detallado más adelante, ya que para cada tipo de dispositivo varía notablemente.

Es importante destacar que, durante la fase de entrenamiento, de las descargas empleadas se conoce su comportamiento disruptivo o no disruptivo. Y en el caso de las descargas disruptivas, también se conoce el instante temporal en el que tiene lugar la disrupción.

Una vez obtenidos, en la fase de entrenamiento, los puntos medios que concentran las descargas disruptivas y no disruptivas; lo que se pretende es analizar cada una de las muestras correspondientes a las descargas que se procesan en una fase posterior (evaluación o test); para determinar si se trata o no realmente de una descarga disruptiva.

Durante la fase de evaluación serán analizados los datos muestreados en determinadas señales cada milisegundo, para determinar el comportamiento de cada una de esas muestras en tiempo real (JET), o simulando que el dispositivo opera en tiempo real (JT-60U y DIII-D). Las señales empleadas fundamentalmente son la intensidad de corriente del plasma y otra señal relacionada con la rotación del plasma.

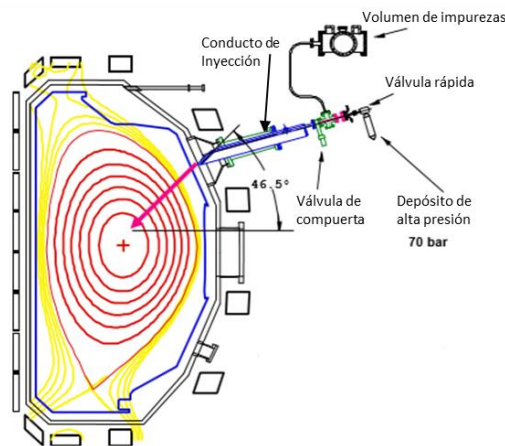
En la siguiente sección se presenta el fenómeno de las disrupciones en los dispositivos de fusión termonuclear, con el fin de poder contextualizar el objetivo de la tesis. Para finalizar el capítulo se presentan de manera muy general, las tareas que se van a realizar para implementar la metodología basada en centroides propuesta en el trabajo de investigación. El estudio en detalle de la metodología y su adecuación a los distintos dispositivos de fusión examinados, serán mostrados en capítulos posteriores.

## 2.2. El fenómeno de las disrupciones.

Uno de los principales problemas de los dispositivos experimentales de fusión termonuclear, es la aparición de un tipo de inestabilidad denominada disrupción. La disrupción se produce por una repentina pérdida del confinamiento del plasma. En ese momento, la enorme cantidad de energía que se libera va hacia las paredes del dispositivo; pudiendo dañarlo de tal forma que lo deje inutilizado. Por tanto, en operación, es deseable el poder predecir si la descarga que se está produciendo, puede dar lugar o no a una disrupción. Una vez que se predice la disrupción, en el caso de que se prediga con suficiente antelación, se puede realizar alguna acción de elusión o de mitigación.

Durante la fase de evaluación de descargas, si se produce una alerta en alguna de sus muestras, y ésta se manifiesta con la suficiente antelación con respecto al instante de la disrupción; se podrían llevar a cabo acciones de elusión o de mitigación a través de un sistema de mitigación de disrupciones (DMS). El ITER DMS se basa en la inyección masiva de impurezas, para irradiar la energía almacenada en el plasma y mitigar los efectos potencialmente dañinos de las disrupciones [Hollmann, 2015].

La Figura 2.1. muestra el sistema de mitigación de disrupciones.

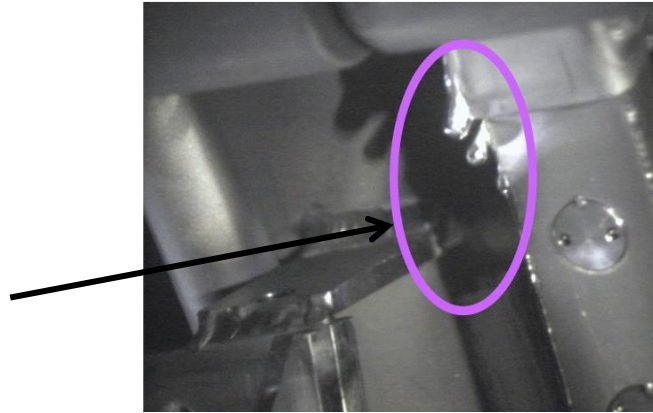


**Figura 2.1.** Imagen que presenta la inyección del gas ante una alerta previa a la disrupción [URL, 41].

Dichas disrupciones son propias de los dispositivos TOKAMAK. Cuando se produce una disrupción se desencadenan consecuencias muy graves, entre las que cabe destacar [Vega, 2016a]:

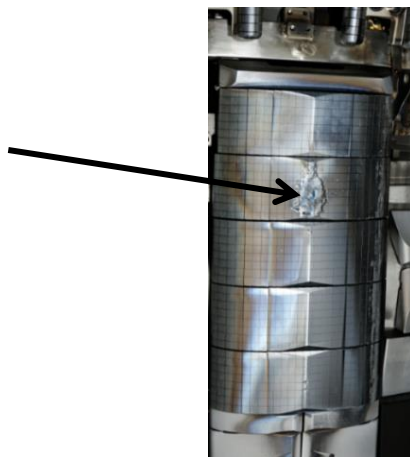
1. Grandes cargas térmicas en componentes del dispositivo. Pueden aparecer temperaturas diez veces superiores al umbral en el que se funden los

materiales. En la Figura 2.2. se puede ver el efecto de las consecuencias térmicas.



**Figura 2.2.** Imagen de los daños ocasionados en el dispositivo, debidos a consecuencias térmicas.

2. Se originan enormes fuerzas de origen electromagnético (del orden de 10 Mega Newtons). Para hacernos una idea de estas consecuencias, son equiparables:
  - a. Al peso de un Airbus A380: 5,5 MN (Mega Newtons).
  - b. O bien al peso de un barco de tamaño medio: 50 MN (Mega Newtons).
3. Se generan grandes corrientes de electrones “runaway” muy energéticos; pudiendo dañar las paredes del dispositivo, tal y como se muestra en la Figura 2.3.



**Figura 2.3.** Imagen de los daños ocasionados en el dispositivo, debidos a grandes corrientes de electrones muy energéticos.

Si se demuestra imposible el control de las disrupciones, la línea TOKAMAK sería inviable como futuro reactor.



El problema que se intenta resolver consiste en predecir los distintos tipos de interrupciones de una manera automática, con el fin de poder actuar en distintos escenarios según el tipo de interrupción. Este problema no es fácil. De hecho, en la comunidad científica que aborda este problema, no hay acuerdo en los intentos de clasificaciones que se han realizado. La señal del Locked Mode (amplitud) es una señal magnética que mide la rotación del plasma, está disponible en la totalidad de los dispositivos experimentales de fusión termonuclear y se suele utilizar para la predicción de las interrupciones.

Se sabe que existen diferentes tipos de interrupciones. Y, aunque algunas de las interrupciones han sido posibles clasificarlas manualmente gracias a personal experto en fusión, todavía no ha sido posible clasificarlas de forma automática.

Una primera aproximación para dicha automatización, sería una clasificación no supervisada [URL, 40]; que permitiera determinar posibles clases de interrupciones. Se trata de un proceso muy complejo, ya que su naturaleza puede ser muy diversa (presencia de campos electromagnéticos, densidad del plasma, corriente del plasma, impurezas en el plasma, etc), pudiendo darse varios factores desencadenantes de tales interrupciones de forma simultánea.

### 2.3. Metodología de centroides en predicción de interrupciones.

En la Figura 2.4 se muestran en forma de diagrama de flujo y de manera muy general, las tareas que se van a realizar para implementar la metodología basada en centroides.

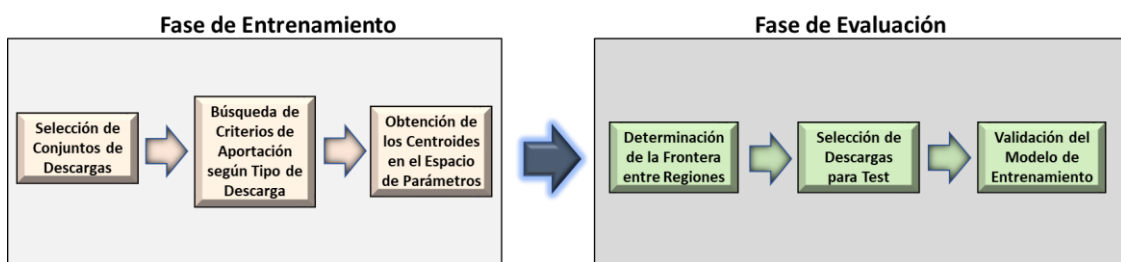


Figura 2.4. Esquema general del método de los centroides.

Se considerarán dos fases, claramente diferenciadas: fase de entrenamiento, y fase de evaluación.

En la fase de entrenamiento, la tarea inicial consistirá en seleccionar dos conjuntos de descargas, cuya naturaleza disruptiva o no disruptiva es conocida. En términos de aprendizaje automático, se va a trabajar con sistemas de clasificación supervisados.

Para cada una de las descargas disruptivas, se analizará su comportamiento en entornos próximos al instante de la disrupción; siendo éste un dato conocido. El comportamiento de cada una de las descargas no disruptivas, será analizado teniendo en cuenta ciertos umbrales que presenta la señal magnética que mide la corriente del plasma. A través del análisis de esos comportamientos en ambos tipos de descargas, se obtendrán los criterios para la contribución de cada una de las descargas a su correspondiente centroide (disruptivo o no disruptivo). De forma que el resultado de la aplicación de dichos criterios para cualquiera que sea el tipo de descarga, debe resumirse en un único punto del espacio de parámetros.

Una vez completado el análisis para los conjuntos de entrenamiento seleccionados, se obtendrán los respectivos centroides; teniendo en cuenta que la contribución a uno u otro centroide, dependerá del tipo de descarga que se trate.

La fase de evaluación comenzará estableciendo la frontera de separación entre regiones, y para ello se tendrán en cuenta los datos ya calculados de los centroides.

Al igual que en la fase de entrenamiento, serán seleccionados dos conjuntos de descargas para evaluar, cuya naturaleza disruptiva o no disruptiva es conocida.

Cada una de las descargas seleccionadas en la fase de evaluación será analizada, teniendo en cuenta la frontera que divide ambas regiones. Además de tener en cuenta la frontera, serán consideradas las variaciones de la señal de amplitud entre las muestras consecutivas que constituyen cada descarga. Si la tasa de aciertos y de falsas alarmas es satisfactoria, quedaría validado el modelo de entrenamiento; pudiendo ser utilizados los centroides obtenidos en la fase de entrenamiento en futuras tareas de clasificación automática.

En los capítulos 5 y 6 se mostrará el procedimiento de los centroides en detalle, presentando también los resultados obtenidos en los distintos dispositivos de fusión estudiados. Los detalles y los resultados más significativos del procedimiento implementado han sido publicados en las siguientes referencias, en las que el doctorando participa como coautor: [\[Vega, 2017a\]](#), [\[Vega, 2017b\]](#), [\[Vega, 2017c\]](#), [\[Vega, 2018a\]](#), [\[Vega, 2018b\]](#), [\[Vega, 2018c\]](#), [\[Vega, 2019\]](#) y [\[Vega, 2020\]](#).

# ENERGÍA DE FUSIÓN Y MARCO DE LA APLICACIÓN

3

## 3.1. Introducción.

### Necesidades energéticas.

Las nuevas demandas en el consumo energético, han contribuido a la búsqueda de nuevas fuentes de energía, que puedan satisfacer esas crecientes necesidades. Una de las recientes demandas, se centra en la sustitución de los combustibles fósiles por otros de naturaleza eléctrica en la industria de la automoción.

Dado que el deuterio (isótopo del hidrógeno) se halla con facilidad en nuestro planeta, desde hace años se viene investigando cómo generar energía mediante fusión (unión de átomos ligeros para formar otros más pesados, con liberación de gran cantidad de energía durante dicho proceso de unión).

### Combustible utilizado para las reacciones de fusión nuclear.

El **Deuterio** y **Tritio** son dos isótopos del hidrógeno, donde:

- El deuterio (D), está formado por un protón y un neutrón. El hidrógeno (H) tiene solamente un protón. El D se trata de un isótopo estable del hidrógeno. Abunda en el agua, existiendo un átomo de D por cada 6.500

átomos de H. Existiendo una concentración de 34 gramos de D (deuterio) por metro cúbico en el agua del mar. El contenido energético de un litro de agua de mar, equivale a la energía que se puede obtener de 250 litros de petróleo. Dada la abundancia de agua de mar en nuestro planeta, el D se consideraría como una fuente de energía inagotable. El D se obtiene por electrolisis del agua.

- El tritio (T). Se trata de un isótopo inestable o radiactivo del hidrógeno, formado por un protón y dos neutrones. El T escasea en la naturaleza, obteniéndose a partir de reacciones con isótopos de litio (Li). El Li si es un elemento que abunda en el agua de mar y en la corteza terrestre, siendo la proporción de T (tritio) a H (hidrógeno) de  $1:10^9$  [URL, 42].

### Fusión nuclear.

Para que se produzca la reacción de fusión, deben darse las siguientes condiciones:

- Temperaturas elevadas, del orden de 150 millones de grados Celsius.
- Suficiente densidad de partículas, para aumentar la probabilidad de colisionar entre ellas.
- Suficiente tiempo de confinamiento del plasma dentro del dispositivo, evitando que éste se expanda.
- Criterio de Lawson, que relaciona la temperatura con la densidad y con el tiempo de confinamiento [URL, 56].

Cuando se alcanzan temperaturas extremas, los gases se convierten en plasma; haciendo que los elementos ligeros puedan fusionarse y producir energía [URL, 43].

En la Figura 3.1. se muestra una de las reacciones del proceso de fusión.

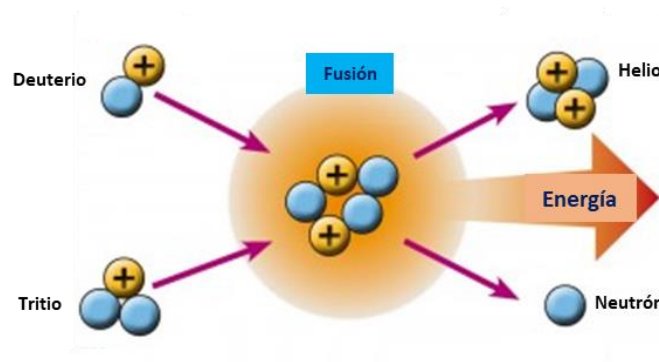
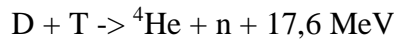


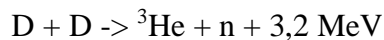
Figura 3.1. Reacción que tiene lugar en el proceso de fusión [URL, 42].

## Reacciones de fusión nuclear.

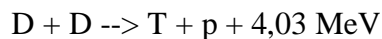
Tal y como se ha descrito anteriormente, los elementos utilizados como combustible para las reacciones de fusión son dos [isótopos](#) de hidrógeno: el Deuterio (D) y el Tritio (T). Las reacciones de fusión más importantes son [\[URL, 42\]](#):



Fusionando un núcleo de Deuterio con un núcleo de Tritio, se obtiene un núcleo de helio formado por dos protones y dos neutrones, liberando 1 [neutrón](#) y 17,6 MeV de energía.



Fusionando dos núcleos de Deuterio, se obtiene un núcleo de helio formado por dos protones y un neutrón, liberando 1 [neutrón](#) y 3,2 MeV de energía.



Fusionando dos núcleos de Deuterio, se obtiene un núcleo de Tritio, un [protón](#) y 4,03 MeV de energía.

Para que dichas reacciones tengan lugar debe suministrarse a los núcleos la [energía cinética](#) necesaria para que se aproximen los núcleos que se van a fusionar, venciendo así las fuerzas de repulsión electrostáticas. Para ello se necesita calentar el gas hasta temperaturas muy elevadas.

El requisito de cualquier reactor de fusión nuclear es confinar dicho plasma con la temperatura y densidad lo bastante elevadas y durante el tiempo justo, a fin de permitir que ocurran suficientes reacciones de fusión nuclear, evitando que se escapen las partículas, para obtener una ganancia neta de energía. Esta ganancia energética depende de que la energía necesaria para calentar y confinar el plasma, sea menor que la energía liberada por las reacciones de fusión nuclear. En principio, por cada miligramo de deuterio-[tritio](#) se pueden obtener 335 MJ [\[URL, 42\]](#).

## Confinamiento para la fusión nuclear.

Los confinamientos convencionales que se utilizan en los [reactores nucleares](#) de fisión no son posibles, debido a las altas temperaturas del plasma que deben soportar. Por esta razón, se han considerado los siguientes métodos de confinamiento [\[URL, 42\]](#):

- FCI: Provoca el choque de las partículas entre sí, creando un medio tan denso como para que dichas partículas no puedan evitar la colisión.

- FCM: Mediante la acción de campos magnéticos, se confinan en un reducido espacio las partículas que forman el plasma. Los dispositivos utilizados para este tipo de confinamiento son el Tokamak y el Stellarator.

## El TOKAMAK.

El Tokamak es un dispositivo experimental que aprovecha la energía de fusión.

La energía producida mediante la fusión de átomos ligeros es absorbida en forma de calor por las paredes del recipiente (dispositivo). Ese calor producirá vapor, y a partir de ese vapor se generará electricidad; mediante generadores y turbinas.

El elemento central del Tokamak lo forma su cámara de vacío (en forma de toroide); donde el combustible gaseoso se convierte en plasma, bajo la influencia de calor y presión extremos. Las partículas del plasma son controladas por las bobinas magnéticas colocadas alrededor de la cámara de vacío. Debido a las altas temperaturas que se alcanzan dentro del vaso de vacío, hay que confinar el plasma caliente lejos de las paredes del vaso de vacío. En caso contrario, podrían llegar a fundirse las paredes del dispositivo [URL, 43]. En la Figura 3.2. se muestra este tipo de dispositivo de fusión.

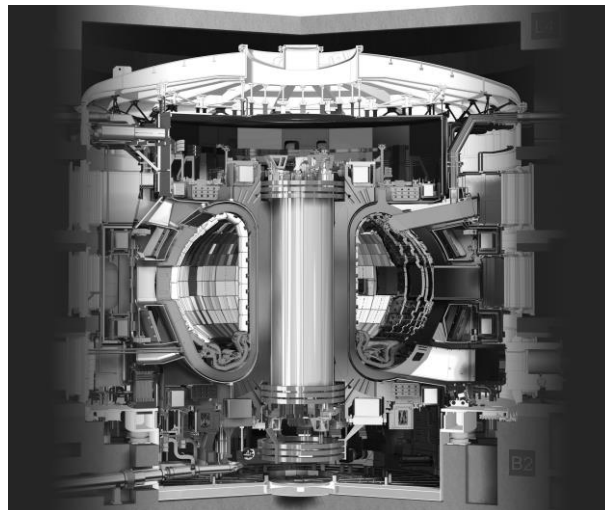


Figura 3.2. TOKAMAK [URL, 43].

La Figura 3.3. ofrece una visión de cada uno de los componentes del dispositivo TOKAMAK de ITER, así como una breve descripción de los mismos.


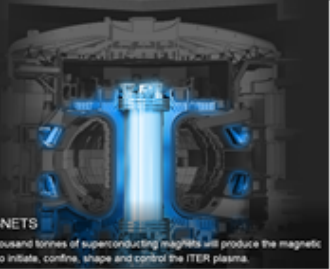
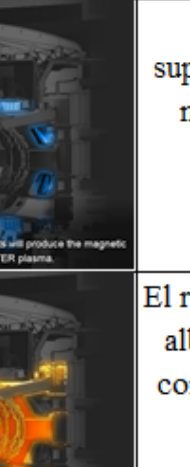
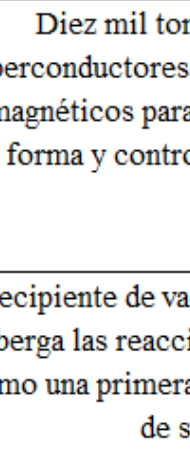
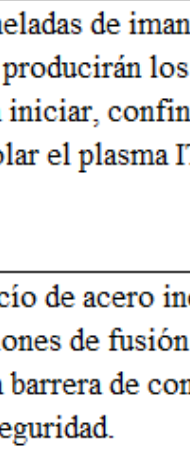
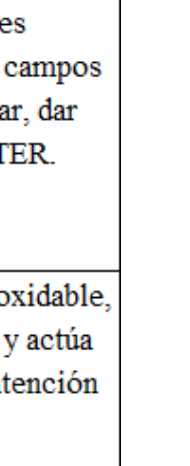
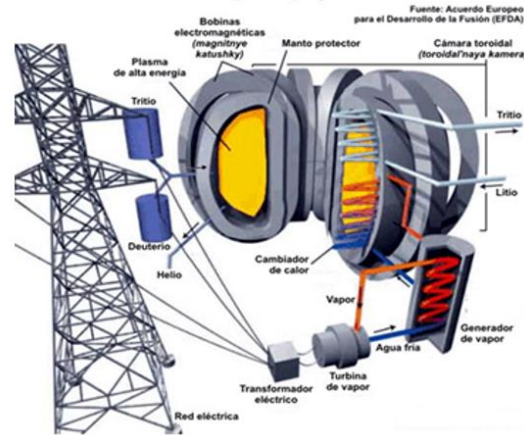
<p><b>A GIANT</b> 23000. Machine weight</p> <p><b>10X THE CORE OF THE SUN</b> 150. Plasma temperature</p> <p><b>FUSION ENERGY</b> 500. MWh</p> <p><b>THE ITER TOKAMAK</b> The tokamak is an experimental machine designed to harness the energy of fusion. ITER will be the world's largest tokamak, with a plasma radius (R) of 6.2 m and a plasma volume of 840 m<sup>3</sup>.</p> 	<p>El tokamak es una máquina experimental diseñada para aprovechar la energía de fusión. ITER será el tokamak más grande del mundo, con un radio de plasma (R) de 6.2 m y un volumen de plasma de 840 m<sup>3</sup>.</p>
<p><b>PACKING A PUNCH</b> 51. GJ Stored magnetic energy</p> <p><b>COLDER THAN PLUTO</b> 4. K Magnet temperature (-269° C)</p> <p><b>10X AROUND THE GLOBE</b> 100000. km MAGNETS Ten thousand tonnes of superconducting magnets will produce the magnetic fields to initiate, confine, shape and control the ITER plasma.</p> 	<p>Diez mil toneladas de imanes superconductores producirán los campos magnéticos para iniciar, confinar, dar forma y controlar el plasma ITER.</p>
<p><b>HEAVIER THAN THE EIFFEL TOWER</b> 8000. t Steel plasma chamber</p> <p><b>LARGEST AMONG TOKAMAKS</b> 840. m<sup>3</sup> Plasma volume</p> <p><b>RECORD RADIUS</b> 6. m VACUUM VESSEL The stainless steel vacuum vessel houses the fusion reactions and acts as a first safety containment barrier.</p> 	<p>El recipiente de vacío de acero inoxidable, alberga las reacciones de fusión y actúa como una primera barrera de contención de seguridad.</p>
<p><b>PROTECTIVE BOUNDARY</b> 440. t Blanket Modules</p> <p><b>1ST ACTIVELY COOLED BLANKET</b> 736. MW Maximal thermal load</p> <p><b>COVERING 600 M<sup>2</sup></b> 180. t BLANKET The blanket shields the steel vacuum vessel and external machine components from high-energy neutrons produced during the fusion reaction.</p> 	<p>El recubrimiento protege el recipiente de vacío de acero y los componentes externos de la máquina, de los neutrones de alta energía producidos durante la reacción de fusión.</p>
<p><b>MODULAR PRECISION</b> 54. Cassettes</p> <p><b>TUNGSTEN COATED</b> 20. MW/m<sup>2</sup> Highest heat load</p> <p><b>EACH CASSETTE</b> 10. t DIVERTOR Positioned at the bottom of the vacuum vessel, the divertor controls the exhaust of waste gas and impurities from the reactor and withstands the highest surface heat loads of the ITER machine.</p> 	<p>Ubicado en el fondo del recipiente de vacío, el desviador controla el escape de gas residual e impurezas del reactor y soporta las cargas de calor superficial más altas de la máquina ITER.</p>
<p><b>CRYOSTAT VACUUM</b> 1. m<sup>3</sup> Times less dense than air</p> <p><b>MASSIVE</b> 3800. t Of steel</p> <p><b>LARGEST STAINLESS STEEL VACUUM CHAMBER IN THE WORLD</b> 16000. m<sup>3</sup> CRYOSTAT The stainless steel cryostat (29 x 29 m) surrounds the vacuum vessel and superconducting magnets and ensures an ultra-cool, vacuum environment.</p> 	<p>El criostato de acero inoxidable (29 x 29 m) rodea el recipiente de vacío y los imanes superconductores y garantiza un entorno de vacío ultrafrío.</p>

Figura 3.3. Componentes del dispositivo TOKAMAK de ITER (en cifras) [URL, 43].

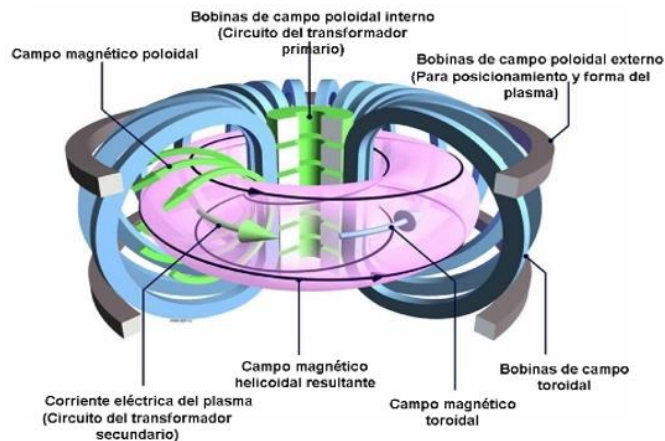
### 3.1. Introducción.

En la Figura 3.4. se muestra el proceso general desde la generación de energía en el dispositivo TOKAMAK, hasta la puesta de dicha energía en la red para su distribución y consumo.



**Figura 3.4.** Esquema general desde la producción de energía hasta su distribución [URL, 44].

En la Figura 3.5., se muestra el modo en el que actúan las bobinas magnéticas, para impedir que el plasma deje de rotar. Además, la acción de los campos magnéticos hará que el plasma guarde cierta distancia con las paredes. Si el plasma dejara de rotar, debido a las altas temperaturas, haría que se fundieran los materiales del dispositivo.



**Figura 3.5.** Esquema de funcionamiento de las bobinas magnéticas en el dispositivo TOKAMAK [URL, 45].

En 1968 se produjo un enorme avance cuando se dio a conocer que, científicos en la Unión Soviética habían logrado generar enormes temperaturas con un reactor y confinar magnéticamente plasma calentado por períodos de tiempos significativamente duraderos. La URSS colocó a la humanidad en el umbral de la primera fusión nuclear controlada.



En 1969 la URSS envió una invitación a un grupo de científicos británicos, los cuales la aceptaron con gran interés, y el Tokamak T-3 fue enviado a Moscú; donde los científicos británicos pudieron confirmar su potencial.

El Tokamak se convirtió entonces en el reactor de fusión más avanzado del mundo y la palabra clave en el estudio de la fusión nuclear con fines civiles. Inmediatamente otros dispositivos Tokamak se comenzarían a construir en varios países.

El Tokamak al igual que sus homólogos, logra generar altas temperaturas mediante un proceso llamado calentamiento óhmico, en el cual corrientes eléctricas crean un campo electromagnético; aumentando la velocidad con que se mueven partículas, lo que se traduce en aumento de temperatura. Este principio es similar al que utilizan los [microondas](#) para calentar las comidas [\[URL, 46\]](#).

### **Fisión Nuclear vs Fusión Nuclear.**

Ambas reacciones liberan la energía almacenada en el núcleo del átomo.

**Fisión nuclear:** [\[URL, 47\]](#) Un núcleo pesado es bombardeado mediante neutrones, convirtiéndolo en inestable; provocando la descomposición en dos núcleos del mismo orden de magnitud en lo que se refiere a sus masas. La suma total de las masas de los nuevos núcleos generados es ligeramente menor a la masa del núcleo pesado, originándose la emisión de dos o tres neutrones y un importante desprendimiento de energía.

Esos dos o tres neutrones generados pueden colisionar a su vez con otros núcleos fisionables, que emitirán nuevos neutrones, y así sucesivamente. De esa manera, se genera una reacción en cadena; liberando una gran cantidad de energía en una pequeña fracción de tiempo.

Para controlar esta reacción de fisión, habría que lograr que sólo uno de los neutrones liberados produzca una fisión posterior, para que sea constante el número de fisiones por unidad de tiempo.

En la Figura 3.6. se puede observar el resultado de la fisión entre un neutrón y un átomo de  $U_{235}$ .

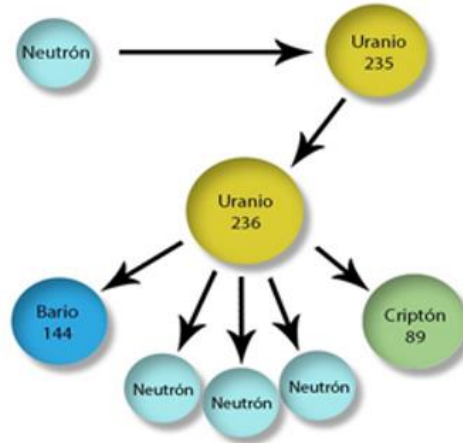


Figura 3.6. El único isótopo fisionable del uranio en la naturaleza es el uranio 235.

Este es el principio en el que se basan los reactores nucleares actuales, para la producción de energía eléctrica.

**Fusión nuclear:** [URL, 47] Consiste en la unión de dos núcleos ligeros, para formar un núcleo más estable y pesado; de masa ligeramente inferior a la suma de los núcleos iniciales. Este tipo de reacción, libera gran cantidad de energía. El origen de la energía producida por el Sol se basa en este tipo de reacción.

Para que la reacción de fusión tenga lugar, los núcleos se aproximan venciendo las fuerzas de repulsión. Dado que en la Tierra no se pueden alcanzar las enormes presiones existentes en el interior del Sol, debe suministrarse energía térmica o utilizar un acelerador de partículas para vencer esas fuerzas de repulsión.

La reacción típica se basa en la unión de dos isótopos del hidrógeno (deuterio y tritio), formando un átomo de helio más un neutrón (véase Figura 3.7.).

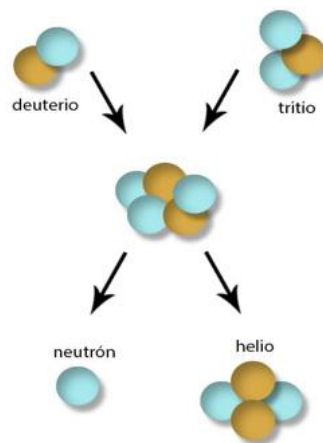


Figura 3.7. Reacción típica de fusión nuclear.

En la reacción de fusión natural producida en el interior del Sol, los átomos de hidrógeno colisionan entre sí, fusionándose a elevadas temperaturas (en torno a 15 millones de grados Celsius). Fusionándose 600 millones de toneladas de hidrógeno cada segundo, y formando helio.

Aunque no existen aún reactores comerciales de fusión, el proyecto experimental ITER tiene como objetivo determinar la viabilidad económica y tecnológica de la fusión nuclear por confinamiento magnético.

### **International Thermonuclear Experimental Reactor (ITER).**

En este proyecto internacional participan siete socios (Unión Europea, Estados Unidos, Federación Rusa, China, India, Japón y Corea del Sur). La Figura 3.8. recoge las nacionalidades de dichos socios. Sus instalaciones se están construyendo en el sureste de Francia, cerca de Marsella.



**Figura 3.8.** Socios participantes en el proyecto ITER.

Se trata del proyecto energético más ambicioso a nivel mundial, participando en la construcción del Tokamak (cámara toroidal con bobinas magnéticas) un total de 35 países. En el diseño de ITER, desde que surgió la idea de un experimento internacional conjunto en 1985, han participado miles de ingenieros y científicos [URL, 43].

El dispositivo ha sido diseñado para [URL, 43]:

1. Producir 500 MW de potencia de fusión. Hasta ahora el récord mundial lo tenía el Tokamak europeo (JET) en 1997, con una producción total de 16 MW a partir de una potencia de calentamiento de entrada de 24 MW; siendo el retorno de producción de  $R = 0,67$ . ITER produciría 500 MW de potencia de

fusión a partir de 50 MW de potencia de calentamiento de entrada, siendo el retorno diez veces mayor ( $R = 10$ ).

2. Los científicos probarán tecnologías como calefacción, control, diagnóstico y mantenimiento remoto.
3. Mantener la reacción de fusión de manera eficiente, para que sea prolongada; manteniendo estable el plasma durante periodos de tiempo más largos.
4. Producir tritio dentro del recipiente de vacío, dado que el suministro de este isótopo del hidrógeno no es suficiente para cubrir las futuras demandas energéticas.
5. Demostrar las características de seguridad de un dispositivo de fusión.

### **ITER – Evolución en el tiempo.**

En diciembre de 2025 estaría programado el encendido y primer plasma de ITER. Los ingenieros y científicos irán ensamblando e integrando los distintos componentes del dispositivo (más de 10 millones de piezas) progresivamente, hasta conseguir que todos los sistemas funcionen juntos.

#### **ITER Timeline**

- 2005** Decisión de ubicar el proyecto en Francia.
- 2006** Firma del Acuerdo ITER.
- 2007** Creación formal de la Organización ITER.
- 2007-2009** Desbroce y nivelación de tierras.
- 2010-2014** Estructura de soporte en tierra y bases para el Tokamak.
- 2012** Hito de la licencia nuclear: ITER se convierte en una instalación nuclear básica según la ley francesa.
- 2014-2021** Construcción del edificio Tokamak (acceso para actividades de montaje en 2019).
- 2010-2021** Construcción de la planta ITER y edificios auxiliares para First Plasma.
- 2008-2021** Fabricación de los principales componentes de First Plasma.
- 2015-2023** Los componentes más grandes se transportan.
- 2020-2025** Montaje principal fase I.

- 2022 Terminación del toro.
- 2024 Cierre de criostato.
- 2024-2025 Fase de puesta en marcha integrada (la puesta en marcha para cada sistema comienza varios años antes).
- Dic 2025 Primer plasma.
- 2026 Comienza la instalación de componentes en el recipiente.
- 2035 Comienza la operación de deuterio-tritio.

A lo largo de la fase de construcción del ITER, el Consejo supervisará de cerca el desempeño de la Organización ITER y las Agencias Nacionales a través de una serie de hitos de proyectos de alto nivel [URL, 43].

## 3.2. JET (Joint European Torus).

CCFE es el laboratorio nacional de investigación de fusión del Reino Unido, siendo propiedad y operado por UKAEA; y ocupando una extensión de 80 hectáreas al sur de Oxford. La imagen de sus instalaciones se puede observar en la Figura 3.9.

CCFE es miembro del consorcio EUROfusion, que cuenta con 30 organizaciones dedicadas a la fusión y universidades de 26 países de la Unión Europea más Suiza y Ucrania, albergando el mayor experimento de fusión magnética del mundo, denominado JET [URL, 45].



Figura 3.9. Imagen aérea del CCFE en Oxfordshire (© CCFE) [URL, 45].

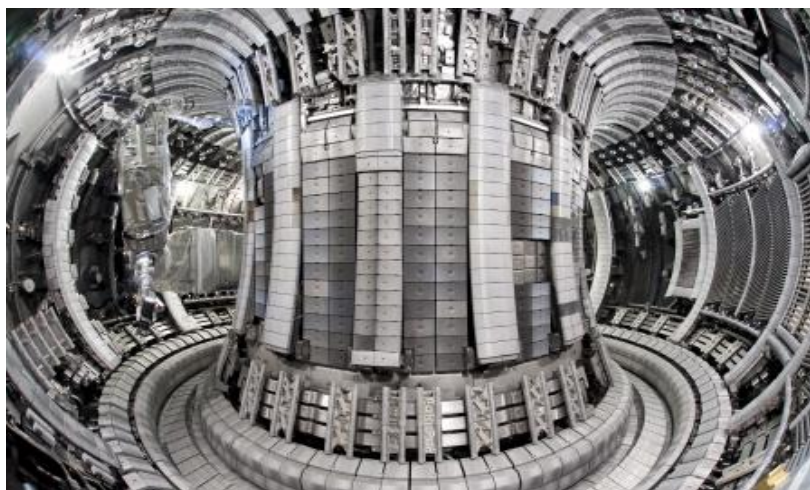
Existe más información disponible en <http://www.culham.org.uk/>.

Para conocer más acerca de los objetivos y planes futuros, dirigirse a [http://www.ccfе.ac.uk/assets/documents/UKAEA Mission & Goals Brochure 2018-19.pdf](http://www.ccfе.ac.uk/assets/documents/UKAEA_Mission_&_Goals_Brochure_2018-19.pdf).

El JET ubicado en el CCFE es el mayor y más poderoso Tokamak que opera actualmente.

Para que la fusión sea útil como fuente de energía, el combustible (deuterio y tritio) deben calentarse a temperaturas superiores a los 100 millones de grados Celsius; convirtiéndose el combustible a dicha temperatura en plasma. Ese plasma es un millón de veces menos denso que el aire, y para evitar que el plasma se contamine y enfríe por contacto con las superficies del dispositivo, está contenido en un sistema de confinamiento magnético.

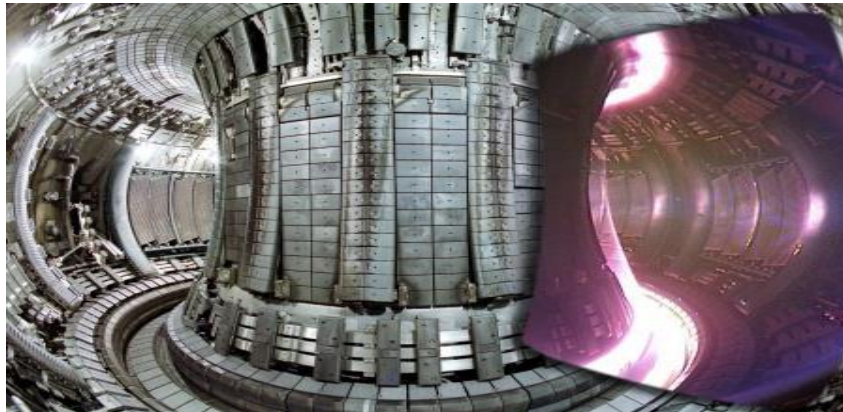
El plasma está contenido en un recipiente al vacío (véase Figura 3.10.), donde el vacío se mantiene gracias a bombas externas. Para generar el plasma se permite la entrada de una pequeña bocanada de gas, que será calentado al hacer circular una corriente a través de él. El plasma caliente queda encerrado por un campo magnético en direcciones toroidal y poloidal, que lo mantiene alejado de las paredes del recipiente de vacío. De esa forma, el campo magnético actúa como una “jaula” magnética, sosteniendo y dando forma al plasma. Para la generación de los campos magnéticos (hasta 4 tesla) y de las corrientes de plasma (5 mega amperios), se utilizan grandes fuentes de alimentación. Para calentar el plasma también se utiliza el calentamiento por radiofrecuencia [URL, 45].



**Figura 3.10.** Vista interna del recipiente, con una pared metálica completa de berilio y tungsteno (© EUROfusion) [URL, 45].

Los 17.6 MeV que se liberan como energía al realizar la reacción de fusión entre el deuterio y tritio para producir helio y neutrones de alta velocidad, representan aproximadamente 10 millones de veces más que la energía que se libera en una reacción química típica [URL, 45].

La imagen de la Figura 3.11. recoge una imagen de la cámara de vacío, donde en su parte derecha aparece el plasma, que es visualizado gracias a una cámara de luz visible.



**Figura 3.11.** Imagen dividida que muestra una vista interior del recipiente de vacío JET, con una imagen superpuesta de un plasma JET real tomada con una cámara de luz visible. Solo se pueden ver los bordes fríos del plasma, ya que el centro está tan caliente que irradia solo en la parte ultravioleta del espectro (© EUROfusion) [URL, 45].

En funcionamiento desde 1983, JET fue explícitamente diseñado para estudiar el comportamiento del plasma en condiciones y dimensiones que se aproximan a las requeridas en un reactor de fusión. Gracias a un diseño original robusto y flexible, la máquina ha podido evolucionar con los avances en la investigación sobre la fusión. Hoy en día, su tarea principal es preparar la construcción y funcionamiento de ITER, actuando como un banco de pruebas para tecnologías ITER y escenarios operativos de plasma [URL, 40].

En la configuración actual los radios mayor y menor del toro de plasma son 3 metros y 0.9 metros respectivamente, y el volumen total de plasma es de 90 metros cúbicos. Un desviador en la parte inferior del recipiente de vacío permite que el calor y el gas de escape se agoten de una manera controlada. Desde 2011 la primera pared del recipiente de vacío se ha hecho de berilio y tungsteno, reflejando las elecciones materiales de ITER. Otras características importantes de JET son [URL, 40]:

- Un sistema de calefacción auxiliar de plasma flexible y potente, que consiste en inyección de haz neutro (34 megavatios), calentamiento por resonancia de

ciclones de ion (10 megavatios) y transmisión de corriente híbrida inferior (7 megavatios).

- Un extenso conjunto de diagnósticos de alrededor de 100 instrumentos individuales, capturando hasta 18 gigabytes de datos brutos por pulso de plasma.
- Un inyector de pellets de alta frecuencia para reabastecimiento de plasma.
- Una válvula de inyección masiva de gas para estudios de interrupción del plasma.
- Capacidades para operar con combustible de tritio, único entre los Tokamaks actuales.
- Instalaciones de manipulación de berilio, que permiten el uso de componentes que se enfrentan al plasma de berilio.
- Instalaciones de manipulación remota que permiten llevar a cabo trabajos avanzados de ingeniería dentro del recipiente de vacío sin necesidad de acceso tripulado.

En la Figura 3.12. se muestran las medidas relativas, que representa el tamaño del recipiente de vacío JET con respecto a una persona.



**Figura 3.12.** Cámara de vacío donde se destacan sus dimensiones (© CCFE) [\[URL, 45\]](#).



### 3.3. JT-60SA.

El 8 de mayo de 2019 se instaló con éxito el solenoide central de 12 metros de altura, siendo filmada la operación por la televisión japonesa [URL, 48]. En la Figura 3.13. se recoge una imagen instantánea de dicho acontecimiento.



Figura 3.13. Instalación del solenoide central de 4 módulos y 100 toneladas de peso [URL, 48].

El experimento, una importante modificación del antiguo Tokamak (JT-60U), está diseñado para abordar los dispositivos de la próxima etapa; que resultarán claves para abordar la física para ITER [URL, 48].

El ensamblaje se completará en marzo de 2020, teniendo previsto el primer plasma para septiembre de ese mismo año [URL, 48].

Los expertos destacan la complementariedad de las máquinas JT-60SA e ITER, que registran superposiciones en áreas que incluyen diseño y prueba de conductores magnéticos, manejo remoto en el recipiente y sistemas de calefacción y accionamiento de corriente. Al entrar en funcionamiento JT-60SA antes de ITER, permitirá la exploración de los regímenes de plasma de alta densidad relevantes para ITER y la optimización de las configuraciones para ITER [URL, 48].

JT-60SA está financiado a través del *Broader Approach Agreement* (conjuntamente entre la Comunidad Europea de Energía Atómica -Euratom- y Japón en 2007, para actividades avanzadas de I + D, que tienen como objetivo complementar el proyecto ITER y acelerar la realización de la energía de fusión), a través de las agencias implementadoras *F4E* y los Institutos Nacionales de Ciencia y Tecnología Cuántica y Radiológica (QST) de Japón [URL, 48].

Se muestra en la Tabla 3.1. la comparativa del dispositivo JT-60SA con respecto a ITER [URL, 49].

**Tabla 3.1.** Datos relativos a JT-60SA en relación con ITER [URL, 49].

	JT-60SA	ITER
Radio mayor (m)	~3.0	~6.2
Radio menor (m)	<1.18	<2.0
Corriente del plasma (MA)	<5.5	<15
Campo toroidal en el eje (T)	<2.3/NbTi	<5.3/Nb <sub>3</sub> Sn
Elongación del plasma	<2.0	<1.8
Triangulación del plasma	<0.5	<0.5
Volumen del plasma (m <sup>3</sup> )	~140	~840
Longitud pulso inductivo (s)	100	>400

### 3.4. DIII-D.

La [DIII-D National Fusion Facility](http://www.ga.com/diii-d) (<http://www.ga.com/diii-d>), operada por General Atomics (GA) para el Departamento de Energía de los EE.UU., es una instalación capaz de llevar a cabo una amplia gama de experimentos para explorar descargas de Tokamak [URL, 50].

En el corazón de la instalación se encuentra el Tokamak DIII-D, que es conocido por su flexibilidad operativa, lo que permite una amplia gama de investigaciones en configuraciones de plasma. El conjunto de diagnósticos DIII-D consta de más de 50 sistemas de diagnóstico, capaces de realizar mediciones definitivas de los parámetros del plasma [URL, 50].

GA ha convertido a DIII-D en una de las instalaciones de Tokamak más productivas del mundo, con una disponibilidad promedio de casi el 80% [URL, 50].

GA también está llevando a cabo un programa de mejora proactiva, diseñado para promover la posición de DIII-D como una instalación de vanguardia para la comunidad de fusión. Este esfuerzo incluye la introducción de avances tecnológicos, nuevas capacidades y la búsqueda de oportunidades para mejorar la eficiencia operativa. Esto asegurará que las contribuciones de DIII-D a la ciencia de la fusión continúen en el futuro [URL, 50].

La misión del grupo de Control DIII-D en GA es desarrollar el conocimiento de control, y las soluciones necesarias para permitir que los Tokamaks operen sin interrupciones con altos niveles de rendimiento. El desarrollo del control de plasma integrado (IPC), un enfoque sistemático para el diseño basado en modelos y la verificación del controlador, ha permitido la aplicación experimental exitosa de algoritmos de control de alta confiabilidad, que requieren un mínimo de tiempo de operación de la máquina para pruebas y ajustes [URL, 51]. En la Figura 3.14. se muestran algunos parámetros relacionados con el estado del plasma.

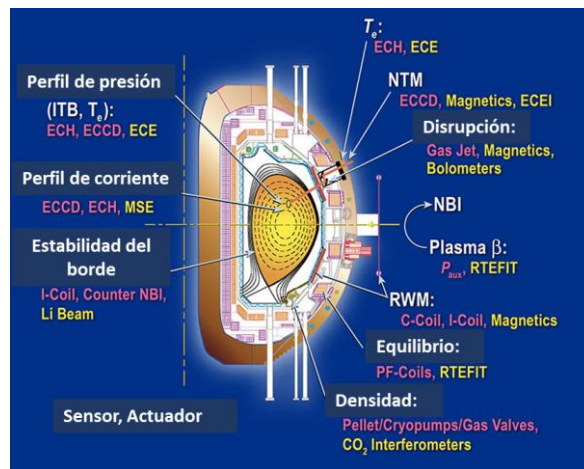


Figura 3.14. Análisis y control del estado del plasma [URL, 51].

El potente y altamente flexible sistema de control de plasma DIII-D es el más utilizado en el mundo, operando actualmente ocho experimentos de fusión en los Estados Unidos, Europa y Asia. En la Figura 3.15. se visualiza el sistema de monitorización para el control del estado del plasma.

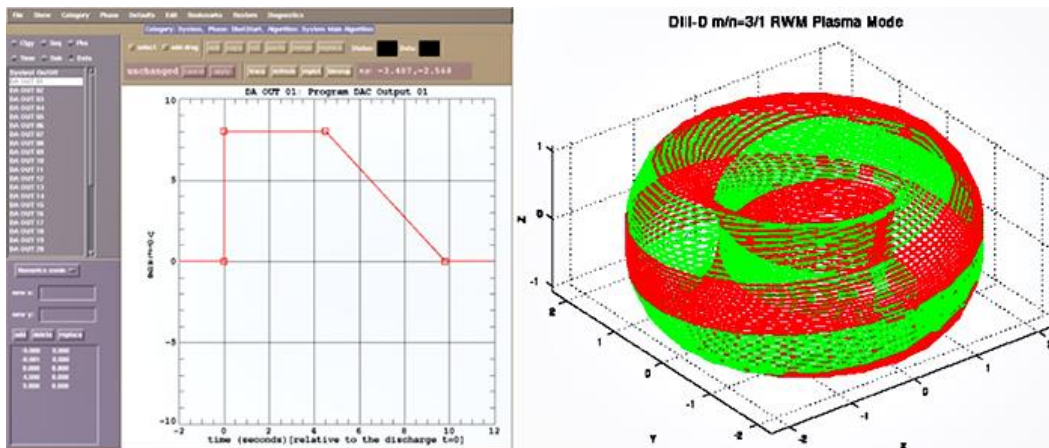


Figura 3.15. Sistema de control del estado del plasma [URL, 51].



# ESTADO DEL ARTE DE LA PREDICCIÓN DE DISRUPCIONES

4

## 4.1. Introducción.

En la sección 4.2. se presentan diferentes trabajos, previos al análisis de esta tesis, y que se hallan relacionados con la predicción de disrupciones. Aunque el objetivo es la predicción del fenómeno de las disrupciones, han sido empleados diferentes métodos. En la sección 4.3. se da una breve descripción del método empleado para este trabajo de investigación.

## 4.2. Evolución histórica de los métodos predictivos de disrupciones.

### 4.2.1. Utilización de redes neuronales.

Las disrupciones en los Tokamaks son eventos de inestabilidad, que pueden dañar los componentes de la máquina. Por tanto, es deseable la elusión y mitigación de las citadas inestabilidades. En ASDEX Upgrade se ha desarrollado una red neuronal, para

predecir la aparición de disrupciones causadas por mecanismos de enfriamiento [URL, 52]. La red neuronal puede predecir la mayoría de las disrupciones (85%), habiendo sido entrenada para predecir el intervalo de tiempo que transcurrirá hasta el instante de la disrupción.

Las máquinas existentes han sido diseñadas para soportar el fenómeno de las disrupciones, ya que uno de sus propósitos es la investigación de límites disruptivos. Sin embargo, la vida útil y limitada de los componentes mecánicos no puede evitar daños en la máquina. Por lo tanto, se recomienda evitar y mitigar las disrupciones en el Tokamak. Un prerrequisito para cualquier acción de elusión (evitación) o mitigación, es un sistema que sea capaz de reconocer una próxima disrupción del plasma. Las disrupciones en ASDEX Upgrade son causadas por mecanismos de enfriamiento [Pautasso, 1998a]. En [Pautasso, 2001] se utiliza la experiencia acumulada durante varios años de operación estudiando las disrupciones, para entrenar una red neuronal, que sea capaz de reconocer una próxima disrupción. Junto con la red neuronal, se ha desarrollado un inyector de pastillas de impurezas sólidas para ASDEX Upgrade. El inyector está configurado para ser activado por red cuando se haya previsto una disrupción inminente. La inyección de gránulos de impureza en el plasma, ya ha demostrado ser un método que reduce las fuerzas mecánicas y las cargas térmicas durante las disrupciones en Tokamaks [NuclFus39, 1999].

Una red neuronal artificial es una función no lineal que mapea un espacio multidimensional  $\{\vec{x}\}$  en otro  $\{\vec{z}\}$ . Para una visión general sobre el tema, ver, por ejemplo, [Bishop]. El entrenamiento consiste en la evaluación de los parámetros, que minimice la diferencia entre la salida prevista  $\{\vec{t}\}$  y la salida que proporcionan los cálculos derivados del entrenamiento de la red neuronal  $\{\vec{z}\}$ .

Ya se ha demostrado que las redes neuronales son adecuadas para la predicción de disrupciones en Tokamaks (DIII-D [Wroblewski, 1997], JET [Milani, 1998], trabajo no publicado realizado con la base de datos COM-PASS-D [Windsor], trabajo exploratorio con disrupciones TEXT [Vannucci, 1999]).

Para el entrenamiento, una red neuronal necesita una gran base de datos de vectores de entrada y salidas asociadas. La base de datos utilizada para entrenar y validar la red de ASDEX Upgrade consta de descargas en el rango 10.000-11.400. La entrada consiste en los historiales de tiempo de varios parámetros relacionados con el plasma. La elección de las variables de entrada fue el resultado de un compromiso entre la física y la

disponibilidad de los datos en tiempo real, ya que la red se utilizará para control en tiempo real de la descarga. Un conjunto preliminar de 30 variables de entrada fue sucesivamente reducido durante el entrenamiento, eliminando las variables que no contribuyeron significativamente a la salida.

El entrenamiento consistió en minimizar la función:

$$E = \sqrt{N^{-1} \sum_{k=1}^N \left[ 1 - \frac{\Delta t_{NN(k)}}{\Delta t_{(k)}} \right]^2} \quad (1)$$

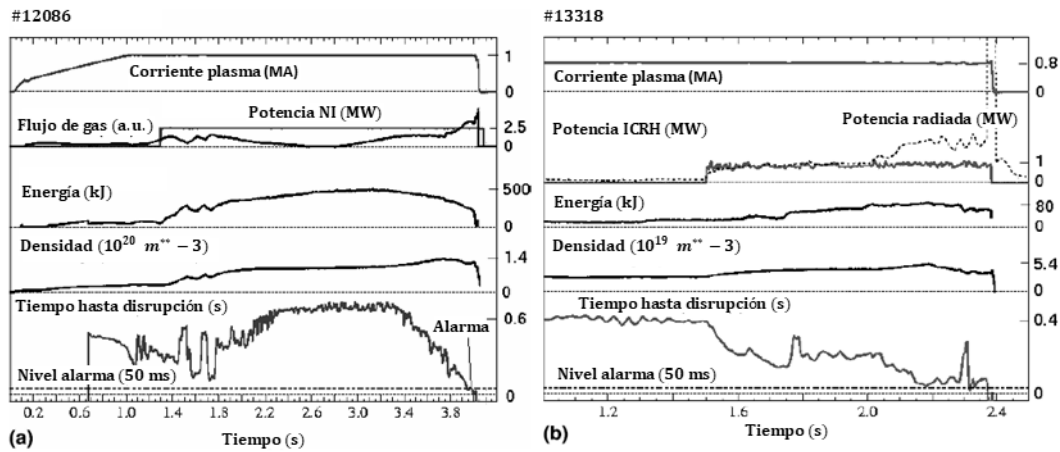
Referido a la expresión (1),  $N$  representa el número de descargas utilizadas en la fase de entrenamiento,  $\Delta t_{(k)}$  representa el instante de la disrupción para la descarga  $k$ , y  $\Delta t_{NN(k)}$  representa el instante en el que la red neuronal predice tal disrupción para dicha descarga  $k$ . El entrenamiento se llevó a cabo con la Toolbox de Matlab que se encarga del tratamiento de redes neuronales y, en particular con una versión de la subrutina que emplea el algoritmo de Levenberg-Marquardt [Demuth, 1998] para la minimización de la expresión (1). Se entrenaron y probaron varias redes con diferentes números de neuronas ocultas. Los datos fueron divididos en un conjunto de entrenamiento (3/4 de los datos) y un conjunto de validación. El entrenamiento de la red neuronal duró típicamente 50-70 ciclos, y se detuvo después de verificar que la función dada por la ecuación (1) había alcanzado un valor mínimo o estacionario.

El rendimiento de varias redes "buenas" se probó fuera de línea con 500 descargas, en el rango de descargas 11.475-12.170, que no se incluyeron en el entrenamiento. La predicción de la disrupción se definió como correcta si era activada en el intervalo de tiempo  $[t_{disr} - 500 \text{ ms}, t_{disr}]$ . Es decir, si la alarma se predecía entre los últimos 500 ms previos al instante de la disrupción. Si se producían alarmas en descargas no disruptivas, o si las alarmas se producían en descargas disruptivas con una antelación superior a los 500 ms, se consideraba una falsa alarma. Si en una descarga disruptiva no se activara la alarma, esa descarga no sería reconocida como disruptiva.

Los resultados del análisis fueron los siguientes:

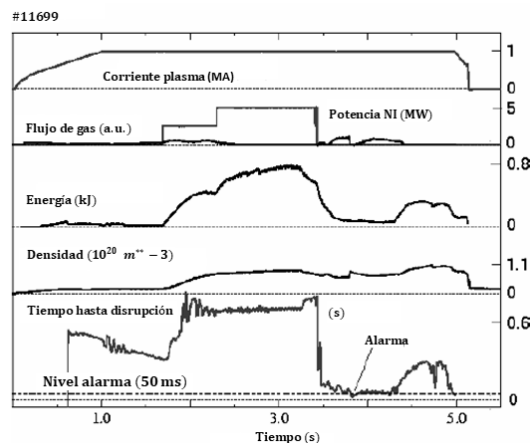
- Se reconoció el 85% de las disrupciones, donde la red no reconoció la disrupción en 10 de las 65 descargas disruptivas (15%).
- La red produjo una o más falsas alarmas en 5 de las 500 descargas (1%).

En la Figura 4.1. se muestran ejemplos de la predicción de la red.



**Figura 4.1.** (a) Predicción de red para una disrupción del límite de densidad; Las condiciones de alarma se alcanzan en  $t = 3.955$  s, 65 ms antes de la disrupción. (b) Predicción de red por una disrupción después de introducir un evento de impureza; las condiciones de alarma se alcanzan en  $t = 2.320$  s, 60 ms antes de la disrupción [Pautasso, 2001].

Las disrupciones que no fueron reconocidas, ocurrieron muy rápido (debido a eventos de impurezas) o mala representación en la base de datos de entrenamiento (acumulación de impurezas o disrupción después de una fase de modo bloqueado en experimentos de 1 TB). Las falsas alarmas fueron causadas por fases previas a la disrupción, de las cuales el plasma se recuperó gracias a una acción tomada por el sistema de retroalimentación (aumentando la potencia de calentamiento o cerrando una válvula de gas) o por plasma cerca de los límites disruptivos, no llegándose a producir la disrupción (ver Figura 4.2.).



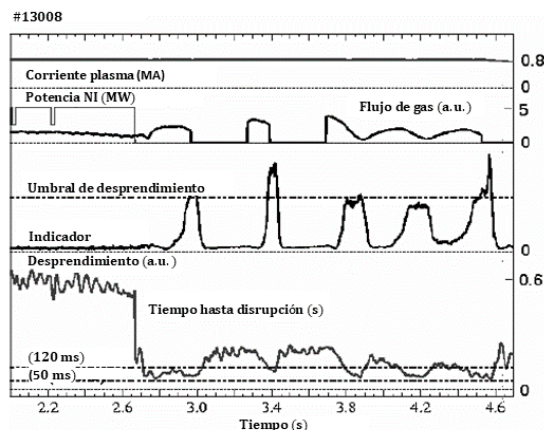
**Figura 4.2.** Falsa alarma en ASDEX Upgrade [Pautasso, 2001].

La red neuronal trabaja en línea, como parte del sistema de control de ASDEX Upgrade. Sus datos de entrada se muestrean, promedian en el tiempo y se normalizan como se hace en el programa fuera de línea. Después de verificar que el sistema en línea



genera el mismo resultado que el programa fuera de línea, la salida de la red se utilizará para evitar o mitigar disrupciones. El rendimiento de la red como sistema para evitar disrupciones, solo puede verificarse usándolo en línea.

La Figura 4.3. muestra un experimento donde, modulando la bocanada de gas, las fases de desprendimiento aparecían periódicamente y donde el algoritmo reaccionó a estas fases cerrando la válvula de gas (en el instante 2.95, 3.4 y 4.55 s). La salida de la red permanece por encima de la alarma de disrupción (configurada en 50 ms) hasta 4.8 s. Desde el instante 4,550 s hasta el instante 4,800 s, existe un intervalo de 250 ms; valor que se encuentra por encima de los 50 ms en los que fue fijada la alarma. También podría haber sido utilizada para controlar la válvula de gas de manera similar, si se hubiera elegido un umbral de evitación de 120 ms. Para la mitigación de disrupciones se prevé el uso de un inyector de gránulos de impurezas. Se ha elegido el límite de 50 ms como alarma de disrupción, para activarlo después de verificar que con esta alarma de umbral el porcentaje de falsas alarmas es aceptablemente bajo; y que el porcentaje de perturbaciones previstas sigue siendo significativo.



**Figura 4.3.** Comparación entre el indicador de desprendimiento (utilizado para controlar una válvula de gas) y la predicción de la red [Pautasso, 2001].

Este valor del umbral para la alarma tiene en cuenta que:

1. El inyector está ahora a 3,5 m del plasma, y el tiempo de retraso entre la descarga y la introducción de impurezas es del orden de 15 ms.
2. La salida de la red debe estar por debajo de 50 ms durante al menos 7,5 ms antes de activar la alarma de disrupción.
3. Se debe permitir un margen de unas pocas decenas de ms para compensar el error que afecta a la predicción de la red.

La inyección de gránulos de impurezas permite una suave terminación del plasma. Aunque la inyección de impurezas mitiga las disrupciones en ASDEX Upgrade y el desarrollo de esta técnica se persigue para proteger a la máquina, su aplicación en ITER no es sencilla. La inyección masiva de gránulos congelados o de un chorro de líquido de deuterio, ha sido propuesta como alternativa para un escenario de apagado ITER. Se desarrolló un nuevo sistema de mitigación de disrupción, una nueva pistola de gránulos para la inyección de gránulos sólidos. El inyector fue diseñado, construido y puesto en funcionamiento en ASDEX Upgrade por el segundo autor de [Pautasso, 2001] en agosto de 1998, y los experimentos se planificaron sobre la base de la experiencia reunida en los años anteriores. Desde entonces, se ha instalado un tubo guía y válvulas rápidas adicionales entre inyector y recipiente, para reducir el gas propulsor que llega al plasma antes del gránulo y ejercer su influencia en el plasma.

Los gránulos utilizados en este experimento estaban hechos de polvo de silicio (Si) o titanio (Ti), mezclado con polietileno fundido (PE) y luego enrollados en forma de cilindro. Fueron obtenidas esferas de 2 mm de diámetro por compresión y calentamiento.

El Si y Ti fueron elegidos por sus "buenas" propiedades radiactivas y su aceptabilidad en la máquina, y el PE fue elegido debido a su energía de sublimación relativamente baja y su composición química aceptable. Estos gránulos se inyectaron con una velocidad en el rango 200-250 m/s; su peso y composición se enumeran en Tabla 4.1.

**Tabla 4.1.** Peso y composición de los gránulos utilizados en ASDEX Upgrade [Pautasso, 2001].

Tipo	Peso (mg)	N <sub>I</sub> (I = Si or Ti)	Gránulos (μm)	N <sub>CH<sub>2</sub></sub>
Si (60%)-PE	7 (4.2 Si)	$0.9 \times 10^{20}$	15-20 (Si)	$1.2 \times 10^{20}$
Ti (72%)-PE	10 (7.2 Ti)	$0.9 \times 10^{20}$	10 (Ti)	$1.2 \times 10^{20}$

Se han llevado a cabo varios experimentos con el inyector, dividiéndose en dos grupos principales:

- (1) Gránulos inyectados en plasmas 'sanos' en un tiempo preprogramado, y
- (2) Gránulos inyectados con el disparador de modo bloqueado en plasmas con modos bloqueados y pre-disrupción.

Grupo (1). Al comienzo, los gránulos se inyectaron a una hora preprogramada para probar el inyector y recopilar datos en una variedad de condiciones plasmáticas. El plasma objetivo no tenía disrupción previa y tenía una energía térmica relativamente alta. Es preferible trabajar con gránulos grandes, que contienen suficientes impurezas, para mitigar la disrupción en el caso de plasmas de alta energía.

Grupo (2). En sucesivas campañas experimentales se inyectaron gránulos utilizando el gatillo de modo bloqueado. Varios de estos gránulos fueron inyectados en presencia de un modo bloqueado y en plasmas con características previas a la disrupción. Algunos de estos gránulos alcanzaron el plasma después del enfriamiento térmico.

[Pautasso, 2001] muestra que es factible un sistema de reconocimiento de disrupciones basado en una red neuronal, siendo relativamente confiable (85% de tasa de aciertos y 1% de falsas alarmas). El cálculo de avance de la salida de una red entrenada es simple y rápido, haciéndolo adecuado para la implementación en tiempo real. Además, la salida elegida de la red será el instante de la disrupción, que es una variable que se puede utilizar para evitar y mitigar disrupciones. Se ha probado exhaustivamente un inyector de gránulos de impurezas con Si o de Ti-PE, demostrándose que funciona de manera confiable y cumple con su propósito de evitar o mitigar disrupciones.

#### 4.2.2. Máquinas de Vectores Soporte.

En los últimos años ha habido un creciente interés en los enfoques de caja negra para la predicción de disrupción. El inconveniente de estos enfoques es que el sistema podría deteriorar su rendimiento. Este podría ser el caso de un predictor de disrupción para JET, donde las nuevas configuraciones de plasma, pueden presentar características completamente diferentes de las observadas en los experimentos utilizados durante la fase de entrenamiento. Esta "novedad" puede ser clasificada incorrectamente por el sistema.

Un método de detección, que determina la novedad de la entrada del sistema de predicción, puede utilizarse para evaluar la fiabilidad del sistema. En [Cannas, 2007a] se presenta un predictor de disrupción de máquinas de vectores soporte para JET, en el que se emiten múltiples señales de diagnóstico de plasma combinadas, para proporcionar un indicador compuesto de anticipación de disrupción inminente. En una máquina de vectores soporte, el análisis de los valores que toma la función de decisión, proporciona información útil sobre la novedad de una entrada y sobre la fiabilidad de la salida del predictor. Los resultados muestran la idoneidad de las máquinas de vectores soporte tanto para las tareas de predicción, como para la detección de novedades en JET.

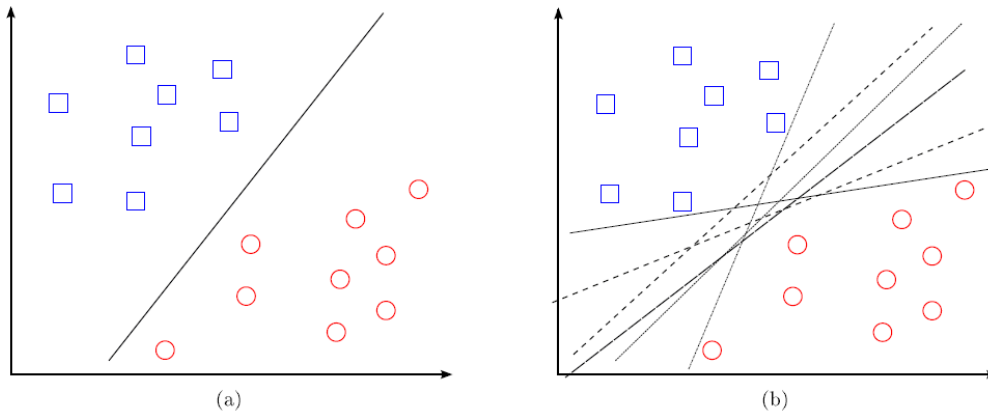
Las disrupciones son endémicas y probablemente inevitables, planteando serios problemas a la integridad y a la vida útil de la máquina. Por esta razón, en los últimos 15 años ha habido varios estudios para la predicción de disrupciones, la mayoría de las cuales utilizan redes neuronales [Cannas, 2002], [Cannas, 2004a] y [Yoshino, 2003].

Uno de los principales inconvenientes del enfoque neuronal, es que el rendimiento de la red normalmente se deteriora cuando se presentan nuevas configuraciones de plasma a la red. De hecho, una red que está entrenada para discriminar entre entradas provenientes de un conjunto de distribuciones, podría producir una salida no confiable cuando los datos de entrada provienen de una distribución completamente nueva. Una posible mejora se puede conseguir mediante la detección de novedades (ND) [Bishop, 1994], [Markou, 2003a] y [Markou, 2003b]. En la aplicación, las novedades detectadas deben usarse para evaluar la confiabilidad de la salida de la red. Es decir, muestras mal representadas en el conjunto de entrenamiento, deben ser descartadas y utilizadas fuera de línea para actualizar automáticamente el predictor de disrupción. En [Cannas, 2005], ha sido implementado un sistema de detección de novedad, basado en un Mapa de Organización Propia, para integrar el predictor de disrupción neuronal propuesto para JET. Los dos sistemas, el sistema de detección de novedades y el predictor presentado, trabajan de forma independiente. Por lo tanto, la identificación de muestras novedosas realizada por el detector de novedad sólo toma en cuenta la información topológica presente en la base de datos, mientras que no tiene relación con el conocimiento adquirido por el predictor. En [Cannas, 2007a], la predicción y la detección de novedades son realizadas por una única máquina de vectores soporte (SVM). Durante el entrenamiento de la SVM, el límite del dominio de datos estimado, está determinado por un subconjunto de puntos de datos de entrenamiento; que está en el límite del dominio estimado, llamado vectores de soporte.

El problema de detección de novedad se formula como un problema de detección de valores atípicos. Muestras que caen fuera del dominio estimado, se declaran como muestras novedosas.

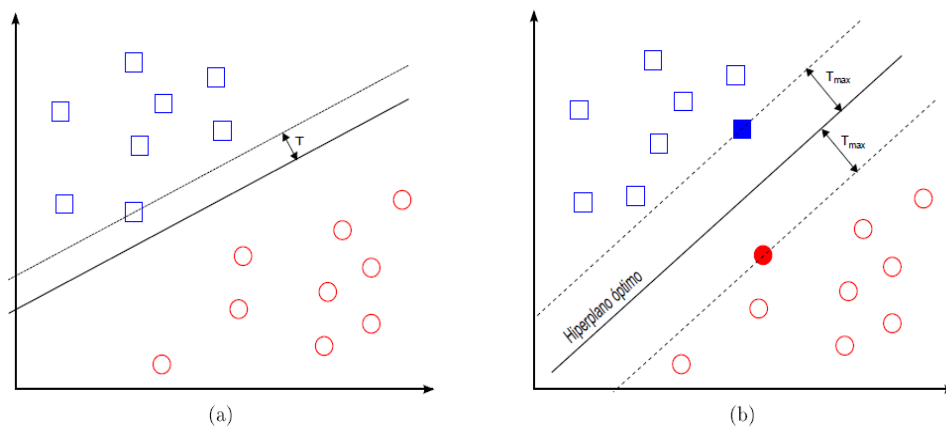
Una SVM es un algoritmo supervisado propuesto por primera vez por [Vapnik, 1998]. El algoritmo aborda el problema general de discriminar entre dos clases de vectores  $n$ -dimensionales. La SVM está construida mediante la detección de un hiperplano de separación óptimo, que maximiza el margen entre sí mismo y el punto más cercano de los datos de entrenamiento. En la Figura 4.4. y Figura 4.5. se ofrece una muestra de ello.

Las máquinas de vectores soporte se utilizan tanto para resolver tareas de clasificación como de regresión [Carmona, 2014].



**Figura 4.4.** Hiperplanos de separación en espacio bidimensional de un conjunto de muestras separables en dos clases: (a) ejemplo de hiperplano de separación (b) otros ejemplos de hiperplanos de separación, de entre los infinitos posibles [Carmona, 2014].

De entre los infinitos hiperplanos posibles en el espacio de características de las dos clases, ha de seleccionarse el hiperplano de separación óptimo.



**Figura 4.5.** Margen de un hiperplano de separación: (a) hiperplano de separación no-óptimo y su margen asociado (no máximo) (b) hiperplano de separación óptimo y su margen asociado (máximo) [Carmona, 2014].

El hiperplano de separación óptimo es aquel que equidista de las muestras más próximas de cada clase. Se puede observar dicho hiperplano de separación óptimo en la Figura 4.5.(b) y Figura 4.6.

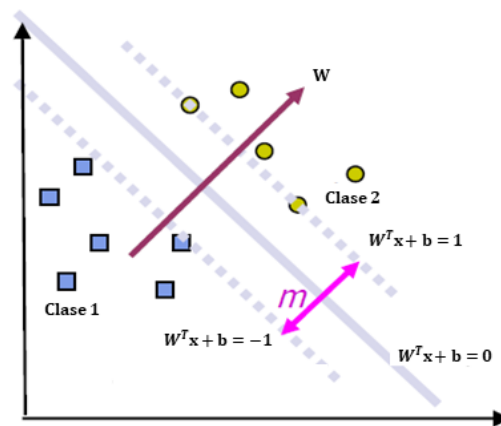


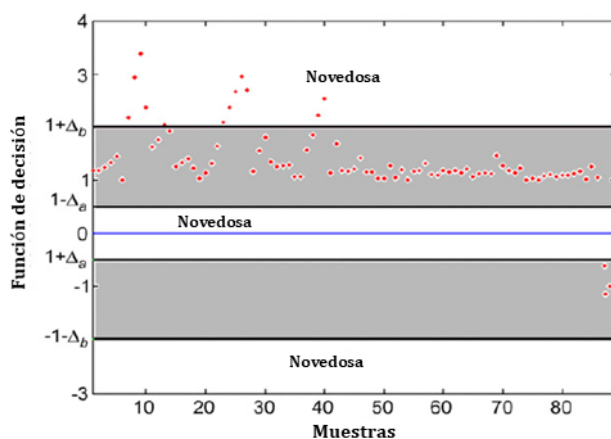
Figura 4.6. La frontera de decisión debe estar tan lejos de los datos de ambas clases como sea posible [Betancourt, 2005].

Actualmente, aplicaciones como el OCR, utilizan técnicas SVM; dada su facilidad para trabajar con imágenes en el espacio de las entradas [Betancourt, 2005].

Las imágenes de los puntos de la entrada siempre se pueden separar linealmente en el espacio de características [Cristianini, 2000].

La detección de novedades consiste en identificar datos desconocidos, que un sistema de aprendizaje automático no conoce durante la fase de entrenamiento. Por lo tanto, la detección de novedades es uno de los requisitos fundamentales de una buena clasificación o sistema de predicción. De hecho, los datos reales pueden contener patrones pertenecientes a regiones operativas no exploradas cuando se desarrolló el sistema de aprendizaje. Este podría ser el caso del predictor de disrupción presentado en [Cannas, 2007a], donde nuevas configuraciones de plasma, podrían presentar características completamente diferentes de las observadas en los experimentos seleccionados para el entrenamiento conjunto. Esta "novedad" puede conducir a un comportamiento incorrecto del predictor SVM. En [Markou, 2003a] y [Markou, 2003b], los autores destacaron que no es posible identificar a priori un solo mejor modelo, dependiendo el éxito de una técnica de detección de novedad, principalmente de propiedades estadísticas de los datos manejados. Ambos métodos, estadísticos y de redes neuronales, pueden ser utilizados para tareas de detección de novedades. En [Cannas, 2007a], la detección de la novedad se realiza mediante un método neuronal. En particular, se analiza el valor de la función de decisión SVM para establecer si una entrada es nueva o no. Como los valores de la función de decisión cercanos a  $+1$  o  $-1$  corresponden a una respuesta correcta del predictor SVM, se pueden introducir dos bandas  $[1 - \Delta_a, 1 + \Delta_b]$  y  $[-1 + \Delta_a, -1 - \Delta_b]$  para

identificar nuevas muestras. En particular, durante la fase de prueba, una muestra se etiquetará como no novedosa si su función de decisión tiene un valor que pertenece a esa banda, y viceversa. Los valores  $\Delta_a$  y  $\Delta_b$  se eligen de tal manera que la función de decisión cae dentro de esas bandas para un determinado porcentaje de las muestras de entrenamiento. Se ha aplicado una aproximación a la optimización, para evitar el ajuste excesivo del conjunto de entrenamiento y, al mismo tiempo, minimice el error en el conjunto de validación; lo que resulta en un valor óptimo del 90%. En la Figura 4.7. se informan los valores de la función de decisión para una prueba. En la Figura 4.7. aparecen las dos bandas resaltadas, donde las muestras que caen fuera de esa banda serán etiquetadas como novedosas por el detector de novedad.



**Figura 4.7.** Función de decisión para un pulso de prueba. Las regiones grises representan las bandas discriminatorias entre muestras novedosas y no novedosas [Cannas, 2007a].

Las señales de diagnóstico para entrenamiento y pruebas de la SVM fueron seleccionadas en el intervalo de pulso 47830–57346, producidas en JET entre marzo de 1999 y octubre de 2002. La base de datos consta de 172 pulsos disruptivos y 102 pulsos no disruptivos. Los pulsos pertenecen a diferentes clases de disrupción, entre las que se encuentran: bloqueo de modo, límite de densidad y potencia de alta radiación. Las descargas incluidas en la base de datos de la red cumplen entre otros, el siguiente requisito: corriente de plasma ( $I_{pla}$ ) > 1.5 MA. De forma que, las descargas con  $I_{pla}$  por debajo de 1.5 MA fueron descartadas; ya que generalmente tienen poco impacto en el acondicionamiento posterior y funcionamiento del dispositivo. Se han seleccionado nueve señales de diagnóstico para describir el régimen plasmático durante la descarga. Estas señales representan la entrada de la SVM. La Tabla 4.2. muestra las señales de diagnóstico seleccionadas. La elección de las señales tiene en cuenta aspectos físicos y la

disponibilidad de datos en tiempo real. El tiempo de muestreo es de 20 ms, para permitir la sincronización entre diferentes sistemas de adquisición [Cannas, 2006a]. Para entrenar, validar y probar el rendimiento de la red [Cristianini, 2000], los pulsos en la base de datos se han dividido en tres conjuntos: el conjunto de entrenamiento, que consta de 69 descargas disruptivas, el conjunto de validación que consta de 17 descargas disruptivas, el conjunto de prueba que consta de 86 descargas disruptivas y 102 descargas no disruptivas. Las descargas no disruptivas se utilizaron solo en fase de prueba.

**Tabla 4.2.** Señales de diagnóstico [Cannas, 2007a].

Nombre de la señal	Unidades
1. Plasma current	[A]
2. Locked mode	[T]
3. Radiated power	[W]
4. Plasma density	[1/m <sup>3</sup> ]
5. Input power	[W]
6. Internal inductance	
7. Safety factor	
8. Poloidal beta	
9. Plasma centroid vertical position	[m]

La estructura del sistema predictivo consta de dos bloques mutuamente conectados: un mapa autoorganizado y una SVM (ver Figura 4.8). Un mapa de organización propia (SOM) [Kohonen, 1982], también conocido como Mapa Auto Organizado de Kohonen, es un tipo de red neuronal artificial, especialmente adecuada para datos encuesta, porque tiene propiedades de visualización prominentes [Vesanto, 1999]. Proyecta el espacio de entrada de alta dimensión en un espacio de baja dimensión (generalmente bidimensional); cuadrícula que se puede utilizar de manera efectiva para visualizar y explorar propiedades de los datos, por ejemplo, la estructura del clúster. Una característica clave de los SOM es que las neuronas en el mapa están dispuestas de tal manera que, las neuronas vecinas representan patrones similares y las neuronas que están bien separadas representan diferentes patrones. Esto significa que los patrones cercanos uno del otro en el espacio de entrada, se asignan a las mismas neuronas o a neuronas vecinas en el espacio de salida. Durante la fase de entrenamiento, un SOM realiza un procedimiento de agrupación y una SVM está entrenada para lanzar la alarma en caso de disrupción inminente. Cuando se completa el procedimiento del entrenamiento, se calculan las bandas introducidas anteriormente, que discriminan entre muestra novedosa y no novedosa.



Para obtener el conjunto de entrenamiento para el predictor SVM, se han construido 86 SOM, uno para cada descarga en los conjuntos de entrenamiento y validación. Cada SOM es utilizado para identificar la fase precursora de la correspondiente disrupción, es decir, para discriminar entre muestras "seguras" y muestras "disruptivas"; que contienen información sobre la proximidad de la disrupción.

Para cada descarga, todas las muestras que pertenecen al grupo que contiene la muestra en el instante de tiempo ( $t_d - 40$  ms), que seguramente se trata de una muestra disruptiva, se considera descarga disruptiva.

Además, se identifica una región ambigua del mapa, que consiste en los grupos que rodean al grupo de muestras disruptivas. Las muestras que pertenecen a esa región se consideran muestras de transición, y son excluidas de la fase de entrenamiento; ya que no pueden ser clasificadas con suficiente confianza, ya sea como disruptivas o bien como muestras no disruptivas. Por el contrario, esas muestras podrían clasificarse correctamente por la SVM, una vez que esté entrenada.

La SVM está entrenada para clasificar muestras no disruptivas y disruptivas, etiquetadas respectivamente como +1 y -1. Durante la aplicación en línea, la SVM se alimenta con todas las muestras de una descarga, y para cada una de ellas, devuelve una etiqueta igual a -1 o +1. En [Cannas, 2007a], el rendimiento del sistema de predicción se evalúa en términos de porcentaje de falsas alarmas (PFA) y porcentaje de alarmas perdidas (PMA), donde PFA se define como la relación entre el número de descargas no disruptivas predichas por el sistema como descargas disruptivas, y el número total de descargas no disruptivas en porcentaje. Mientras que PMA se define como la relación entre el número de descargas disruptivas predichas como descargas no disruptivas y el número de descargas disruptivas en porcentaje. Además, para descargas disruptivas, el porcentaje de alarmas prematuras (PPA) se define como la relación entre la cantidad de alarmas prematuras y cantidad de descargas disruptivas en porcentaje. Una predicción de disrupción es prematura si el sistema además de predecirla, lo hace con mucha antelación con respecto al instante de la disrupción. En JET se considera la predicción prematura si el sistema activa la alarma al menos 1 s antes del instante de disrupción. Por lo tanto, una predicción de disrupción se considera exitosa si el sistema puede predecir correctamente la disrupción no más de 1 s, y no menos de 100 ms, antes del instante de disrupción. Finalmente, la tasa de éxito de predicción (PSR) se define como la predicción correcta tanto en descargas disruptivas como en no disruptivas. Además, el valor de la función de decisión es utilizado para evaluar la confiabilidad de la salida SVM: la salida de la SVM

está relacionada con una muestra de prueba, que presenta el valor de la función de decisión. Si el valor de la función cae fuera de las dos bandas  $[1 - \Delta_a, 1 + \Delta_b]$  y  $[-1 + \Delta_a, -1 - \Delta_b]$ , tiene que ser rechazado. Se reconoce que esa muestra es nueva, y puede ser utilizada para actualizar el predictor de interrupción.

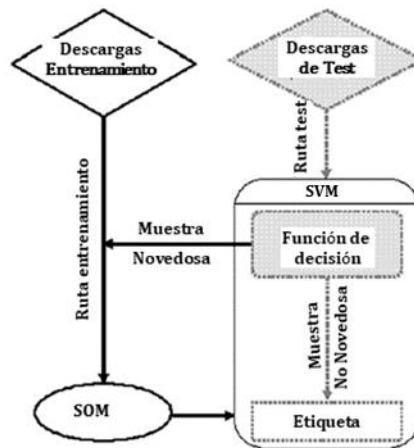


Figura 4.8. Arquitectura de la ruta del predictor / novedad; ruta de prueba [Cannas, 2007a].

La SVM ha sido entrenada usando la Toolbox Matlab Clasificador OSU SVM (ver. 3.00) basada en la versión 2.33 de LIBSVM [Chang, 2001]. Se han realizado varias sesiones de entrenamiento para seleccionar los parámetros SVM. Los mejores resultados se han obtenido para el Radial Basis Function Kernel con valor  $\sigma = 18.7$  y penalización de error  $C = 40$ .

La Tabla 4.3. muestra el rendimiento del predictor SVM para el entrenamiento, la validación y la prueba conjunta, respectivamente en términos de PFA, PMA, PPA y PSR; donde para cada descarga del entrenamiento y conjuntos de validación, los errores se han calculado considerando las secuencias completas de las muestras en lugar de las muestras seleccionadas por el SOM durante la fase de entrenamiento. El rendimiento de la red es excelente en términos de falsas alarmas, y muy bueno en términos de PMA, y también si se compara con los datos reportados en [Cannas, 2005]. En particular, el sistema puede predecir correctamente 63 sobre 86 descargas disruptivas del conjunto de prueba; incluso si 9 de ellas son predichas de forma prematura. De todos modos, la capacidad de la SVM para predecir la interrupción 100 ms antes la ocurrencia del fenómeno es muy prometedora para poder aplicar procedimientos de mitigación. Cuando el predictor está integrado con el bloque de detección de novedades, el rendimiento se modifica de la siguiente manera: 14 de las 23 alarmas perdidas son etiquetadas como novedosas por el detector de

novedades (ND), 5 de las 9 descargas disruptivas que activan una alarma prematura están etiquetadas como novedosas por el ND, 18 de las 63 descargas disruptivas predichas correctamente por el predictor están etiquetados como novedosas por el ND.

**Tabla 4.3.** Rendimiento del predictor SVM en términos de porcentaje de falsas alarmas (PFA), porcentaje de alarmas perdidas (PMA), porcentaje de alarmas prematuras (PPA) y tasa de éxito en la predicción (PSR) [Cannas, 2007a].

	Conj. Entrenamiento (%)	Conj. Validación (%)	Conj. Test (%)
PMA	0	24	27
PFA	–	–	0
PPA	0	0	10
PSR	100	76	83

La Tabla 4.4. muestra una comparación de PMA, PFA, PPA y PSR obtenido por el predictor SVM y por el predictor SVM con ND. En el segundo caso las actuaciones se calculan en un conjunto de prueba reducido, obtenido por el conjunto de prueba anterior, descartando las descargas etiquetadas como novedosas por el detector de novedades. Una muestra se considera novedosa si  $1 - \Delta_a < |f(\underline{x})| < 1 + \Delta_b$  donde  $\Delta_a = 0.5$  y  $\Delta_b = 1$ .

Como se puede observar, el número de falsas alarmas disminuye considerablemente, el PPA se reduce, y el éxito de la tasa de predicción calculada en todo el conjunto de prueba va desde 83% al 91%.

**Tabla 4.4.** SVM Predictor y detector predictor / novedad. Rendimiento en términos de porcentaje de falsas alarmas (PFA), porcentaje de alarmas perdidas (PMA), porcentaje de alarmas prematuras (PPA) y tasa de acierto en la predicción (PSR) [Cannas, 2007a].

	Conjunto de test para el predictor SVM (%)	Conjunto de test para el detector predictor/novedades SVM (%)
PMA	27 (23/86)	18 (9/49)
PFA	0 (0/102)	0 (0/102)
PPA	10 (9/86)	8 (4/49)
PSR	83 (156/188)	91 (138/151)

Resumiendo, el predictor con detector de novedad alcanza mejores rendimientos en comparación con respecto al simple predictor sin detector de novedades. Ello se debe principalmente porque el detector de novedad puede justificar muchas de las alarmas perdidas del predictor, ya que se reconoce que pertenecen a regiones inexploradas del espacio operacional. El detector de novedad influye en el comportamiento del predictor sólo en caso de alarma de disrupción. En particular, si la red activa la alarma para una muestra considerada "novedosa", la alarma será rechazada. Incluso si la alarma es

rechazada por el ND, para una operación no disruptiva de la máquina, la descarga debe terminarse; especialmente en condiciones donde las disrupciones no pueden ser toleradas.

El envejecimiento de un sistema de predicción neuronal es un problema importante para máquinas experimentales como JET, debiendo ser evitado. Para contrarrestar o evitar este problema, es crucial tener un sistema capaz de medir la fiabilidad de la salida de la red, e identificar configuraciones de entrada que sean novedosas. Las muestras que tienen etiquetado como 'novedosas' podrían agregarse a la base de datos original y ser utilizadas para actualizar automáticamente la red. En [Cannas, 2007a] se propone una SVM, que integra capacidades de predicción y detección de novedades en un mismo sistema. Resultando prometedora dicha técnica para mejorar la fiabilidad del predictor. El sistema propuesto muestra un porcentaje nulo de falsas alarmas, un porcentaje de alarmas perdidas igual al 18%, y un porcentaje de predicción correcta igual al 91%. Usando el conocimiento adquirido durante la fase de entrenamiento del predictor, el sistema puede detectar las novedades de nuevas muestras en las descargas, aumentando el rendimiento y la robustez del sistema completo. En particular, el detector de novedades es capaz de justificar muchas de las alarmas perdidas del predictor, al reconocer que pertenecen a regiones inexploradas del espacio operacional.

Ha de tenerse en cuenta que, el aumento del rendimiento conduce a una reducción de la capacidad de discriminación, porque algunas descargas disruptivas correctamente predichas por la SVM, están etiquetadas como novedosas por el ND.

#### 4.2.3. Predictor para JET simulando entornos en tiempo real.

En [Rattá, 2010], se presenta un predictor en tiempo real específicamente desarrollado para JET y basado en máquinas de vectores soporte; siendo su principal objetivo, obtener altas tasas de reconocimiento en un entorno simulado en tiempo real. Para este fin, el predictor ha sido probado en los intervalos de tiempo de descargas completas, exactamente como en la operación del mundo real. Desde el año 2000, los experimentos en JET se han organizado en campañas, que comienzan secuencialmente con la campaña C1. En [Rattá, 2010] se informa de los resultados de la campaña C1 (año 2000) hasta C19 (año 2007). El predictor ha sido entrenado con datos de las campañas de JET hasta C7, con especial atención en la reducción de la cantidad de alarmas perdidas, que son inferiores al 1%, para un conjunto de prueba de descargas de las mismas campañas (C1 a C7) utilizadas para el entrenamiento. Las falsas alarmas más las alarmas

prematuras son del orden del 6,4%, obteniéndose una tasa de aciertos total de más del 92%. La robustez del predictor se ha demostrado probándolo con un amplio subconjunto de descargas de las campañas más recientes (de C8 a C19), sin ningún tipo de reentrenamiento. La tasa de éxito durante el período entre C8 y C14 obtiene como promedio el 88%, y en ningún caso cae por debajo del 82%; lo que confirma las buenas capacidades de generalización de la técnica desarrollada. Después de C14, se implementaron modificaciones significativas en JET y sus diagnósticos y, en consecuencia, el éxito las tasas del predictor entre C15 y C19 decaen a un promedio de 79%. Finalmente, el rendimiento del sistema de detección desarrollado se ha comparado con las predicciones del sistema de protección JET (JPS). El nuevo predictor supera claramente a JPS en el tiempo de anticipación a la disrupción, hasta en aproximadamente 180 ms.

Se ha extraído un gran conjunto de datos de la base de datos de JET, utilizando en [Rattá, 2010] descargas en el rango de 42815 a 70722. Ha sido utilizado un subconjunto de la base de datos para construir el predictor y probarlo. En la Tabla 4.5. se indica la parte de la base de datos utilizada para entrenamiento y pruebas, con selección aleatoria de descargas hasta la campaña C7. Una vez que el sistema ha sido entrenado y probado, se ha demostrado su robustez con un gran número de descargas desde C1 (año 2000) hasta C19 (año 2007). La calidad de las predicciones finalmente se compara con los datos proporcionados por el JPS (descargas de C1 a C19). Asimismo, la cantidad de descargas disruptivas y no disruptivas utilizadas en las etapas de prueba, se detallan en el resto de la Tabla 4.5.

**Tabla 4.5.** Lista de la base de datos según campañas y número total de descargas utilizadas para el análisis [Rattá, 2010].

Descargas (Campañas)	Número de descargas
Disruptivas para entrenamiento (C7 y anteriores)	263
No disruptivas para entrenamiento (C7 y anteriores)	175
Disruptivas para test (C7 y anteriores)	66
No disruptivas para test (C7 y anteriores)	44
Disruptivas para test (C1 a C19)	1245
No disruptivas para test (C8 a C19)	331
Total disruptivas (entrenamiento y test)	1574
Total no disruptivas (entrenamiento y test)	550
Total descargas (entrenamiento y test)	2124

En JET, se adquieren miles de señales en cada descarga. Resulta fundamental seleccionar aquellas señales que representan las cantidades físicas más representativas,

para identificar adecuadamente una disrupción. Por un lado, demasiadas señales podrían sobrecargar la capacidad de aprendizaje de un sistema automático. Por otro lado, con un número escaso de señales no se podría proporcionar suficiente información para realizar un trabajo confiable de predicción. Para estudiar las disrupciones, [Rattá, 2010] ha basado su análisis en trece señales para entrenar y probar el clasificador (ver Tabla 4.6.).

**Tabla 4.6.** Lista de las señales analizadas en cada descarga [Rattá, 2010].

Nombre de la señal	Unidades
(1) Plasma current	A
(2) Poloidal beta	
(3) Poloidal beta time derivative	s <sup>-1</sup>
(4) Mode lock amplitude	T
(5) Safety factor at 95% of minor radius	
(6) Safety factor at 95% of minor radius time derivative	s <sup>-1</sup>
(7) Total input power	W
(8) Plasma internal inductance	
(9) Plasma internal inductance time derivative	s <sup>-1</sup>
(10) Plasma vertical centroid position	m
(11) Plasma density	m <sup>-3</sup>
(12) Stored diamagnetic energy time derivative	W
(13) Net power (total input power minus total radiated power)	W

Los resultados mostrados en un trabajo anterior [Rattá, 2008] demostraron que, con esta selección de señales, se pueden obtener buenos resultados. Esas mismas señales también han sido utilizadas en trabajos anteriores sobre disrupciones, como en [Murari, 2008] o [Cannas, 2007b]. Para cada una de dichas señales se ha aplicado un algoritmo de interpolación, estandarizando la frecuencia de muestreo a 1 kHz. El porcentaje de descargas que ha sido descartado ha sido aproximadamente del 8.5%, principalmente debido a la falta de estimaciones confiables de la señal potencia radiada total. El sistema de predicción desarrollado, activará una alarma para detener la descarga si se pierde una señal durante la operación en tiempo real. Otro problema potencial es la diferencia de amplitud extrema entre algunas señales, en muchos casos de varios órdenes de magnitud. Este aspecto es un factor de complicación para cualquier sistema de clasificación. El sistema podría automáticamente asignar pesos más altos a ciertas señales, debido a sus valores absolutos y no por su relevancia para la predicción. En consecuencia, ha sido aplicada una fórmula de normalización estándar (I):

$$\text{Señal Normalizada} = \frac{\text{Señal} - \text{Mín}}{\text{Máx} - \text{Mín}} \quad (\text{I})$$

Donde *Máx* y *Mín*, representan respectivamente, el valor máximo y valor mínimo de cada señal, en el conjunto de entrenamiento. Esta normalización asigna a la señal

valores dentro del intervalo  $[0,1]$ , preservando la información concerniente a sus respectivas magnitudes.

Para demostrar que un predictor proporciona rendimiento suficiente y confiable, es necesario entrenarlo y probarlo usando un número considerable de descargas; disponiendo de una gran base de datos de señales almacenadas. Para evaluar adecuadamente el potencial del predictor para aplicaciones de control, el predictor ha sido diseñado para usar los datos secuencialmente; como si se convirtieran a medida que las muestras van estando disponibles durante un experimento real. Para ello, se consideran ventanas de tiempo de 30 ms para cada uno de los parámetros que representa el comportamiento del plasma. Se procesa cada ventana de tiempo, condensando la información disruptiva en vectores característicos. Cada vector de características contiene 13 valores (uno por señal). Para cada señal, el procedimiento de extracción de características consiste en el cálculo de la transformada rápida de Fourier. La parte positiva del espectro es retenido, descartando el primer componente (off-set); y calculándose la desviación estándar del espectro retenido, obteniéndose un valor por cada señal. Finalmente, los 13 valores se concatenan para obtener el vector de características. Ha sido elegido este procedimiento de extracción de características, ya que proporciona el mejor rendimiento, tal y como se detalla en [Rattá, 2008]. Usando estos vectores de características, la acción final de tomar una decisión se basa en activar o no una alarma; tal y como si los datos estuvieran siendo adquiridos y procesados en tiempo real.

Se han entrenado varios clasificadores basados en la SVM para identificar las futuras disrupciones. El entrenamiento ha sido realizado proporcionando a los sistemas de aprendizaje dos clases de entradas: características pertenecientes a descargas disruptivas y características de descargas no disruptivas. Para ello, se han utilizado de forma aleatoria 263 descargas disruptivas y 175 descargas no disruptivas. La notación utilizada para identificar los modelos puede ser resumida de la siguiente manera:

$M(i)$  es el modelo entrenado para un intervalo de tiempo específico  $i$  antes de la disrupción. donde  $i = 1, 2, 3, \dots, 8$ .

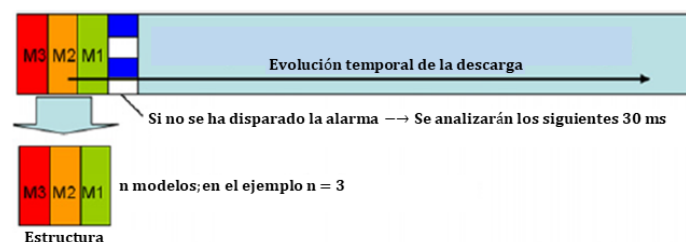
Donde  $M(1)$  corresponde al período  $[-60, -30]$  ms antes de la disrupción.

El tiempo mínimo para realizar acciones de mitigación en JET es de 30 ms [de Vries, 2009]. El sistema de detección debe ser entrenado para reconocer disrupciones al menos 30 ms antes de su ocurrencia. En consecuencia, el período más cercano a la disrupción  $[-30, 0]$  ha sido descartado para el entrenamiento. Los otros  $M(i)$  se refieren

a los intervalos  $[-30(i + 1), -30i]$  ms antes de la disrupción. Las entradas para el entrenamiento de cada modelo  $M(i)$ , han sido las siguientes:

- Como características disruptivas: ventanas de tiempo  $[-30(i + 1), -30i]$  ms antes de la disrupción, para todas las descargas disruptivas que intervienen en el conjunto de datos de entrenamiento.
- Como características no disruptivas:
  - Todas las ventanas de tiempo desde el comienzo de la descarga hasta 1 s antes del instante de la disrupción, para todas las descargas que intervienen en el conjunto de datos de entrenamiento disruptivo. Teniendo en cuenta [Rattá, 2008], aun tratándose de descargas disruptivas, los datos muestreados 1 s antes de la disrupción no llegan a mostrar de forma clara comportamientos anormales; pudiendo considerarse su comportamiento como no disruptivo. Además, si las alarmas se dispararan 1 s antes del instante de la disrupción, podría considerarse como "prematura", interrumpiendo la continuidad de la operación [Cannas, 2007a].
  - Todas las ventanas de tiempo de las descargas no disruptivas disponibles en el conjunto de datos de entrenamiento no disruptivo.

Para optimizar las tasas de aciertos, se han considerado diferentes combinaciones de los clasificadores entrenados. La elección de implementar más de un clasificador, se debe a la necesidad de tomar en cuenta la evolución temporal de la descarga. Con este fin, secuencias de clasificadores consecutivos optimizados para varios intervalos de tiempo de 30 ms antes de las disrupciones, han demostrado lograr un mayor rendimiento que un solo modelo entrenado durante un período más largo. Los clasificadores están destinados a funcionar en paralelo, en ventanas de tiempo consecutivas como se muestra en la Figura 4.9.



**Figura 4.9.** Esquema del análisis en tiempo real de una descarga, utilizando una serie de clasificadores que analizan en ventanas consecutivas de tiempo de forma paralela [Rattá, 2010].



En el ejemplo de la Figura 4.9. se muestran  $n$  predicciones, proporcionadas por la concatenación de  $n$  modelos; siendo  $n = 3$ .

Llegados a este punto, la cuestión a plantearse sería ¿cuál debería ser el número óptimo de clasificadores a utilizar en paralelo? Debido a la novedad de este enfoque y siendo  $n$  un parámetro de valor desconocido, se han probado diferentes combinaciones. Habiéndose desarrollado ocho clasificadores. En los clasificadores más alejados del instante de la disrupción (M8, M7, M6, ...), los rendimientos son muy bajos; debido a que la información contenida es insuficiente, por encontrarse muy alejadas esas muestras del instante de la disrupción. Con estos modelos, se incluyen siete secuencias de clasificadores consecutivos que han sido creadas concatenándolos de la siguiente manera:

Secuencia 1: M2, M1.

Secuencia 2: M3, M2, M1.

Secuencia 3: M4, M3, M2, M1.

Secuencia 4: M5, M4, M3, M2, M1.

Secuencia 5: M6, M5, M4, M3, M2, M1.

Secuencia 6: M7, M6, M5, M4, M3, M2, M1.

Secuencia 7: M8, M7, M6, M5, M4, M3, M2, M1.

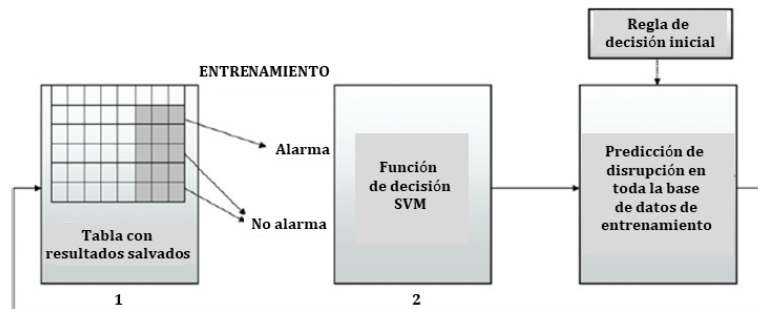
Teniendo en cuenta, tal y como se mostrará más adelante, que la secuencia 2 es la secuencia óptima en cuanto a resultados:

- El primer clasificador de la secuencia (M3) analiza los primeros 30 ms de cada descarga, el segundo clasificador (M2) los segundos 30 ms y el tercero (M1) los terceros 30 ms (ver de nuevo Figura 4.9.).
- Las tres salidas de la secuencia (una por cada clasificador incluido en la secuencia) se analizan con una función de decisión (DF), que consiste en otro clasificador SVM (Máquina de Vectores Soporte). Esta DF determina si la combinación de salidas proporcionadas por la secuencia, indica la proximidad o no de una disrupción.
  - Si se aprecia un comportamiento disruptivo, se activa de forma inmediata una alarma, almacenándose los datos en un archivo similar a la Tabla 4.7. y a la Figura 4.10., para futuros análisis.
  - Si no se reconoce ningún comportamiento disruptivo, no se activa alarma alguna, tampoco se almacenan datos; y se pasan a analizar los siguientes 30 ms de la descarga. De esta forma, el modelo M3 analizará la ventana de tiempo previamente analizada por el modelo M2, el

modelo M2 analizará la ventana de tiempo previamente analizada por el modelo M1, y el modelo M1 analizará los próximos 30 ms de las muestras que componen la descarga.

**Tabla 4.7.** Ejemplos de los seis casos posibles de clasificaciones correctas e incorrectas que pueden ocurrir en la práctica [Rattá, 2010].

Nº	Descarga	Instante alarma (s)	Instante disrupción (s)	Margen (ms)	Salida de M3	Salida de M2	Salida de M1
1	56658	23.911	24.021	110	-0.910	0.448	0.0852
2(a)	54827	0	13.866	0	0	0	0
2(b)	55253	23.101	22.979	-122	-1.800	-1.391	-1.154
3	53740	5.881	10.22	4339	-0.896	-1	0.445
4	56782	0	0	0	0	0	0
5	52641	0.916	0	-916	-1.176	-1.202	-1.008



**Figura 4.10.** Esquema general de entrenamiento de la DF [Rattá, 2010].

1. Se almacenan los datos relevantes.
2. Se entrena una nueva DF que reemplaza a la anterior.

Los  $n$  valores de salida de los clasificadores no siempre coinciden, es decir, algunos modelos predicen una disrupción y otros no. Por tanto, se debe implementar una DF para determinar automáticamente si una alarma tiene que ser activada o no. El desarrollo de esta función es crucial, para lograr las mejores tasas de reconocimiento posibles para el predictor global. La solución óptima, que relaciona los  $n$  resultados con los mejores rendimientos, se puede realizar automáticamente por un sistema de clasificación. Como consecuencia, se aplicó nuevamente la SVM para este propósito. En el procedimiento adoptado, los resultados proporcionados por los clasificadores individuales descritos, se usan para entrenar el sistema de clasificación que implementa la DF. La DF final se obtiene a través de un proceso iterativo. Primero, ha sido formulado un conjunto de condiciones empíricamente, para determinar la regla de decisión inicial (IDR), que permite realizar la primera discriminación de las descargas (disruptivas y no disruptivas) La IDR solo se aplica en el primer paso. Luego se ha implementado un

procedimiento iterativo para converger en una DF optimizada. Este procedimiento realiza el análisis de las descargas, para obtener resultados refinados con el fin de entrenar a la siguiente DF. El proceso continúa hasta que una condición final sea activada.

La IDR ha sido formulada empíricamente. Se ha realizado un análisis que requiere mucho tiempo para interpretar el resultado de los  $n$  clasificadores y para establecer las condiciones empíricas que deciden la activación de una alarma. También vale la pena mencionar, que hay una IDR diferente para cada secuencia de clasificadores. Para la secuencia 2 (M3, M2, M1) se describe la IDR a continuación, donde  $V_i$  representa las distancias de salidas  $M_i$  a los hiperplanos de separación.

Si ( $V_3 > 0$  y  $V_2 > -0.4$  y  $V_1 > -0.8$ )  $\rightarrow$  La alarma debe ser activada.

En cambio, para la secuencia 1 (M2, M1) las condiciones son:

Si ( $V_1 + V_2 > 0$ ) o Si ( $V_2 > 0.3$ )  $\rightarrow$  La alarma debe ser activada.

Obviamente, esta combinación de reglas empíricas no es el procedimiento más preciso para lograr el reconocimiento más alto de tasas (aplicando la IDR a la secuencia 2, se ha obtenido una tasa de aciertos del 80%). Al final de la primera iteración, los resultados almacenados son empleados para entrenar a la próxima DF (Figura 4.10.). El criterio consiste en guardar los datos relevantes cada vez que se activa una alarma. Los resultados llenan una tabla que contiene un número de filas igual al número de descargas analizadas. En la Tabla 4.7., se resumen los seis tipos posibles de resultados que pueden ocurrir. La primera columna representa un índice, la segunda el número de descarga, la tercera el tiempo en el cual el sistema ha detectado un comportamiento disruptivo, la cuarta el instante real de la disrupción (con valor 0 si la descarga es no disruptiva), la quinta es la diferencia entre las columnas tres y cuatro y las últimas tres columnas informan los valores proporcionados por cada modelo de secuencia.

Estos resultados son solo ejemplos obtenidos de una iteración del proceso de optimización, que conduce al desarrollo de la DF final. Las filas representan:

- (1) Una disrupción reconocida correctamente 110 ms antes de que ocurra.
- (2) Dos posibles casos de alarmas perdidas:
  - (a) Un comportamiento disruptivo no reconocido en una descarga disruptiva.
  - (b) Una disrupción que se reconoce después de la ocurrencia de la misma.

(3) Una alarma prematura: se detecta un comportamiento disruptivo en una descarga disruptiva con demasiada anticipación (más de 1 s). En el ejemplo, 4339 ms antes de que ocurra la disrupción.

(4) Una descarga no disruptiva donde el comportamiento ha sido reconocido también como no disruptivo.

(5) Una falsa alarma: una descarga no disruptiva donde, se ha reconocido de forma incorrecta un comportamiento disruptivo.

Las tres últimas columnas del ejemplo en la Tabla 4.7., que se corresponden con las salidas de cada clasificador, se asignan al sistema de aprendizaje de la SVM para crear la DF. Los resultados para entrenar a la DF se pueden dividir en dos grupos o clases:

- La primera clase (activación de alarma) contiene todos los casos donde el sistema ha determinado de forma correcta la ocurrencia de una disrupción. De esta manera el sistema aprenderá cuando debe activarse una alarma.
- La segunda clase (sin activación de alarma) contiene los casos que contemplan las alarmas prematuras y las falsas alarmas. En estos casos, el sistema no está entrenado para activar una alarma demasiado pronto en descargas disruptivas ni en descargas no disruptivas.

Después de cada iteración se entrena una nueva DF, reemplazando a la anterior y obteniéndose una nueva tabla similar a la Tabla 4.7. Estos nuevos datos se utilizan para entrenar a la próxima DF. De esta manera, en cada nueva iteración, se van introduciendo más datos para refinar la DF anterior. El proceso de optimización continúa hasta que la tasa de aciertos con la DF no mejora después de cinco iteraciones consecutivas en el proceso de entrenamiento (véase Figura 4.11.).

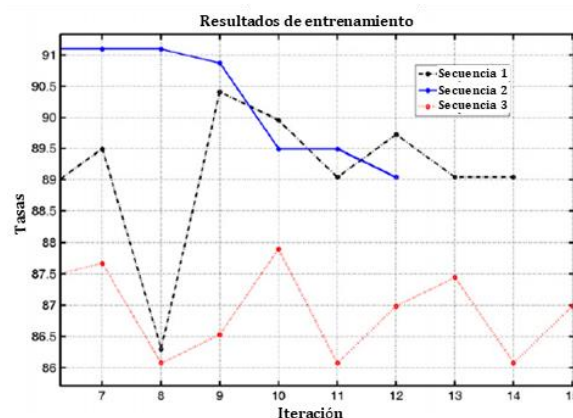


Figura 4.11. Resultados de entrenamiento de tres secuencias. El proceso de entrenamiento se detiene después de cinco iteraciones sin ninguna mejora [Rattá, 2010].

Para la secuencia 2 se observa que en la iteración número 7, la tasa de aciertos está por encima del 91%, en la iteración número 8 la tasa de aciertos se sigue manteniendo por encima del 91%, en la iteración 9 desciende la tasa por debajo del 91%, en la iteración 10 la tasa desciende al 89,5%, se mantiene la tasa del 89,5% para la iteración 11, y en la iteración 12 la tasa termina descendiendo al 89%. Como en las iteraciones 8-12 no ha habido mejora en cuanto a la tasa de aciertos para detección de alarmas, finalizará el proceso de optimización del entrenamiento para esa secuencia de clasificadores.

La fase de prueba se realizó con un subconjunto diferente de descargas, pertenecientes al mismo período con el que fue entrenado el predictor. Como en el caso del entrenamiento, las descargas han sido divididas en segmentos de tiempo de 30 ms desde el inicio de la descarga. Para cada una de las siete secuencias indicadas con anterioridad, la DF con mayor rendimiento es utilizada para identificar comportamientos disruptivos en la fase de prueba. Finalmente, los resultados obtenidos se emplean para calcular estadísticas en cada una de las secuencias. La máxima tasa de aciertos se obtiene con la secuencia 2, permitiendo la compensación adecuada entre complejidad y simplicidad. El rendimiento de los clasificadores sólo puede ser evaluado de forma fiable en la fase de prueba. Cuando se dan al predictor los nuevos casos a clasificar, los resultados muestran de nuevo que la mejor secuencia permanece como la número 2, tal y como se presenta en la Tabla 4.8.

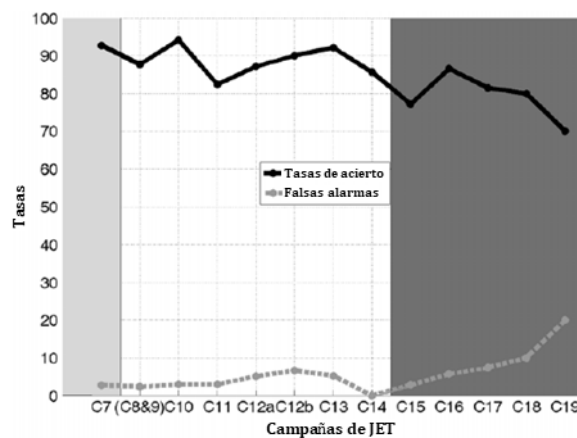
**Tabla 4.8.** Resultados generales de la prueba [Rattá, 2010].

	MA	FA	PA	TE	SR	AVG
Predictor 1 [Sequence 1, DF1]	0	5.4545	5.4545	10.909	89.091	128.01
<b>Predictor 2 [Sequence 2, DF2]</b>	<b>0.909</b>	<b>4.5455</b>	<b>1.8182</b>	<b>7.2727</b>	<b>92.727</b>	<b>146.23</b>
Predictor 3 [Sequence 3, DF3]	0	4.5455	5.4545	10	90	132.23
Predictor 4 [Sequence 4, DF4]	0	5.4545	9.0909	14.545	85.455	128.84
Predictor 5 [Sequence 5, DF5]	0	4.5455	7.2727	11.818	88.182	136.45
Predictor 6 [Sequence 6, DF6]	0	5.4545	5.4545	10.909	89.091	136.23
Predictor 7 [Sequence 7, DF7]	0	5.4545	6.3636	11.818	88.182	121.01

La secuencia 2, incluso permitiendo un pequeño porcentaje del 0.909% de alarmas perdidas, presenta una tasa considerablemente más baja de alarmas prematuras que en el resto de los casos. Demuestra que la secuencia prioriza el logro de los mejores resultados generales. También se pueden apreciar rendimientos superiores en términos de tiempos de reconocimiento promedio (AVG). En la tabla, de izquierda a derecha, cada columna representa los clasificadores y las tasas de alarmas pérdidas (MA), falsas alarmas

(FA), alarmas prematuras (PA), errores totales (TE) y tasas de acierto (SR); expresadas en porcentajes. AVG representa el tiempo medio, expresado en ms, entre la activación de las alarmas y la ocurrencia de las interrupciones reales, teniendo en cuenta todo el conjunto de descargas de prueba. Los tiempos de reacción típicos en JET, para realizar acciones de mitigación, se encuentran entre 30 ms y 200 ms [de Vries, 2009].

Para confirmar el potencial del predictor desarrollado, se ha probado con descargas de las campañas C8 – C19; donde hay que destacar que después de C14 se han implementado cambios significativos en el dispositivo JET. Los resultados finales se resumen en la Figura 4.12.



**Figura 4.12.** Tasas de acierto y falsas alarmas de los modelos previamente entrenados en campañas posteriores. Han sido analizados de forma independiente tres períodos diferentes [Rattá, 2010].

Primero (izquierda), el periodo que incluye el entrenamiento y validación de la campaña C7 y las anteriores.

En segundo lugar (en la zona intermedia), el período de prueba entre las campañas C8 a C14.

Y tercero (derecha), después de los cambios estructurales del dispositivo, Período comprendido entre C15 y C19.

El rectángulo sombreado claro de la Figura 4.12. indica las campañas en que el modelo ha sido previamente entrenado y validado.

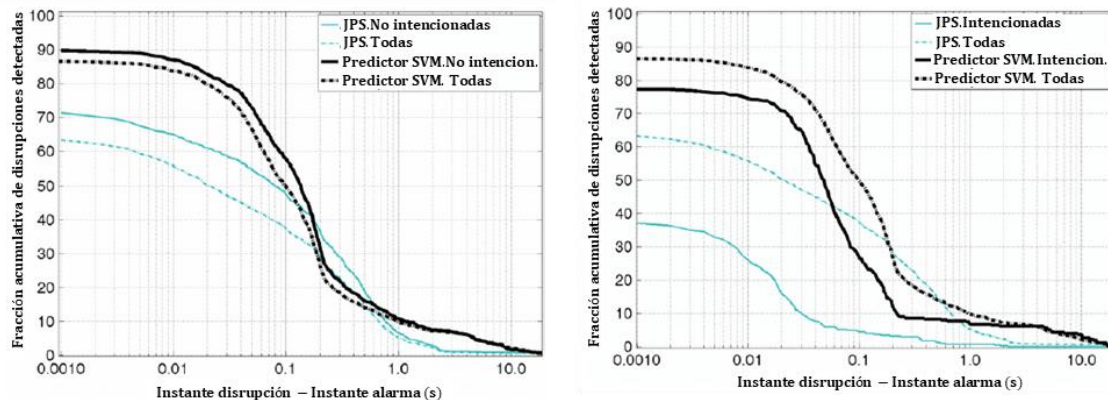
La parte media no sombreada de la Figura 4.12. muestra los porcentajes que se han calculado sobre 376 descargas (50% de ellas eran disruptivas). La variación en la tasa de aciertos está por debajo del 11.6% (entre 94.12% y 82.35%). En ese intervalo (descargas procedentes de campañas JET desde C8 a C14), la tasa mínima de aciertos ha sido obtenida en C11, debido a las restricciones operativas en experimentos con tritio en JET; habiéndose producido un menor número de interrupciones. No se han producido interrupciones típicas que generalmente ocurren con otras campañas. Muchas de las interrupciones producidas durante la campaña C11 fueron rápidas y difíciles de clasificar.

Finalmente, la parte sombreada oscura de la Figura 4.12. muestra los resultados de las campañas C15 – C19 (246 descargas, la mitad de las cuales eran disruptivas). La causa de muchas de las falsas alarmas después de C14, se puede justificar fácilmente con el uso de los cambios estructurales introducidos al dispositivo, afectando a la señal de modo bloqueado y confundiendo así al sistema de detección de la SVM. Si se toman medidas para evitar tales falsas alarmas, el rendimiento se vería mejorado considerablemente. Pero se ha decidido dejar el sistema tal y como está, para dar una tendencia realista del rendimiento para un dispositivo de trabajo, en el que se agregan a menudo nuevos componentes. Además, desde C15 hasta principios de C16, la máquina fue sometida a un incremento inusual de problemas operacionales; como plasmas sucios e incluso fugas en el recipiente de vacío, interrumpiéndose el programa de investigación. Las peores tasas se han alcanzado en C19, principalmente debido al incremento notable de falsas alarmas. Alrededor del 47% de esas alarmas son causadas por anomalías detectadas por el predictor. En estos experimentos, fueron utilizados los EFCC [Barlow, 2001], y probablemente fuera afectada la señal de modo bloqueado. Aproximadamente el 44% de las otras falsas alarmas se produjeron por la detección de grandes picos de radiación y, por tanto, el sistema de detección SVM identificó erróneamente estos incrementos no aprendidos de la radiación como precursores de la disrupción.

Los rendimientos del predictor se han comparado con los resultados del JET Protection System (JPS) [Santagiustina, 1995], que es el sistema utilizado en tiempo real en JET para la generación de alarmas debidas a disrupciones. La mayoría de las alarmas activadas por el JPS se deben a valores elevados en la señal de modo bloqueado [Yoshino, 2003]. Este sistema ha sido implementado en JET para estrategias de elusión en todas las descargas y, por lo tanto, es la principal herramienta de referencia. Para una comparación adecuada de ambos sistemas, deben tenerse en cuenta algunos problemas. El JPS interviene durante la ejecución de un experimento cada vez que considera que se ha detectado un comportamiento disruptivo [de Vries, 2009]. Debido a esta intervención directa sobre la evolución del plasma, es imposible calcular estadísticas de falsas alarmas y, por lo tanto, solo se han tenido en cuenta las descargas disruptivas para la comparación. También para fines de comparación, las alarmas prematuras no han sido medidas como hasta ahora, haciéndose de la misma manera en que son calculadas por el JPS. Las tasas se calculan en relación a los "tiempos de anticipación" (diferencia entre el instante de la disrupción y el instante en el que se produce la alarma). Esos tiempos de anticipación son

relevantes, porque representan el margen temporal que el actuador tiene para poder realizar acciones de evitación o mitigación durante el procesamiento de una descarga.

Cada punto en las curvas trazadas en la Figura 4.13. representa el porcentaje acumulado de interrupciones reconocidas, con un tiempo de anticipación igual o superior al especificado en el eje x correspondiente.



**Figura 4.13.** Tiempos de anticipación (tiempo de interrupción menos tiempo de alarma) para todas las descargas disruptivas desde la campaña C1 hasta la campaña C19 [Rattá, 2010].

Parte izquierda figura: los porcentajes acumulativos de interrupciones no intencionadas y todas las detectadas. Se comparan las interrupciones para el método SVM con los resultados de JPS.

Figura derecha: los porcentajes acumulativos de interrupciones intencionadas y todas las interrupciones. Se comparan las interrupciones detectadas para el método SVM con los resultados de JPS.

Utilizando la base de datos disruptiva de JET, ha sido posible probar el sistema SVM sobre todas las descargas disruptivas incluidas en el período comprendido entre el comienzo de la campaña C8 hasta el fin de la campaña C19 de JET (483 interrupciones, 347 de ellas no intencionadas). En un escenario en tiempo real, la detección de interrupciones no intencionadas (las que ocurren durante la operación normal), requiere un mayor nivel de atención que el reconocimiento de interrupciones intencionadas (las activadas explícitamente durante sesiones dedicadas a estudiar su física). Para verificar los resultados obtenidos en cada caso, fueron calculadas estadísticas separadas y combinadas sobre interrupciones no intencionadas e intencionadas. Los resultados representados en la Figura 4.13. muestran que se logran notables porcentajes de reconocimiento más altos con el predictor SVM para ambos tipos (interrupciones no intencionadas e interrupciones intencionadas). El sistema SVM proporciona tasas significativamente más altas que el JPS en el intervalo de tiempo que se encuentra a más de 180 ms antes de la interrupción. En el intervalo entre ~180 ms y ~0.9 s, el JPS muestra



una mejor precisión. Este comportamiento es coherente con el entrenamiento del sistema, basado en ejemplos que cubren el intervalo hasta 200 ms antes de la disrupción. Por otro lado, alguna de las alarmas perdidas del JPS puede haberse debido a malas adquisiciones de las muestras en la señal de modo bloqueado, una de las principales medidas de referencia analizadas por JPS [de Vries, 2009]. Se han calculado las estadísticas del sistema JPS sobre todas las descargas de C1 a C19, incluso con posibles datos corruptos. Para el predictor SVM se descartaron el 8.5% de las descargas; debido a la falta de algunas señales requeridas en las descargas analizadas.

Resumiendo, para conseguir predecir disrupciones con una alta tasa de aciertos, se han utilizado varias secuencias, que analizan la evolución completa de cada descarga. El predictor final ha sido entrenado y probado sobre todo tipo de disrupciones hasta la campaña C7 de JET. Posteriormente el predictor se ha probado en una amplia gama de descargas en campañas más recientes no utilizadas para la fase de entrenamiento. Finalmente, el rendimiento del predictor SVM se ha comparado con el JPS. En la etapa de prueba que incluyó la campaña C7 y campañas anteriores, con el mejor clasificador se consiguió una tasa de aciertos general del 92,73%, con un 0,91% de alarmas pérdidas. Este porcentaje de alarmas perdidas es el resultado de un error en más de 110 descargas (66 disruptivas y 44 no disruptivas). En general, las alarmas perdidas se han minimizado y la suma de las falsas alarmas y alarmas prematuras se redujo a menos del 6.4% (ver Predictor 2 en Tabla 4.8.). La robustez del predictor se ha demostrado mediante pruebas con un número significativo de descargas, pertenecientes a más campañas recientes que con las que fue entrenado. La relevancia de este aspecto es significativa, porque la falta de la capacidad de generalización en todas las campañas fue una de las principales debilidades de predictores anteriores. Los resultados muestran claramente que los altos rendimientos permanecen bastante constantes hasta que fue implementada una gran modificación en el dispositivo de fusión JET. Después de dicha implementación, la tasa de aciertos sigue siendo elevada, pero aumenta el porcentaje de errores al 9%. Finalmente, el método desarrollado ha sido comparado con JPS, utilizado en JET durante muchos años. Los mayores rendimientos se han alcanzado en el intervalo de hasta 150 ms antes de la disrupción. Estos resultados son especialmente relevantes, debido al hecho de que el tiempo típico requerido por los actuadores para realizar acciones efectivas de mitigación en JET es de aproximadamente 30 ms. Para concluir, vale la pena mencionar que en [Rattá, 2010] se ha asignado la máxima prioridad a reducir el número de alarmas perdidas, para preservar la integridad del dispositivo.

#### 4.2.4. Predictor adaptativo para ASDEX Upgrade.

En [Cannas, 2010] se ha construido un predictor neuronal utilizando descargas de plasma, seleccionadas de los experimentos realizados en el periodo comprendido entre julio de 2002 y julio de 2004 y procedentes de ASDEX Upgrade. Para probar la capacidad de predicción del sistema en tiempo real, se ha evaluado su rendimiento utilizando descargas procedentes de diferentes campañas experimentales comprendidas entre junio de 2005 y julio de 2007. Se incluyeron todas las disrupciones que ocurrieron en las campañas experimentales seleccionadas, con la excepción de: aquellas que ocurren en la fase de aceleración, en la fase de desaceleración (si la disrupción no ocurre en los primeros 100 ms), las causadas por inyección masiva de gas y exceptuando también disrupciones que se originan debido a eventos de desplazamiento vertical del plasma.

Para cada descarga, se seleccionaron siete señales de diagnóstico de plasma, entre las numerosas señales disponibles en tiempo real. Durante el proceso de entrenamiento se ha utilizado un mapa autoorganizado, para reducir el tamaño de la base de datos y así mejorar el entrenamiento de la red neuronal. Además, se ha realizado un proceso de optimización para discriminar entre fases no disruptivas y pre disruptivas. La tasa de aciertos en la predicción ha sido optimizada, realizando un entrenamiento adaptativo de la red cada vez que el predictor activa una alarma perdida.

En los Tokamaks actuales, las disrupciones son generalmente tolerables. Sin embargo, en ITER que tendrá mucha más energía magnética y térmica almacenada que en los Tokamaks actuales, solo se permitirá un número limitado de disrupciones [Hender, 2007] e [ITER, 1999].

Recientemente, ha habido un progreso significativo en evitar o mitigar las disrupciones [Esposito, 2009] y [Pautasso, 2009]. Estas acciones de control deben ser apoyadas por sistemas predictivos adecuados, que proporcionen suficiente anticipación ante la ocurrencia de una disrupción inminente; pudiendo tomar acciones preventivas de evitación y/o mitigación.

Durante la última década, se ha investigado en varios Tokamaks el tema de la implementación confiable de predictores disruptivos en tiempo real. El análisis de la experiencia previa adquirida con redes neuronales (NN) como método de predicción de disrupción [Hender, 2007], [Cannas, 2006b] y [Cannas, 2007b], permite resaltar algunos inconvenientes que deben resolverse para desarrollar un predictor efectivo de disrupciones en tiempo real. Los resultados afirman que el rendimiento de predicción es

bastante bueno en los modelos NN, pero la mayoría de ellos están relacionados con un conjunto de prueba de descargas seleccionadas en la misma campaña utilizada para entrenar el modelo en sí. Desafortunadamente, el rendimiento de NN disminuye considerablemente cuando se aplica a diferentes campañas. Este es el caso de las máquinas experimentales actuales, donde el espacio operacional del plasma evoluciona para producir plasmas con un rendimiento superior. Para limitar este inconveniente, las descargas utilizadas para entrenar el modelo neuronal tienen que pertenecer a un espacio operativo lo más grande posible, o bien el modelo neuronal debe actualizarse tal y como se expande el dominio operativo del plasma.

En [Cannas, 2010] se aborda el problema de predicción de disrupción como un problema de clasificación. Se formaliza el entrenamiento de la red neuronal, como un problema de identificación de la pertenencia de una muestra a una clase específica: la clase no disruptiva o la clase pre disruptiva. Por lo tanto, para construir el conjunto de entrenamiento para el modelo NN, es necesario identificar el instante de tiempo que discrimina entre la fase no disruptiva y pre disruptiva del plasma. Ese instante temporal no tiene un valor prefijado y su identificación, a menudo, resulta ser una tarea muy difícil. Para resolver este problema, se han propuesto varios criterios físicos y estadísticos [Pautasso, 2002], [Esposito, 2009], [Cannas, 2006b], [Giannone, 2004a] y [Murari, 2009]. Sin embargo, ninguna de las propuestas demostró ser considerada como un criterio final.

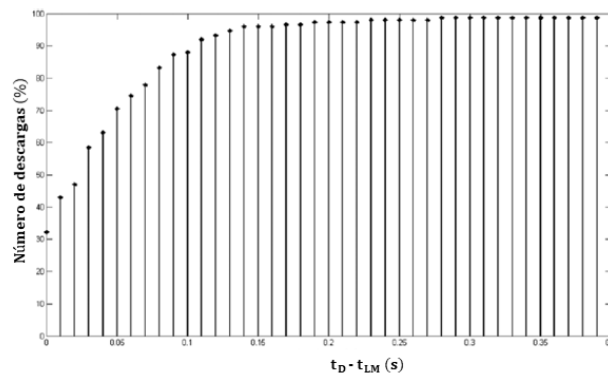
Finalmente, usando modelos basados en datos, como por ejemplo redes neuronales (NNs), requiere un conjunto de señales de entrada con alta confiabilidad, resolución de tiempo adecuada y confiabilidad a largo plazo. Para cumplir con estos requisitos, se debe prestar especial atención a la generación del conjunto de datos. En [Cannas, 2010] se propone un sistema de predicción de disrupción basado en NNs para ASDEX Upgrade (AUG). Teniendo en cuenta los requisitos en tiempo real y la superación de no pocos inconvenientes, se ha prestado especial atención a la construcción de la base de datos, tomándose varios pasos para desarrollar el predictor de disrupción:

1. Para poder extraer información útil de las señales, se ha empleado la minería de datos utilizando los archivos que representan las descargas de AUG.
2. Se han implementado algoritmos para obtener información automáticamente, en casos en los que algunas señales de diagnóstico no estén disponibles para alguna de las descargas.

3. Se realizó un procedimiento de optimización para identificar el tiempo instantáneo  $t_{pre-disr}$ , que discrimina entre comportamiento no disruptivo o pre disruptivo del plasma.
4. Basándose en mapas autoorganizados (SOM), un procedimiento de agrupamiento permite seleccionar automáticamente un número limitado de muestras significativas de entre todas las muestras que componen cada descarga.
5. Para evaluar el rendimiento en tiempo real del predictor de disrupción, se han tenido en cuenta dos conjuntos de datos diferentes: uno de ellos se ha empleado para entrenamiento y validación, y el otro se ha empleado para probar. Ambos conjuntos de datos contienen descargas pertenecientes a diferentes campañas experimentales y realizadas en años diferentes.
6. Para mejorar la capacidad de predicción del modelo neuronal, se realiza un entrenamiento adaptativo de la red cada vez que se activa una alarma perdida.

Se pueden definir diferentes fases de disrupción en una descarga de plasma. En la fase pre precursora, el equilibrio evoluciona hacia una configuración inestable y en el precursor aparecen inestabilidades magnetohidrodinámicas (MHD). En la fase de enfriamiento térmico, se pierde la energía térmica del plasma. Esta fase es seguida por la fase de apagado, durante la cual la energía magnética es disipada principalmente por la radiación [Wesson, 2004] y [Schuller, 1995]. Un predictor de disrupción debe producir una alarma antes o durante la fase precursora, ya que en las dos últimas fases ni la elusión ni la mitigación pueden ser efectivas. Se han considerado diferentes criterios físicos para discriminar entre la fase no disruptiva y la fase previa a la disrupción, en base a la Transición H – L [Giannone, 2004b], la señal de voltaje de bucle [Esposito, 2009] y la señal LM [Pautasso, 1998b]. Sin embargo, estos criterios no eran aplicables en AUG, debido a la falta de fiabilidad o la falta de disponibilidad de algunas señales para muchas de las descargas disruptivas en archivos de descargas de AUG. Para construir un predictor capaz de detectar rápidamente una disrupción, la distinción entre la fase no disruptiva y la fase pre disruptiva debe realizarse en cada descarga perteneciente al conjunto de entrenamiento. Es decir, cada muestra perteneciente a la descarga debe etiquetarse como disruptiva o no disruptiva. La reconstrucción de algunas señales no confiables o no disponibles que son precursoras de disrupción, podría ser una solución que requiere mucho tiempo. En [Cannas, 2006b] los autores manejan la identificación de la fase previa

a la disrupción por medio de procedimientos heurísticos. En [Cannas, 2007b] los autores renunciaron a identificar como precursor al tiempo instantáneo, y fue introducida una banda de incertidumbre (o región de transición), que separa las fases no disruptivas y disruptivas, agrupando muestras de cada descarga por medio de un mapa autoorganizado SOM [Kohonen, 1982]. En [Cannas, 2010], el tiempo instantáneo  $t_{pre-disr}$ , que discrimina entre la fase no disruptiva y pre disruptiva de una descarga disruptiva, se identifica eligiendo su valor óptimo en el intervalo de tiempo  $[t_D-160, t_D]$  ms. Se encontró que el valor óptimo de  $t_{pre-disr}$  es igual a 40 ms, obteniéndose sobre una base estadística. Para definir la ventana de tiempo donde se ha buscado el valor óptimo  $t_{pre-disr}$ , se ha utilizado la señal de modo bloqueado (LM). La Figura 4.14. muestra la distribución del intervalo de tiempo  $t_D - t_{LM}$ , donde el tiempo LM ( $t_{LM}$ ) es el instante donde la frecuencia de oscilación de la señal LM disminuye y/o su valor absoluto aumenta rápidamente. En el histograma, la altura de cada elemento es igual al número acumulado de descargas en porcentaje, para cada uno de los posibles valores de  $t_D - t_{LM}$ .



**Figura 4.14.** Histograma acumulativo del intervalo de tiempo ( $t_D - t_{LM}$ ) entre el tiempo LM y la ocurrencia de la disrupción [Cannas, 2010].

Cuando se identifica  $t_{pre-disr}$ , las muestras en el intervalo de tiempo  $[t_{pre-disr}, t_D]$  se consideran muestras disruptivas, y las anteriores a  $t_{pre-disr}$  se consideran muestras no disruptivas.

En [Cannas, 2010], ha sido entrenado un perceptrón multicapa (MLP) [Hertz, 1991] para pronosticar una disrupción inminente, teniendo en cuenta un conjunto de parámetros relacionados con el plasma, y medidos durante el procesamiento de descargas disruptivas. El MLP entrenado, captura la relación entre los parámetros seleccionados del plasma y los estados disruptivos o no disruptivos del mismo. Durante la fase de entrenamiento no se han utilizado descargas no disruptivas, usándose este tipo de

descargas solamente durante la fase de prueba. Como el número de muestras en la fase no disruptiva de una descarga disruptiva es mucho mayor que el número de muestras en la fase disruptiva de esa misma descarga disruptiva; para entrenar el predictor MLP hay que realizar una selección de un número limitado de muestras no disruptivas para equilibrar el número de muestras que describen la fase disruptiva y las que describen la fase no disruptiva. Para este propósito, se utiliza un procedimiento de agrupamiento basado en mapas autoorganizados (SOM), que permite seleccionar automáticamente un número limitado de muestras significativas de toda la longitud de la descarga, con el fin de reducir el tiempo computacional.

El problema de entrenamiento del MLP se ha formalizado en [Cannas, 2010] como la identificación de la pertenencia de una muestra a una determinada clase: la clase "no disruptiva" o la clase "disruptiva". Las muestras no disruptivas pertenecen a la fase no disruptiva de una descarga disruptiva o a una descarga no disruptiva. Las muestras disruptivas pertenecen a la fase pre disruptiva de una descarga disruptiva. La salida del MLP se ha establecido igual a 0 para muestras que pertenecen a la clase no disruptiva, mientras que ha sido establecida igual a 1 para las muestras que pertenecen a la clase disruptiva. Por lo tanto, el objetivo del entrenamiento es resolver un problema de clasificación de dos clases. El esquema del predictor de interrupción propuesto consiste de dos partes, un bloque de agrupamiento (SOM) y un predictor neuronal de interrupciones (MLP). Durante la fase de entrenamiento, el SOM se encarga de la reducción de los datos y alimenta al predictor MLP. Durante la predicción en tiempo real, todas las muestras de la descarga son introducidas consecutivamente al predictor MLP. Para cada muestra, el MLP devuelve un número real entre 0 y 1, que indica respectivamente la pertenencia de la muestra a la clase no disruptiva o disruptiva. Como ejemplo, en la Figura 4.15. se muestra la alarma proporcionada por el predictor para la descarga de prueba # 16513. Se muestran también otros parámetros relativos al plasma (corriente de plasma  $I_{pla}$ , la densidad de electrones/densidad de Greenwald  $ne_{Greenwald}$ , beta poloidal  $\beta_{pol}$ , inductividad interna  $l_i$ , factor de seguridad  $q_{95}$ , entrada total de energía  $P_{inp}$ , potencia radiada / potencia de entrada total  $P_{frac}$ ). El predictor activará una alarma solo cuando la salida de la NN permanezca por encima de un umbral de alarma predeterminado (AT), y para un número de muestras consecutivas prefijado (K).

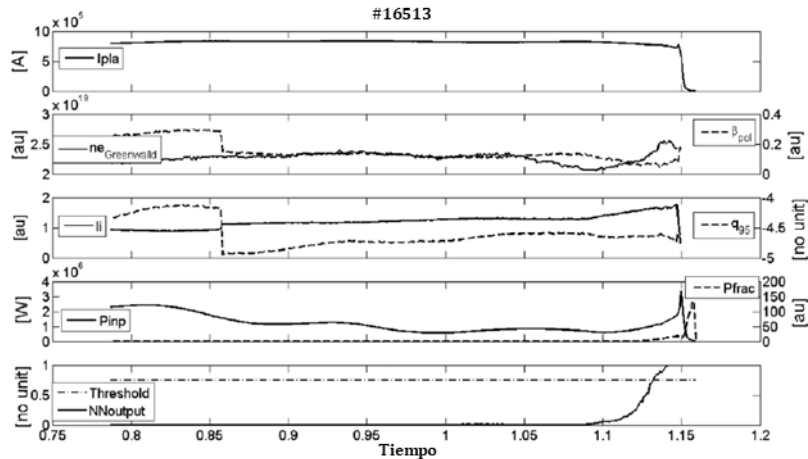


Figura 4.15. Algunos parámetros de plasma y función de alarma correspondiente a descarga #16513 [Cannas, 2010].

El rendimiento de un predictor de disrupción es típicamente cuantificado en términos de la tasa de aciertos, es decir, la fracción de descargas de plasma donde el predictor predice con éxito la ocurrencia de disrupción o, alternativamente, la tasa de fallos, en la que el predictor no puede detectar una disrupción inminente. La tasa de falsas alarmas se ve incrementada por producirse una alarma disruptiva en muestras que no lo son [Hender, 2007].

En [Cannas, 2010], el rendimiento del sistema de predicción es evaluado en términos de:

- Predicciones exitosas (SP): la fracción de descargas (disruptivas y/o no disruptivas) cuya predicción ha sido correcta.
- Alarmas perdidas (MA): la fracción de descargas disruptivas predicha como no disruptiva.
- Detecciones prematuras (PD): la fracción de descargas disruptivas que activan la alarma con demasiada antelación con respecto al instante en el que ocurre la disrupción.
- Detecciones tardías (TD): la fracción de descargas disruptivas que activan la alarma demasiado tarde con respecto al instante de la disrupción.
- Falsas alarmas (FA): la fracción de descargas no disruptivas predicha por el sistema como disruptiva.

Hay que tener en cuenta que una predicción de disrupción se considera exitosa si el sistema puede predecir correctamente la disrupción hasta algunos milisegundos antes del instante en el que ocurre la disrupción. El tiempo necesario para que intervengan los sistemas de mitigación, dependen de las técnicas de evitación y/o mitigación empleadas

y sobre la inherente escala de tiempo de la máquina Tokamak considerada. Las técnicas de mitigación introducidas recientemente en AUG [Pautasso, 2009] son capaces de mitigar una disrupción dentro de 2 ms. Además, se supone que una alarma es prematura si se da 160 ms antes del instante de la disrupción. Por lo tanto, en [Cannas, 2010], una predicción de disrupción se considera exitosa si el sistema es capaz de activar la alarma en un instante de tiempo dentro del ventana temporal  $[t_D-160, t_D-2]$  ms; mientras que la respuesta del predictor para una descarga no disruptiva se considera exitosa si la alarma no es activada en absoluto. Una acción exitosa de elusión o mitigación, reduce el peligro de los potenciales daños estructurales sobre el dispositivo de fusión. Por lo tanto, el predictor de la disrupción debe caracterizarse por una alta tasa de éxito en la predicción SP y bajas tasas en PD, TD, MA y FA.

Los datos para [Cannas, 2010] se seleccionaron de las campañas experimentales realizadas desde agosto de 2002 a julio de 2007. Se han construido dos bases de datos, correspondientes a dos diferentes períodos:

- DB 1: de julio de 2002 a abril de 2005, incluyendo 80 descargas no disruptivas y 149 disruptivas en el rango de descargas 16200–19999.
- DB 2: de junio de 2005 a julio de 2007, incluyendo 540 descargas no disruptivas y 81 disruptivas en el rango de descargas 20000–22146.

Los conjuntos de entrenamiento y de validación se han creado seleccionando descargas disruptivas solo desde DB 1, mientras que los rendimientos en tiempo real se evalúan en descargas disruptivas y no disruptivas de DB 2.

Las bases de datos están compuestas por siete valores que representan parámetros del plasma (véase Tabla 4.9.) durante la parte superior plana de la descarga, es decir, durante el periodo comprendido después de la fase de aceleración y anterior a la fase de deceleración de la descarga.

**Tabla 4.9.** Señales en la base de datos [Cannas, 2010].

Nombre de la señal	Acrónimo
Safety factor at 95% of the major radius	$q_{95}$
Total input power (W)	$P_{\text{tot}}$
Total radiated power/total input power	$P_{\text{frac}}$
Internal inductivity	$l_i$
Polooidal beta	$\beta_{\text{pol}}$
Electron density/Greenwald density	$ne_{\text{Greenwald}}$
Locked mode signal (V)	LM



Los parámetros han sido seleccionados en base a resultados anteriores, y teniendo en cuenta consideraciones físicas; así como la disponibilidad de datos en tiempo real. Las primeras seis señales enumeradas en la Tabla 4.9. se han utilizado como entradas al predictor neuronal, mientras que la señal LM se ha utilizado para el análisis y comparación. Todas las señales se muestrean haciendo referencia a la base de tiempo de la corriente de plasma. Como el tiempo de muestreo de los diferentes sistemas de diagnóstico son desiguales, las señales en los conjuntos de entrenamiento y en los conjuntos de validación se han vuelto a muestrear mediante un algoritmo de interpolación, para lograr una resolución de 1 ms. Para cumplir los requisitos en tiempo real, durante la prueba se realiza una regresión lineal en el conjunto de prueba; utilizando una función de extrapolación. Cuando el predictor neuronal se utilice en línea, se supone que el muestreo de 1 ms estará disponible para todas las señales. Cuando se pierdan algunos datos, las muestras perdidas serán reconstruidas usando la misma función de extrapolación. Las señales presentan amplitudes que difieren en varios órdenes de magnitud, que es contraproducente para el predictor basado en el perceptrón multicapa (MLP) [Hertz, 1991]. Por lo tanto, se han normalizado en el intervalo [0, 1], usando la siguiente expresión:

$$x_{j,norm} = \frac{x_j - m_j}{M_j - m_j}; \text{ con } j = 1, \dots, 7. \quad (2)$$

Referido a la ecuación (2),  $M_j$  y  $m_j$  son respectivamente los valores máximo y mínimo de la  $j$ -ésima señal de entrenamiento  $x_j$ .

El objetivo de [Cannas, 2010] es desarrollar un sistema capaz de predecir disrupciones que ocurren durante la fase estacionaria del plasma. Por lo tanto, se incluyeron todas las disrupciones que ocurrieron en el período experimental elegido, con la excepción de las que ocurren en las fases de aceleración y desaceleración, las causadas por inyección masiva de gas y disrupciones que ocurren después de un evento de desplazamiento vertical (VDE). Además, se han incluido en la base de datos algunas disrupciones menores con pérdida parcial de la corriente de plasma. En particular, se han observado dos tipos de disrupciones menores. El primero se acompaña de una pérdida parcial de corriente del plasma ( $I_{pla}$ ), debida a una disminución en la energía térmica ( $W_{mhd}$ ) y por la pérdida de la estabilidad vertical (aumento en la posición vertical del centroide  $Z_{curr}$ ); como se muestra en la Figura 4.16. en el instante 3,373 s.

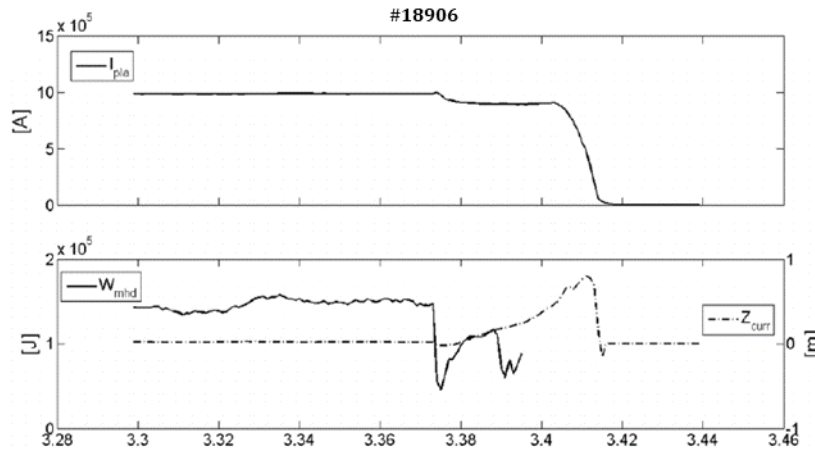


Figura 4.16. Algunos parámetros de plasma que intervienen en la disrupción de la descarga # 18906, donde la disrupción es provocada por un evento de desplazamiento vertical (VDE) [Cannas, 2010].

En este caso, el instante de la disrupción  $t_d$  se establece cuando se produce la disrupción menor, para predecir la disrupción menor en lugar de la mayor.

En el segundo tipo de disrupción menor no hay grandes cambios en la corriente de plasma, o bien se observa pérdida de inestabilidad vertical, pero apreciándose solo una disminución repentina en  $W_{mhd}$ ; tal y como puede apreciarse en la Figura 4.17. para el instante  $t = 2,971$  s.

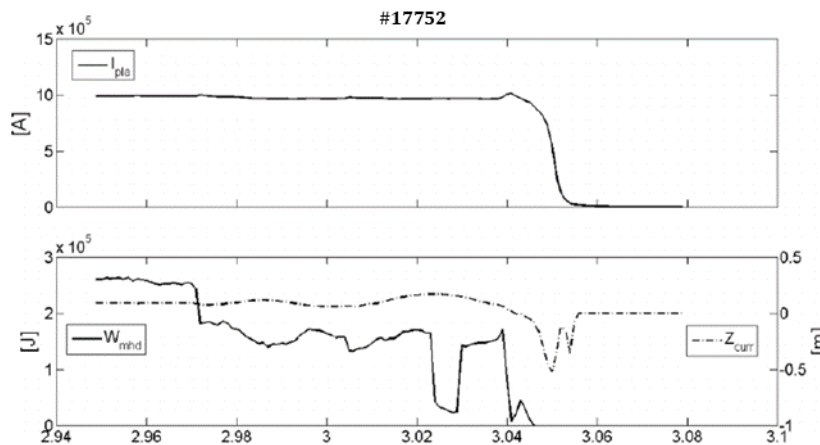


Figura 4.17. Algunos parámetros de plasma para la descarga disruptiva # 17752, donde el VDE es provocado por la disrupción [Cannas, 2010].

En este caso, como se desea predecir una disrupción mayor, el instante de la disrupción  $t_d$  se establece en el momento en el que ocurre la disrupción.

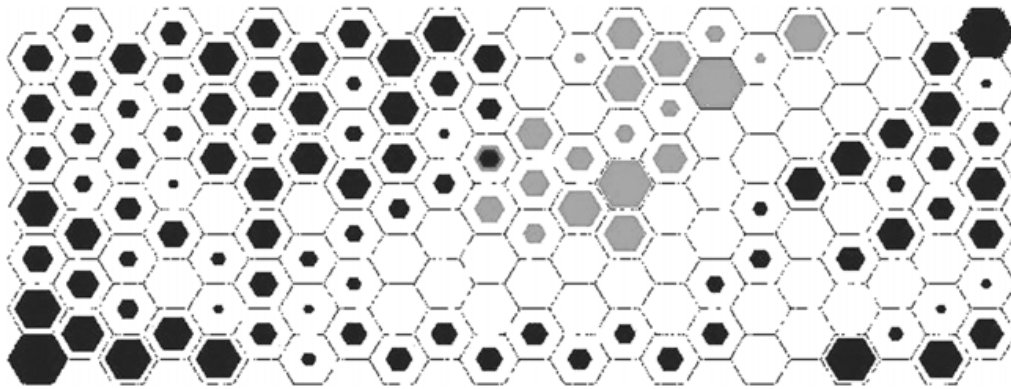
Para verificar el funcionamiento del predictor en términos de falsas alarmas, se han seleccionado 80 descargas no disruptivas, pertenecientes al mismo rango de

descargas disruptivas utilizadas en la fase de aprendizaje; es decir, pertenecientes al rango de descargas 16200–19999. Para investigar el envejecimiento del sistema predictivo, se han seleccionado otras 540 descargas no disruptivas, pertenecientes al rango 20000–22146, disponibles en las campañas experimentales de junio de 2005 hasta julio de 2007.

Se realizó un análisis visual preliminar de las señales, para encontrar aquellas que tenían mediciones anómalas. Luego, para cada una de las descargas, fueron verificados mediante inspección visual los parámetros relevantes para la predicción de la disrupción, y que fueron enumerados en la Tabla 4.9. Este análisis detallado permitió encontrar descargas que tuvieron que ser descartadas, porque varios parámetros estaban corrompidos.

Como se describió anteriormente, el esquema del predictor consiste en dos NN en cascada: el SOM realiza una reducción de datos, lo cual es crucial para el entrenamiento del MLP. El asunto a considerar en la generación del conjunto de entrenamiento, es la selección de un número limitado de muestras para cada descarga. Esa selección de muestras debe ser tal que permita describir inequívocamente el dominio operativo de los experimentos. Sin embargo, el número de muestras en una descarga no disruptiva es demasiado grande para ser utilizado en el conjunto de entrenamiento. Además, la cantidad de muestras en la fase no disruptiva de una descarga disruptiva es mucho mayor con respecto al número de muestras disponibles en la fase disruptiva de dicha descarga disruptiva. Se necesita un equilibrio entre el número de muestras seleccionadas para describir la fase disruptiva y aquellas muestras seleccionadas para describir la fase no disruptiva. En el aprendizaje neuronal, si los datos están desbalanceados entre diferentes clases, las características que representan las clases que tienen un número pequeño de muestras en el conjunto de entrenamiento pueden ser ignoradas por el sistema. En [Cannas, 2010] se desarrolla un SOM para cada descarga [Cannas, 2006b] y [Cannas, 2007b]. Los patrones de entrada de seis dimensiones son asignados a dos dimensiones, permitiendo la visualización de la distribución de los datos. El SOM realiza una agrupación de muestras de la descarga, preservando las similitudes del espacio de la entrada. Por lo tanto, se supone que cada grupo contiene muestras pertenecientes a estados plasmáticos similares. Las muestras pertenecientes al intervalo  $[0, t_{pre-disr}]$ , se consideran muestras no disruptivas, mientras que las correspondientes a los instantes de tiempo posteriores hasta  $t_d$ , se consideran muestras disruptivas. Como ejemplo, en la Figura 4.18. se muestra el SOM bidimensional relacionado con una descarga disruptiva de plasma. Las muestras no disruptivas están resaltadas en gris oscuro, mientras que las

muestras disruptivas están marcadas en gris claro. Además, para cada unidad del mapa, la densidad del color es proporcional al número de muestras contenidas dentro de las propias unidades. Esta asignación se puede aplicar para seleccionar automáticamente un número limitado de muestras. En particular, como la mayoría de las muestras en una descarga disruptiva pertenecen a la fase no disruptiva, solo es considerada una muestra no disruptiva para cada unidad de mapa que contiene muestras seguras, mientras que son retenidas todas las muestras disruptivas.



**Figura 4.18.** Visualización de un SOM para una descarga disruptiva a partir de un espacio de parámetros de seis dimensiones en un espacio bidimensional [Cannas, 2010].

Para entrenar el MLP, se han subdividido las descargas de DB 1 en tres juegos, tal y como se indica a continuación:

1. Conjunto de entrenamiento: 100 descargas disruptivas.
2. Conjunto de validación: 16 descargas disruptivas.
3. Conjunto de prueba: 33 descargas disruptivas y 80 descargas no disruptivas.

Las entradas de red están constituidas por todas las señales de la Tabla 4.9., siendo crucial la importancia de la señal LM.

Para cada conjunto de entrada, se han realizado 13 sesiones de entrenamiento variando  $t_{pre-disr}$ . Para cada valor de  $t_{pre-disr}$  se realiza una reducción de datos por medio de los SOM. Las muestras seleccionadas a través de la reducción de datos se han utilizado como entrada para los MLP, cuya topología de red ha cambiado durante el entrenamiento, para alcanzar un rendimiento prescrito. En particular, se ha adoptado el método de cultivo [Hertz, 1991], que consiste en aumentar progresivamente el número de neuronas ocultas, hasta que el rendimiento de la red durante la fase de validación alcance el valor deseado. Durante el proceso de entrenamiento se podría producir una

penalización debida a la presencia de falsas alarmas (FA) [Cannas, 2004b], que en la aplicación prevista terminaría la descarga. Ha de tenerse en cuenta una solución de compromiso que permita optimizar los objetivos de maximizar SP y minimizar PD y MA.

Para seleccionar la mejor configuración de red, han de ser optimizados varios parámetros: el umbral de alarmas (AT), el tiempo de seguridad (K), el tiempo previo a la disrupción ( $t_{pre-disr}$ ) y la topología del MLP. Los dos primeros parámetros están relacionados con la activación de la alarma disruptiva. El umbral de alarma (AT) ha sido optimizado en el rango de 0.55 a 0.8, teniendo en cuenta incrementos de 0.05 unidades. El tiempo de seguridad (K) ha sido optimizado en el rango de 1 a 15. En [Cannas, 2010], la red seleccionada ofrece la mejor solución de compromiso en términos de PD, MA, SP y FA en el conjunto de validación.

Uno de los principales inconvenientes de los enfoques de predicción NN, es que el rendimiento de la red normalmente se deteriora cuando se presentan nuevas configuraciones de plasma a la red. De hecho, en nuevos experimentos las configuraciones de plasma podrían tener características completamente diferentes de las observadas en los experimentos seleccionados para el conjunto de entrenamiento. Esta "novedad" puede conducir a un comportamiento incorrecto de los predictores del MLP (envejecimiento de la NN) [Pautasso, 2002]. En la Tabla 4.10. se presentan las prestaciones del predictor NN para el conjunto de prueba de DB 1. Para verificar la robustez del sistema predictivo, en [Cannas, 2010] se realizó una prueba fuera de línea en la base de datos DB 2, que contiene descargas no disruptivas y disruptivas fuera del rango del conjunto de entrenamiento. Durante la simulación de las operaciones en línea, el predictor del sistema neuronal ha sido alimentado con todas las muestras consecutivas de 540 descargas no disruptivas y de las 81 descargas disruptivas de DB 2.

**Tabla 4.10.** Prestaciones de NNP en el conjunto de prueba de DB 1 [Cannas, 2010].

	FA (%)	PD (%) $t_{Alarma} < (t_{Disr} - 160 \text{ ms})$	TD (%) $(t_{Disr} - 2 \text{ ms}) < t_{Alarma} < t_{Disr}$	MA (%)	SP (%) $(t_{Disr} - 160 \text{ ms}) < t_{Alarma} < (t_{Disr} - 2 \text{ ms})$
Descargas disruptivas (33)	—	3.03	9.09	6.06	81.82
Descargas no disruptivas (80)	1.25	—	—	—	98.75
Descargas totales (113)	0.88	0.88	2.65	1.77	93.82

La Tabla 4.11. informa los mismos índices de rendimiento, pero en este caso se aprecia que el desempeño predictivo general se deteriora. Sin embargo, la tasa de aciertos de predicción en descargas disruptivas sigue siendo superior al 80%, si se evalúa como se propone en [Hender, 2007], es decir,  $SP = 1 - MA$ .

Tabla 4.11. Rendimiento de NNP en DB 2 [Cannas, 2010].

	FA (%)	PD (%) $t_{\text{Alarma}} < (t_{\text{Disr}} - 160 \text{ ms})$	TD (%) $(t_{\text{Disr}} - 2 \text{ ms}) < t_{\text{Alarma}} < t_{\text{Disr}}$	MA (%)	SP (%) $(t_{\text{Disr}} - 160 \text{ ms}) < t_{\text{Alarma}} < (t_{\text{Disr}} - 2 \text{ ms})$
Descargas disruptivas (81)	—	9.87	7.41	18.52	64.20
Descargas no disruptivas (540)	4.81	—	—	—	95.19
Descargas totales (621)	4.18	1.29	0.97	2.42	91.14

Además, el predictor sigue siendo robusto con respecto a las falsas alarmas. Para investigar las razones de ese deterioro, se han mapeado de forma conjunta en un SOM, las muestras de seis dimensiones de DB 1 (conjuntos de entrenamiento y validación) y las muestras de los últimos 160 ms de MA de DB 2. La Figura 4.19. muestra el SOM 2D. En este mapeo, se asignan diferentes colores a los diferentes tipos de muestras.

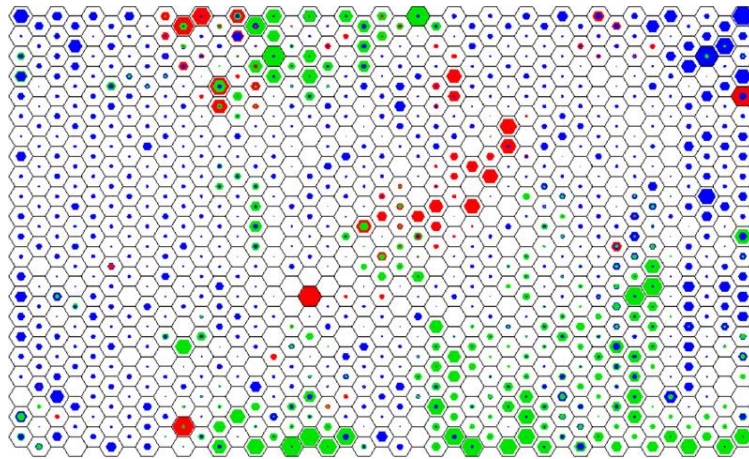


Figura 4.19. SOM 2D que muestra el espacio de parámetros: en azul las muestras no disruptivas de DB 1; en verde las muestras disruptivas de DB 1; en rojo las muestras en los últimos 160 ms de las MA de DB 2 [Cannas, 2010].

Para cada clúster, la densidad de color es proporcional al número de muestras contenidas dentro de los grupos. Este mapa destaca que el mayor número de MA se debe a múltiples razones. En primer lugar, un grupo de MA que aparece en el centro del mapa, pertenece a una nueva combinación de variables de entrada no presente en el conjunto de entrenamiento. Un análisis separado mostró que  $|q_{95}|$  y  $l_i$  son, respectivamente más grandes y más pequeños que los valores actualmente conocidos. En segundo lugar, aparecen grupos de alarmas perdidas (MA) en la parte superior derecha del mapa, donde se agrupan muestras no disruptivas; lo que sugiere que las variables de entrada utilizadas en [Cannas, 2010] no permiten discriminar completamente entre estados plasmáticos no disruptivos y disruptivos. En tercer lugar, varias MA ocupan en la parte superior central parte del mapa SOM, regiones del espacio de entrada ya poblado por muestras disruptivas

del conjunto de entrenamiento; lo que hace pensar en el ineficiente aprendizaje previo de la NN.

La capacidad de predicción del predictor neuronal también ha sido comparada con el rendimiento de las alarmas de modo bloqueado (LM). Dicha comparativa se ilustra en la Tabla 4.12.

**Tabla 4.12.** Rendimientos de predicción en descargas disruptivas de DB 2 para LM y para el predictor NN [Cannas, 2010].

	PD (%) $t_{\text{Alarma}} < (t_{\text{Disr}} - 160 \text{ ms})$	TD (%) $(t_{\text{Disr}} - 2 \text{ ms}) < t_{\text{Alarma}} < t_{\text{Disr}}$	MA (%)	SP (%) $t_{\text{Alarma}} < (t_{\text{Disr}} - 2 \text{ ms})$
Alarmas LM (81)	16.05	4.94	55.56	23.46
NNP (81)	9.87	7.41	18.52	64.20

Usando el activador LM, la alarma se manifiesta cuando se ha excedido el valor de un cierto umbral. Los valores de la Tabla 4.12. destacan la superioridad de rendimiento del NNP con respecto al activador LM. Cabe señalar que 10 de las 15 MA del activador LM también se pierden en el predictor neuronal. Además, para las 26 descargas caracterizadas por perfiles de corriente con valores altos de  $q_{95}$  ( $|q_{95}| > 5$ ), la alarma del activador LM falló 24 veces, mientras que en el predictor neuronal la alarma falló solo tres veces.

Se define el tiempo de disrupción como la diferencia entre el instante en el que tiene lugar la disrupción  $t_D$  y el instante de predicción de la alarma. La Figura 4.20. muestra el tiempo hasta la disrupción del activador LM con respecto al valor absoluto de  $q_{95}$ . Los tiempos de disrupción se informan por separado en las sub gráficas que se refieren a la detección prematura, predicción exitosa, alarmas tardías y alarmas perdidas. Como se puede observar, para los perfiles de corriente con alto valor de  $q_{95}$  ( $|q_{95}| > 5$ ), el activador LM concentra una alta tasa de alarmas perdidas.

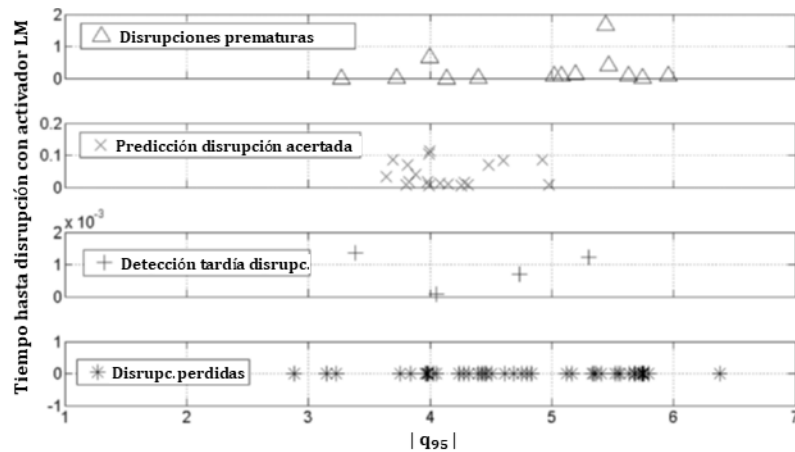


Figura 4.20. Intervalo de tiempo ( $t_D - t_{LM}$ ) entre la ocurrencia de la disrupción y el activador LM para las descargas en DB 2, frente a los valores de  $|q_{95}|$  [Cannas, 2010].

La Figura 4.21. proporciona la misma información que la Figura 4.20., pero referida al predictor que utiliza redes neuronales (NNP); donde para valores altos de  $q_{95}$ , NNP solo perdió tres alarmas.

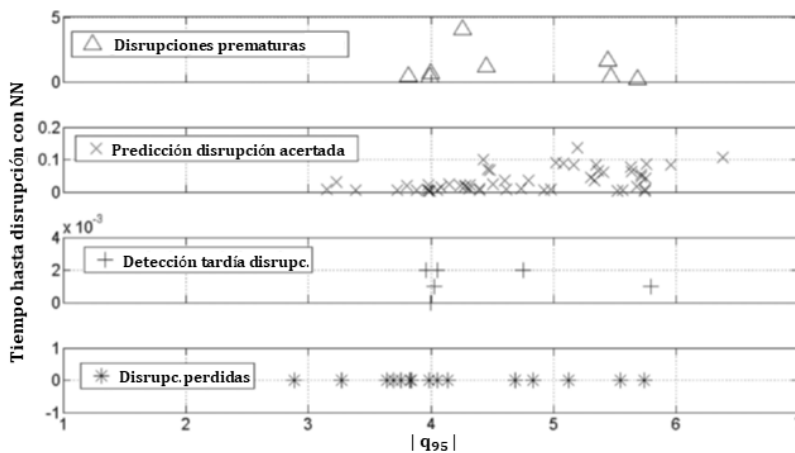


Figura 4.21. Intervalo de tiempo ( $t_D - t_{LM}$ ) entre la ocurrencia de la disrupción y el activador NNP para las descargas en DB 2, frente a los valores de  $|q_{95}|$  [Cannas, 2010].

En resumen, el predictor neuronal propuesto, demostró ser una herramienta válida para ser utilizado en un sistema de control en tiempo real. Teniendo en cuenta que es independiente de la señal LM, puede ser utilizado en el sistema de protección AUG como complementario al activador LM.

Para mejorar la capacidad de extrapolación del predictor neuronal, que es intrínsecamente pobre cuando se encuentra en regiones del espacio de parámetros en las que el predictor no ha sido entrenado, se debe proporcionar más conocimiento sobre estas



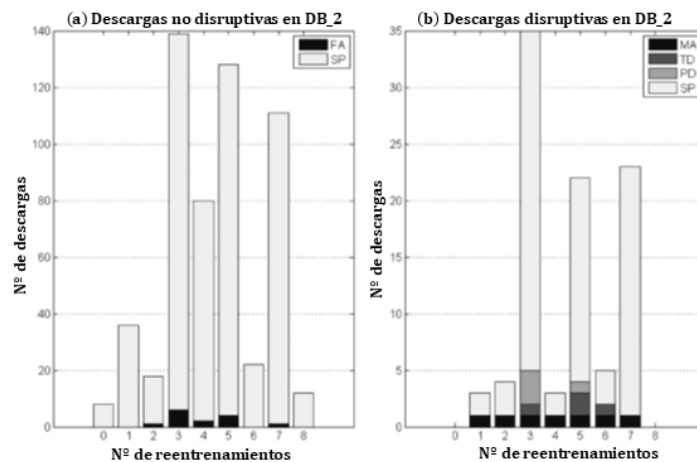
regiones. Esto puede realizarse ocasionalmente reentrenando el modelo neuronal, por ejemplo, cada vez que se produce una respuesta incorrecta (FA, PD, TD o MA). En [Cannas, 2010] el procedimiento se realiza solo cuando es activada una alarma perdida (MA) por el predictor. Durante el reentrenamiento, el modelo puede aprender de forma incremental a partir de nuevos datos, ya que el nuevo espacio operativo puede ampliarse con respecto a las campañas anteriores. Existen muchas propuestas para algoritmos incrementales de reentrenamiento [Tanprasert, 2002], siendo uno de los enfoques típicos el adoptado en [Polikar, 2001]. Este enfoque implica descartar la red existente y reentrenar la red utilizando los datos que han sido recogidos hasta ahora. Se ha simulado la aplicación en línea del predictor adaptativo durante las campañas experimentales de DB 2. Cada vez que el predictor pierde una alarma en una descarga disruptiva, se realiza un procedimiento de reentrenamiento, agregando la correspondiente descarga al conjunto de entrenamiento. Este procedimiento también proporciona al predictor nuevos conocimientos sobre la fase no disruptiva; debido a la presencia de las muestras de la fase no disruptiva en descargas que si son disruptivas.

Cuando el predictor pierde la  $i$ -ésima alarma, se realiza el procedimiento del  $i$ -ésimo reentrenamiento, utilizando los siguientes conjuntos de datos:

1.  $i$ -ésimo conjunto de entrenamiento: el conjunto de entrenamiento de DB 1 ampliado con todas las descargas que dan lugar a MA.
2.  $i$ -ésimo conjunto de validación: el conjunto de validación de DB 1 y como máximo 14 descargas disruptivas de DB 2, que preceden a la  $i$ -ésima MA, sin incluir las descargas que ya se hayan incluido en el conjunto de entrenamiento.
3.  $i$ -ésimo conjunto de prueba: las descargas disruptivas y no disruptivas en DB 2 que tienen lugar entre la  $i$ -ésima y la  $(i + 1)$ -ésima MA. Esta prueba proporciona la información sobre el rendimiento del predictor adaptativo, durante la operación en tiempo real. De hecho, las descargas de este conjunto son aquellas vistas de forma efectiva por el predictor adaptativo antes de que se actualice debido a un MA.

Después de que se haya realizado un nuevo reentrenamiento, es posible que el viejo predictor tenga un mejor rendimiento que el nuevo. Nuevamente, para simular el funcionamiento en tiempo real del predictor, la elección entre las dos redes se puede hacer solo usando descargas de plasma disponibles en el momento actual. Para este propósito, las dos redes se prueban en otro conjunto de pruebas, que contiene todas las descargas

disruptivas en el conjunto de prueba de DB 1, y todas las descargas disruptivas de entrenamiento y validación de DB 2 que se encuentren disponibles en el momento actual; siendo seleccionada como red actual aquella con el mejor rendimiento en esta prueba. La Figura 4.22. resume los resultados obtenidos utilizando la simulación en tiempo real del predictor adaptativo.



**Figura 4.22.** Histograma apilado del rendimiento del sistema adaptativo en relación con los sucesivos procedimientos de reentrenamiento, para las descargas de DB 2 [Cannas, 2010].

El gráfico de columnas apiladas en la Figura 4.22. (a) muestra la distribución de las respuestas del predictor a descargas no disruptivas durante el tiempo real de operación: en gris claro se representan predicciones correctas, mientras que en gris oscuro se representan las FA. El gráfico de columnas apiladas en la Figura 4.22. (b) muestra la distribución de las respuestas del predictor a descargas disruptivas: gris claro representa predicciones correctas, mientras que las grises más oscuras representan PD, TD y MA, tal y como indican las leyendas del gráfico. Como se puede observar en los cuadros, el número de descargas en cada conjunto de prueba es muy diferente de una iteración a la otra, pero el número de predicciones incorrectas siempre es limitado. En particular, el sistema adaptativo solo pierde ocho alarmas, mientras que NNP perdió 15 alarmas. Por lo tanto, el sistema adaptativo solo tiene que ser reentrenado ocho veces. Sin embargo, el procedimiento de reentrenamiento selecciona una red con mejor rendimiento que la anterior solo tres veces. Además, el sistema adaptativo detecta prematuramente una alarma cuatro veces, y las alarmas tardías se detectan en cuatro ocasiones. El número de PD y TD de NNP fue de ocho y seis, respectivamente. El número de FA del sistema adaptativo disminuye a 14, mientras que las FA de NNP fueron 26.

La Tabla 4.13. muestra que la tasa de éxito de predicción del predictor adaptativo en descargas disruptivas es superior al 80%, y el predictor es muy robusto con respecto a las falsas alarmas.

**Tabla 4.13.** Rendimiento del predictor neuronal adaptativo en DB 2 [Cannas, 2010].

	FA (%)	PD (%) $t_{\text{Alarma}} < (t_{\text{Disr}} - 160 \text{ ms})$	TD (%) $(t_{\text{Disr}} - 2 \text{ ms}) < t_{\text{Alarma}} < t_{\text{Disr}}$	MA (%)	SP (%) $(t_{\text{Disr}} - 160 \text{ ms}) < t_{\text{Alarma}} < (t_{\text{Disr}} - 2 \text{ ms})$
Descargas disruptivas (81)	—	4.94	4.94	9.88	80.25
Descargas no disruptivas (540)	2.59	—	—	—	97.41
Descargas totales (621)	2.25	0.64	0.64	1.29	95.18

#### 4.2.5. Predictor de disrupciones en tiempo real para JET.

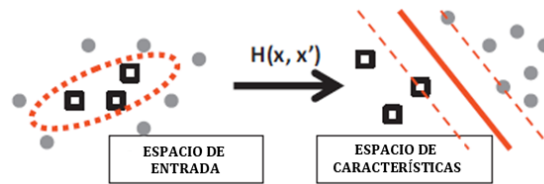
La sustitución de la pared anterior del dispositivo, compuesta por fibra de carbono (CFC), por una pared similar a ITER (ILW); donde la nueva pared es completamente metálica y más frágil, hace que el impacto de las disrupciones en JET tenga aún más repercusiones.

Un requisito previo importante para cualquier método de mitigación, es que el predictor de disrupciones en tiempo real sea confiable. El Predictor Anticipado de Disrupciones (APODIS), ha sido instalado en la red de JET para el reconocimiento de disrupciones en tiempo real. Aunque este predictor funciona en pared metálica (ILW), ha sido entrenado solamente con descargas pertenecientes a campañas experimentales con pared de carbono (CFC).

[Vega, 2013] muestra los resultados del predictor de disrupción APODIS en JET, que ha estado en funcionamiento en la red en tiempo real de JET durante las tres campañas iniciales de ILW (C28 – C30, entre agosto 2011 y julio de 2012). El objetivo ha sido evaluar sus capacidades de predicción, para ser usado en campañas posteriores como desencadenante en tareas de mitigación.

[Vega, 2013] se centra en el desarrollo de un clasificador binario en JET que permita distinguir entre dos tipos de comportamiento del plasma: disruptivo y no disruptivo. Para ello, dado un conjunto de datos cuya clase es conocida (conjunto de entrenamiento), la función de decisión debe determinar la clase a la que pertenecen nuevas muestras (conjunto de prueba) cuya clase está por determinar. Cuanto mayor sea la tasa de aciertos con el conjunto de prueba, más confiable será el clasificador. En general, una descarga se compone de un conjunto de muestras, y cada una de esas muestras puede ser descrita por un conjunto de características. A ese conjunto de

características de la muestra se le llama vector de características. La Figura 4.23. ilustra la clasificación binaria basada en máquinas de vectores soporte (SVM).



**Figura 4.23.** El conjunto de datos de entrenamiento está formado por círculos y cuadrados. La función de decisión (línea punteada roja) separa ambas clases. Con clasificadores SVM, la función de decisión en el espacio de características divide ambas clases por medio de un hiperplano (línea roja), donde el hiperplano se define por las muestras más cercanas de cada clase. Estas muestras aparecen en las líneas discontinuas y se denominan vectores soporte [Vega, 2013].

Los clasificadores basados en máquinas de vectores soporte (SVM) se tratan con mayor profundidad en [Betancourt, 2005], [Carmona, 2014] y su formulación matemática también se puede encontrar en [Cherkassky, 1998].

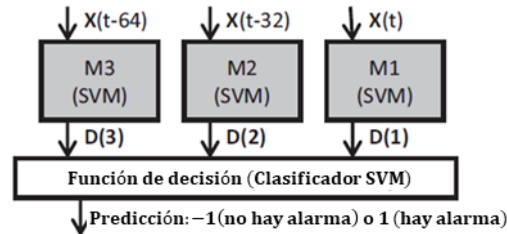
El proyecto APODIS comenzó en 2008 con el objetivo de desarrollar un predictor de disrupción para JET [Rattá, 2008] y [Rattá, 2010] que solventa los problemas de predictores anteriores. La primera versión de APODIS para JET realizó la simulación del procesamiento en tiempo real del predictor [Rattá, 2010]. La versión de APODIS a la que se refiere [Vega, 2013], para las campañas JET ILW, difiere de la primera en seis aspectos: aplicabilidad en tiempo real, la cantidad de descargas considerada para el entrenamiento, las señales (véase Tabla 4.14.), las representaciones de señales, un exhaustivo procesamiento previo de datos para eliminar descargas con señales atípicas y el uso de computación de alto rendimiento.

**Tabla 4.14.** Lista de señales que determinan el estado del plasma en JET [Vega, 2013].

Nombre de la señal	Unidades
Plasma current	A
Mode lock amplitude	T
Total input power	W
Plasma internal inductance	
Plasma density	$m^{-3}$
Stored diamagnetic energy time derivative	W
Radiated power	W

Las señales utilizadas por APODIS para identificar un comportamiento disruptivo son más o menos las mismas que las utilizadas en predictores anteriores. Sin embargo,

APODIS ofrece una novedad importante en relación con otros predictores de disrupción. Sigue una arquitectura de varios niveles (Figura 4.24.) en la que los tres clasificadores de primer nivel se combinan con otro clasificador en el segundo nivel.



**Figura 4.24.** Arquitectura de varios niveles de APODIS. La función de decisión está determinada combinando los modelos M1, M2 y M3. La predicción final identifica un comportamiento no disruptivo (-1) o un comportamiento disruptivo (1) [Vega, 2013].

Las unidades básicas APODIS son ventanas temporales de 32 ms. Todas las señales utilizan un período de muestreo de 1 ms, por lo que cada ventana temporal contiene 32 muestras. Ha sido demostrado en [Rattá, 2012] que la combinación de diferentes representaciones de señal en la arquitectura APODIS, permite reducir el número de señales necesarias.

El proceso de entrenamiento APODIS se describe completamente en [Rattá, 2010] sin embargo, es importante resumir cómo funciona APODIS durante el procesamiento de una descarga. Una vez que la corriente del plasma llega al umbral de 750 kA, se forman ventanas temporales de 32 ms cada una. Teniendo en cuenta que intervienen 7 señales y que cada una de ellas recoge dos características, cada una de las muestras queda registrada en el vector característico  $\mathbf{X}$ , formado por 14 componentes (7 señales x 2 características que tiene cada señal). Esas dos características son: la desviación estándar del espectro de Fourier y el valor medio.

Después de completar una ventana de tiempo de 32 ms en un instante  $t$ , los tres vectores de características más recientes: ( $\mathbf{X}(t-64)$ ,  $\mathbf{X}(t-32)$  y  $\mathbf{X}(t)$ ), son utilizados por APODIS (Figura 4.24.) como entradas respectivas a M3, M2 y M1. La salida de cada clasificador proporciona el estado (disruptivo/no disruptivo) de su vector de características. Debido al hecho de que los tres clasificadores pueden no estar de acuerdo en sus respectivas predicciones, es necesario un clasificador adicional para combinar la clasificación individual en predicciones coherentes. Esto se logra mediante el clasificador binario basado en SVM de segundo nivel, que es el que proporciona la función de decisión, para identificar el estado del plasma en el instante  $t$ .

El predictor APODIS para la operación en pared metálica (ILW) ha tenido que ser entrenado/probado con descargas procedentes de campañas experimentales realizadas con pared de carbono. Para ello, han sido tomadas en cuenta un número total de 10845 descargas, de las cuales se realizó un filtrado de las mismas por razones de amplitud fuera de rango o por no resultar suficientemente informativas. Después de este filtro, el número de descargas disponibles fue de 8407: 7648 no disruptivas, 521 disruptivas no intencionadas y 238 disruptivas intencionadas. Es importante tener en cuenta que las disrupciones intencionadas no son válidas para entrenar a un predictor; ya que en este tipo de descargas el plasma evoluciona de manera segura hasta que sea provocada la disrupción por el sistema de control. La mayoría de las disrupciones intencionadas en JET se activan induciendo un evento de desplazamiento vertical repentino (VDE) o por un colapso radiactivo por exceso de gas o inyección de impurezas [de Vries, 2011]. En el caso de VDE, las disrupciones son eventos muy rápidos y no aparecen precursores "naturales" en las señales.

Dados los buenos resultados a lo largo del tiempo [Rattá, 2010] mediante la arquitectura APODIS, han sido considerados para la fase de entrenamiento dos posibles conjuntos de descargas:

- Campañas C15 – C18.
- Campañas C19 – C22.

Dada la diferencia entre el número de descargas no disruptivas (7648) frente al número de descargas disruptivas no intencionadas (521) se verificó en [Vega, 2013] que, un equilibrio de los datos durante la fase de entrenamiento, proporcionaba mejores tasas de predicción. Como el número de descargas disruptivas no intencionadas entre C19 y C22 era de 125, en [Vega, 2013] se optó por el entrenamiento de 125 descargas disruptivas y 100 descargas no disruptivas aleatorias de entre las 2312 descargas no disruptivas correspondientes entre C19 y C22.

Para evitar cualquier sesgo en la selección de pulsos seguros, han sido elegidos 50 conjuntos de datos de entrenamiento diferentes (la diferencia entre ellos está en la selección aleatoria de las descargas no disruptivas).

Se han calculado diferentes predictores con miles de muestras de entrenamiento. Para lograr estos cálculos en un período de tiempo razonable, ha sido utilizada computación de alto rendimiento (HPC). Típicamente en cada uno de los cálculos se han

utilizado: 128 nodos del Clúster CIEMAT HPC (240 nodos procesadores: 2 Xeon de cuatro núcleos 3.0 GHz, memoria RAM 16 GB), calculándose un total de 7500 predictores; siendo el tiempo de CPU de 900 h para los clasificadores de primer nivel y 0,5 h para la función de decisión.

Para probar los predictores han sido seleccionadas las descargas correspondientes a las campañas C23 – C27b; habiendo sido utilizadas 3578 descargas no disruptivas y 228 descargas disruptivas no intencionadas. Siendo la tasa de aciertos de la prueba del 93,42% (213/228), la tasa de alarmas perdidas del 6,58% (15/228) y la tasa de falsas alarmas del 5,11% (183/3.578).

La implementación APODIS en el JET RTDN (Real-Time Data Network) se realizó bajo el entorno de ejecución en tiempo real de aplicaciones multiproceso (MARTE) [Neto, 2010] en una arquitectura x86 de seis núcleos. APODIS se ha implementado utilizando dos hilos [López, 2012], donde el primero recoge las muestras de las fuentes de entrada, y el segundo recibe los datos del primer hilo, los organiza para encajar en la arquitectura de tres ventanas de la Figura 4.24. y evalúa los 4 clasificadores SVM.

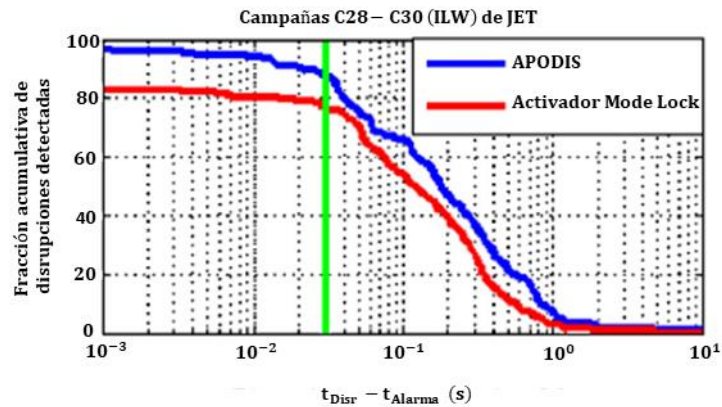
Como se ha mencionado, la versión APODIS instalada para operar en las campañas de pared metálica (ILW), ha sido entrenada con datos en pared de carbono, resumiéndose en la Tabla 4.15. sus resultados.

**Tabla 4.15.** Evaluación APODIS: tasa de aciertos, alarmas perdidas y falsas alarmas en las campañas JET C28–C30 [Vega, 2013].

	Clasificación off – line de JET	Predicciones APODIS
Descargas no disruptivas	651	645 (99.08%)
Falsas alarmas	n/a	6 (0.92%)
Disruptivas no intencionadas	305	300 (98.36%)
Alarmas perdidas	n/a	5 (1.64%)
Total	956	956
Disruptivas intencionadas	35	n/a

Cabe señalar que el conjunto de entrenamiento contiene múltiples tipos de disrupciones Por otro lado, con la pared metálica (ILW), la mayoría de las disrupciones pertenecen a un solo tipo que APODIS reconoce bien. Esta es la razón de las diferentes tasas de éxito entre el conjunto de pruebas. (93,4%) y los resultados ILW (99,1%). APODIS ha estado trabajando incluso con plasmas de alto rendimiento. (corriente de plasma 3,5 MA, potencia de entrada 25 MW, temperatura de electrones 1 keV y densidad  $2 \times 10^{20} m^{-3}$ ).

La Figura 4.25. muestra el porcentaje acumulado en las disrupciones no intencionadas, del tiempo de anticipación (instante donde se produce la alarma) con respecto al instante de la disrupción.



**Figura 4.25.** Comparación de los tiempos de predicción entre APODIS y el sistema activador de bloqueo del modo de JET [Vega, 2013].

Se establece una comparación. entre la predicción de APODIS y el sistema de activación de bloqueo del modo JET, donde el activador de bloqueo de modo se basa en un umbral aplicado en una señal magnética. APODIS supera claramente al anterior sistema de predicción JET. Cabe mencionar que el tiempo mínimo estimado para realizar acciones de mitigación en JET es de 30 ms (línea vertical verde). En ese instante, APODIS ha reconocido el 90% de todas las disrupciones no intencionadas, mientras que el sistema de activación de bloqueo de modo identifica casi el 80% de ellas. El valor medio del tiempo de predicción de APODIS es 426 ms antes de la disrupción (entre 1 ms y 10,323 s). Solo 13 disrupciones se han pronosticado con menos de 30 ms.

APODIS no es un sistema basado en la física. Se trata de un sistema de ingeniería capaz de aprender de situaciones pasadas, demostrando una alta capacidad de generalización. Dicha capacidad ha sido demostrada mediante los resultados obtenidos en campañas experimentales con pared metálica, habiendo basado el entrenamiento en campañas con pared de carbono. Esos resultados le confieren una alta confiabilidad como predictor, pudiéndose usar en futuras campañas de JET como activador en labores de mitigación.

Entre las ventajas que tiene el predictor APODIS sobre otros predictores previos, se encuentran:



- Opera en tiempo real, pudiendo seguir el comportamiento del plasma durante todo el periodo de procesamiento de la descarga.
- Se emplea el dominio del tiempo y de la frecuencia.
- Para cada predicción se tienen en cuenta los 96 ms más recientes (3 ventanas temporales de 32 ms cada una de ellas), teniendo en cuenta la evolución temporal del plasma.
- La combinación de clasificadores le proporciona un mejor conocimiento del espacio de parámetros.

#### 4.2.6. Predictor en tiempo real desde cero para JET.

Las primeras operaciones en JET con la nueva pared metálica del tipo ITER, constituyen una buena oportunidad para probar el desarrollo de nuevos predictores desde cero, basándose esas metodologías en la arquitectura APODIS; referenciado también por [Vega, 2013].

Las disrupciones se producen debido a la pérdida de estabilidad y/o confinamiento, que causa la terminación abrupta de la descarga [Wesson, 1989] y [Hender, 2007]. En las campañas de JET con pared metálica (C28 – C30, de julio de 2011 a junio de 2012), el número de descargas disruptivas no intencionadas ha aumentado con respecto a anteriores campañas [de Vries, 2012]. Tal y como se han presentado en otros trabajos anteriores, se han aplicado técnicas de aprendizaje automático, basadas principalmente en redes neuronales y máquinas de vectores soporte con el objetivo de predecir disrupciones [Yoshino, 2005], [Windsor, 2005], [Cannas, 2007b], [Murari, 2008], [Rattá, 2010], [Pautasso, 2001], [Cannas, 2007a], [Cannas, 2010] y [Vega, 2013]. Una vez que los sistemas fueron entrenados, los modelos eran utilizados para la detección de fenómenos disruptivos. Dichos fenómenos deben ser detectados en tiempo real, y además los predictores de disrupciones deben proporcionar, de forma simultánea, una alta tasa de aciertos con bajas tasas de falsas alarmas y alarmas perdidas. De ahí la importancia de predecir las disrupciones tan pronto como sea posible. El predictor APODIS activa las alarmas en JET con una anticipación media de 426 ms antes del instante de la disrupción. ITER necesita predictores de disrupción que sean confiables sin tener que esperar cientos de disrupciones. Por lo tanto, es fundamental el desarrollo de sistemas de predicción de disrupción, que consigan altas tasas de aprendizaje con pequeños conjuntos de entrenamiento; debiendo entrenarse predictores rápidos. El principal problema es la

selección de un conjunto suficientemente bueno de descargas, para lograr altas tasas de aciertos con mínimas tasas de falsas alarmas y de alarmas perdidas. En definitiva, se deben buscar predictores de disrupción que permitan resolver los siguientes objetivos:

- Identificación en tiempo real.
- Predicción temprana, para ser capaz de aplicar acciones de mitigación.
- Alta tasa de aciertos, baja tasa de falsas alarmas y baja tasa de alarmas perdidas.

[Dormido-Canto, 2013] se centra en el desarrollo de una metodología, para crear predictores de disrupción bajo las condiciones relevantes de ITER. Esas condiciones imponen que se resuelva el problema que comprende los tres objetivos vistos con anterioridad comenzando desde cero, sin ningún conocimiento previo sobre disrupciones.

En [Dormido-Canto, 2013] las metodologías se particularizan para tareas de clasificación binaria. El objetivo es crear un predictor desde cero, para poder clasificar el comportamiento disruptivo o no disruptivo del plasma en cualquier momento de una descarga. Dado el desbalanceo existente entre el número de descargas disruptivas y no disruptivas en cualquier dispositivo Tokamak, siendo mucho mayor el número de descargas no disruptivas; desde un punto de vista metodológico, la predicción de disrupciones desde cero requiere un análisis en profundidad para determinar la influencia de este hecho en el proceso de entrenamiento.

Teniendo en cuenta que los predictores de disrupción desde cero se van generando a medida que se producen las descargas, se debe tener en cuenta el orden cronológico de dichas descargas; permitiendo predictores adaptativos que estén aprendiendo continuamente. Para determinar cuándo debe aprender el sistema, se deben definir los criterios que posibilitan el aprendizaje de nuevos conocimientos.

En [Dormido-Canto, 2013] un primer criterio, ha sido generar un nuevo predictor justo después de la aparición de cada nueva descarga disruptiva. La razón para esto es incorporar cualquier información nueva que aparezca en las nuevas descargas disruptivas; utilizándose dicha información en la predicción de futuras disrupciones. Para explicar este criterio, [Dormido-Canto, 2013] considera el ejemplo de la Figura 4.26.



Figura 4.26. Ejemplo simplificado [Dormido-Canto, 2013].

En el ejemplo, está representado un conjunto de 50 descargas procesadas en orden cronológico. El ejemplo considera 45 descargas no disruptivas (rodeadas por círculos) y 5 descargas disruptivas (encerradas por cuadrados). Por motivos didácticos, las disrupciones han sido espaciadas uniformemente, aunque en la práctica, la metodología empleada es exactamente la misma; independientemente del número de descargas disruptivas o el número variable de descargas no disruptivas entre las que son disruptivas. A mayor número de descargas disruptivas, habrá que entrenar más veces; ya que el sistema debe aprender y adaptarse cada vez que se produce una disrupción. Si nos referimos al ejemplo de la Figura 4.26., se requerirá la generación de cinco predictores (uno por cada disruptiva, siendo estas descargas: 10, 20, 30, 40 y 50).

Según [Dormido-Canto, 2013], para conjuntos de datos de entrenamiento desbalanceados (diferencia entre el número de descargas disruptivas y no disruptivas), se emplean métodos de aprendizaje automático para resolver problemas de clasificación binaria. El grado del desbalanceo puede representarse por la relación entre el número de muestras de la clase minoritaria con respecto de la clase mayoritaria. Refiriéndonos a la Figura 4.26., es posible suponer que la proporción de las descargas minoritarias con respecto a las mayoritarias pueden ser aproximadamente de 1 a 10. Durante el último período de operación del dispositivo JET en pared de carbono, el desbalanceo fue menor (4% de descargas disruptivas frente a las no disruptivas) [de Vries, 2011]. Por otro lado, durante las primeras campañas de JET en pared metálica (ILW), el porcentaje de disrupciones aumentó significativamente [de Vries, 2012]. Por lo tanto, en la Figura 4.26., el grado de desbalanceo considerado es 1 a 10.

En [Dormido-Canto, 2013], la notación utilizada para identificar los clasificadores entrenados (modelos), se puede resumir de la siguiente manera:

- $d$  es el número de descargas disruptivas (DD).
- $d(i)$  es la posición ordinal para la  $i$ -ésima descarga disruptiva en la secuencia de todas las descargas (en la Figura 4.26. para  $i = 1 \Rightarrow d(1) = 10$ , para  $i = 2 \Rightarrow d(2) = 20$ , para  $i = 3 \Rightarrow d(3) = 30$ , para  $i = 4 \Rightarrow d(4) = 40$  y para  $i = 5 \Rightarrow d(5) = 50$ ).
- $M(i)$  es el modelo entrenado cuando se produce la  $i$ -ésima descarga disruptiva (en el ejemplo simplificado  $i = 1, 2, 3, 4$  y  $5$ ).
- $nd(i)$  es la posición ordinal para la  $i$ -ésima descarga no disruptiva en la secuencia de todas las descargas (en el ejemplo simplificado para  $i = 1 \Rightarrow nd(1) = 1$ , para  $i = 18 \Rightarrow nd(18)$  y así sucesivamente).
- $NDD(i)$  es el número de descargas no disruptivas (NDD) en el modelo  $M(i)$ . Por ejemplo, en el ejemplo simplificado de la Figura 4.26.,  $M(4)$  corresponde a un modelo compuesto por cuatro descargas disruptivas (números 10, 20, 30 y 40) y  $NDD(4) = 36$  descargas no disruptivas.

La Tabla 4.16. muestra los cinco modelos de entrenamiento teniendo en cuenta conjuntos desbalanceados según las descargas de la Figura 4.26.

**Tabla 4.16.** Modelos de entrenamiento con conjuntos de datos desbalanceados para el ejemplo simplificado de la Figura 4.26. [Dormido-Canto, 2013].

Modelo Entren.	Descargas	$i$	$NDD(i)$
$M(1)$	DD: {10} NDD: {[1, 9]}	1	9
$M(2)$	DD: {10, 20} NDD: {[1, 9] $\cup$ [11, 19]}	2	18
$M(3)$	DD: {10, 20, 30} NDD: {[1, 9] $\cup$ [11, 19] $\cup$ [21, 29]}	3	27
$M(4)$	DD: {10, 20, 30, 40} NDD: {[1, 9] $\cup$ [11, 19] $\cup$ [21, 29] $\cup$ [31, 39]}	4	36
$M(5)$	DD: {10, 20, 30, 40, 50} NDD: {[1, 9] $\cup$ [11, 19] $\cup$ [21, 29] $\cup$ [31, 39] $\cup$ [41, 49]}	5	45

La primera columna representa el modelo de entrenamiento.

La segunda, los números de descarga que pertenecen a cada modelo de entrenamiento; teniendo en cuenta las descargas disruptivas (DD) y las descargas no disruptivas (NDD). Por ejemplo, para el modelo  $M(1)$ , la única descarga disruptiva hasta el momento sería  $DD = \{10\}$ , y el número de descargas no disruptivas hasta el momento sería  $NDD = \{[1, 9]\}$  (conjunto de las nueve descargas comprendidas entre 1 y 9

inclusive). A medida que van surgiendo nuevos modelos, el número de descargas en DD y NDD va aumentando, de acuerdo a la notación mostrada en la Tabla 4.16.

La tercera columna muestra el número de descargas disruptivas que se han tenido en cuenta en el modelo de entrenamiento actual.

La cuarta columna muestra el número de descargas no disruptivas que intervienen en el modelo de entrenamiento actual.

Según [Dormido-Canto, 2013], en problemas de clasificación binaria, el entrenamiento es balanceado cuando el número de elementos es aproximado en ambas clases; pudiendo ser visto ese balanceo como garantía de equidad. Existen otros factores, como la complejidad, que también pueden afectar el rendimiento de los predictores; correspondiendo esa complejidad al nivel de separabilidad existente entre los datos. Para conjuntos de datos que son linealmente separables, los rendimientos de los clasificadores no son susceptibles al desbalanceo que pueda haber en cuanto a cantidad de datos [Japkowicz, 2002].

En el predictor adaptativo que se presenta en [Dormido-Canto, 2013], a medida que aumenta el desbalanceo, la complejidad es mayor. Como el número de descargas no disruptivas es significativamente mayor que el número de descargas disruptivas, el enfoque de entrenamiento balanceado requiere una selección aleatoria de descargas no disruptivas.

La Tabla 4.17. muestra los cinco modelos de entrenamiento teniendo en cuenta conjuntos de datos balanceados para el ejemplo simplificado de la Figura 4.26.

**Tabla 4.17.** Modelos de entrenamiento con conjuntos de datos balanceados para el ejemplo simplificado de la Figura 4.26. [Dormido-Canto, 2013].

Modelo Entren.	Descargas	i	NDD(i)
$M(1)$	DD: {10} NDD: 1-random from {[1, 9]}	1	1
$M(2)$	DD: {10, 20} NDD: 2-random from {[1,9] $\cup$ [11, 19]}	2	2
$M(3)$	DD: {10, 20, 30} NDD: 3-random from {[1, 9] $\cup$ [11, 19] $\cup$ [21, 29]}	3	3
$M(4)$	DD: {10, 20, 30, 40} NDD: 4-random from {[1, 9] $\cup$ [11, 19] $\cup$ [21, 29] $\cup$ [31, 39]}	4	4
$M(5)$	DD: {10, 20, 30, 40, 50} NDD: 5-random from {[1, 9] $\cup$ [11, 19] $\cup$ [21, 29] $\cup$ [31, 39] $\cup$ [41, 49]}	5	5

Tal y como resumía la tabla referente a los conjuntos con desbalanceo en los datos (Tabla 4.16.), las columnas primera y tercera contienen exactamente los mismos valores. También contienen los mismos valores para los números que representan las descargas

disruptivas (DD) en la segunda columna. Sin embargo, las descargas no disruptivas (NDD) que pasen a formar parte del conjunto de entrenamiento, han de ser seleccionadas de forma aleatoria de entre todas las descargas existentes hasta el momento y cuyo comportamiento sea no disruptivo; seleccionando aleatoriamente tantas descargas no disruptivas como número de descargas disruptivas existan en el conjunto DD. De esa manera, la cuarta columna tendrá el mismo número de representantes que la tercera.

Una vez obtenidos los modelos de entrenamiento, teniendo en cuenta balanceo y desbalanceo en el conjunto de descargas que definen los comportamientos disruptivos y no disruptivos, la fase de prueba de cada modelo se realiza con todas las descargas producidas con posterioridad a la fase de entrenamiento.

La Tabla 4.18. muestra los conjuntos de prueba utilizados para cada modelo entrenado según el ejemplo simplificado de la Figura 4.26.

**Tabla 4.18.** Conjuntos de prueba para el ejemplo simplificado de la Figura 4.26. [Dormido-Canto, 2013].

Modelo Entren.	Conj. Test
$M(1)$	$\{\{11, 50\}\}$
$M(2)$	$\{\{21, 50\}\}$
$M(3)$	$\{\{31, 50\}\}$
$M(4)$	$\{\{41, 50\}\}$
$M(5)$	—

La Tabla 4.18. muestra los conjuntos de prueba utilizados para cada modelo entrenado en el ejemplo que se viene utilizando. Las métricas de evaluación juegan un papel importante en el aprendizaje automático; utilizándose para evaluar y guiar los algoritmos de aprendizaje.

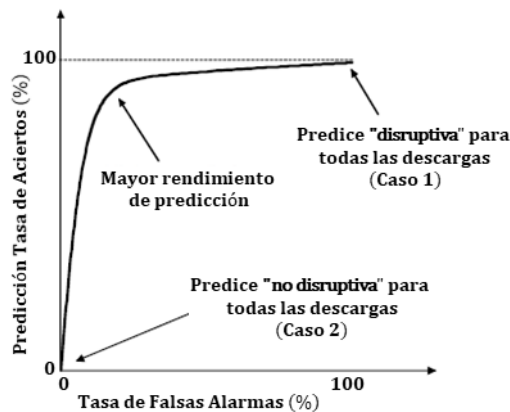
En el caso de conjuntos de datos desbalanceados, si se elige una métrica particular sin evaluar adecuadamente a la clase minoritaria, entonces los algoritmos de aprendizaje no podrán tratar de forma eficiente el problema del desbalanceo. En un conjunto de datos desbalanceado, la clasificación general (basada en la precisión) no es una métrica adecuada, ya que la clase minoritaria tiene menos efecto con respecto a la clase prevalente sobre la precisión [Weiss, 2003]. En [Dormido-Canto, 2013] un problema complejo de optimización es considerado como un multiobjetivo: conseguir alta tasa de aciertos en las predicciones de disrupción y a la vez que se consiga la minimización de las falsas alarmas; y todo ello teniendo en cuenta predicciones en tiempo real. Expresándolo de forma matemática, se podría resumir como en la expresión (3).

$$\text{Maximizar } \left\{ \frac{\text{Predicciones Disruptivas}}{N^{\circ} \text{ de Descargas Disruptivas}} \right\} \text{ y } \text{Minimizar } \left\{ \frac{\text{Predicciones Disruptivas}}{N^{\circ} \text{ de Descargas No Disruptivas}} \right\} \quad (3)$$

Pudiéndose considerar los siguientes casos extremos:

- **Caso 1.** Se predice la ocurrencia de disrupciones para todas las descargas, pudiéndose obtener una tasa de éxito de predicción del 100% junto a una tasa de falsas alarmas cercana al 100%.
- **Caso 2.** La aparición de disrupciones no está prevista para ninguna descarga, pudiéndose obtener una tasa de falsas alarmas del 0% junto a una tasa de éxito de predicción del 0%.

El rendimiento de la predicción se puede representar mediante la tasa de aciertos frente a la tasa de falsas alarmas, tal y como aparece reflejado en la Figura 4.27.



**Figura 4.27.** Rendimiento de predicción de la tasa de aciertos frente a la tasa de falsas alarmas [Dormido-Canto, 2013].

La Figura 4.27. muestra que un alto rendimiento general de predicción, vendría dado por una predicción positiva de tasa de aciertos de  $\sim 100\%$  y de forma simultánea, una baja tasa de falsas alarmas de  $\sim 0\%$ .

Las metodologías propuestas por [Dormido-Canto, 2013], para entrenar desde cero han sido implementadas con el APODIS, al que ya se hizo referencia en [Vega, 2013], y durante el periodo de publicación de [Dormido-Canto, 2013], se estaba ejecutando en la red de datos en tiempo real de JET [López, 2012] dentro del entorno de ejecución multiproceso de aplicaciones en tiempo real (MARTe) [Neto, 2010].

La Figura 4.28. muestra la arquitectura. de APODIS basado en dos capas de clasificadores consistentes en máquinas de vectores soporte (SVM). La primera capa

sigue la evolución temporal del plasma a través de tres clasificadores secuenciales SVM (M1, M2 y M3), y la segunda capa implementa un nuevo clasificador SVM para activar o no la alarma. La explicación del funcionamiento de dicha arquitectura ya fue realizada según [Vega, 2013].

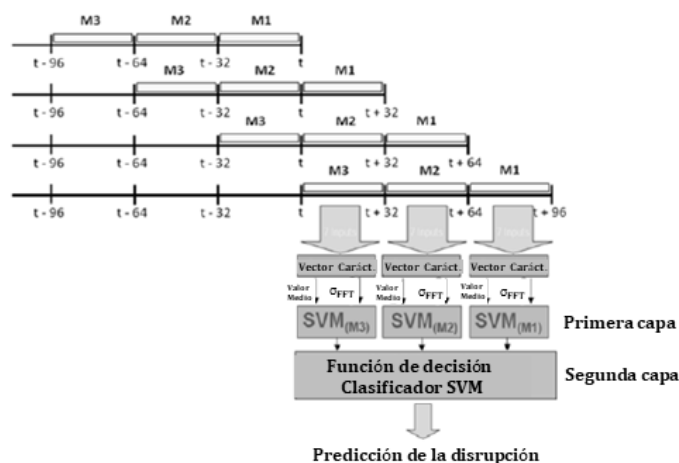


Figura 4.28. Arquitectura APODIS [Dormido-Canto, 2013].

APODIS se puede configurar para usar diferentes conjuntos de señales dependiendo de las necesidades. Tal y como se indicó en referencia a [Vega, 2013], estas señales son procesadas utilizando ventanas de tiempo de 32 ms con una frecuencia de muestreo de 1 kHz. Para cada ventana de tiempo, se calculan dos características por señal: (1) el valor medio y (2) la desviación estándar de la transformada rápida de Fourier (FFT) que elimina la frecuencia de componente nula.

Para demostrar la robustez y la generalización [Dormido-Canto, 2013] ha utilizado el conjunto de datos C28-C30, correspondiente a campañas experimentales en pared metálica (ILW) de JET. La Tabla 4.19. resume el número y el rango de descargas utilizadas, del que han sido excluidas descargas incompletas o con valores corruptos.

Tabla 4.19. Base de datos de campañas JET C28 – C30 [Dormido-Canto, 2013].

	Número de descargas	Rango descargas
Disruptivas	201	[81852, 83793]
No disruptivas	1036	
Grado desequilibrio	$\cong 1$ to 5	

En JET, se adquieren miles de señales en cada pulso. Aunque en JET se adquieren miles de señales por descarga, es fundamental la selección de las cantidades físicas más representativas para identificar adecuadamente la ocurrencia de una disrupción. Por un lado, demasiadas señales podrían sobrecargar la capacidad de aprendizaje. Por otro lado,



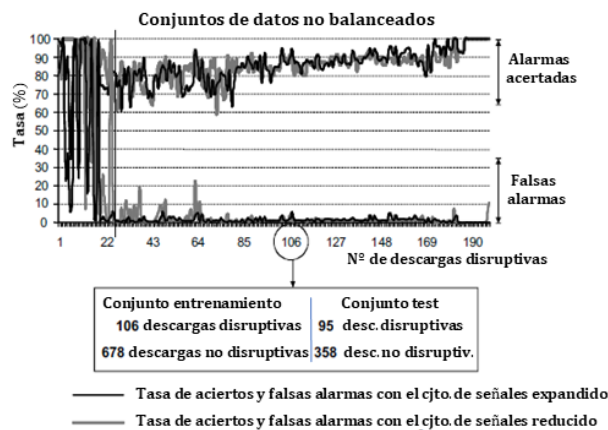
con pocas señales no se proporcionaría suficiente información para realizar un trabajo confiable en cuanto a predicciones. [Dormido-Canto, 2013] eligió doce señales para los modelos de entrenamiento y pruebas, que son recogidas en la Tabla 4.20.

**Tabla 4.20.** Lista de señales [Dormido-Canto, 2013].

Nº Id.	Nombre de la señal
(1)	Plasma current
(2)	Mode lock amplitude
(3)	Plasma internal inductance
(4)	Plasma density
(5)	Stored diamagnetic energy time derivative
(6)	Radiated power
(7)	Total input power
(8)	Poloidal beta
(9)	Plasma vertical centroid position
(10)	Plasma internal inductance time derivative
(11)	Poloidal beta time derivative
(12)	Plasma vertical centroid position time derivative

Todas o la mayoría de estas señales también se ha utilizado en investigaciones anteriores sobre disrupciones, con bastante buenos resultados [Cannas, 2007b], [Murari, 2008], [Rattá, 2010], [Rattá, 2008] y [Rattá, 2012]. En [Dormido-Canto, 2013] se considera un conjunto, que incluye todas las señales de la Tabla 4.20. Y también se considera un conjunto reducido, que abarca de las señales (1) a la (7); siendo este último conjunto el que emplea la arquitectura APODIS en la red de datos en tiempo real de JET.

La Figura 4.29. muestra la predicción de disrupción para el enfoque basado en un conjunto de descargas desbalanceado, tanto si se tienen en cuenta las 12 señales de la lista (representado por líneas negras), como si se consideran sólo las siete primeras señales de la lista (representado por líneas grises). El eje X de la Figura 4.29. representa el número de descargas a medida que los modelos van creciendo con la incorporación de nuevas descargas.



**Figura 4.29.** Resultados de predicción de disrupción desde cero con conjuntos de datos no balanceados [Dormido-Canto, 2013].

Por ejemplo, el valor 106 en el eje de abscisas, significa que el modelo 106 está compuesto de las primeras 106 descargas disruptivas (en orden cronológico) y todas las descargas no disruptivas que han ocurrido hasta ese instante (para el caso del ejemplo, exactamente 678 descargas no disruptivas). Cada clasificador se evalúa con el resto de las descargas que se producen después de que se genere dicho clasificador. Para el clasificador 106, el conjunto de prueba está formado por 95 descargas disruptivas y 358 descargas no disruptivas. A medida que se van incorporando nuevas descargas al conjunto de entrenamiento, aumenta el rendimiento del sistema; ofreciendo mayor tasa de aciertos y menor tasa de falsas alarmas. Hasta el modelo 24, las predicciones son bastante erráticas e inestables; por lo que se requiere más entrenamiento, dotando al sistema de mayor conocimiento del que actualmente dispone. A partir del modelo 24, donde se determina una frontera clara, los modelos ofrecen un mejor comportamiento; mostrando una tendencia positiva en la tasa de aciertos, junto con la disminución de la tasa de falsas alarmas. La Tabla 4.21. muestra las tasas promedio de aciertos y falsas alarmas con la desviación estándar del modelo 24 en adelante, con ambos conjuntos de señales: expandido (las 12 señales) y reducido (7 primeras señales de la tabla de señales 4.20.).

**Tabla 4.21.** Tasas promedio de éxito y falsas alarmas con la desviación estándar para conjuntos de datos no balanceados para los conjuntos de señales expandido y reducido. Los datos se calculan a partir del modelo 24

[Dormido-Canto, 2013].

	Tasa aciertos (%)	Tasa falsas alarmas (%)
Cjto. expandido señal.	85.99 ± 8.43	1.37 ± 1.19
Cjto. reducido señales	85.65 ± 7.78	2.27 ± 3.06

La Figura 4.30. muestra la predicción de interrupciones para el enfoque de descargas balanceado, teniendo en cuenta también el conjunto de señales expandido (representado por líneas negras) y el conjunto de señales reducido (representado por líneas grises).

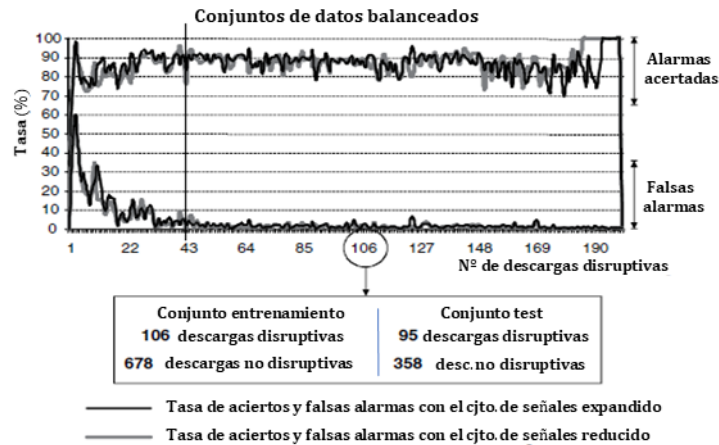


Figura 4.30. Predicción de disrupciones desde cero teniendo en cuenta datos balanceados [Dormido-Canto, 2013].

Como en el caso de datos desbalanceados, cada clasificador se evalúa con el resto de las descargas producidas después de que se genera dicho clasificador. A diferencia de lo que sucedió con los conjuntos de datos no balanceados, el enfoque actual (balanceado) obtiene mejores resultados en los primeros modelos en términos de estabilidad. Es decir, el número de oscilaciones se ha reducido significativamente destacando que, en términos de rendimiento, la tasa de aciertos y tasa de falsas alarmas desde el modelo 42 son muy estables.

La Tabla 4.22. muestra las tasas promedio de aciertos y falsas alarmas con la desviación estándar del modelo 42 en adelante, con ambos conjuntos de señales: expandido (las 12 señales) y reducido (7 primeras señales de la tabla de señales 4.20.).

Tabla 4.22. Tasas promedio de éxito y falsas alarmas con la desviación estándar para conjuntos de datos balanceados para los conjuntos de señales expandido y reducido. Los datos se calculan a partir del modelo 42 en adelante [Dormido-Canto, 2013].

	Tasa aciertos (%)	Tasa falsas alarmas (%)
Cjto. expandido señal.	87.63 ± 5.27	1.76 ± 1.01
Cjto. reducido señales	88.80 ± 5.41	1.48 ± 0.98

Teniendo en cuenta los resultados ofrecidos por la Tabla 4.21. y la Tabla 4.22., los conjuntos de datos desbalanceados muestran menor tasa de falsas alarmas y tasa de aciertos que el enfoque balanceado. Tratando de satisfacer simultáneamente una alta tasa de aciertos (como la obtenida con conjuntos de datos balanceados) y una baja tasa de falsas alarmas (como la obtenida con conjuntos de datos no balanceados), podría ser implementado un enfoque híbrido.

El enfoque híbrido se basa inicialmente en la generación de predictores con conjuntos de datos balanceados (para evitar las predicciones inestables y erráticas de los conjuntos de datos desbalanceados) y, con el objetivo de mantener la tasa de falsas alarmas tan baja como sea posible, cambiar a entrenamientos desbalanceados en un instante determinado. El punto de conmutación establece el número de descargas no disruptivas que deben ser utilizadas en los conjuntos de datos de entrenamiento. A partir de ese instante, los conjuntos de datos de entrenamiento estarán desbalanceados, siempre con el mismo número de descargas no disruptivas (elegidas al azar de la base de datos) y un número creciente de descargas disruptivas (cada vez que se produce una disrupción, se incluye dicha descarga disruptiva en el nuevo conjunto de datos de entrenamiento). Según la Figura 4.30., después del modelo 42, los resultados son bastante estables; siendo esta la cantidad de descargas no disruptivas para ser utilizadas en un enfoque híbrido. La Figura 4.31. muestra el rendimiento de un enfoque híbrido, donde los predictores están entrenados con conjuntos de datos balanceados hasta la disrupción número 42. A partir de ese momento, el entrenamiento será desbalanceado a favor de las descargas disruptivas que vayan surgiendo.

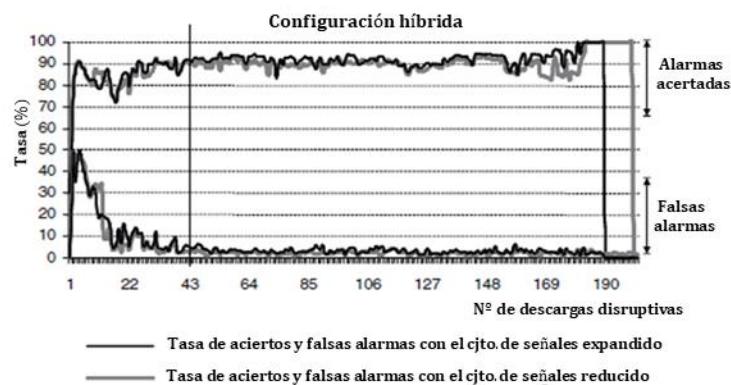


Figura 4.31. Resultados de predicción de disrupción desde cero con un enfoque híbrido [Dormido-Canto, 2013].

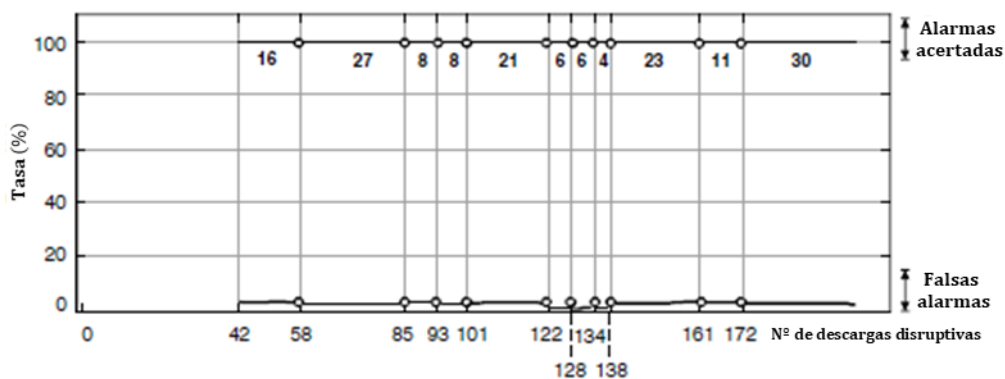
La Tabla 4.23. muestra las tasas promedio de aciertos y falsas alarmas con la desviación estándar del enfoque híbrido del modelo 42 en adelante, con ambos conjuntos de señales: expandido (las 12 señales) y reducido (7 primeras señales de la tabla de señales 4.20.).

**Tabla 4.23.** Tasas promedio de éxito y falsas alarmas con la desviación estándar para un enfoque híbrido con señal expandida y reducida. Los datos se calculan a partir del modelo 42 en adelante [Dormido-Canto, 2013].

	Tasa aciertos (%)	Tasa falsas alarmas (%)
Cjto.expandido señal.	93.03 ± 2.91	3.14 ± 1.14
Cjto.reducido señales	91.27 ± 3.99	2.25 ± 0.80

Con respecto a la Figura 4.31., la fuerte caída en la tasa de éxito en el extremo derecho de la figura debe interpretarse correctamente. Se deriva claramente del hecho de que los últimos modelos utilizan conjuntos de prueba muy reducidos. Por ejemplo, para el modelo 200, solo hay una descarga disruptiva en el conjunto de prueba y, por lo tanto, si falla la predicción, la tasa de aciertos cae al 0%.

La Figura 4.32. muestra la tasa de aciertos de un enfoque híbrido con el conjunto de señales ampliado, cuyo criterio para reentrenar el sistema se produce solo cuando se pierde una alarma disruptiva. Teniendo esto en cuenta para las campañas en pared metálica (ILW) de JET, sólo se producirían 10 reentrenamientos después de la disrupción 42; en lugar de los 159 que aparecen en la Figura 4.31.



**Figura 4.32.** Predicción de disrupciones desde cero con un enfoque híbrido. Los reentrenamientos se llevan a cabo después de perderse una disrupción, que sucede en las disrupciones 58, 85, 93 y así sucesivamente [Dormido-Canto, 2013].

En la Figura 4.32., cada reentrenamiento está representado por un círculo. Los números en la parte superior de la Figura 4.32., muestran el número de descargas disruptivas reconocido con éxito por cada predictor antes de perder una alarma. Ocurrieron un total de 160 disrupciones, y habiéndose reconocido 150 de ellas (93,75%). La tasa promedio de falsas alarmas es del 2,79%.

[Dormido-Canto, 2013] trata un método específico, para desarrollar predictores de disrupción desde cero. El término "desde cero" significa que el espacio de parámetros del Tokamak es completamente desconocido y, por lo tanto, al principio no hay información disponible para entrenar a un predictor. Esto sucede cuando una nueva máquina inicia su funcionamiento, por ejemplo, JET con pared metálica (ILW) o ITER. Las primeras campañas proporcionarán un excelente conjunto de datos para aplicar los predictores adaptativos que se describen en [Dormido-Canto, 2013]. A medida que aumenta el rendimiento y las disrupciones se vuelven más peligrosas, el aprendizaje adaptativo proporcionará un método óptimo para minimizar disrupciones y posibles daños a la máquina. El análisis de las tres primeras campañas en pared metálica (ILW) de JET muestra los buenos resultados de una metodología híbrida; teniendo en cuenta que se debe generar un nuevo predictor después de cada alarma perdida (no cuando se produce una nueva disrupción). Esto último permite evitar excesivos reentrenamientos, a la vez que obtiene excelentes tasas de rendimiento.

#### 4.2.7. Predictor APODIS con entorno MARTe.

[López, 2014] también describe la implementación en tiempo real de un predictor de disrupción basado en clasificadores SVM, pero esta vez realizado bajo el marco de la aplicación MARTe, con una arquitectura x86 de seis núcleos. El sistema está conectado a través de la red de datos en tiempo real de JET (RTDN).

Una vez más, se recalca que las disrupciones en los dispositivos Tokamak son inevitables, pudiendo tener un impacto significativo en la integridad de la máquina [Boozer, 2012]. El predictor APODIS ha sido diseñado para funcionar en el entorno de ejecución en tiempo real de aplicaciones multiproceso (MARTe) [Neto, 2010]. Actualmente, la salida de alarmas se almacena como una señal en la base de datos de JET; pero está lista para ser enviada al sistema secuenciador de protección en tiempo real (RTPS) [Stephen, 2011]. Como ya se ha indicado en referencias anteriores, APODIS se basa en una combinación de clasificadores SVM [Cherkassky, 1998], habiéndose discutido sus capacidades en [Rattá, 2010]; y mostrando APODIS unos rendimientos superiores que sus predecesores.

La implementación realizada en [López, 2014] utiliza siete señales para caracterizar el estado de plasma disruptivo/no disruptivo, siendo éstas las presentadas por [Vega, 2013]. Una vez más, en la Figura 4.33. se muestra la arquitectura del predictor

APODIS, pero esta vez se resaltan como entradas a cada uno de los clasificadores SVM de la primera capa, las 7 señales de entrada y las dos características por señal (valor medio de la señal y la desviación estándar de la transformada rápida de Fourier (FFT)). La Figura 4.33. difiere de la Figura 4.28. en las siete entradas que sí se indican en la Figura 4.33.

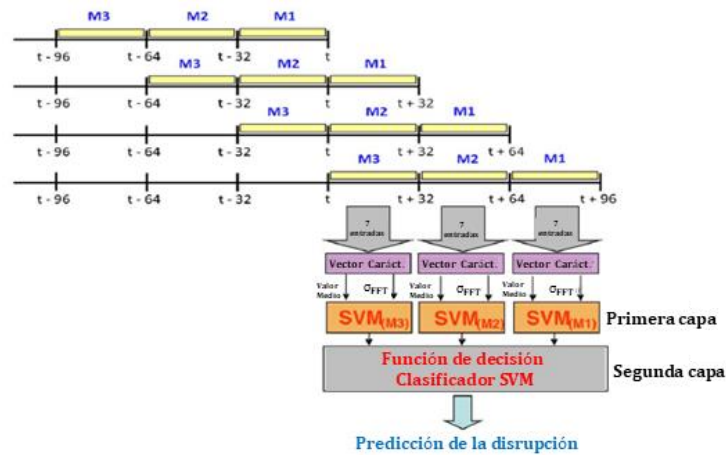


Figura 4.33. Arquitectura APODIS, resaltando las 7 señales de entrada a los clasificadores SVM [López, 2014].

Como se ha mencionado en referencias anteriores, la primera capa de APODIS está compuesta por tres clasificadores SVM con núcleos de función de base radial (RBF). La segunda capa también es un clasificador SVM, pero basado en un núcleo lineal [Burges, 1998]. El funcionamiento de la arquitectura APODIS ya fue referenciado en [Vega, 2013] y [Dormido-Canto, 2013]. El tamaño de 32 ms de la ventana, se elige para tener una potencia de 2 y así mejorar el tiempo de ejecución de la FFT [Ruiz, 2011].

Para filtrar el conjunto de datos a procesar, [López, 2014] ha considerado la Tabla 4.24.

Tabla 4.24. Rango de señales para selección de descargas [López, 2014].

Señal	Rango
Plasma density	$0 \leq \text{to} \leq 10^{19} \text{ m}^3$
Mode lock amplitude	$0 \leq \text{to} \leq 10^{-4} \text{ T}$
Plasma internal inductance	$-1 \leq \text{to} \leq 10^{-6} \text{ H}$
Poloidal beta	$-1 \leq \text{to} \leq 30$

De forma que, si alguna de esas cuatro señales presenta muestras fuera de los rangos de la Tabla 4.24., la descarga correspondiente sería descartada. Después de este proceso de selección, se consideran válidas 8407 descargas; distribuidas en la Tabla 4.25. de acuerdo al número de campaña experimental.

Tabla 4.25. Descargas seleccionadas [López, 2014].

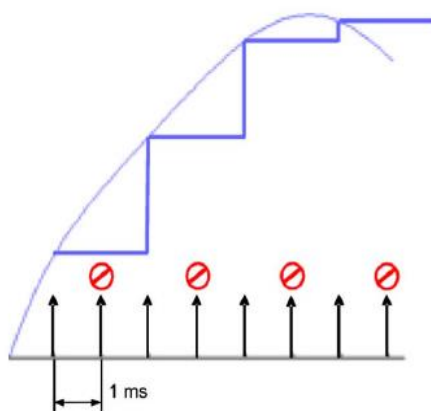
Campaña	Disr.No Intencion.	Disrupt. Intencion.	No Disruptivas	Total
C15a	21	6	256	283
C15b	10	0	158	168
C16	3	0	10	13
C17	73	0	1032	1105
C18	14	13	302	329
C19	47	41	585	673
C20	28	12	703	743
C21	16	3	573	592
C22	34	3	451	488
C23	24	8	490	522
C24	14	12	362	388
C25	19	22	570	611
C26	58	49	1323	1430
C27a	43	10	320	373
C27b	70	59	513	642
<b>Total</b>	<b>521</b>	<b>238</b>	<b>7648</b>	<b>8407</b>

Se deben manejar dos juegos de descargas disruptivas: intencionadas y no intencionadas. En descargas con disrupciones intencionadas, el plasma no muestra ningún comportamiento anómalo hasta el instante en el que se provoca la disrupción, eliminando este tipo de descargas del proceso de entrenamiento; pero si serán utilizadas en la fase de prueba. Para el conjunto de entrenamiento [López, 2014] ha considerado las campañas C15a-C22, seleccionado aleatoriamente 100 descargas no disruptivas de entre un grupo de 2312 posibles, y 125 descargas disruptivas no intencionadas. Los conjuntos de datos de prueba consisten en 3578 descargas no disruptivas y 228 descargas disruptivas no intencionadas procedentes de las campañas C23-C27b. La cantidad significativa de los datos involucrados en la fase de entrenamiento hizo necesario el uso de computación de alto rendimiento (HPC), habiéndose utilizado para este propósito el entorno HPC del CIEMAT, que está formado por 240 nodos de 2 procesadores Quad-Core Xeon (X5450 y X5570) en 3 GHz y con 16 GB de memoria RAM. A este tipo de computación de alto rendimiento ya recurrió [Vega, 2013].

El marco MARTe se utiliza para crear aplicaciones optimizadas en arquitecturas multinúcleo. Este marco ha sido utilizado en JET durante los últimos años, y ha demostrado ser una plataforma altamente configurable para desarrollar aplicaciones en tiempo real [Alves, 2011], [Alves, 2012], [Neto, 2012]. El marco se basa en una biblioteca C++ orientada en tiempo real, llamada BaseLib2. El desarrollador tiene que escribir el código software que implementa su propia aplicación, siendo MARTe básicamente un entorno de ejecución secuencial de módulos de aplicación genéricos (GAM) en un contexto en tiempo real. Los GAM se ejecutan de acuerdo con las transiciones de estado, y la comunicación entre los GAM se realiza escribiendo y leyendo señales del buffer



dinámico de datos (DDB). El modo de transferencia asincrónica (ATM) de JET se basa en la red de datos en tiempo real (RTDN), que es el medio de comunicación para todos los sistemas en tiempo real en JET. Todas las señales utilizadas por APODIS están disponibles en la RTDN. Las señales provienen de diferentes sistemas en tiempo real, siendo estos sistemas muy heterogéneos; y teniendo cada señal una frecuencia de muestreo diferente. Sin embargo, APODIS funciona a una frecuencia de muestreo de 1 kHz. Si la muestra no está disponible, entonces la RTDN proporciona la muestra más reciente. Esta situación es equivalente a un efecto de muestra y retención, que se muestra en la Figura 4.34.



**Figura 4.34.** Efecto en la lectura de la señal cuando no hay muestra disponible en la RTDN [López, 2014].

Donde las muestras no disponibles (a 1 ms de la muestra anterior existente) están marcadas en rojo. Para evitar discontinuidades, se proporciona una muestra con el valor en azul.

APODIS ha sido implementado usando dos hilos de aplicaciones. El primer hilo tiene dos funciones:

- a) Recolectar muestras para las fuentes de entrada.
- b) Calcular los vectores de características para las señales de entrada.

El segundo hilo recopila los datos del primer hilo, y los organiza para encajar en la arquitectura de tres ventanas.

La Figura 4.35. muestra los GAM, implementados de acuerdo con el marco de aplicación MARTE.

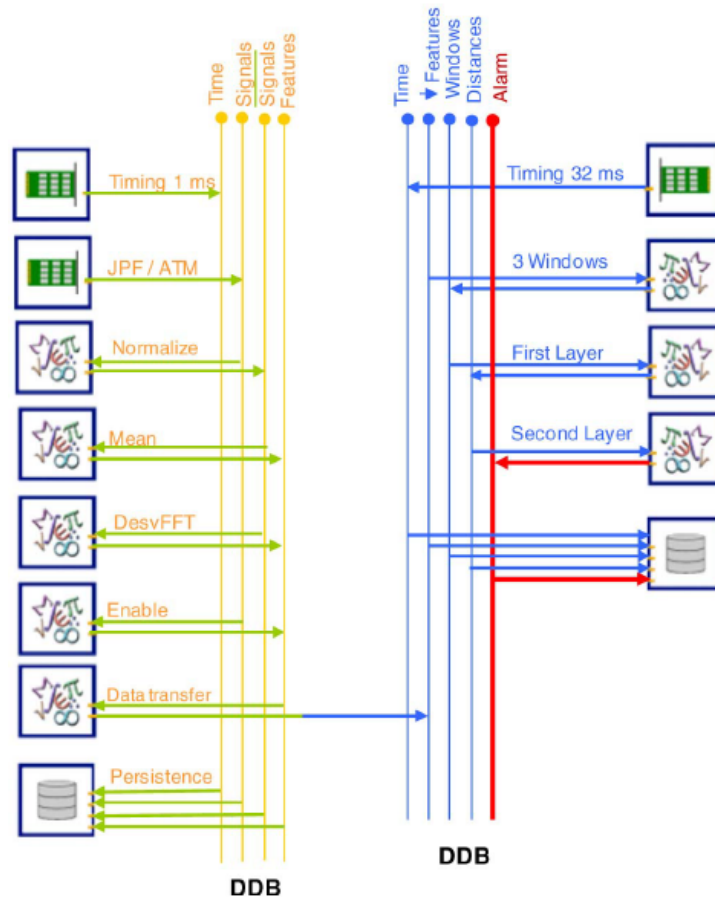


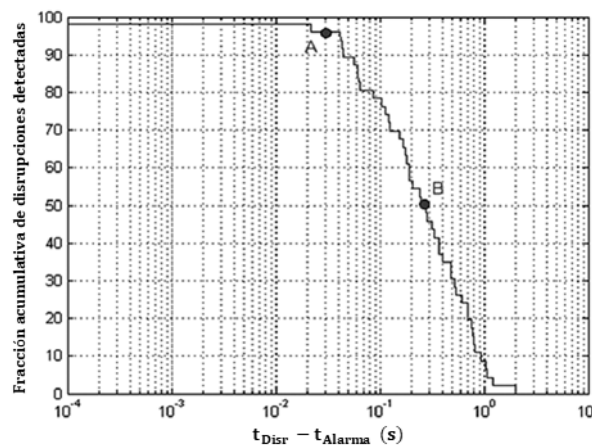
Figura 4.35. Arquitectura APODIS utilizando el marco MARTe [López, 2014].

1. JPF/ATM: este GAM hace los paquetes basados en el modo de transferencia asincrónica (ATM) o señales de archivo disponibles como señales del buffer dinámico de datos (DDB).
2. Normalize: normaliza las señales de entrada de acuerdo con el rango de entrenamiento utilizado.
3. Mean: calcula la media de 32 muestras.
4. Enable: monitorea si la corriente del plasma alcanza cierto umbral, para iniciar el predictor.
5. Data transfer: se utiliza para transferir los resultados de bloques de 32 muestras al segundo hilo.
6. 3 Windows: es una pila FIFO que almacena muestras, para formar las ventanas deslizantes.
7. Persistence: forman parte del marco MARTe, y se usan para almacenar los valores intermedios y los resultados generados por la aplicación en la base de datos JET.

8. El resto de GAM (DesvFFT, First Layer, Second Layer) obedecen a ecuaciones matemáticas, pudiendo consultarse en [López, 2014].

El sistema se ha implementado en una arquitectura x86 de seis núcleos con una tarjeta de interfaz de red (NIC) ethernet para administración remota, y una NIC en modo de transferencia asíncrona (ATM), que maneja todas las E / S en tiempo real dentro de la RTDN de JET.

Desde que APODIS se instaló en JET, en [López, 2014] se han analizado desde la descarga 82429 a la 82905 de la campaña C28. La Figura 4.36. muestra la fracción acumulativa de disrupciones detectadas con respecto al instante de disrupción. El sistema puede predecir el 97% de las disrupciones con una anticipación de 30 ms (punto A, en la Figura 4.36.), con una tasa de falsas alarmas del 5%.



**Figura 4.36.** Disrupciones acumulativas detectadas de APODIS versus tiempo de disrupción para descargas 82429 a 82905 en la campaña C28 [López, 2014].

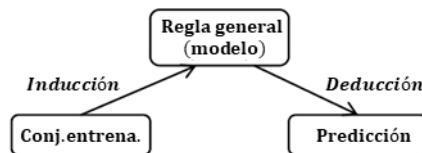
Se estima que 30 ms es un tiempo suficiente para emprender acciones de mitigación. Hay que tener en cuenta que el 50% de las disrupciones se pronostica con 280 ms de anticipación (punto B, en la Figura 4.36.). Otra consideración a tener en cuenta es que el sistema fue entrenado en pared de carbono (campañas C15a a C27b), mientras que en la campaña C28 la pared es metálica (ILW). Esta situación muestra que el procedimiento de diseño es muy robusto, y el predictor obtenido es preciso y confiable (alta tasa de aciertos en la predicción y baja tasa de falsas alarmas). Los resultados obtenidos muestran que esta implementación es válida, incluso en comparación con otras arquitecturas que usan implementaciones de hardware, utilizando FPGAs [Ruiz, 2011].

#### 4.2.8. Predictores adaptativos probabilísticos.

[Vega, 2014] analiza los requisitos para crear predictores adaptativos desde cero, y aprender de los datos de una máquina individual desde el comienzo de la operación; desarrollando un algoritmo basado en clasificadores probabilísticos.

El punto clave de cualquier predictor de disrupción desde cero (DPFS), es determinar la cantidad de disrupciones que son necesarias para tener un predictor confiable. [Vega, 2014] desarrolla un clasificador basado en predictores Venn [Vovk, 2005] desde cero, que comienza a hacer predicciones desde la primera disrupción. Los predictores Venn son clasificadores probabilísticos que determinan la probabilidad de cada predicción individual. Esto implica que no hacen predicciones simples, sino que además proporcionan un intervalo de probabilidad de cuán precisa y confiable es cada predicción. En términos de predictores de disrupción, esto significa que cada predicción disruptiva/no disruptiva se califica con un intervalo de probabilidad; que se usa como el nivel de confianza de la predicción. El predictor se ha aplicado a 1237 descargas de JET, correspondiente a las tres primeras campañas experimentales en pared metálica (ILW). Se ha vuelto a entrenar de forma adaptativa, después de cada alarma perdida, siguiendo el orden cronológico de las descargas. El predictor usa un número reducido de señales comunes del Tokamak, y los resultados muestran una tasa de éxito del 94%, una tasa de falsas alarmas de aproximadamente el 4% y un tiempo de anticipación promedio de 654 ms. La probabilidad promedio global de cada predicción es  $0.811 \pm 0.189$ , lo que proporciona un alto grado de confianza en el clasificador. En este punto, se debe especificar el significado de varios términos utilizados en [Vega, 2014] con respecto a descargas disruptivas. El término detección tardía se aplica a aquellas disrupciones que son reconocidas con un tiempo de anticipación inferior a 30 ms. El término de alarma prematura está reservado para predicciones cuya anticipación es superior a 1 s. La alarma válida se refiere a predicciones con tiempos de anticipación entre 30 ms y 1 s. El término tasa de aciertos se usa como la tasa de reconocimiento de comportamientos disruptivos, independientemente del número de alertas. Es decir, la tasa de éxito es la suma de la frecuencia de detección tardía, la frecuencia de alarma prematura y la frecuencia de alarma válida. Esta definición de la tasa de éxito es necesaria para poder comparar los resultados del predictor Venn con predictores anteriores desde cero ya publicados, que utilizan el término éxito como sinónimo de reconocimiento de comportamiento disruptivo.

En problemas de clasificación automática, una muestra es representada por un par ordenado  $(x_i, y_i)$ , donde  $x_i \in R^m$  es el vector de características (es decir, el conjunto de  $m$  características que caracterizan la muestra  $i$ -ésima, e  $y_i$  es su etiqueta correspondiente ( $y_i \in \{C_1, C_2, \dots, C_L\}$ , donde  $C_j, j = 1, \dots, L$  son las clases existentes). Para la predicción de la disrupción, los clasificadores son binarios. ( $L = 2$ ), que corresponden a las etiquetas "disruptivo" y "no disruptivo". De ahora en adelante, toda referencia a clasificadores significará clasificadores binarios, a menos que se indique lo contrario. Cualquier sistema de clasificación automática requiere un proceso de entrenamiento. En el caso de clasificadores supervisados, el proceso de entrenamiento conoce las etiquetas de las muestras. El proceso de predicción está compuesto por dos pasos: inducción y deducción, tal y como muestra la Figura 4.37.



**Figura 4.37.** Pasos que implican el proceso de clasificación: inducción y deducción [Vega, 2014].

El primero es una fase inductiva, mediante la cual se lleva a cabo el proceso de aprendizaje real. Las muestras de entrenamiento  $(x_i, y_i)$  con  $i = 1, \dots, N$ , se utilizan para obtener una regla general (también llamada modelo o función de decisión). A partir de ese momento, se utiliza el paso de deducción para hacer predicciones. Dada cualquier nueva muestra representada por el par  $(x, y)$ , con el vector de características conocido  $x$ ; pero con etiqueta desconocida  $y$ ; el objetivo es estimar el valor de la etiqueta. El vector de características se usa como entrada a la regla general, y la salida del modelo es la etiqueta predicha. Algunos métodos comunes utilizados para los pasos inductivos en problemas de clasificación supervisada son: redes neuronales artificiales [Bishop, 2004], máquinas de vectores soporte (SVM) [Cherkassky, 2007],  $k$  vecinos más próximos [Duda, 2001], mapas autoorganizados [Kohonen, 1998] y mapas topográficos generativos [Bishop, 1998] entre otros.

Después de terminar el proceso de entrenamiento, se utiliza un conjunto de datos de prueba con las etiquetas conocidas para estimar la calidad del clasificador. Los vectores de características del conjunto de datos de prueba se utilizan como entradas para el modelo y las predicciones se comparan con las etiquetas reales. Los vectores de características tienen que proporcionar información sobre el espacio de configuración del

problema, y el conjunto de componentes puede verse como una "parametrización" de tal espacio.

[Vega, 2014] remarca nuevamente que un predictor de disrupción es un requisito previo para cualquier sistema de mitigación. Esta característica fundamental determina la lista de requisitos generales, que debe cumplirse para cualquier DPFS basado en datos. Deben tenerse en cuenta algunos requisitos operativos como: aprender desde cero, operación en tiempo real, alta tasa de éxito, alta tasa de aprendizaje, reconocimiento temprano de disrupciones, baja tasa de falsas alarmas, efecto controlado de "envejecimiento", simplicidad del predictor, proceso de entrenamiento rápido y predicciones confiables.

*R1. Aprender desde cero:* los modelos basados en datos requieren un conjunto de muestras de entrenamiento con etiquetas conocidas, para distinguir entre diferentes clases. Cuanto mayor sea el número de muestras, mejor será el aprendizaje. Sin embargo, en ITER habrá una ausencia total de datos experimentales anteriores, y el entrenamiento debe realizarse de forma adaptativa, a medida que se van produciendo las descargas. Esta falta de información previa no solo se limita a los nuevos dispositivos de fusión, porque ante cambios estructurales importantes en dispositivos ya existentes, puede ser requerido ese tipo de entrenamiento adaptativo (como por ejemplo en JET, al sustituir la pared de carbono por pared metálica).

*R2. Operación en tiempo real:* hace referencia a una respuesta determinista, para satisfacer las necesidades de tiempo del dispositivo de fusión. En general, el paso de inducción de la Figura 4.37. puede ser muy costoso en términos de recursos computacionales, llevándose a cabo fuera de línea. Sin embargo, el requisito de operación en tiempo real está relacionado con el paso de deducción de Figura 4.37.

*R3. Alta tasa de éxito:* equivale a exigir una baja tasa de alarmas perdidas. Esto es consecuencia de las implicaciones negativas que una alarma perdida puede tener en un dispositivo. ITER necesita una tasa de éxito del 95%.

*R4. Alta tasa de aprendizaje:* el requisito anterior no es suficiente para un DPFS, que debe usarse desde el principio de la operación del dispositivo. El desarrollo de predictores con bajas tasas de aprendizaje es de bajo interés práctico. El requerimiento R3 debe lograrse en el menor tiempo posible, y esto significa que el predictor tiene que demostrar una alta tasa de aprendizaje.

*R5. Reconocimiento temprano de disrupciones:* este requisito es esencial en la perspectiva de las acciones de mitigación. Uno de los objetivos de un sistema de mitigación es minimizar el tiempo de reacción, y está claro que el tiempo de reacción tiene que ser menor que el tiempo de anticipación (véase Figura 4.38.).



**Figura 4.38.** Si el tiempo de reacción es mayor que el tiempo de anticipación, la acción de mitigación es tardía [Vega, 2014].

*R6. Baja tasa de falsas alarmas:* es necesario un compromiso entre el riesgo de disrupción y la disrupción de la operación, de tal manera que las falsas alarmas se minimicen.

*R7. Efecto controlado de "envejecimiento":* el efecto de "envejecimiento" en el contexto de un predictor de alta tasa de aprendizaje desde cero, significa que el predictor tiene que ser un predictor adaptativo desde el principio. Una estrategia continua de reentrenamiento es fundamental para mantener un sistema actualizado, capaz de funcionar correctamente en cualquier momento.

*R8. Simplicidad:* significa desarrollar lo más simple posible el clasificador, para distinguir en cualquier momento entre comportamientos disruptivos y no disruptivos. Para este fin, el espacio de configuración de la disrupción tendrá que caracterizarse con un número reducido de características que sean compatibles con las mejores capacidades de generalización.

*R9. Proceso de entrenamiento rápido:* debido a que es necesario un clasificador adaptativo para incorporar continuamente nueva información relevante, el proceso de entrenamiento debe ser lo suficientemente rápido como para permitir entrenamientos entre descargas cuando sea necesario.

*R10. Predicciones confiables:* cualquier proceso de entrenamiento con un número bajo de muestras es un problema. Por esta razón, cada predicción individual debe calificarse con una estimación de su fiabilidad.

En general, el nivel de idoneidad de un candidato como predictor desde cero, depende del cumplimiento de los 10 requisitos anteriores.

[Vega, 2014] describe el diseño y desarrollo de un DPFS particular. Su objetivo inicial es proporcionar un predictor de disrupciones que sea confiable, en situaciones en las que no existen antecedentes con información previa. Se trata de aprender de forma continua para comenzar las predicciones lo antes posible. En [Vega, 2014] el primer predictor incluirá características de descargas disruptivas y no disruptivas; de forma que ese primer predictor se desarrollaría teniendo al menos una descarga de cada tipo; además de tener en cuenta todos los requisitos generales (R1 – R10). Según [Vega, 2014], el requisito R10 (predicciones confiables) ha tenido un gran impacto; debido a la necesidad que supone la búsqueda de enfoques matemáticos, para calificar cada predicción con una medida de fiabilidad. Los predictores probabilísticos son particularmente adecuados para estos propósitos. Sin embargo, teniendo en cuenta que el primero de los predictores se basa en un conjunto muy reducido de descargas de entrenamiento, la probabilidad en la predicción podría resultar engañosa; siendo necesario determinar un intervalo de probabilidad para cada predicción. El intervalo de probabilidad proporcionará una idea clara sobre la eficiencia de la predicción. Cuanto más pequeño sea el intervalo de predicción, más eficiente será la predicción. [Vega, 2014] pone el ejemplo de que una clasificación con un intervalo de predicción [0, 1] es completamente informativo, y un intervalo de predicción [0.61, 0.62] es mejor que un intervalo de predicción [0.6, 0.7].

[Vega, 2014] considera un clasificador basado en la teoría de decisión de Bayes, donde se tendrían dos posibles clases de salida para el clasificador:  $w_1$  y  $w_2$ ; determinando estas clases el comportamiento disruptivo o no disruptivo de cada una de las muestras a clasificar  $(x, y)$ . Recuérdese que, para una muestra a clasificar, el vector de características  $x$  es conocido, pero la etiqueta  $y$  es desconocida.

$$\text{Regla de Bayes: } P(w_k / x) = \frac{p(x | w_k) P(w_k)}{p(x)}, \text{ con } k = 1, 2. \quad (4)$$

Esta fórmula de Bayes, expresada según la ecuación (4), da la probabilidad a posteriori de pertenencia de un vector de características conocidas  $x$  a la clase  $w_k$ . La función de distribución de probabilidad  $p(x | w_k)$  es la función de probabilidad de  $w_k$  con respecto a  $x$ .  $P(w_k)$  es la probabilidad a priori de la clase  $w_k$ , y  $p(x) = \sum_{i=1}^2 p(x | w_k) P(w_k)$ .

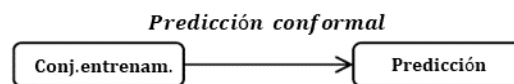
Para aplicar la regla de Bayes, debe ser conocida la probabilidad previa de cada clase. Sin embargo, en el caso particular de predictores de disrupción desde cero, la



estimación de la probabilidad es un problema; residiendo éste en el hecho de que no hay evidencias previas para el cálculo de la probabilidad. En estas circunstancias, se utilizan enfoques estimadores de densidad de probabilidad, siendo el más popular el método de ventana Parzen [Duda, 2001]; pero necesita un número mínimo de muestras para producir estimaciones confiables. Debido al hecho de que los primeros predictores no tendrán disponibles un número suficiente de muestras, en [Vega, 2014] han sido descartados los clasificadores bayesianos; descartándose también otros clasificadores probabilísticos porque no siempre proporcionan salidas bien calibradas [Lambrou, 2012].

[Vega, 2014] ha considerado finalmente los predictores Venn para DPFS, que son algoritmos de aprendizaje automático, y proporcionan un intervalo de predicción de probabilidad para cada clasificación; además de proporcionar salidas bien calibradas [Vovk, 2005].

Los predictores Venn pertenecen a la familia de predictores conformales [Vovk, 1999] y [Saunders, 1999]. Este tipo de predictores, a diferencia de otros métodos de aprendizaje de última generación, facilitan información sobre su propia precisión y fiabilidad. Los predictores conformales no siguen los pasos inductivos/deductivos de la Figura 4.37. Para hacer una predicción con una nueva muestra, se toma un acceso directo en la Figura 4.37., y el proceso de clasificación se mueve directamente de las muestras de entrenamiento a la predicción (véase Figura 4.39.).



**Figura 4.39.** En los predictores conformales todas las muestras de entrenamiento se usan con cada nueva predicción [Vega, 2014].

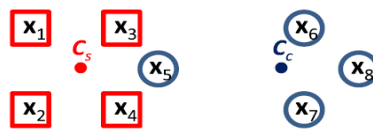
Centrando la atención en la formulación de un predictor Venn [Vovk, 2005], el punto de partida es el siguiente.

Sea  $\{z_1, \dots, z_{n-1}\}$  un conjunto de entrenamiento donde cada  $z_i \in Z = X \times Y$  es un par  $(x_i, y_i)$ , que consiste en la muestra o el vector de características  $x_i$  y su correspondiente clase  $y_i$ . Dado un nuevo vector de características  $x_n$ , el objetivo de un predictor Venn es estimar la probabilidad de pertenecer a una clase  $Y_j$  de un conjunto de clases  $J$ , con  $Y_j \in \{Y_1, \dots, Y_J\}$ .

La predicción Venn asigna el nuevo vector de características  $x_n$  a cada una de las posibles clases  $Y_j$ , dividiendo todas las muestras  $\{(x_1, y_1), \dots, (x_{n-1}, y_{n-1}), (x_n, Y_j)\}$  en

una serie de categorías. Para llevar a cabo esta división, los predictores Venn utilizan una función de taxonomía  $A_n$ , con  $n \in \mathbb{N}$ , que clasifica la relación entre una muestra y el conjunto de las otras muestras:  $\tau_i = A_n((x_1, y_1), \dots, (x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1}), \dots, (x_n, y_n))$ .

Los valores  $\tau_i$  se denominan categorías, y se toman de un conjunto finito  $T = \{\tau_1, \tau_2, \dots, \tau_T\}$ , asignando una función taxonómica a cada muestra  $(x_i, y_i)$  con su correspondiente categoría  $\tau_i$ , es decir, agrupando todas las muestras en un conjunto finito de categorías. Las taxonomías se analizan en [Papadopoulos, 2013], donde los predictores Venn se basan en redes neuronales. [Lambrou, 2012] muestra una aplicación de predictores conformales inductivos, para desarrollar un predictor inductivo Venn, con una taxonomía derivada de un clasificador SVM multiclase. [Nouretdinov, 2012] describe la taxonomía logística, que se crea a partir del método de probabilidad de regresión logística. En [Vega, 2014] la taxonomía seleccionada ha sido la que se basa en el centroide más próximo (NCT) [Dashevskiy, 2008]. Debido al hecho de que hay dos clases para el predictor de disrupción  $Y = \{Y_{disruptivo}, Y_{no-disruptivo}\}$ , el número de categorías en la taxonomía Venn NCT también es 2; que corresponde a los casos "disruptivo" y "no disruptivo". Esto significa que la NCT establece la categoría  $\tau_i$  de una muestra  $z_i$  igual a la etiqueta de su centroide más próximo (Figura 4.40.).



**Figura 4.40.** Las cuatro muestras más a la izquierda de este ejemplo tienen etiqueta "Cuadrado" y las cuatro más a la derecha tienen la etiqueta "círculo".  $C_s$  y  $C_c$  son respectivamente los centroides de las clases "cuadrado" y "círculo" [Vega, 2014].

La formulación matemática de la NCT, así como el procedimiento completo de clasificación se describen en la sección 6.2. de [Vega, 2014].

El uso de predictores Venn está justificado atendiendo a tres hechos principales:

1. No realiza suposiciones adicionales a las que son propias del aprendizaje automático típico sobre las muestras necesarias.
2. Los predictores Venn proporcionan un intervalo de probabilidad para cada predicción, en lugar de una sola probabilidad. Este intervalo de probabilidad se puede ver como una barra de error de probabilidad.

- Las predicciones están bien calibradas, lo que significa que la precisión cae entre el intervalo de probabilidad mínimo y máximo.

Para superar el obstáculo del tiempo de cálculo empleado, inicialmente se toma un conjunto limitado de características que concentren la información relativa a los comportamientos disruptivos y no disruptivos.

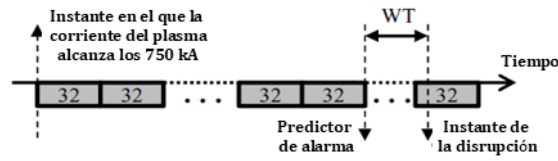
El primer proceso de entrenamiento del DPFS requiere al menos una muestra disruptiva y una no disruptiva. A medida que el sistema se va ampliando, el reentrenamiento del mismo sólo se hará cuando se produzca una alarma perdida. Para hacer una nueva predicción, los predictores Venn usan conjuntamente todas las muestras anteriores y la nueva muestra a clasificar. Por consiguiente, dependiendo del número total de muestras y del número de características que contenga cada muestra, el predictor Venn puede ser computacionalmente muy costoso; pudiendo ser un problema en lo relativo a necesidades computacionales en tiempo real.

Para satisfacer el requisito de simplicidad, se ha tomado un conjunto reducido de características en el DPFS, siendo las señales las mostradas en la Tabla 4.26. Se trata del mismo conjunto de señales que APODIS emplea en JET.

**Tabla 4.26.** Lista de señales para caracterizar el comportamiento disruptivo/no disruptivo en JET [Vega, 2014].

Nombre de la señal	Acronimo	Unidades
Plasma current	Ip	A
Mode locked amplitude	ML	T
Plasma internal inductance	LI	
Plasma density	Ne	m <sup>-3</sup>
Stored diamagnetic energy time derivative	dW/dr	W
Radiated power	Pout	W
Total input power	Pin	W

Tratando de aprovechar conocimientos previos de APODIS, donde APODIS ha demostrado que los valores medios (en el dominio del tiempo) y desviaciones estándar (en el dominio de la frecuencia), correspondientes a ventanas temporales de 32 ms, proporcionan una poderosa parametrización de la configuración de las disrupciones; el nuevo predictor Venn también usará ventanas temporales de 32 ms con períodos de muestreo de 1 ms para todas las señales. Evaluándose el comportamiento disruptivo/no disruptivo de la descarga cada 32 ms, desde el inicio del plasma hasta su extinción. Y teniendo en cuenta que la intensidad de corriente del plasma supere los 750 kA, tal y como muestra la Figura 4.41.



**Figura 4.41.** Ventanas temporales de 32 ms formadas desde el instante en que la corriente alcanza el umbral de 750 kA hasta que abandona dicho umbral [Vega, 2014].

Las predicciones se hacen cada 32 ms. En la prueba de una descarga disruptiva con una predicción también disruptiva, se calcula el tiempo de anticipación (WT) cuando el predictor activa la alarma. Si  $WT < 30$  ms, se trata de una detección tardía. Si  $30 \text{ ms} \leq WT \leq 1$  s, se trata de una alarma válida, y si  $WT > 1$  s, se trata de una alarma prematura.

En [Vega, 2014] los predictores de disrupción tratarán una serie de señales (entre 2 y 7), calculando el valor de dichas señales cada 32 ms, para predecir el comportamiento del plasma. El DPFS utiliza un enfoque balanceado de descargas disruptivas y no disruptivas en los conjuntos de entrenamiento. Dadas una serie de señales entre 2 y 7, el primer conjunto de entrenamiento estará compuesto por un solo vector de características disruptivas y un solo vector de características no disruptivas. La Figura 4.42. muestra el conjunto de vectores de características de 32 ms de longitud correspondientes a la primera descarga disruptiva.

...	...	...	...	
Caract.1	Caract.2	...	Caract.n	$[t_D - 96, t_D - 65]$
Caract.1	Caract.2	...	Caract.n	$[t_D - 64, t_D - 33]$
Caract.1	Caract.2	...	Caract.n	$[t_D - 32, t_D - 1]$
<b>D I S R U P C I Ó N (<math>t_D</math>)</b>				

**Figura 4.42.** Cada fila corresponde a un vector de n características durante la descarga [Vega, 2014].

Cada fila tendrá entre 2 y 7 características (columnas). Han sido ordenadas temporalmente hacia atrás desde el instante de disrupción  $t_D$ . La columna de la derecha muestra los intervalos de tiempo de las ventanas temporales. La fila que caracteriza el primer vector de características disruptivas en el primer "conjunto de entrenamiento" corresponde a la ventana de tiempo anterior a la disrupción (fondo gris en la Figura 4.42.). Esta selección se basa en el hecho de que todas las señales muestran los patrones morfológicos más claros de una disrupción entrante, justo en el segmento de tiempo más próximo a la disrupción.

La selección del primer vector de características no disruptivo para el primer "conjunto de entrenamiento" no es tan sencillo como el vector característico disruptivo. Al haber muchas más descargas no disruptivas que disruptivas (y, por supuesto, muchos más vectores de características no disruptivas que disruptivas), el objetivo es utilizar un vector de características capaz de condensar de alguna forma el carácter no disruptivo en un solo vector. Una elección aleatoria de un vector de características no disruptivo no garantiza la selección de un vector de características que represente un "comportamiento promedio" como comportamiento no disruptivo. El vector de características de "comportamiento promedio" se ha definido como el valor medio de todos los vectores de características de descargas anteriores no disruptivas, tal y como muestra la Figura 4.43.

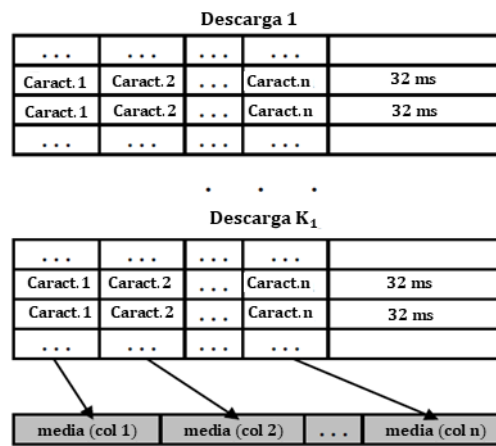


Figura 4.43. Vector característico promedio en descargas no disruptivas [Vega, 2014].

El ejemplo que describe un comportamiento no disruptivo en el primer conjunto de entrenamiento, es un vector fila cuyas componentes son los valores medios de las diferentes características entre la primera descarga no disruptiva y la última descarga no disruptiva previa a la primera disrupción.

Después de definir el primer "conjunto de entrenamiento", y una vez que se ha producido la primera disrupción, ya es posible hacer predicciones. De acuerdo con la estrategia de reentrenamiento, este "conjunto de entrenamiento" puede ser usado hasta la primera alarma perdida. Cuando esto sucede, ha de ser incorporado al predictor el nuevo conocimiento; sin perder de vista la utilización de un sistema balanceado de descargas disruptivas y no disruptivas. El nuevo conjunto de entrenamiento incluirá, además de los vectores de características procedentes del entrenamiento previo, conocimiento sobre la alarma perdida y las descargas no disruptivas del primer "conjunto de entrenamiento". La

información que se debe incluir en relación con la alarma perdida, es el último vector de características anterior a la disrupción; tal y como se muestra en la fila sombreada de la Figura 4.42. Con respecto al nuevo vector de características correspondiente al comportamiento no disruptivo, hay que añadir un vector de características de "comportamiento promedio", cuyas componentes son el valor medio de las componentes de todas las descargas no disruptivas desde el entrenamiento anterior hasta la última descarga no disruptiva previa a la disrupción perdida. Este criterio garantiza conjuntos de entrenamiento balanceados entre muestras disruptivas y no disruptivas. Sin embargo, es necesario tener en cuenta que este balanceo puede ser alterado bajo una condición específica. Si dos o más descargas consecutivas son disruptivas y el predictor no acierta en el reconocimiento, no hay descargas no disruptivas entre las descargas disruptivas contiguas y, por lo tanto, solo serán incluidos los vectores característicos de cada descarga disruptiva previa a las disrupciones.

Idealmente, los comportamientos disruptivos/no disruptivos deben condensarse en dos regiones del espacio de parámetros que están bien separadas entre ellas. Esta condición busca lograr suficiente capacidad de generalización con el predictor, independientemente del número de vectores de características en el conjunto de entrenamiento. Se trata de concentrar tanto como sea posible la información de ambos comportamientos en solo dos puntos, donde el predictor Venn usa la taxonomía del centroide más próximo (NCT).

En relación con el DPFS, es necesario enfatizar la importancia de hacer predicciones lo antes posible, es decir con el mínimo número de descargas. Teniendo en cuenta la decisión de usar "conjuntos de entrenamiento" balanceados, idealmente el primer predictor debe ponerse en funcionamiento después de la primera disrupción. Normalmente, cuando un Tokamak inicia la operación, se produce un número de  $k_1$  descargas no disruptivas antes de la primera descarga disruptiva. Estas  $k_1 + 1$  descargas se utilizarán para generar el conjunto de entrenamiento inicial. El predictor Venn se usa por primera vez con la descarga  $k_1 + 2$ . Durante esta descarga, se calculan cada 32 ms los vectores característicos para evaluar el comportamiento plasmático, y cada vector de características junto con el conjunto de entrenamiento inicial, se utiliza para hacer predicciones. El "conjunto de entrenamiento" inicial, con descargas tanto disruptivas como no disruptivas, sigue siendo válido hasta que no se produzca una nueva alarma. En

ese momento es necesario reentrenar al predictor, lo que significa generar un nuevo conjunto de entrenamiento.

En [Vega, 2014] se enfatiza que las selecciones hechas sobre las muestras de entrenamiento, el número de características y la taxonomía del centroide más próximo (NCT); permiten calcular cada predicción individual en menos de 32 ms. Esto es necesario para asegurar que no hay predicciones pendientes, cuando un nuevo vector de características está listo para ser clasificado. De hecho, el tiempo de cálculo para cada predicción tarda menos de 0.1 ms [Acero, 2014].

DPFS con NCT ha sido aplicado por [Vega, 2014] a una base de datos de 1237 descargas JET (1036 descargas no disruptivas y 201 descargas disruptivas no intencionadas), correspondientes a las tres primeras campañas experimentales JET después de la instalación de la pared metálica. El primer predictor se obtiene después de primera disrupción y, desde ese momento todas las descargas son analizadas en orden cronológico. Cada descarga es analizada simulando un procesamiento de datos en tiempo real. Los vectores de características son creados cada 32 ms y se clasifican como disruptivos o no disruptivos. Después de una alarma perdida, se crea un nuevo "conjunto de entrenamiento", para incorporar nuevos conocimientos, tal y como ya fue comentado con anterioridad. La Tabla 4.27. muestra las 14 características utilizadas con los predictores Venn.

**Tabla 4.27.** Identificación de características [Vega, 2014].

Caracterist.	Definición
1	mean (Ip)
2	std ( fft(Ip) )
3	mean (ML)
4	std ( fft(ML) )
5	mean (LI)
6	std ( fft(LI) )
7	mean (Ne)
8	std ( fft(Ne) )
9	mean (dW/dr)
10	std ( fft(dW/dr) )
11	mean (Pout)
12	std ( fft(Pout) )
13	mean (Pin)
14	std ( fft(Pin) )

Los acrónimos están relacionados con la Tabla 4.26., donde *mean* (.) representa el valor medio durante la ventana de tiempo de 32 ms, y *std* (|fft(.)|) significa la desviación estándar del espectro de Fourier durante el intervalo de tiempo de 32 ms. Habiendo sido probadas todas las combinaciones posibles entre 2 y 7 señales, se desarrollaron 9893 predictores. El primer predictor es generado después de la primera disrupción y se utiliza con todas descargas posteriores, sujetas a reentrenamiento después de cada alarma

pérdida. Una descarga se clasifica como no disruptiva cuando todos sus vectores de características se clasifican como no disruptivos. Además, una descarga se reconoce como disruptiva cuando alguno de sus vectores de características se clasifica como disruptivo. Por último, un solo vector de características clasificado como disruptivo en una descarga no disruptiva es suficiente para identificar el evento como una falsa alarma. La Tabla 4.28. muestra la tasa promedio de aciertos y las tasas de falsas alarmas correspondientes a los diferentes predictores desarrollados, con diferente número de características.

**Tabla 4.28.** Resultados de análisis de DPFS con NCT [Vega, 2014].

Caract.	$C_{14,Caract.}$	ASR (%)	AFAR (%)
2	91	96.68 ± 2.18	70.30 ± 30.68
3	364	96.79 ± 2.13	67.82 ± 31.37
4	1001	96.37 ± 1.97	58.19 ± 30.68
5	2002	96.34 ± 1.76	54.94 ± 27.81
6	3003	96.38 ± 1.56	52.97 ± 24.67
7	3432	96.47 ± 1.39	51.53 ± 21.51

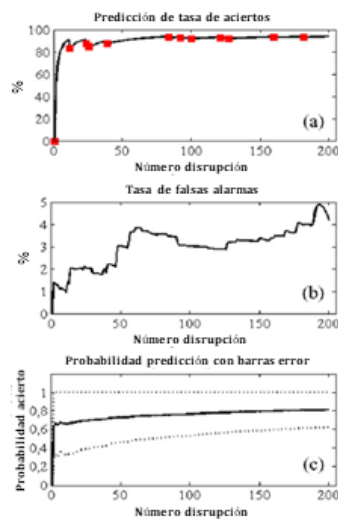
Donde *Caract.* representa el número de señales consideradas en los predictores en los predictores,  $C_{14,Caract.}$  representa el número de predictores desarrollados con 14 características tomadas *Caract.* en un instante, *ASR* es la tasa de promedio de aciertos correspondiente a todos los predictores con *Caract.* características, y *AFAR* es la tasa promedio de falsas alarmas correspondiente a todos clasificadores con *Caract.* características. A nivel global, las tasas de éxito están por encima del 96%, pero las tasas de falsas alarmas son demasiado altas. Esta tasa disminuye a medida que aumenta el número de características, pero no lo suficiente. Según estos resultados, la selección de un predictor específico no debe basarse únicamente en la tasa de aciertos, aunque sean porcentajes elevados en este caso; sino en una combinación de alta tasa de aciertos y mínima tasa de falsas alarmas. Cuanto menor sea la tasa de falsas alarmas, mejores serán los resultados. La Tabla 4.29. proporciona diferentes resultados en función de las características seleccionadas. Cada etiqueta de las columnas numeradas del 1 al 14, hace referencia a la identificación de las características que aparece en la Tabla 4.27.

**Tabla 4.29.** Resultados basados en distintas combinaciones de características [Vega, 2014].

Identificación de Característica														SR (%)	FA (%)	AVP	
1	2	3	4	5	6	7	8	9	10	11	12	13	14				
		x	x												94.00	4.70	0.813 ± 0.187
	x		x												92.50	5.09	0.831 ± 0.169
	x	x	x												94.00	4.31	0.809 ± 0.191
		x	x							x					94.00	4.70	0.813 ± 0.187
	x	x	x	x											94.00	4.21	0.811 ± 0.189
	x	x	x							x					94.00	4.21	0.810 ± 0.190
	x	x	x	x						x					94.00	4.21	0.811 ± 0.189
	x	x	x			x	x								94.00	4.21	0.803 ± 0.197
	x	x	x			x	x			x					94.00	4.21	0.803 ± 0.197
	x	x	x	x						x	x				94.00	4.31	0.810 ± 0.190
	x	x	x	x		x	x			x					94.00	4.31	0.802 ± 0.198
	x	x	x			x	x			x	x				94.00	4.31	0.802 ± 0.198



Donde SR representa la tasa de aciertos, FA representa la tasa de falsas alarmas y AVP representa la probabilidad media de predicción para distintas combinaciones de vectores de características. El primer resultado a destacar es el logro de una misma tasa de aciertos del 94% para casi todos los predictores. El límite más bajo de falsas alarmas es 4.21% y la probabilidad media de predicción es alta; siendo el valor mínimo de la barra de error de 0.604 (60,4%) y siendo el máximo siempre igual a 1 (100%). Esta alta probabilidad en la predicción da suficiente confianza sobre la fiabilidad de los predictores. El uso de clasificadores probabilísticos ha sido contemplado solo para validar los resultados de la predicción, siempre que las probabilidades de predicción sean lo suficientemente altas. Según la Tabla 4.29., los mejores candidatos a predictores son los que tienen tasa de falsas alarmas del 4,21%. Entre los cinco candidatos, y atendiendo al requisito de simplicidad (R8), son preferibles aquellos predictores con menor número de características, reduciendo los candidatos tan solo a dos. Finalmente, el mejor predictor será aquel cuyo intervalo de probabilidad sea menor, es decir, el predictor con menor probabilidad de error (el predictor con las características 2, 3, 4 y 5), porque proporciona una predicción más precisa. La Figura 4.44. corresponde a los resultados del predictor con características 2, 3, 4 y 5.



**Figura 4.44.** Resultados del predictor Venn con las características 2, 3, 4 y 5: evolución de la tasa de aciertos, tasa de falsas alarma y probabilidad de predicción según el número de disrupciones [Vega, 2014].

La Figura 4.44.(a) muestra la evolución de la tasa de aciertos a medida que aumenta el número de disrupciones. Es importante observar la rápida tasa de aprendizaje lograda en el proceso de aprendizaje, ya que se alcanza una tasa de éxito del 90,91% con un conjunto de entrenamiento compuesto por tan solo 2 descargas disruptivas y 2

descargas no disruptivas Los cuadrados rojos en la Figura 4.44.(a), muestran las alarmas perdidas.

La Figura 4.44.(b) representa la evolución de la tasa de falsas alarmas. La tasa de falsas alarmas tiende a aumentar ligeramente con el número de interrupciones

La Figura 4.44.(c) muestra la evolución de la probabilidad de predicción según va aumentando el número de interrupciones. A medida que la probabilidad de predicción aumenta, las barras de error disminuyen, haciendo más precisas las predicciones. La comparación de estos resultados obtenidos por [Vega, 2014] con respecto a los obtenidos con la versión APODIS desde cero y descritos en [Dormido-Canto, 2013]; es favorable para los predictores Venn. Los predictores Venn no necesitan decenas de interrupciones para obtener buenos rendimientos, como ocurría en el caso de APODIS desde cero [de Vries, 2011]; y representados en la Figura 4.45.

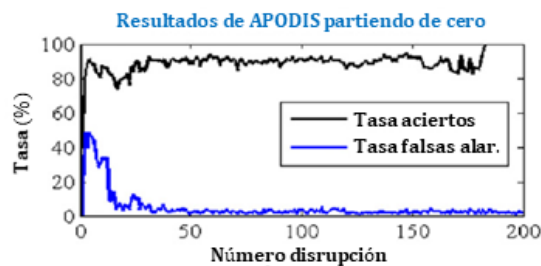


Figura 4.45. Resultados de la versión APODIS desde cero [Dormido-Canto, 2013].

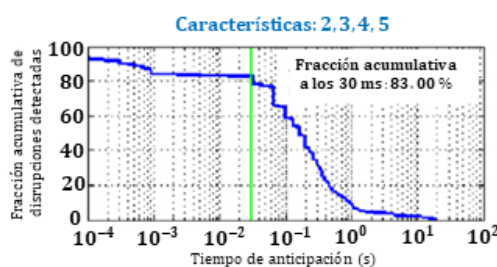
Cabe destacar las diferentes escalas en el eje Y entre la Figura 4.45. y la Figura 4.44.(b) [Vega, 2014].

La Tabla 4.30. identifica las características de la Tabla 4.27., AWT es el tiempo de anticipación promedio de los diferentes predictores y AF30 es la fracción acumulativa de interrupciones detectadas 30 ms antes de la interrupción.

Tabla 4.30. Tiempos promedio de anticipación AWT con predictores Venn [Vega, 2014].

Características	AWT (ms)	AF30 (%)
3, 4	654.564	83.0
2, 4	732.281	79.0
2, 3, 4	641.798	83.0
3, 4, 11	654.564	83.0
2, 3, 4, 5	654.394	83.0
2, 3, 4, 11	653.884	83.0
2, 3, 4, 5, 11	654.394	83.0
2, 3, 4, 7, 8	672.947	83.5
2, 3, 4, 7, 8, 11	672.947	83.5
2, 3, 4, 5, 11, 12	707.330	83.0
2, 3, 4, 5, 7, 8, 11	673.628	83.5
2, 3, 4, 7, 8, 11, 12	725.883	83.5

La comparación de estos tiempos de anticipación con los obtenidos por APODIS durante las mismas campañas experimentales, debe interpretarse en el sentido de que los predictores desde cero son mucho más sensibles al comportamiento disruptivo. La Figura 4.46. muestra los tiempos de anticipación obtenidos con el predictor Venn para las características 2, 3, 4 y 5 de la Tabla 4.27.



**Figura 4.46.** Tiempos de anticipación del predictor Venn con características 2, 3, 4 y 5 [Vega, 2014].

Tal y como se puede observar en la Figura 4.46., el 83% de las alarmas disruptivas se detectan con una anticipación de 30 ms, tiempo de antelación suficiente para realizar tareas de mitigación. La sección 8 de [Vega, 2014] recoge los detalles que explican estos resultados.

#### 4.2.9. Predictor basado en detección de anomalías (SPAD).

En [Esquembri, 2018] se presenta un nuevo tipo de predictor de disrupciones, basado en la detección de anomalías en la señal de modo bloqueado (SPAD). El análisis sobre las campañas (C28 -C34) en pared metálica (ILW) de JET muestra que SPAD pudo predecir el 83.57% de las disrupciones, con suficiente tiempo de antelación como para poder aplicar técnicas de mitigación; siendo la anticipación promedio de 389 ms.

[Esquembri, 2018] presenta la implementación de un predictor de disrupción en tiempo real basado en la detección de anomalías de señal. El predictor de señal única basado en detección de anomalías (SPAD), también conocido como predictor basado en detección atípica [Vega, 2015a] y [Vega, 2015b], aprende el comportamiento normal de una descarga desde el principio de la misma. SPAD usa una sola señal, señal de modo bloqueado (LM), y desencadena la alarma disruptiva cuando se detecta un comportamiento anormal en esta señal. SPAD presenta buenos resultados de detección, comparables con los predictores basados en el aprendizaje automático, pero sin la necesidad de emplear un proceso de entrenamiento. [Esquembri, 2018] describe una

implementación particular de SPAD, mediante el uso del entorno de ejecución de marco de aplicación multiproceso en tiempo real (MARTe) [Neto, 2010]. Este desarrollo permite que SPAD se integre completamente en la red de datos en tiempo real de JET (RTDN) [Felton, 1999].

Como se mencionó anteriormente, el predictor SPAD detecta el comportamiento anormal en una señal, para activar la alarma de interrupción. La señal LM está estrechamente relacionada con el comportamiento disruptivo, y el predictor muestra una buena relación entre anomalías en la señal y las interrupciones. Dados los buenos resultados de detección de APODIS, procesando bloques de 32 muestras adquiridas cada ms, SPAD sigue la misma metodología; pero procesa las últimas 32 muestras cada 2 ms. El predictor usa información del tiempo y la frecuencia de la señal LM obtenida por medio de una transformada wavelet [URL, 53]. En particular, los coeficientes de aproximación de la transformada wavelet de Haar (ver sección 6 de [URL, 53]). Estos coeficientes se usan como vectores de características. La representación de este vector como un punto en el espacio multidimensional, muestra que en descargas no disruptivas y en las fases no disruptivas de una descarga disruptiva, los vectores de características se distribuyen en un clúster compacto; mientras que en la fase disruptiva de una descarga, estos vectores están lejos del grupo original [Vega, 2015a], [Vega, 2015b]. La Figura 4.47. muestra la representación. de estos vectores para el caso de los vectores de características 2-D.

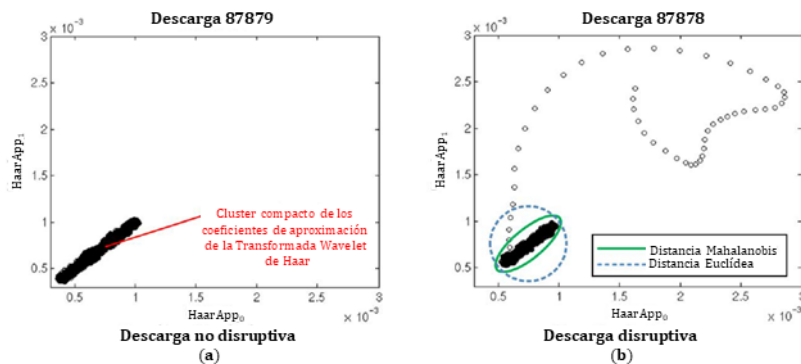


Figura 4.47. Representación de los dos coeficientes de aproximación de la transformada wavelet de Haar para (a) descargas no disruptivas y (b) disruptivas como un vector de características en el espacio bidimensional [Esquembrí, 2018].

La Figura 4.47. muestra el nivel de transformación Wavelet 4 aplicado a la señal LM, muestreada a 1 kHz en 32 ventanas de muestras, actualizadas cada 2 ms. (a) Durante las fases no disruptivas de una descarga, los vectores se distribuyen en un grupo compacto

con covarianza positiva. (b) Las fases disruptivas presentan valores atípicos con respecto al grupo compacto. Debido a la covarianza entre los miembros del grupo, los valores atípicos se pueden detectar mejor utilizando la distancia de Mahalanobis [URL, 54].

Debido a la obvia covarianza presente entre los puntos característicos del cluster, se eligió la distancia de Mahalanobis para medir la distancia de un punto con respecto al centroide del grupo. Finalmente, se calcula un factor atípico utilizando medidas estadísticas de la distancia de Mahalanobis durante la descarga, como la desviación estándar y la media. La alarma disruptiva se activará cuando el factor atípico supera un cierto umbral. El análisis sobre todas las campañas ILW de JET muestra buenos resultados de detección, al establecer este umbral en 10 para todas las descargas [Vega, 2015a] y [Vega, 2015b].

Aunque el entorno de ejecución MARTe ya fue introducido en [López, 2014], se tendrán en cuenta las nuevas aportaciones realizadas por [Esquembrí, 2018]. El principal objetivo de MARTe es abstraer una implementación de la interfaz hardware y software. También proporciona herramientas para la programación de subprocesos en tiempo real, validación fuera de línea y en línea, y una arquitectura que permite la reutilización del código y su fácil mantenimiento.

Las aplicaciones MARTe consisten en un conjunto de módulos de aplicaciones genéricas (GAM), conectados por medio de los buffers dinámicos de datos (DDB). Cada GAM puede leer y escribir una cantidad arbitraria de datos del DDB. Se pueden usar GAM especiales para realizar tareas de entrada-salida, como muestras de señales procedentes de la RTDN o envío del valor de alarma a la RTDN. Los GAM son ejecutados dentro de hilos gestionados por el planificador en tiempo real. En la Figura 4.48. se muestra la estructura de la aplicación MARTe, y de cómo se integra con la RTDN de JET y con el sistema de persistencia.

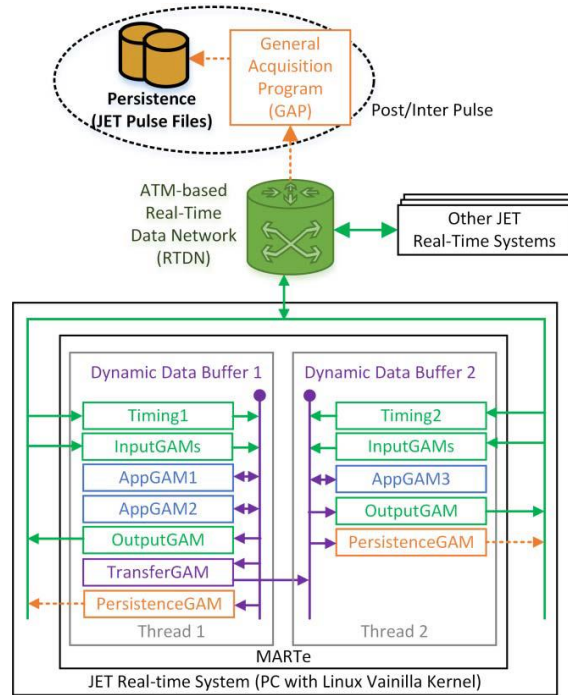


Figura 4.48. Esquema de la aplicación típica de MARTe y de su integración con RTDN de JET y con el sistema de persistencia [Esquembri, 2018].

El ejecutor en tiempo real ejecutará periódicamente los GAM de cada hilo en el orden especificado. La implementación del predictor SPAD para MARTe se ilustra en la Figura 4.49.

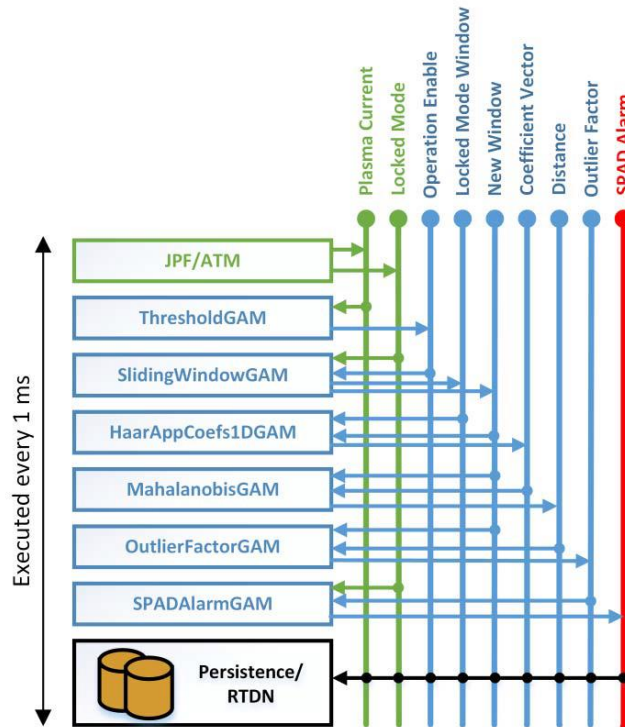


Figura 4.49. Diagrama de implementación de SPAD en MARTe [Esquembri, 2018].

Los GAM se ejecutan en orden de arriba abajo, requiriendo esta implementación el desarrollo de los seis GAMS que aparecen en la Figura 4.49. También utiliza GAMs MARTE estándar, para leer la corriente de plasma y señales LM desde la RTDN; así como para escribir el valor de la alarma también en la RTDN. En las fases de verificación y validación, estas señales fueron leídas de los archivos de pulso de JET (JPF). Los JPF también se utilizarán para almacenar todos los datos intermedios, utilizados en el algoritmo para el análisis del futuro comportamiento del predictor.

Todos estos GAM se ejecutan secuencialmente en un solo hilo. La señal LM es leída cada 1 ms, y el hilo debe procesar por tanto todos los GAM SPAD en menos de 1 ms. La configuración de los GAM se realiza en la fase de inicialización, no permitiéndose cambiar su tiempo de ejecución. Eso permite realizar algunos cálculos y toda la asignación de memoria en la fase de inicio.

- **ThresholdGAM:** este GAM controla que el resto de los GAM no realiza ninguna operación hasta que la fase inicial del plasma haya terminado. La fase de inicio se considera terminada una vez que la corriente de plasma haya alcanzado un cierto umbral, típicamente 750 kA. Por lo tanto, este GAM solo compara el valor de entrada con un umbral configurable; permitiendo una señal de salida al DDB cuando se cumpla la condición. Esta salida será una entrada para el resto de los GAMS, que no hará nada más que restablecerse al estado inicial cuando la señal esté desactivada.
- **SlidingWindowGAM:** este GAM producirá la entrada para la transformada wavelet de Haar. El GAM almacena la entrada internamente, en este caso la señal LM, y produce una salida con un conjunto de los últimos valores de la señal. Este conjunto de valores se denominará "ventana". Este GAM está configurado con un tamaño de ventana de 32 elementos, que se actualizan cada dos ciclos (cada 2 ms). La Figura 4.50. muestra un ejemplo de cómo se agrupará la señal en estos deslizamientos de ventanas.

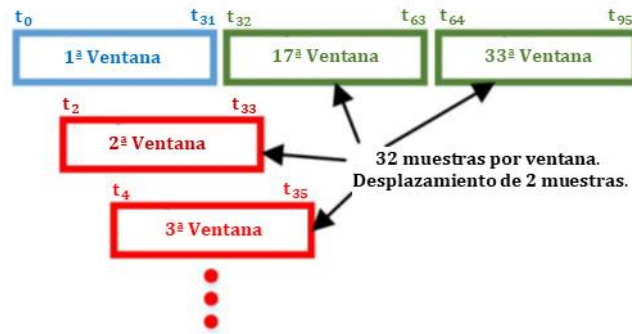


Figura 4.50. Explicación de las muestras de señales en ventanas deslizantes [Esquemabri, 2018].

La Figura 4.50. muestra la distribución de muestras en ventanas de 32 elementos deslizantes cada dos elementos (dos muestras consecutivas). Como el hilo se ejecuta cada 1 ms, SlidingWindowGAM producirá una nueva ventana cada 2 ms que contiene las últimas 32 muestras de la señal LM.

- **HaarApp1DCoefGAM:** GAM que calculará los coeficientes de aproximación de la transformada wavelet de Haar de cada nueva ventana devuelta por SlidingWindowGAM. La aplicación de la transformada wavelet de Haar ( $H$ ) a un vector con número de coeficientes pares  $X$  de tamaño  $n$ , produce  $n$  coeficientes; siendo la mitad de ellos coeficientes de aproximación ( $H_{App}$ ), y la otra mitad coeficientes de detalle ( $H_{Det}$ ) como se muestra en la siguiente ecuación:

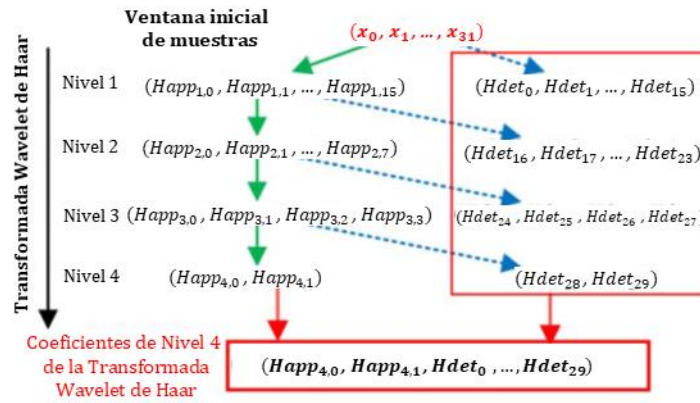
$$X = (x_0, \dots, x_{n-1})$$

$$H(X) = H = (H_{App}, H_{Det}) = ((H_{App,0}, \dots, H_{App,n/2}), (H_{Det,0}, \dots, H_{Det,n/2}))$$

$$\text{Donde } H_{App,i} = \frac{\sqrt{2}}{2} (x_{2i} + x_{2i+1}); H_{Det,i} = \frac{\sqrt{2}}{2} (x_{2i} - x_{2i+1}).$$

La Figura 4.51. muestra cómo es la transformada wavelet de Haar aplicada a un vector de 32 elementos, para obtener los coeficientes de nivel 4.





**Figura 4.51.** Proceso de obtención de coeficientes de la transformada wavelet de Haar sobre varias iteraciones de la transformada [Esquembri, 2018].

En SPAD solo se utilizarán coeficientes de aproximación. El nivel de transformación aplicado es configurable, pero hay que tener en cuenta que el siguiente nivel de la transformada solo se puede aplicar si el número de nivel anterior contiene los coeficientes de aproximación pares.

- **MahalanobisGAM:** GAM que calcula la distancia de Mahalanobis entre el vector de características actual y el centroide de todos los vectores de características pasados en un espacio multidimensional. La distancia de Mahalanobis (DM) se define según la expresión (5).

$$D_M[k] = \sqrt{(H_{app}[k] - \mu(H_{app}[0..k-1]))^T \Sigma^{-1} (H_{app}[k] - \mu(H_{app}[0..k-1]))} \quad (5)$$

- En este caso, el vector de características utilizado como entrada será el vector que contiene los coeficientes de aproximación de Haar ( $H_{App}$ ). Teniendo en cuenta que se calculará un nuevo vector de características para cada ventana de muestras, siendo  $H_{App}[i]$  el vector de características calculado para la  $i$ -ésima ventana, siendo  $H_{App}[k]$  el vector de características calculado para la última ventana disponible, y  $D_M[k]$  es el último valor de la distancia de Mahalanobis calculada. La ecuación anterior incluye la media  $\mu$  y la matriz de covarianza  $\Sigma$  de un conjunto de vectores característicos, cuyo número va aumentando durante la descarga. Si se calcularan el valor medio y la matriz de covarianza desde cero en cada ciclo, aumentará la memoria utilizada y también el tiempo de cálculo en cada ciclo; haciendo que la implementación no sea adecuada para

descargas largas. [Esquembri, 2018] ha solventado los problemas anteriores limitando el vector de características para todas las descargas.

- **OutlierFactorGAM:** GAM que calcula el valor del factor atípico ( $O_F$ ) del vector de características actual (ver ecuación (3) de la sección III.C. de [Esquembri, 2018]). Este GAM es similar al MahalanobisGAM, donde la desviación estándar ( $\sigma$ ) y la media ( $\mu$ ) que aparece en dicha ecuación, debe calcularse con respecto al conjunto de todos los valores de distancia de Mahalanobis anteriores en la descarga. El enfoque implementado en este GAM es comparable con el de MahalanobisGAM, que consiste en el almacenamiento de sumas parciales y productos, para hacer el tiempo de ejecución independiente del número de distancias de Mahalanobis calculadas previamente.
- **SPADAlarmGAM:** GAM que activa la alarma en caso de producirse detección de anomalías. Las anomalías se detectan cuando el factor atípico es mayor que un umbral configurable, teniendo los valores previos muestreados para la señal LM un factor inferior al valor umbral configurado.

La implementación presentada ha sido validada usando los datos de todas las descargas no disruptivas y disruptivas no intencionadas de JET, en el rango de 82460–87918; pertenecientes a todas las campañas con ILW desde 2011 hasta 2014. [Esquembri, 2018] ha analizado 1738 descargas no disruptivas, y 566 descargas disruptivas no intencionadas. La comparación de los resultados de detección de la implementación de MARTE en tiempo real y una implementación equivalente en MATLAB [Vega, 2015a] y [Vega, 2015b] sin optimizaciones, mostró que las optimizaciones no afectaron el comportamiento de la alarma activada. El entorno de prueba para la validación fue una computadora con una CPU i7 4790 (cuatro núcleos, dos hilos / núcleo, y 3.6 GHz) y 16 GB de RAM con Red Hat Enterprise Linux 6.5 y con un Kernel Linux Vanilla. En la validación, cada descarga se simuló usando los datos obtenidos de LM y de la corriente del plasma de JPFs. Para todas las descargas, los GAM se configuran de manera que el umbral de corriente de plasma para iniciar el predictor sea de 750 kA. El último valor disponible de la señal LM se leyó cada 1 ms, y se procesó cada 2 ms en ventanas de tiempo que contienen las últimas 32 muestras. En [Esquembri, 2018] se probaron varios niveles de la transformada wavelet de Haar, incluso desde el nivel 1 (16 coeficientes de

aproximación) al nivel 4 (dos coeficientes de aproximación). El umbral del factor atípico se estableció en 10 para todas las descargas.

La Tabla 4.31. muestra los resultados de detección para la configuración anteriormente descrita.

**Tabla 4.31.** Comparación de las Tasas de Detección con SPAD [Esquembri, 2018].

Tamaño Vector Características	Falsas Alarmas	Alarmas Perdidas	Detecciones Tardías	Alarmas Válidas	Alarmas Prematuras
16	18.24 %	11.31 %	3.00 %	79.68 %	6.01 %
8	8.98 %	10.60 %	3.18 %	83.57 %	2.65 %
4	9.55 %	10.95 %	3.36 %	82.86 %	2.83 %
2	9.84 %	11.31 %	3.53 %	81.98 %	3.18 %

Las detecciones de disrupción que se activaron 1,5 s antes la disrupción son consideradas "alarmas prematuras", ya que no se consideran predicciones verdaderas. Considerándose "alarmas válidas" las que se activaron entre 1,5 s y 10 ms; por ser 10 ms el tiempo mínimo en JET para aplicar técnicas de mitigación con inyección masiva de gas. Las "detecciones tardías" son aquellas activadas con menos de 10 ms de anticipación, ya que en esos instantes es demasiado tarde para aplicar las técnicas de mitigación mencionadas anteriormente. Si la alarma se activó después de la disrupción o no se activó en absoluto en una descarga disruptiva, se considera una "alarma perdida". Aquellos casos en los que se activa la alarma y no hay disrupción, se clasifican como "falsas alarmas". En estos últimos casos, si SPAD estaba directamente conectado a los sistemas de mitigación de disrupciones, esas falsas alarmas habrían terminado las descargas de plasma no disruptivas prematuramente, sin que existiera razón alguna para hacerlo; con la pérdida de tiempo y recursos que ello conlleva. La Tabla 4.32. representa la comparación de SPAD con LMPT (Predictor con criterio de umbral basado en la señal LM) y APODIS utilizando los mismos criterios de clasificación.

**Tabla 4.32.** Comparación de Tasas de Detección entre SPAD, APODIS y LMPT [Esquembri, 2018].

Predictor	Falsas Alarmas	Alarmas Perdidas	Detecciones Tardías	Alarmas Válidas	Alarmas Prematuras
SPAD	7.42 %	10.60 %	3.18 %	83.57 %	2.65 %
APODIS	<5 %	15.38 %	2.47 %	79.15 %	3.00 %
LMPT	---	30.39 %	3.00 %	63.69 %	2.65 %

Con más del 83% de alarmas válidas, SPAD supera ligeramente los resultados de detección de APODIS (79,15%), mientras que LMPT permanece con solo el 63,96% de las alarmas válidas. Esta misma comparativa se muestra en la Figura 4.52.

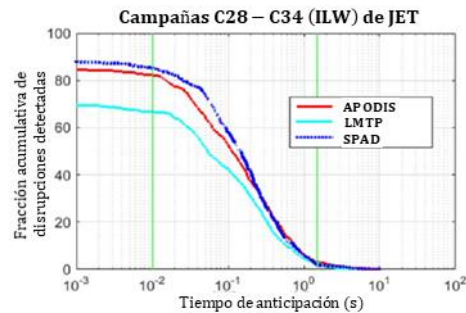


Figura 4.52. Representación [Neto, 2010] de la fracción acumulativa de detección de disrupciones con respecto a las disrupciones totales durante todas las campañas ILW de JET [Esquembri, 2018].

La Figura 4.53. muestra la evolución del factor atípico SPAD para la descarga disruptiva 82960.

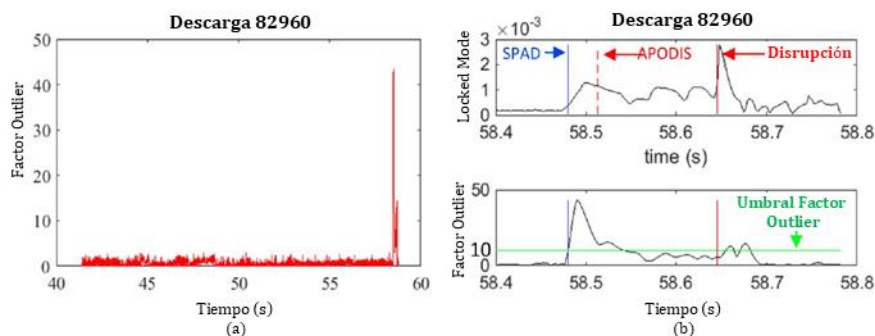
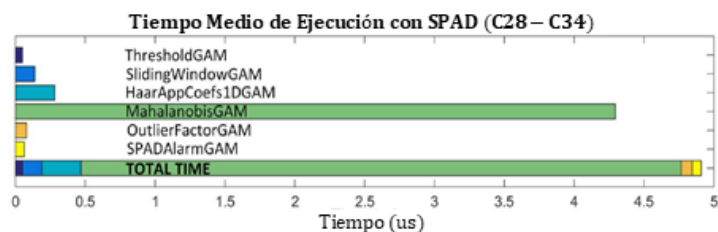


Figura 4.53. Representación [Neto, 2010] del comportamiento del factor atípico SPAD en la descarga 82960 [Esquembri, 2018].

Donde (a) muestra el valor del factor atípico durante toda la descarga, y (b) realiza la comparación entre la señal LM y el factor atípico momentos antes de la disrupción, así como el momento en que SPAD y APODIS predicen la disrupción. Con el predictor basado en el criterio de umbral para la señal LM (LMPT) se perdió esta disrupción. Se puede observar que el valor atípico permanece muy bajo hasta el final de la descarga, instante en el que se produce la disrupción. En la Figura 4.53.(b) se muestran los valores de la señal LM y el factor atípico en detalle, incluido el tiempo de detección SPAD y APODIS. Con LMPT se perdió esta disrupción porque el aumento en la señal LM ocurre muy cerca de la disrupción y se estableció el umbral demasiado alto.

El tiempo de ejecución se midió con las herramientas MARTE. La Figura 4.54. muestra un desglose por GAM del tiempo promedio de ejecución para todo el conjunto de pruebas, con un nivel de confianza del 97%.



**Figura 4.54.** Tiempo promedio de ejecución del ciclo para los GAM de SPAD con un nivel de confianza del 97% [Esquembri, 2018].

El nivel de confianza del tiempo de ejecución se obtiene agregando tres veces la desviación estándar a la media. En la Figura 4.54. se puede observar que el tiempo total de ejecución del ciclo está lejos del máximo permitido para la ejecución adecuada del algoritmo (1 ms). El GAM que requiere más tiempo de ejecución es el MahalanobisGAM, que es razonable debido al cálculo de la matriz inversa de la covarianza y la multiplicación vector-matriz-vector necesaria para obtener la distancia de Mahalanobis. El tiempo máximo de ejecución medido durante el análisis realizado en [Esquembri, 2018] fue de 26,9280  $\mu$ s.

El algoritmo implementado no necesita datos de descargas anteriores. Esto puede ser muy útil no solo para dispositivos nuevos, sino también después de una actualización en una máquina existente, o cuando los parámetros de operación de plasma cambian drásticamente. Otro hecho importante es que no hay necesidad de cambiar la configuración dependiendo de las características de las descargas. El rendimiento del algoritmo ha sido medido en un entorno de prueba, probando su idoneidad para un sistema en tiempo real. [Esquembri, 2018] propone como una de las más importantes mejoras para el predictor, reducir el número de falsas alarmas; trabajando para identificar las causas de estas detecciones falsas y tratando de disminuir su número al menos al porcentaje válido estimado para ITER (< 5%). Sin embargo, las disrupciones pueden dañar mucho la máquina, y SPAD podría detectar numerosas disrupciones que el predictor LMPT no podría detectar. Otra línea de trabajo apuntada por [Esquembri, 2018] es la de identificar nuevas señales para poder ser añadidas en el algoritmo SPAD. La inclusión de señales relevantes, relacionadas con las causas fundamentales de la disrupción, podría mejorar los resultados de detección o reducir la proporción de falsas

alarmas. Otra posible mejora propuesta por [Esquembri, 2018] es la estimación de forma dinámica del umbral óptimo para el factor atípico durante la descarga. Con respecto al tiempo de ejecución, la implementación actual es lo suficientemente buena para el proceso de una señal con la extracción de ocho coeficientes de aproximación de Haar. Sin embargo, la adición de una segunda o tercera señal con sus correspondientes coeficientes de aproximación, podrían hacer que la dimensionalidad de la matriz de covarianza sea demasiado grande para la implementación, corriendo el riesgo de que no se pueda procesar en menos de 1 ms. Para resolver esto, [Esquembri, 2018] propone estudiar e implementar mejores enfoques para el cálculo de la matriz inversa y el producto vector-matriz-vector, requerido para la distancia Mahalanobis.

### **4.3. Métodos de predicción empleados en la Tesis.**

Los predictores de interrupción tratados en la tesis actual, se basan en la obtención de una ecuación lineal en el plano o de un punto en la recta, que sirve como frontera entre las regiones disruptiva y no disruptiva. Dicha frontera se obtiene a partir del entrenamiento de un número significativo de descargas, cuya naturaleza disruptiva o no disruptiva es conocida.

Una vez obtenida dicha ecuación lineal o punto frontera, se toma otro conjunto significativo de descargas de distinta naturaleza, para estudiar la validez del modelo que ha permitido obtener la zona de separación o frontera.

Si los resultados de dicha evaluación son aceptables, es decir, si se obtiene una alta tasa de aciertos para las descargas disruptivas, y una baja tasa de falsas alarmas para las descargas no disruptivas; el modelo encontrado se podría aplicar en entornos de explotación, para evaluar la naturaleza de las descargas en tiempo real.

Se han analizado tres dispositivos de fusión diferentes, donde cada uno de ellos tiene sus particularidades; de ahí que resulte muy complejo trasladar los descubrimientos y/o conclusiones de uno de los dispositivos a los otros dos. Uno de los motivos de esa complejidad se encuentra en que, los distintos dispositivos no utilizan el mismo conjunto de señales. Además, se puede dar el caso de que, utilizando algunas señales iguales, éstas no estén calibradas de la misma manera. En este sentido, es importante volver a resaltar la complejidad de alcanzar una solución general con la metodología propuesta. De ahí

que se haya buscado una solución *ad hoc* para cada dispositivo, a partir de la metodología desarrollada.

En el siguiente capítulo, se desarrollará de forma detallada, el procedimiento que aquí se ha introducido; con el fin de predecir las disrupciones. Para ello, esta tesis se apoyará también en las publicaciones realizadas por [Vega, 2019] y [Vega, 2020].





# PREDICCIÓN BASADA EN EL MÉTODO DE LOS CENTROIDES

5

## 5.1. Metodología matemática general del método de los centroides.

Desde un punto de vista de alto nivel, el procedimiento comenzará seleccionando dos conjuntos de descargas cuya naturaleza es conocida (aprendizaje supervisado). Uno de los conjuntos tendrá en cuenta las descargas disruptivas, incluyendo su instante de disrupción; aunque en determinados casos será preciso calcular dicho instante, como tarea previa a la propia selección de descargas. El otro conjunto tendrá en cuenta las descargas no disruptivas.

Una vez realizada la selección anterior, habrá que estudiar distintos criterios para caracterizar a cada uno de los conjuntos. Cada descarga está compuesta por distintas señales, entre las que cabe destacar: la intensidad de corriente del plasma y la amplitud (relacionada con la rotación del plasma); y donde cada señal está muestreada a razón de 1 ms. El criterio de contribución de cada descarga no disruptiva al centroide no disruptivo, estará relacionado con las amplitudes que se encuentran bajo ciertos umbrales de corriente del plasma. En el caso de las descargas disruptivas, la contribución a su correspondiente centroide disruptivo, será estudiada en muestras próximas al instante de la disrupción.

Cada descarga, sea del tipo que sea, deberá caracterizarse mediante un único punto en el espacio de parámetros; resumiendo dicho punto el comportamiento general de la descarga.

Finalizado el análisis para los dos conjuntos de entrenamiento seleccionados, se hallará el valor medio de todos los puntos encontrados en la fase anterior, para cada tipo de descargas. De esa forma, se obtendrán dos únicos puntos; resumiendo cada uno de ellos el comportamiento de cada tipo de descarga.

Las posiciones ocupadas por ambos centroides, determinarán la frontera de separación entre las regiones disruptiva y no disruptiva.

Una vez obtenida la frontera que divide ambas regiones, se seleccionarán dos conjuntos de descargas, de naturaleza también conocida, para ser evaluados. Análogamente a los seleccionados para la fase de entrenamiento, uno de los conjuntos tendrá en cuenta las descargas disruptivas, incluyendo su instante de disrupción; y el otro tendrá en cuenta las descargas no disruptivas.

Teniendo en cuenta los dos conjuntos seleccionados para la fase de evaluación, junto con la frontera calculada a partir de las posiciones de los centroides; se validará el modelo de entrenamiento. De forma que, si los resultados en la fase de evaluación no fueran satisfactorios, se descartaría el modelo de entrenamiento; debiéndose entrenar nuevos conjuntos de datos, que mejoren los resultados producidos en la fase de evaluación.

Además, en la fase de evaluación, es muy importante tener en cuenta la variación de la amplitud entre muestras consecutivas. Esa variación de la amplitud, proporcionará la velocidad de la descarga; de forma que grandes variaciones de amplitud entre dos muestras consecutivas, alertarán de la existencia de un precursor de comportamiento disruptivo.

Se trata de una metodología general, que habrá que adaptar a cada dispositivo de fusión; ya que las señales son dependientes de los dispositivos sobre los que se aplica el método de los centroides. Como ya se comentó en la sección 4.3. del capítulo anterior, la metodología desarrollada en el presente trabajo de investigación, es muy sensible al dispositivo de fusión termonuclear en el que se aplique.

En las siguientes secciones se explica con más detalle el procedimiento que se acaba de comentar, adaptado a cada dispositivo objeto de estudio.

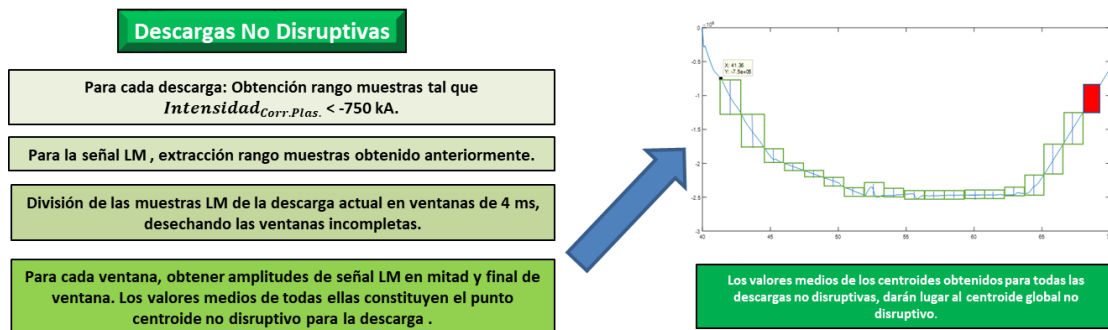
## 5.2. Metodología matemática de predicción en JET.

Como base de datos a utilizar, se ha tomado un gran volumen de descargas, cuyo comportamiento disruptivo y no disruptivo es conocido. Con dichas descargas se han calculado los centroides que representan ambas regiones (disruptiva y no disruptiva) durante la fase de entrenamiento; donde se han tomado las muestras cuya intensidad de corriente del plasma cumple ciertas condiciones de valores umbrales.

Una vez calculados los centroides anteriores, se ha analizado otro conjunto significativo de descargas (distintas a las consideradas en la fase de entrenamiento); y cuya naturaleza disruptiva o no disruptiva es conocida. De forma que, al evaluar las muestras de estas últimas descargas, se pueda determinar la validez o no del modelo predictor de disrupciones que se ha obtenido.

### 5.2.1. Fase de entrenamiento para descargas no disruptivas.

El centroide no disruptivo se calcula según muestra el proceso de la Figura 5.1.



**Figura 5.1.** Procedimiento para la obtención del centroide no disruptivo en la fase de entrenamiento de JET.

Para cada una de las descargas no disruptivas de esta fase:

1. Se determina el rango de muestras donde la señal IPLA (Intensidad de Corriente del Plasma) comienza estando por debajo de los -750 kA (kiloamperios), y donde finaliza estando por debajo de esos -750 kA.
2. Una vez encontrado el rango temporal anterior, se extraen las muestras para la señal LM (Amplitud de Modo Bloqueado) que corresponden a dicho rango temporal.

3. Con las muestras extraídas de la señal LM, se forman ventanas temporales de longitud 4 ms. Aquellas ventanas que resulten incompletas al final del rango temporal, no serán consideradas para el cálculo del centroide a nivel de descarga.
4. Para cada ventana completa (de 4 ms), se obtienen sus amplitudes correspondientes a la mitad y al final de la misma.
5. Una vez obtenidas las amplitudes de la mitad y final de todas las ventanas, se hallan las amplitudes medias para la mitad y final de las mismas, obteniendo un único punto para cada descarga:

$$(A, B) = (\mathbf{Amplitud\_Media}_{mitad\_ventana}, \mathbf{Amplitud\_Media}_{final\_ventana}).$$

6. Una vez procesadas todas las descargas no disruptivas de esta fase de entrenamiento:

Se hallan las amplitudes medias de todas las descargas. De forma que, se obtiene un único punto:

$(X, Y) = (\mathbf{Amplitud\_Media}_A, \mathbf{Amplitud\_Media}_B)$ . Ese único punto (X, Y) será una de las referencias fundamentales durante la Fase de Evaluación (Test).

### 5.2.2. Fase de entrenamiento para descargas disruptivas.

La Figura 5.2. muestra distintas alternativas empleadas durante la fase de entrenamiento para las descargas disruptivas.

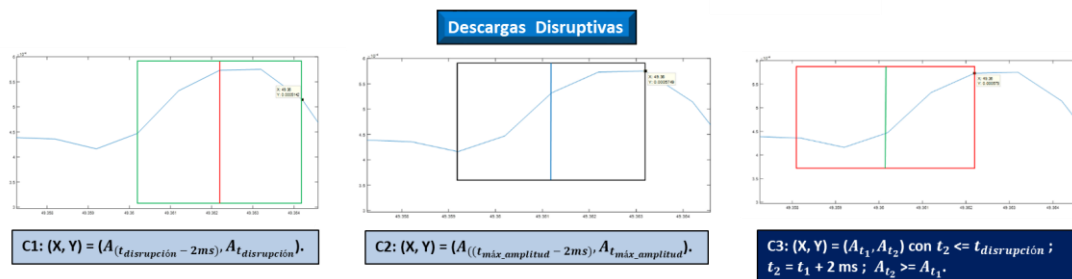


Figura 5.2. Procedimiento para la obtención del centroide disruptivo en la fase de entrenamiento de JET.

Para cada una de las descargas disruptivas de esta fase:

1. Se analizaron tres criterios diferentes (C1, C2 y C3 según la Figura 5.2.), donde cada uno de ellos sólo considera un punto con dos componentes para cada descarga.
  - 1.1. Según el criterio C1: Se ha tomado como coordenada  $X$  del punto, el valor que adopta la amplitud 2 ms antes del instante en el que se produce la disrupción. Y como coordenada  $Y$  del punto, se ha tomado el valor de la amplitud en el instante donde se produce la disrupción.
  - 1.2. Según el criterio C2: Se ha elegido como coordenada  $X$  del punto, el valor que adopta la amplitud 2 ms antes de que la amplitud sea máxima; teniendo en cuenta que dicho instante se encuentra antes de producirse la disrupción. Y como coordenada  $Y$  del punto, se ha seleccionado el valor máximo de la amplitud; también teniendo en cuenta que ese instante se da antes de producirse la disrupción.
  - 1.3. Según el criterio C3: Se ha considerado como coordenada  $X$  del punto, el valor que adopta la amplitud 2 ms antes del instante en el que se produce la disrupción. Como coordenada  $Y$  del punto, se ha tomado el valor de la amplitud en el instante donde se produce la disrupción. Si el valor de la componente  $Y$  del punto viene representado por una amplitud menor que la de la componente  $X$ , se van recorriendo las muestras en orden cronológico inverso hasta encontrar que  $Y \geq X$ .  
El recorrido cronológico inverso tiene en cuenta las muestras que se van encontrando cada 2 ms.
2. Por ser la solución más óptima en cuanto a resultados, se adoptó finalmente aquella etiquetada como C3, donde los centroides de las descargas se sitúan por encima de la bisectriz del primer cuadrante en el plano bidimensional.
3. Una vez procesadas todas las descargas disruptivas de esta fase de entrenamiento mediante el criterio C3, se han calculado las amplitudes medias de todas las descargas. De forma que, se ha obtenido un único punto:  
 $(X, Y) = (\text{Amplitud\_Media}_{\text{Primera\_Componente}}, \text{Amplitud\_Media}_{\text{Segunda\_Componente}})$ .  
Ese único punto  $(X, Y)$  será otra de las referencias fundamentales durante la Fase de Evaluación.

### 5.2.3. Fase de evaluación.

En la Figura 5.3. se puede observar la posición relativa de una muestra perteneciente al procesamiento de una descarga en tiempo real, con respecto a los puntos medios que representan a los centroides disruptivo y no disruptivo (obtenidos en la fase de entrenamiento previa).

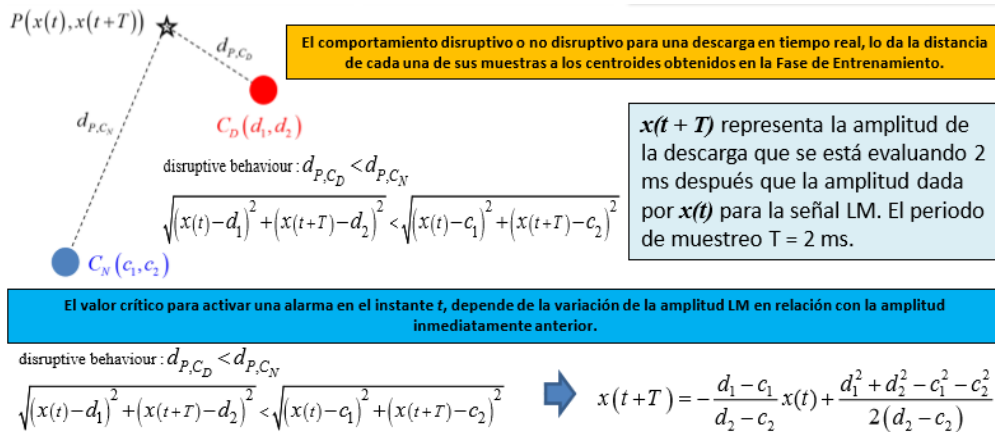


Figura 5.3. Procedimiento que determina la naturaleza de una muestra en fase de evaluación de JET.

La distancia euclídea de la muestra actual con respecto a dichos centroides, determinará la pertenencia a cada uno de los dos espacios (disruptivo o no disruptivo). Así, si la muestra que está siendo evaluada dista menos del centroide disruptivo que del no disruptivo, dicha muestra presentará comportamiento disruptivo. Por el hecho de presentar ese comportamiento una de las muestras en particular que componen la descarga, ese mismo comportamiento se hace extensivo para la descarga en general.

Casos que se pueden presentar durante la evaluación (fase de test) de una descarga:

1. Tratándose de la evaluación de una descarga con comportamiento no disruptivo, la descarga presenta para casi la totalidad de sus muestras ese mismo comportamiento no disruptivo, pero puede darse el caso de que exista un número limitado de muestras con comportamiento disruptivo; lanzándose una alarma que alerta acerca de dicho comportamiento. En este caso se estaría ante la presencia de una **falsa alarma**. La suma total de falsas alarmas, contribuirá a la **tasa de falsas alarmas** que presenta la evaluación de descargas; cuyo comportamiento general es no disruptivo.
2. Conocido el comportamiento disruptivo de una descarga, en la evaluación de la misma:

- a. Puede ocurrir que no se presente ningún tipo de alarma; estando ante el caso de una **alarma completamente perdida**.
  - b. Si se presenta una alarma muy próxima y anterior al instante de la disrupción, durante el instante de la disrupción, o posterior a la misma; se estaría ante la presencia de una **alarma tardía**.
  - c. En caso de presentarse una alarma con una notable antelación al instante de la disrupción, se estaría ante una **alarma prematura**.
  - d. Cuando la alarma se produce con una antelación próxima al instante de la disrupción, pero con el margen suficiente para poder realizar tareas de elusión o mitigación, se estaría ante la presencia de una **alarma válida**.
3. La suma de alarmas válidas y de alarmas prematuras, proporcionará la **tasa de aciertos** que arroja la evaluación de descargas, cuyo comportamiento es disruptivo.
  4. La suma de alarmas completamente perdidas y de alarmas tardías, dará la **tasa de alarmas perdidas** en la evaluación de descargas disruptivas.

Para la activación de distintos tipos de alarmas, se considera la variación de la amplitud entre una muestra y la anterior.

Toda la casuística anterior se presentará con mayor detalle en el siguiente capítulo.

Atendiendo a la Figura 5.4., se han considerado distintos desplazamientos relativos del centroide disruptivo y del no disruptivo.

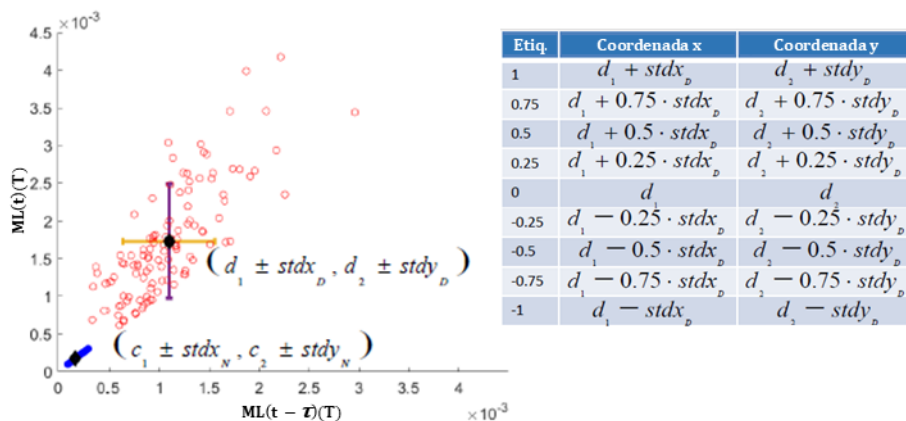


Figura 5.4. Procedimiento del desplazamiento relativo de los centroides en la fase de evaluación de JET.

Los centroides disruptivo y no disruptivo irán variando sus posiciones con respecto a la posición original calculada, y que aparece en la tabla de la Figura 5.4. como

Etiqu. con valor igual a 0. Aunque en la Figura 5.4. aparecen referenciadas las componentes  $(d_1, d_2)$  que forman el centroide disruptivo, esos desplazamientos relativos también son aplicados para las componentes  $(c_1, c_2)$  que forman el centroide no disruptivo. El desplazamiento relativo del centroide no disruptivo sigue el mismo procedimiento que muestra la Figura 5.5. para el centroide disruptivo.

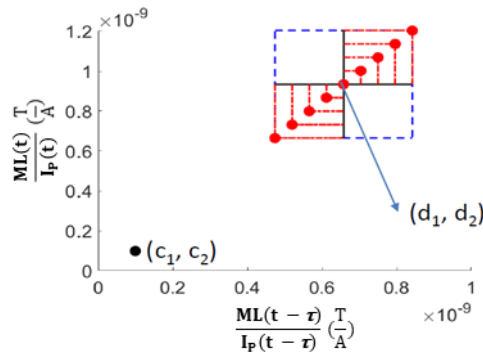


Figura 5.5. Detalle del desplazamiento relativo del centroide disruptivo en la fase de evaluación de JET.

Se observará que dichos desplazamientos relativos para el centroide no disruptivo son más pequeños, porque las muestras en el caso de descargas no disruptivas, se encuentran mucho más concentradas. No presentando la dispersión de muestras que sí se da en las descargas disruptivas.

Por tanto, la fase de evaluación comentada anteriormente para las posiciones originales de los centroides (Etiqu. = 0), hay que hacerla extensiva para el resto de posiciones relativas de los mismos (Etiqu. = -1, -0.75, -0.50, -0.25, 0.25, 0.50, 0.75, 1).

De esta forma, a la hora de obtener los resultados, se tendrán en cuenta 9 evaluaciones diferentes para el mismo conjunto de descargas. Tomando como mejor resultado aquel en el que la posición relativa de los centroides, permita obtener la mejor combinación entre tasa de aciertos y tasa de falsas alarmas; tratando de buscar la mayor tasa de aciertos y la menor tasa de falsas alarmas.

Ese desplazamiento relativo de los centroides se podría hacer extensivo a 81 posiciones diferentes, en lugar de las 9 posiciones que aparecen referenciadas en la Figura 5.5. En la Figura 5.6. se pueden observar las posiciones relativas a las que se podrían desplazar cada uno de dichos centroides.



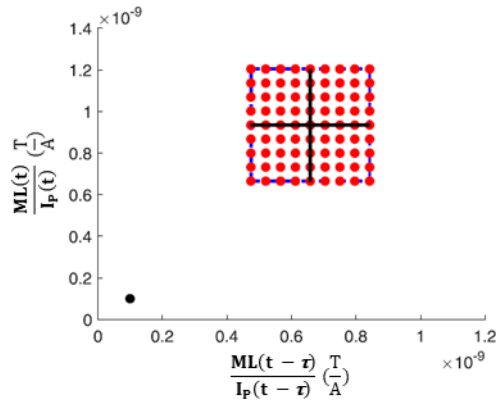


Figura 5.6. Desplazamiento relativo del centroide disruptivo de JET con 81 posiciones.

Además del centroide disruptivo inicial ( $d_1, d_2$ ), representado en el centro de la cruz en negro, que se observa en la Figura 5.6.; se obtienen 80 nuevos centroides disruptivos con incrementos positivos/negativos de sus coordenadas en relación con las barras de error en cada dimensión. Los 81 diferentes centroides disruptivos vienen representados por cada uno de los puntos en rojo. Análogamente, se tendrán otros tantos centroides no disruptivos en relación con las barras de error que representan la región no disruptiva. Al estar concentradas las contribuciones no disruptivas en un espacio más reducido que las contribuciones disruptivas, las barras de error que representan esas 81 diferentes ubicaciones del centroide no disruptivo, suponen menores desplazamientos relativos en relación con la posición del centroide original no disruptivo ( $c_1, c_2$ ). Como la dispersión de las contribuciones al centroide no disruptivo es menor, esa matriz de 81 puntos en rojo estaría más concentrada; lo que se muestra en la Figura 5.7. para dar una idea aproximada.

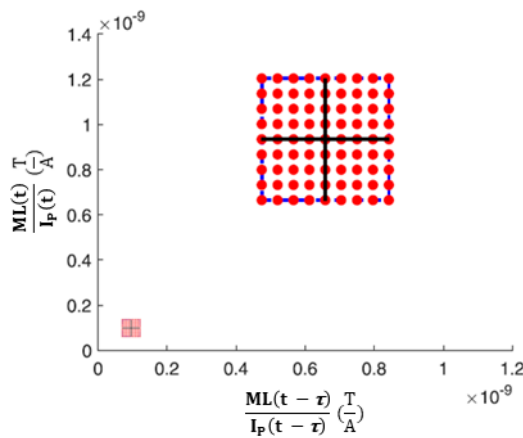


Figura 5.7. Desplazamiento relativo de los dos centroides de JET con 81 posiciones.

## 5.3. Metodología matemática de predicción en JT-60U.

Igual que en el caso de JET, se ha tomado un número lo suficientemente significativo de descargas, cuyo comportamiento disruptivo y no disruptivo es conocido; aplicando a las mismas, una fase de entrenamiento, y a continuación una fase de evaluación.

### 5.3.1. Fase de entrenamiento para descargas no disruptivas.

El procedimiento de cálculo para el centroide no disruptivo, se recoge en la Figura 5.8.

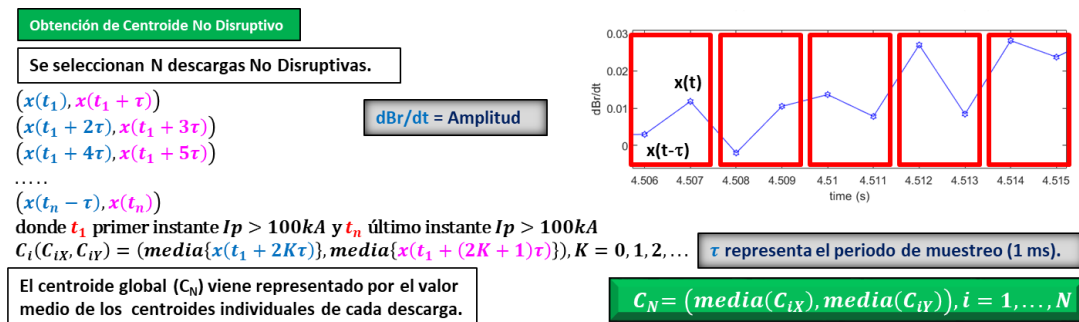


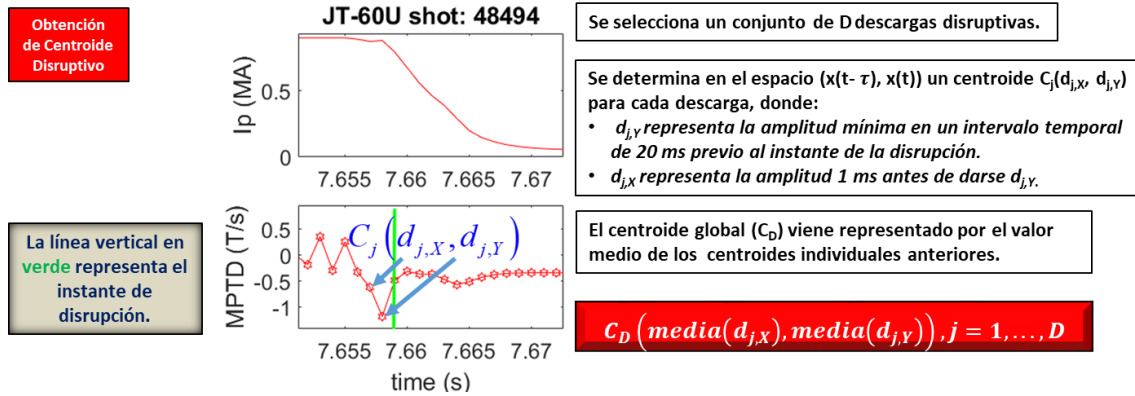
Figura 5.8. Proceso de obtención del centroide no disruptivo en fase entrenamiento de JT-60U.

El procedimiento de obtención del centroide no disruptivo es similar al descrito para JET en la sección 5.2.1., pero teniendo en cuenta las siguientes consideraciones:

1. El rango de muestras que en JET tenía un valor umbral para la intensidad de corriente del plasma inferior a -750 kA, para el dispositivo JT-60U pasa a considerarse un umbral de corriente superior a los 100 kA.
2. La señal de la amplitud en este dispositivo es una señal distinta a la empleada en JET. Se trata de la señal dBr/dt.
3. En lugar de considerar ventanas de muestras de 4 ms, en este caso se emplean ventanas de 2 ms; de donde se extraen las componentes de la amplitud que contribuyen al cálculo del centroide.
4. El resto de los pasos que componen el procedimiento para el cálculo del centroide no disruptivo, es análogo al empleado para JET en la sección 5.2.1.

### 5.3.2. Fase de entrenamiento para descargas disruptivas.

En la Figura 5.9. se puede observar el procedimiento de cálculo para el centroide disruptivo.



**Figura 5.9.** Proceso de obtención del centroide disruptivo en fase de entrenamiento de JT-60U.

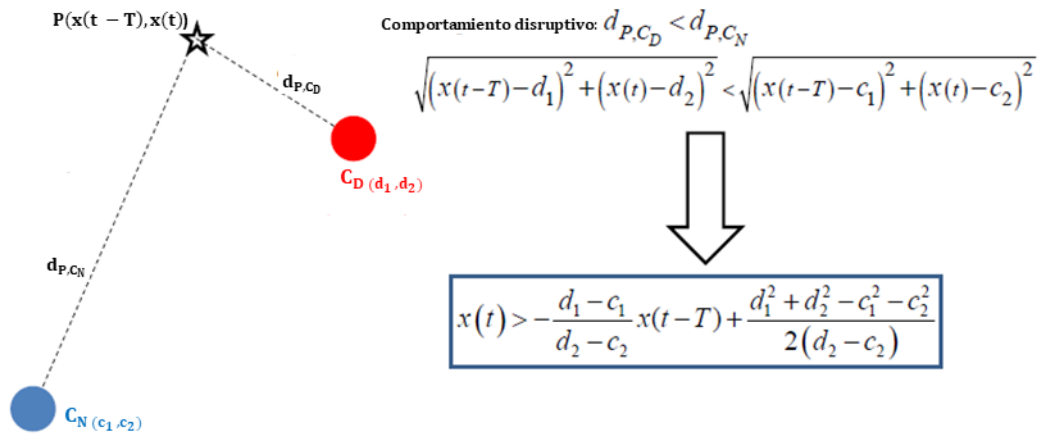
Para cada una de las descargas disruptivas  $j$ , tomadas en la fase de entrenamiento:

- Se toma como coordenada  $Y$  del punto, la amplitud mínima encontrada en el intervalo de los últimos 20 ms antes de producirse la interrupción  $d_{j,y}$ .
- Tomándose como coordenada  $X$  del punto, el valor de la amplitud que se obtiene 1 ms antes de la muestra que representa la coordenada  $Y$  del punto  $d_{j,x}$ .

Una vez procesadas todas las descargas disruptivas  $j$  de esta fase de entrenamiento, se obtendrá un único punto  $(X, Y) = (\text{media}(d_{j,x}), \text{media}(d_{j,y}))$ . Pudiéndose utilizar tanto este punto como el obtenido para el centroide no disruptivo, durante la fase de evaluación.

### 5.3.3. Fase de evaluación.

El comportamiento disruptivo o no disruptivo, igual que en la sección 5.2.3., viene marcado por la proximidad de la muestra que está siendo evaluada en la fase de test, con respecto a los centroides obtenidos en la fase de entrenamiento previa. La Figura 5.10. resume el proceso de evaluación.



**Figura 5.10.** Obtención de la distancia de una muestra que está siendo evaluada en tiempo real con respecto a los centroides procedentes de JT-60U.

En caso de darse:

$$x(t) > -\frac{d_1-c_1}{d_2-c_2} x(t-T) + \frac{d_1^2 + d_2^2 - c_1^2 - c_2^2}{2(d_2-c_2)}$$

Se manifestará un comportamiento disruptivo para la muestra que está siendo evaluada.

Cuando la descarga que se está evaluando tiene comportamiento NO DISRUPTIVO, todas sus muestras (de la señal de Amplitud) se concentran en una única región; estando dicha región en torno al punto del plano (0, 0).

Cuando se producen grandes variaciones de la amplitud entre muestras consecutivas de la descarga magnética, estamos ante un comportamiento DISRUPTIVO. Donde no todas sus muestras se concentran en una región, sino que debido a las grandes variaciones que sufre la señal de la amplitud, se puede apreciar la dispersión de dichas muestras con respecto al origen de coordenadas (0, 0). La Figura 5.11. recoge ambos comportamientos.

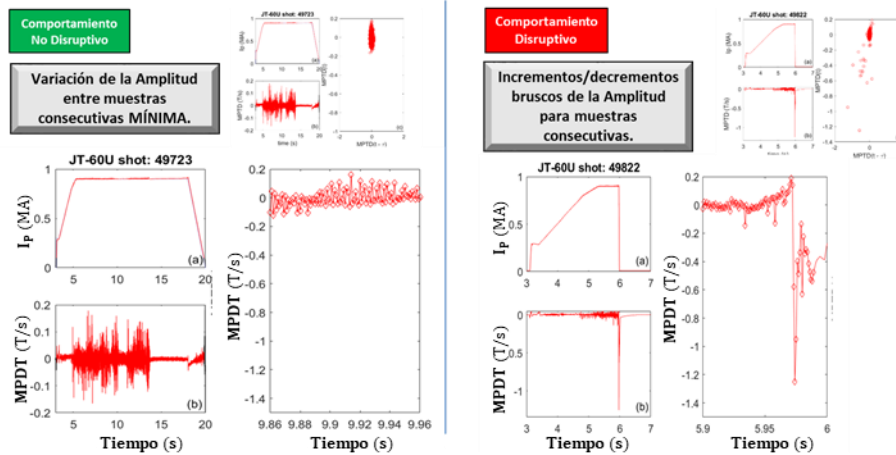


Figura 5.11. Distintos comportamientos de la señal Amplitud para JT-60U.

**Comportamiento NO DISRUPTIVO:** A pesar de todos los “picos” que muestra la Amplitud, en la Figura 5.11. ampliada, se observa que la variación entre muestras consecutivas es mínima. Se puede apreciar en la mitad izquierda de la Figura 5.11.

**Comportamiento DISRUPTIVO:** En este caso, sin embargo, se aprecian los incrementos/decrementos bruscos de la Amplitud para muestras consecutivas. Ver mitad derecha de la Figura 5.11.

Los casos que pueden presentarse durante la fase de evaluación para los distintos tipos de alarmas, son los mismos que ya fueron descritos en la sección 5.2.3., y que se presentarán con mayor detalle en el siguiente capítulo.

## 5.4. Predicción de disrupciones en DIII-D.

Como en los dos casos anteriores, se describirán a continuación los procedimientos seleccionados tanto para la fase de entrenamiento como para la fase de evaluación, asociados al dispositivo de fusión DIII-D.

Debido a la polaridad que presentan los valores de la señal amplitud para este dispositivo, la forma de proceder con el análisis de sus datos presenta ciertas variaciones con respecto a los dos dispositivos anteriores: JET y JT-60U. De ahí que se haya tomado como señal representativa la derivada de la amplitud. Ésta proporcionará las diferencias de amplitud entre dos muestras consecutivas (instantes consecutivos). Esa variación de la amplitud entre dos instantes consecutivos, nos da idea del comportamiento que ofrece la señal; siendo el espacio de representación unidimensional.

Para el análisis subsiguiente, se hablará de la señal valor absoluto de la derivada de la amplitud, y que se denotará como  $abs(diff(Amplitud))$ .

#### 5.4.1. Fase de entrenamiento para descargas no disruptivas.

Para cada descarga no disruptiva que interviene en la fase de entrenamiento, se toma un rango de muestras de la señal  $abs(diff(Amplitud))$ . Ese rango de muestras se obtiene a partir de los instantes temporales, donde la intensidad de corriente del plasma se sitúa por encima del umbral de los 750 kA; siendo el valor que determina la contribución de cada descarga al centroide no disruptivo el siguiente:

- $C_i(c_{xi}) = Valor\ Medio(abs(diff(Amplitud)))$  , teniendo en cuenta para la contribución a ese valor medio, todo el rango de muestras seleccionado según el valor umbral de la señal intensidad de corriente del plasma.

Una vez tratadas todas las descargas no disruptivas, que intervienen en la fase de entrenamiento, el valor global para el centroide no disruptivo se calcula como:

- $C_N(c_x) = \frac{\sum_{i=1}^n C_i(c_{xi})}{n}$  ; siendo  $n$  el número total de descargas no disruptivas que intervienen en la fase de entrenamiento.

#### 5.4.2. Fase de entrenamiento para descargas disruptivas.

En este apartado se considerarán varios casos, debido a los distintos resultados encontrados para cada una de las opciones.

##### 1. Consideración de los últimos 100 ms previos a la disrupción.

Tomando el rango de muestras de la señal  $abs(diff(Amplitud))$ , que se encuentra en el intervalo de los 100 ms previos al instante de la disrupción, el valor que determina la contribución de cada descarga al centroide disruptivo viene dado por:

- $C_j(c_{xj}) = Valor\ Máximo(abs(diff(Amplitud)))$  , teniendo en cuenta para la contribución de ese valor máximo, el rango de muestras que se encuentra en los últimos 100 ms previos al instante de la disrupción.

Una vez tratadas todas las descargas disruptivas, que intervienen en la fase de entrenamiento, el valor global para el centroide disruptivo se calcula como:

- $C_D(c_{xj}) = \frac{\sum_{j=1}^k C_j(c_{xj})}{k}$  ; siendo  $k$  el número total de descargas disruptivas que intervienen en la fase de entrenamiento.

En la Figura 5.12. se representan de forma conjunta ambos centroides para este caso.

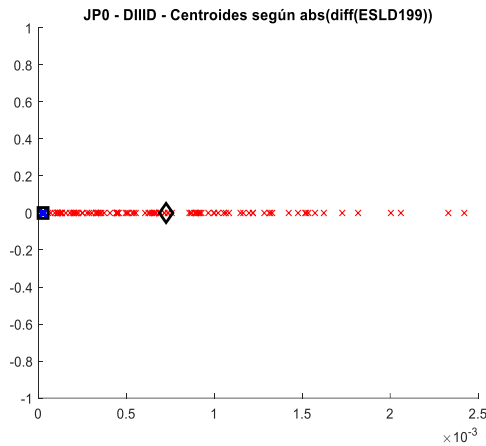


Figura 5.12. Centroides DIII-D con centroide disruptivo considerando 100 ms previos a la interrupción.

2. Otras opciones consideradas para el centroide disruptivo.

Desde la Figura 5.13. a la Figura 5.16. se muestra la representación de otras alternativas, teniendo en cuenta distintas variaciones en cuanto a la obtención del centroide disruptivo.

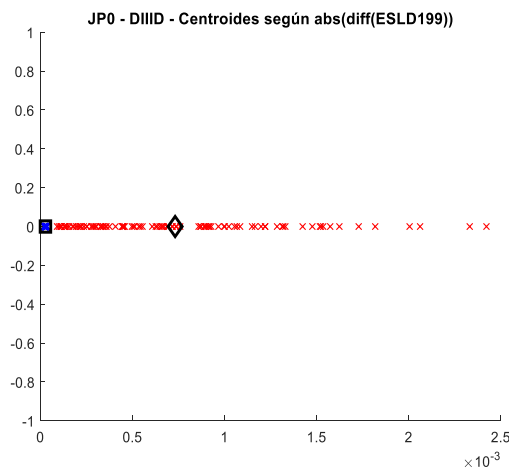


Figura 5.13. Centroides DIII-D con centroide disruptivo considerando el valor máximo en los 200 ms previos a la interrupción.

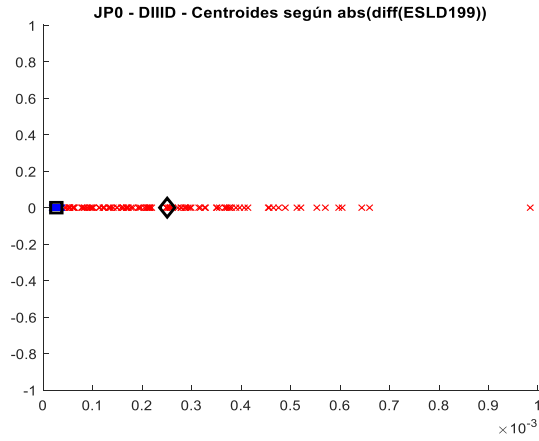


Figura 5.14. Centroides DIII-D con centroide disruptivo considerando el valor medio de los dos valores máximos en los 100 ms previos a la disrupción.

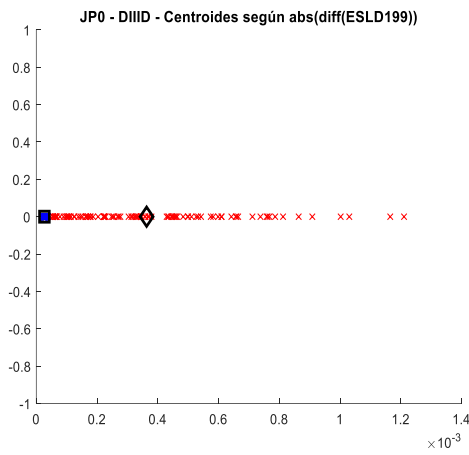


Figura 5.15. Centroides DIII-D con centroide disruptivo considerando el valor medio de los valores máximo y mínimo en los 100 ms previos a la disrupción.

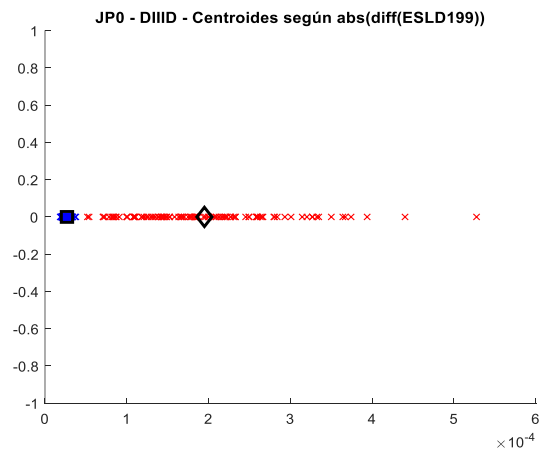


Figura 5.16. Centroides DIII-D con centroide disruptivo considerando el valor medio de los diez valores máximos en los 100 ms previos a la disrupción.



Indicar que, en relación a las gráficas previas, las señales en rojo representan las contribuciones de cada una de las descargas al centroide disruptivo (representado en forma de diamante en negro). Y que, las señales en azul representan las contribuciones de cada una de las descargas al centroide no disruptivo (representado en forma de cuadrado en negro).

Para observar mejor las posiciones relativas de cada uno de los centroides en referencia a sus contribuciones, se dará una imagen ampliada de uno de los casos anteriores en Figura 5.17. y Figura 5.18.

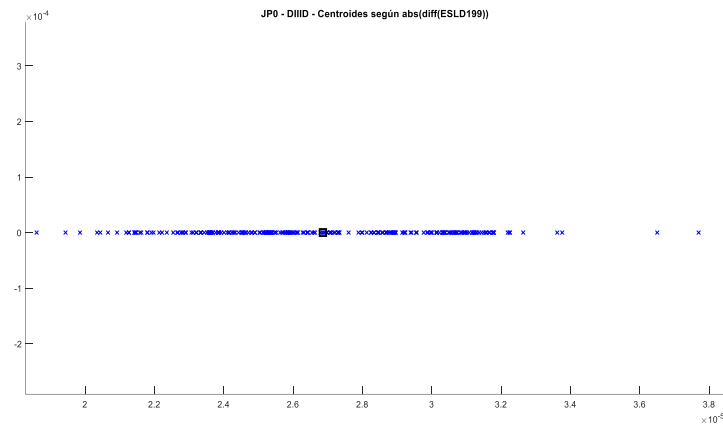


Figura 5.17. Centroide no disruptivo y contribución al mismo de cada una de las descargas.

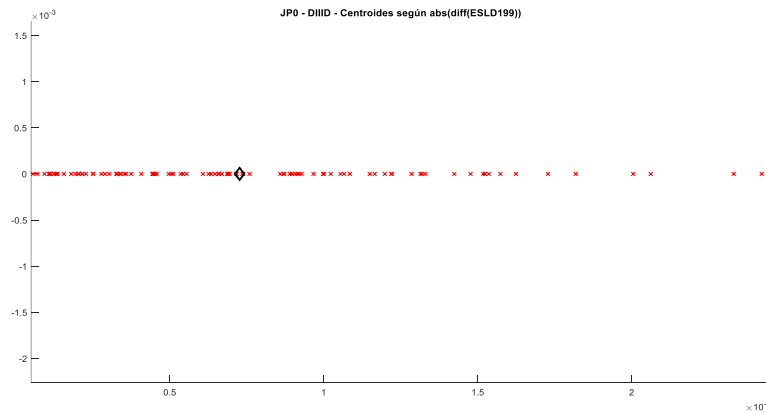


Figura 5.18. Centroide disruptivo y contribución de sus descargas.

### 5.4.3. Fase de evaluación.

El comportamiento disruptivo o no disruptivo, igual que se describe en las secciones 5.2.3. y 5.3.3., viene marcado por la proximidad de la muestra que está siendo

evaluada en la fase de test, con respecto a los centroides obtenidos en la fase de entrenamiento previa.

En este caso, al trabajar en el espacio de una sola dimensión, la frontera entre ambas regiones no vendrá marcada por una recta; sino por un punto. Esa frontera queda determinada mediante la ecuación:

$$\mathbf{Punto}_{frontera} = \frac{C_D(c_{xj}) + C_N(c_{xi})}{2};$$

es decir, la frontera queda determinada por el

valor medio que ocupan los centroides obtenidos durante la fase de entrenamiento.

Como en los casos de los dos dispositivos anteriores, durante la fase de evaluación también se tendrá en cuenta el desplazamiento relativo de los centroides; para tratar de obtener mejores resultados. No obstante, hay que tener en cuenta que mientras para JET y JT-60U esos desplazamientos relativos se hacían a lo largo de la diagonal del plano coordenado o bien a partir de las 81 posiciones diferentes; aquí los desplazamientos relativos siguen el movimiento en relación a un solo eje. Para entenderlo mejor, es como si se trabajara variando sólo la coordenada X, manteniendo la coordenada Y con el valor  $Y = 0$ .

En las Figuras 5.19. y 5.20. se observan ambos comportamientos durante la evaluación de las muestras que componen una descarga, y cuya intensidad de corriente del plasma se encuentra por encima del umbral de 750 kA.



Figura 5.19. Descarga con todas sus muestras no disruptivas.



Figura 5.20. Descarga con alguna de sus muestras disruptivas.

En ambos casos, el punto que aparece marcado como **X** en negro, representa la frontera entre las regiones disruptiva y no disruptiva, situándose a la izquierda de dicha frontera la región no disruptiva. La región disruptiva se sitúa a la derecha del punto frontera. El centroide no disruptivo aparece marcado como un símbolo **O** en negro, y el centroide disruptivo aparece como un símbolo **O** en rojo.

En la Figura 5.19. se observa que todas las muestras que están siendo objeto de la evaluación para la descarga actual, se sitúan en la región no disruptiva. De ahí que, el comportamiento de dicha descarga durante la fase de evaluación, sea no disruptivo.

No ocurre lo mismo para la Figura 5.20., donde existe más de una muestra (marcadas como puntos en rojo), que se sitúan a la derecha del punto que establece la frontera entre ambas regiones. Y, por tanto, al existir alguna muestra disruptiva en la descarga actual; el comportamiento general de dicha descarga será disruptivo.



# APLICACIÓN DEL MÉTODO DE LOS CENTROIDES

---

## 6

---

### 6.1. Justificación en el uso de los centroides.

La predicción del instante de tiempo a la disrupción (TTD) es un problema complejo y de difícil solución hasta el momento. Los predictores de disrupción necesitan ser implementados como ecuaciones simples, utilizando magnitudes físicas que tengan sentido. En esta tesis se describe un método simple y generalizado de cómo desarrollar predictores basados en un principio muy general.

Se considera un espacio de parámetros multidimensionales  $S \subset \mathbb{R}^m$  donde cada dimensión viene representada por una magnitud física. El método general presentado en este trabajo de investigación para desarrollar predictores genéricos, se puede resumir en dos reglas simples:

1. El conocimiento físico de las características disruptivas y no disruptivas en el espacio de parámetros  $S$ , se puede comprimir en dos únicos puntos: centroide disruptivo ( $C_D$ ) y centroide no disruptivo ( $C_N$ ). De forma gráfica, son mostrados en la Figura 6.1.

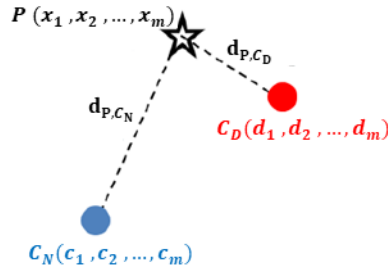


Figura 6.1. Técnica de predicción basada en el centroide más próximo.

2. En cualquier instante  $t$ , el estado del plasma viene representado por un punto  $P$ , que contiene  $m$  características.

En la Figura 6.1.,  $C_D$  y  $C_N$  son centroides, que condensan respectivamente la información de los comportamientos disruptivos y no disruptivos. La técnica basada en el centroide más próximo consiste en determinar la pertenencia de una muestra, que contiene  $m$  características, a una de dichas regiones. La pertenencia de la muestra a una de las dos regiones se basa en el concepto de distancia Euclídea [URL, 55], de forma que si la distancia  $d_{P,C_D} < d_{P,C_N}$ , entonces la muestra representada en el espacio de parámetros como punto  $P$  con sus  $m$  características; tendrá comportamiento disruptivo. En caso contrario, la muestra representada mediante el punto  $P$ , presentará un comportamiento no disruptivo. Evaluar la expresión  $d_{P,C_D} < d_{P,C_N}$  es equivalente a evaluar:

$$\sqrt{\sum_{i=1}^m (x_i - d_i)^2} < \sqrt{\sum_{i=1}^m (x_i - c_i)^2} \quad (6)$$

Operando con la ecuación (6):

$$2 \sum_{i=1}^m (d_i - c_i) \cdot x_i > \sum_{i=1}^m (d_i^2 - c_i^2) \quad (7)$$

Y teniendo en cuenta que las coordenadas de los centroides son fijas, la ecuación (7) se puede expresar como:

$$\sum_{i=1}^m A_i \cdot x_i > K \quad (8)$$

Donde  $A_i$ ,  $i = 1, 2, \dots, m$  y  $K$  son constantes que se obtienen durante el proceso de cálculo de los centroides. Para hacer más intuitiva la explicación anterior, véase Figura 6.2.

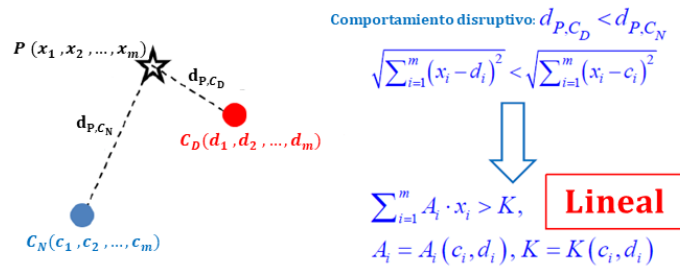


Figura 6.2. Fundamentos del predictor SIDTIP.

Los incrementos en los valores de la señal ML ayudarán a predecir el instante de la interrupción (SIDTIP). Aplicando la técnica del centroide más próximo, se determinará el comportamiento que corresponde a una muestra cualquiera dentro de la descarga que se esté procesando.

El modo en el que se calculan esos centroides ya fue presentado en el capítulo 5, donde además se hacía énfasis en las particularidades de cada máquina. Es decir, no se calculaban de la misma forma dependiendo si se trataba de JET, JT-60U o DIII-D.

Las ventajas de este predictor con respecto a otros, son las siguientes:

1. No se parte de ningún tipo de hipótesis adicional, no se requiere neuro computación ni otro tipo de estructuras.
2. La ecuación (8) que representa las magnitudes físicas es lineal, facilitando cualquier interpretación física relacionada con el plasma.
3. Además, la ecuación (8) permite un reconocimiento muy simple del tipo de comportamiento (disruptivo o no disruptivo), atendiendo a las coordenadas del punto  $P(x_1, x_2, \dots, x_m)$  en cualquier instante.
4. Debido a la validez general de la ecuación, podría ser utilizada en los procesos de elusión y mitigación.

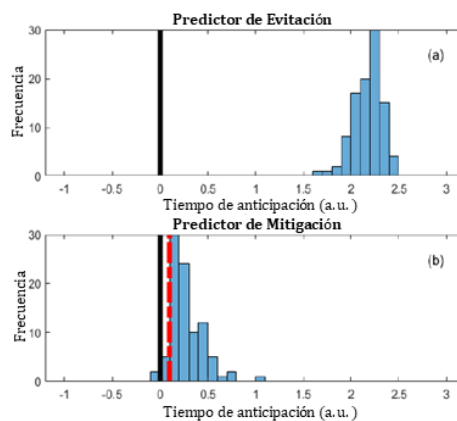


Figura 6.3. Vista conceptual de las distribuciones de tiempo de anticipación.

La Figura 6.3. (a) corresponde a la anticipación para tareas de elusión, y la Figura 6.3. (b) para tareas de mitigación. La línea discontinua en rojo que presenta la Figura 6.3. (b) representa el tiempo mínimo de anticipación para poder aplicar técnicas de mitigación.

5. La predicción se logra a través de una desigualdad lineal, siendo ésta muy eficiente desde el punto de vista computacional, incluso para valores grandes de  $m$ . Implementar esta técnica bajo los requisitos en tiempo real de JET no es un problema, teniendo en cuenta el cálculo de predicciones mediante la ecuación (8).

## 6.2. Base de datos utilizada para JET.

La amplia base de datos utilizada comprende las campañas C15-C37. Estas campañas experimentales tuvieron lugar entre el 24 de abril de 2006 y el 15 de noviembre de 2016. El conjunto de descargas experimentales para dichas campañas, se encuentra en el rango 66080-92504; habiéndose tomados distintos subconjuntos de las mismas, según los diferentes análisis realizados.

Antes de seleccionar el conjunto de datos definitivo con el que realizar el análisis, ha sido preciso un filtrado de los mismos. De todas las campañas de JET consideradas (C15a – C34), tan sólo fueron descartadas 21 descargas no disruptivas. Siendo el motivo de ese descarte, no contener al menos cuatro muestras. Se recuerda que las muestras a considerar, deben tener su intensidad de corriente del plasma ( $I_p$ ) cumpliendo con los requisitos de umbral descritos en la sección 5.2.1., que las ventanas temporales con las que se va a trabajar en JET para este caso son de 4 ms; y que el muestreo se realiza a razón de 1 muestra/ms. Al no completarse ni tan siquiera una ventana, habrá que descartar dicha descarga, porque:

- No realizará ninguna aportación al correspondiente centroide durante la fase de entrenamiento, por no contener los valores de la señal LM que corresponden a la mitad y al final de dicha ventana.
- No podrá ser evaluada en la fase de test, ya que al considerar en esta fase periodos de muestreo de 2 ms (correspondientes a la mitad y al final de la ventana, como en el caso anterior); no se podría evaluar el comportamiento para dicha descarga.



Inicialmente, la base de datos utilizada para el análisis de predictores en JET comprende las campañas y descargas disruptivas y no disruptivas que se muestran respectivamente en Tabla 6.1. y Tabla 6.2. Además del número de descargas por tipo de comportamiento (disruptivo/no disruptivo) y campaña; también se incluyen las posiciones relativas para esos centroides (por tipo de comportamiento + campaña), y los posibles motivos por los que se pueden haber descartado descargas.

**Tabla 6.1.** Información de campañas seleccionadas con descargas disruptivas.

Centroides correspondientes a las Descargas Disruptivas No Intencionadas			
Campaña	Número Descargas Válidas	Coordenada en X	Coordenada en Y
C15a	21	0.000557432	0.000517354
C15b	10	0.000384446	0.000345343
C16	3	0.000270396	0.00025664
C1617	73	0.000631637	0.000544987
C18	14	0.00081872	0.000628731
C19	47	0.000689151	0.0005649
C20	28	0.0009308	0.00118584
C21	16	0.000792075	0.000942733
C22	34	0.000983223	0.00149209
C23	24	0.000978245	0.00140746
C24	14	0.00104833	0.00147978
C25	19	0.000978053	0.0013894
C26	58	0.000857899	0.00125743
C27a	43	0.000934131	0.00140658
C27b	70	0.00112703	0.0016002
C2830	201	0.000942271	0.00122864
C3132	200	0.000901228	0.00131531
C33	151	0.00113712	0.0012921
C34	37	0.00156678	0.0016895

**Tabla 6.2.** Información de campañas seleccionadas con descargas no disruptivas.

Centroides correspondientes a las Descargas No Disruptivas				
Campaña	Número Descargas Válidas	Coordenada en X	Coordenada en Y	Observaciones
C15a	256	0.000169749	0.000169741	
C15b	158	0.000170239	0.000170229	
C16	10	0.0001959	0.000195912	
C1617	1028	0.000176288	0.00017628	4 descargas han sido descartadas por no completar ventana.
C18	302	0.000157861	0.000157855	
C19	585	0.000183791	0.000183783	
C20	702	0.000221778	0.000221769	1 descarga ha sido descartada por no completar ventana.
C21	572	0.000239567	0.000239559	1 descarga ha sido descartada por no completar ventana.
C22	451	0.000241994	0.000241986	
C23	490	0.000171623	0.000171618	
C24	361	0.000166594	0.000166587	1 descarga ha sido descartada por no completar ventana.
C25	568	0.000168572	0.000168565	2 descargas han sido descartadas por no completar ventana.
C26	1322	0.000205865	0.000205859	1 descarga ha sido descartada por no completar ventana.
C27a	319	0.000275821	0.000275812	1 descarga ha sido descartada por no completar ventana.
C27b	513	0.000233892	0.000233884	
C2830	1031	0.000183809	0.000183801	5 descargas han sido descartadas por no completar ventana.
C3132	398	0.000197207	0.000197201	3 descargas han sido descartadas por no completar ventana.
C33	489	0.000176485	0.000176479	2 descargas han sido descartadas por no completar ventana.
C34	139	0.000200392	0.000200386	

Sobre fondo verde se recogen los datos relativos a las campañas de pared de carbono, en fondo naranja los relativos a las campañas de pared metálica (ILW), y sobre fondo rojo se recogen los datos relativos a la campaña de hidrógeno.

Los valores para los centroides representados en la Tabla 6.1. y en la Tabla 6.2. anteriores no son todavía definitivos, ya que corresponden a la fase inicial de la investigación; queriendo poner de manifiesto en este momento, el estado de los datos de entrada para su posterior análisis.

Además, para la fase de evaluación, también han sido utilizados los datos correspondientes a las campañas C35-C37 (descargas 88930-92504), procedentes de JET.

## 6.3. Método de los centroides aplicado a JET.

El objetivo es la implementación de un predictor de disrupciones para JET, a partir de la ecuación (8) del apartado 6.1. de este mismo capítulo. En dicha implementación, se han tenido en cuenta diversas consideraciones:

1. Para facilitar la interpretación física de las predicciones, solo se deben utilizar magnitudes en el dominio del tiempo ya que; si se utilizaran funciones en el dominio de la frecuencia, además de ser más difíciles de interpretar, también impondrían la necesidad de procesamiento de datos en ventanas de tiempo con un número mínimo de muestras.
2. Hasta ahora, los predictores basados en métodos de aprendizaje automático de propósito general no han alcanzado altas tasas de éxito. Para ITER se requiere una tasa de aciertos superior al 95%, a la vez que también se requiere de forma simultánea una tasa de falsas alarmas inferior al 5%. Siendo el objetivo, encontrar el predictor lineal más simple basado en el método del centroide que pueda acercarse a los requisitos impuestos por ITER.
3. El uso de la ecuación (8) de la sección 6.1. establece una frontera de separación entre comportamientos disruptivos y no disruptivos. De forma que, si se tratara de un espacio mono dimensional, la ecuación de la recta que representa la frontera sería:  $A_1 \cdot x_1 = K$ . En espacios mono dimensionales, la frontera está ligada a la existencia de umbrales simples en señales individuales, para identificar condiciones disruptivas. Sin embargo, la experiencia dice que esto no es óptimo, como se menciona en la sección 3 de [Vega, 2020]. Por lo tanto, el predictor más simple no está vinculado a

un umbral basado en la ecuación (8) de un espacio mono dimensional; siendo la frontera de separación entre comportamientos disruptivos y no disruptivos de la forma:  $A_1 \cdot x_1 + A_2 \cdot x_2 = K$ . Lo que significa que el espacio de parámetros correspondiente  $S$ , tiene dimensión 2. En la Figura 6.4. se muestra la representación bidimensional en el dominio del tiempo para la señal ML.

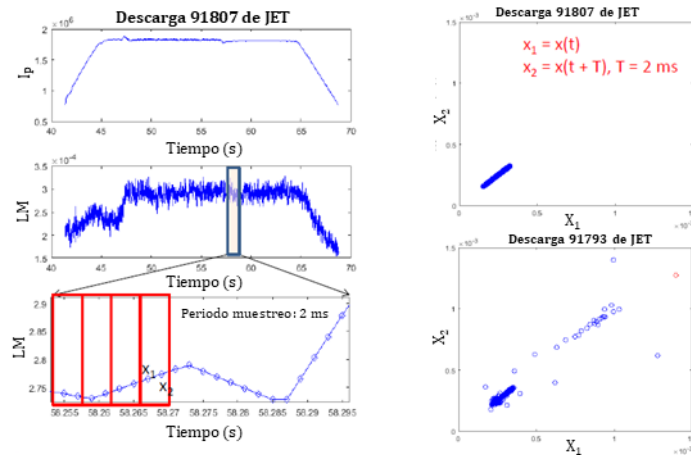


Figura 6.4. Representación de la señal ML en el dominio del tiempo.

En la Figura 6.4. se representa el proceso de selección de las muestras hasta que pasan a formar parte del espacio de parámetros:

- En primer lugar, se seleccionan las muestras que cumplan las condiciones de umbral para la intensidad de corriente del plasma  $I_p$ , estableciéndose un rango temporal dentro de la descarga.
- Teniendo en cuenta el rango temporal obtenido anteriormente, se obtienen los valores de la señal ML.
- Se obtienen las ventanas (de longitud 4 ms) que corresponden a las muestras de la señal ML.
- Como se comprobará a continuación a través de un ejemplo, a partir de los valores contenidos en dichas ventanas, se formarán los puntos en el plano coordenado, que constituirán el espacio de parámetros.
- Una vez que los valores de las muestras son transformados al espacio de parámetros, los puntos que representan esas muestras pueden estar más o menos dispersos; conduciendo respectivamente a un estado disruptivo o no disruptivo del plasma.

4. El predictor desarrollado ha sido optimizado, con propósitos de mitigación, para reconocer comportamientos disruptivos próximos a las disrupciones.
5. La señal ML es una de las magnitudes más relevantes en JET para crear predictores para la mitigación. Por un lado, se sabe que la mayoría de las disrupciones JET generan una ML previa a la disrupción [de Vries, 2016]. Por otro lado, la falta de la señal ML produce predicciones muy malas en JET [Rattá, 2014]. Por lo tanto, la señal ML estará presente en la aplicación del presente predictor. Las consideraciones anteriores motivan el desarrollo de un predictor en un espacio 2D y el uso de la señal ML. Siendo el siguiente paso la búsqueda de una segunda señal para ser utilizada junto con el ML. Sin embargo, si fuera posible, a menor número de señales más simple sería el predictor. Por lo tanto, la idea es usar solo la señal ML mapeándola en un espacio de características bidimensional. Este mapeo es una evolución de un enfoque reciente, encaminado a la detección de anomalías para reconocer eventos disruptivos [Vega, 2015a], [Vega, 2015b] y [Esquembri, 2018]. En el caso del predictor actual (que se basa en los cálculos de los centroides), y suponiendo que la señal ML tiene un período de muestreo  $\tau$ ; el espacio de parámetros 2D está definido por las amplitudes de muestras consecutivas de la señal ML. La Figura 6.5. es un ejemplo de cómo las amplitudes de muestras consecutivas definen puntos en este espacio de características bidimensionales.

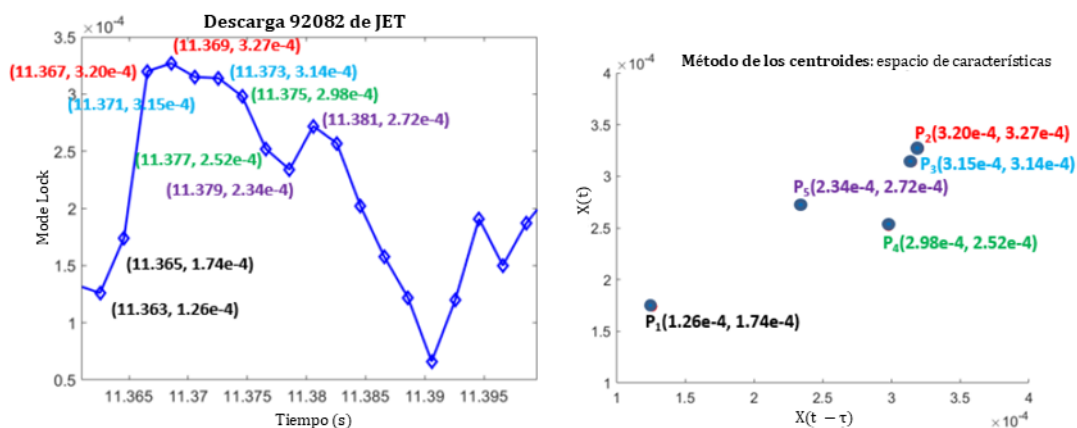


Figura 6.5. Amplitudes de ML en JET.

Atendiendo a la Figura 6.5., y fijando la atención en la parte izquierda de la misma:

- La primera componente de cada uno de los puntos, representa el instante de tiempo en el que se produce la muestra. El periodo de muestreo es  $\tau = 2$  ms. Siendo 10 el número total de muestras etiquetadas.

- La segunda componente, representa el valor de la amplitud para la señal ML. Igual que en el caso anterior, se trata de 10 muestras etiquetadas.
- Como ya se indicó en un capítulo anterior, se tienen en cuenta ventanas temporales de longitud 4 ms. De esa forma, cada par de puntos representados en la parte izquierda de la Figura 6.5., formará un único punto en el espacio de características bidimensional, y que se representa en la parte derecha de la Figura 6.5. Nótese, que atendiendo al número de muestras etiquetadas (10), le corresponden 5 puntos en la parte de la derecha; por existir 5 ventanas de 4 ms. Ha de tenerse en cuenta, además que, de las 10 muestras etiquetadas en la Figura 6.5., las muestras {1, 3, 5, 7, 9} ocuparían la mitad de la ventana, y estarían representadas por la coordenada  $x$  en el espacio de características bidimensional. Las muestras {2, 4, 6, 8, 10} ocuparían el final de la ventana y vendrían representadas por la coordenada  $y$  en el espacio de características.

Fijando ahora la atención en la parte derecha de la Figura 6.5.:

- Cada uno de los 5 puntos que representan el espacio de características de las muestras etiquetadas, tienen como componente  $x$  el valor de la amplitud representada mediante la señal ML en las muestras impares {1, 3, 5, 7, 9}, y como componente  $y$  el valor de la señal ML en las muestras pares {2, 4, 6, 8, 10}.

Cuando se ha logrado una representación de las muestras en el espacio de parámetros, éstas pueden adoptar diferentes disposiciones en el plano bidimensional. Muestra de ello se ofrece en la Figura 6.6.

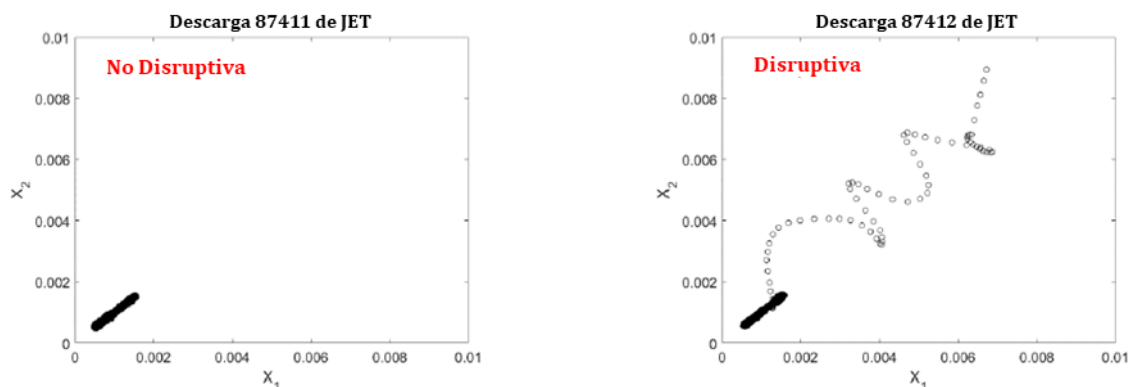


Figura 6.6. Comportamiento de las muestras en el espacio de parámetros.

A la vista de las gráficas de dispersión de las señales ML en el espacio de parámetros de dos muestras consecutivas correspondientes a diferentes descargas, se observa que para descargas cuyo comportamiento es no disruptivo, los puntos aparecen agrupados y compactos. Sin embargo, para una descarga con comportamiento disruptivo, se produce una dispersión de los puntos que representan su espacio de parámetros.

En el caso de una descarga no disruptiva, el intervalo de tiempo considerado para el entrenamiento comienza cuando la corriente de plasma está por encima de cierto umbral, digamos  $I_{T1}$ , hasta que la corriente de plasma disminuye por debajo de un umbral que podría ser diferente  $I_{T2}$ . En la presente tesis se trata con  $I_{T1} = 0.9$  MA e  $I_{T2} = 0.7$  MA.

El entrenamiento de descargas disruptivas en  $\mathcal{S}$ , se determina de diferente modo. En general, la selección de una serie de vectores de características para entrenar predictores no es única, porque los precursores de disrupción no siempre aparecen con la misma anticipación. Sin embargo, para fines de mitigación, es razonable suponer que cuanto más cerca esté el vector de características de la disrupción, más confiable será esa selección. Sólo se considera en  $\mathcal{S}$  una muestra por cada descarga disruptiva y tiene que estar lo más cerca posible de la disrupción.

El cálculo de los centroides para JET ya fue tratado en el capítulo 5.

Los conjuntos de datos de descargas corresponden a operaciones JET con ILW, desde comienzos de 2011 hasta noviembre de 2016. La información específica aparece en la Tabla 6.3.

**Tabla 6.3.** Conjuntos de datos de descargas disruptivas (D) y no disruptivas (ND) para probar un predictor basado en centroides en JET.

Tipo/Fase	Nº de descargas	Rango
Disr/Entrenam	113	80181–82504 SEP 2011–MAR 2012
No disr/Entren	1397	80176–82550 SEP 2011–MAR 2012
Disr/Test	277	82569–92410 MAR 2012–NOV 2016
No disr/Test	3027	82552–92504 MAR 2012–NOV 2016

Cuando se hace referencia a proceso de "entrenamiento", significa proceso de cálculo de los centroides. El desbalanceo entre descargas disruptivas y no disruptivas reflejan el hecho de que el número de disrupciones en JET es aproximadamente el 8% de las descargas.

Los tiempos de disrupción se han definido como los instantes en que comienza el enfriamiento de la corriente. Aproximadamente un tercio de las descargas disponibles se

han utilizado para entrenamiento (cálculos del centroide), y el resto para la prueba. Es importante tener en cuenta que no existe sesgo en la selección de descargas. Prácticamente se han elegido todas las descargas disruptivas y no disruptivas, con la única excepción de eliminar las descargas disruptivas intencionadas y las mitigadas. Por lo tanto, las descargas disruptivas de la Tabla 6.3. son no intencionadas y no mitigadas. El predictor basado en centroides ha sido probado con el bloqueo de modo normalizado a la corriente de plasma. A este respecto, es importante tener en cuenta que el razonamiento llevado a cabo mediante la explicación de la Figura 6.3. sobre la señal ML, es válido para  $ML/I_p$ . La señal  $ML/I_p$  se usa típicamente en el sistema de control de JET, para activar alarmas de disrupción cuando su amplitud está por encima de un umbral seleccionado. El umbral típico en JET se establece cuando  $ML/I_p$  está por encima de un valor generalmente entre 0.400 – 0.520 mT/MA [Reux, 2013].

La Figura 6.7. muestra los centroides disruptivos y no disruptivos obtenidos con las descargas de entrenamiento de la Tabla 6.3.:

centroide disruptivo:  $(d_1 \pm stdx_D, d_2 \pm stdy_D)$

centroide no disruptivo:  $(c_1 \pm stdx_N, c_2 \pm stdy_N)$

dónde:

$(stdx_D, stdy_D) = \text{std}(\delta_{1,i}, \delta_{2,i}), i = 1, \dots, n_D$

$(stdx_N, stdy_N) = \text{std}(\chi_{1,i}, \chi_{2,i}), i = 1, \dots, n_N$

Y donde  $\text{std}()$ , devuelve un vector que contiene las desviaciones estándar.

La línea discontinua en la Figura 6.7. representa la frontera de separación lineal entre comportamientos disruptivos y no disruptivos, en el espacio de parámetros de los centroides.

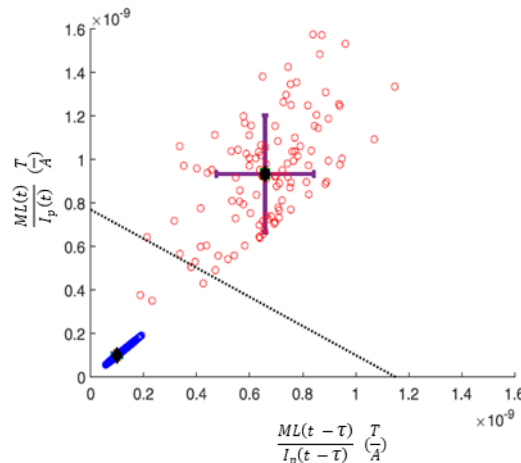


Figura 6.7. Representación de los centroides en JET atendiendo a los datos de la Tabla 6.3.

Los puntos negros con barras de error son los centroides. Los puntos azules y rojos son muestras no disruptivas y disruptivas de descargas individuales, respectivamente. El gráfico muestra los centroides para el modo bloqueo normalizado a la corriente de plasma. El período de muestreo de la señal  $ML/I_p$  ha sido  $\tau = 0.002$  s.

La Figura 6.7. también pone de manifiesto que las muestras no disruptivas forman grupos muy compactos, pero las muestras disruptivas están dispersas. Además, debe enfatizarse que las barras de error del centroide disruptivo no se superponen con el conjunto de muestras no disruptivas. Las barras de error en el caso del centroide no disruptivo hacen que no tengan ninguna influencia en la estimación de la frontera, debido a la compacidad de sus muestras. Esto es una consecuencia del hecho de que  $stdx_N \approx stdy_N \approx 0$ . Sin embargo, en el desplazamiento del centroide disruptivo dentro del espacio cuadrado definido por las barras de error de la Figura 6.8., cualquier pequeño cambio puede tener notables consecuencias. Por lo tanto, para tener en cuenta el impacto potencial de las barras de error, estas se han dividido en cuadrantes; considerándose 9 puntos en cada dirección (9 en la dirección del eje  $x$ , y 9 en la dirección del eje  $y$ ). De esta manera, se estarían considerando 81 ubicaciones diferentes para el centroide disruptivo, tal y como se muestra en la Figura 6.8.

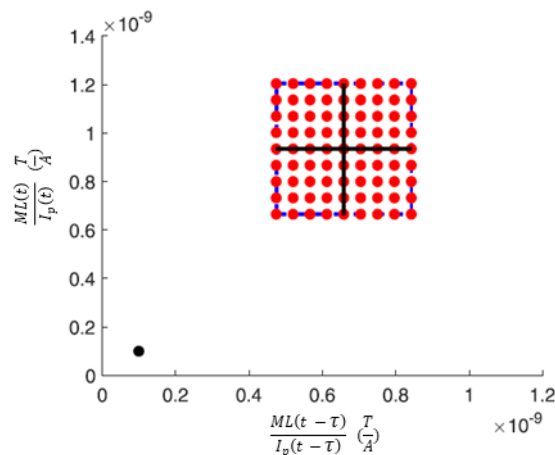


Figura 6.8. Posiciones del  $C_D$  debidas a desplazamientos relativos según las barras de error.

Este hecho ya se puso de manifiesto en el capítulo 5, pero se vuelve a mostrar de nuevo para indicar la razón de uso de esas 81 posibles ubicaciones del  $C_D$ ; debido a las barras de error originadas a partir de la dispersión de muestras disruptivas. Además del  $C_D$  inicial ( $d_1, d_2$ ), se obtienen 80 nuevos centroides disruptivos con incrementos



positivos/negativos de sus coordenadas, en relación con las barras de error en cada dimensión. Los 81 centroides vienen representados por los puntos rojos de la Figura 6.8.

En consecuencia, 81 modelos diferentes tienen que ser probados, debiéndose establecer un criterio para elegir el mejor. Ha sido implementado el análisis de las características operativas del receptor (ROC). Dados diferentes modelos para el mismo conjunto de datos de entrenamiento y prueba, la curva ROC [Cios, 2007] traza la tasa de aciertos de cada modelo frente a sus correspondientes tasas de falsas alarmas. El mejor modelo está definido por el punto (falsa alarma, tasa de éxito) que está más cerca del punto (0, 100).

### 6.3.1. Determinación de los centroides no disruptivos en JET.

Resumiendo lo ya presentado en el capítulo 5 relativo a este punto:

- Se elige un conjunto de datos de descargas no disruptivas  $N_N$ .
- Se determina un centroide por cada descarga ( $C_{Descarga}$ ) en el espacio de parámetros  $(x(t - \tau), x(t))$ .
- El centroide global no disruptivo ( $C_N$ ) estará representado por el valor medio de los centroides individuales a nivel de descarga.

De forma matemática y gráfica, el procedimiento se recoge en la Figura 6.9.

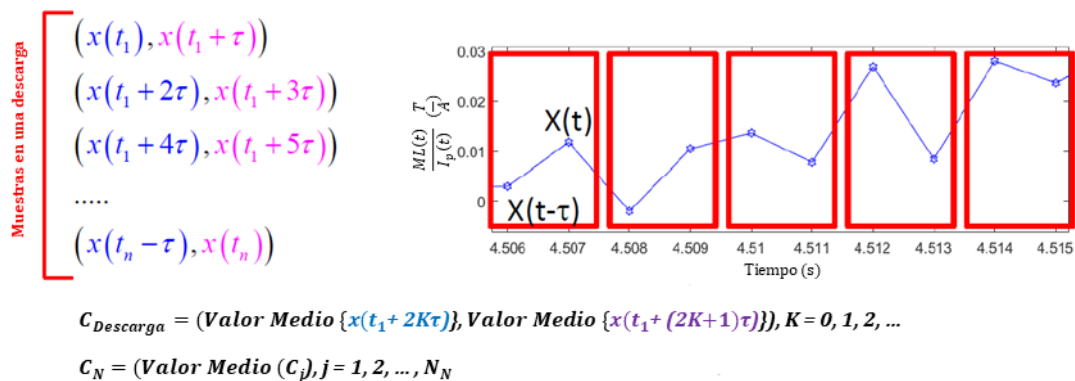


Figura 6.9. Procedimiento de obtención del centroide global no disruptivo en JET.

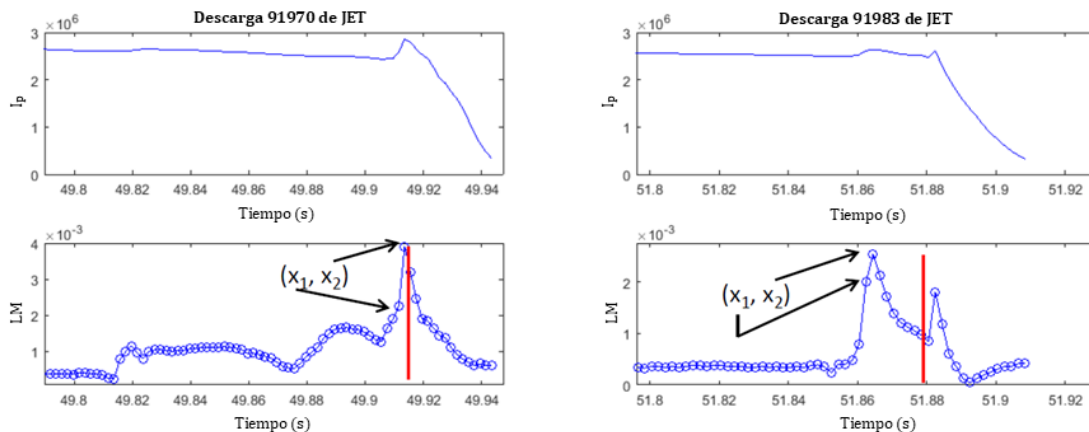
Debe considerarse que, de las descargas seleccionadas, se tendrá en cuenta el rango de muestras cuya intensidad de corriente del plasma ( $I_p$ ) comience estando por debajo de los -750 kA (kiloamperios), y finalice estando por debajo de esos -750 kA.

### 6.3.2. Determinación de los centroides disruptivos en JET.

Al igual que sucedía con el centroide no disruptivo, el centroide disruptivo ya fue tratado en el capítulo 5; mostrando distintas alternativas o criterios para el mismo. Finalmente, fue adoptado el siguiente criterio:

- Se selecciona un conjunto de datos procedente de descargas disruptivas  $N_D$ .
- Se determina un centroide por cada descarga ( $C_{Descarga}$ ) en el espacio de parámetros  $(x(t - \tau), x(t))$ , teniendo en cuenta lo siguiente:
  - Se toma el último punto  $(x(t - \tau), x(t))$ , antes de la disrupción.
  - Se debe tener en cuenta la restricción  $x(t - \tau) < x(t)$ .
- El centroide global disruptivo ( $C_D$ ) estará representado por el valor medio de los centroides individuales a nivel de descarga.

De forma gráfica, el procedimiento se recoge en la Figura 6.10., a través de dos descargas que son tomadas como ejemplos:



**Figura 6.10.** Ejemplos de descargas disruptivas en JET y su contribución al centroide disruptivo.

En la Figura 6.10., la descarga 91970 tiene un valor para las componentes del punto de  $x_2 > x_1$ , por lo que se está cumpliendo la restricción  $x(t - \tau) < x(t)$ . Además  $x_1$  y  $x_2$  se encuentran antes del instante de la disrupción (representado en la Figura 6.10. mediante una línea vertical en rojo).

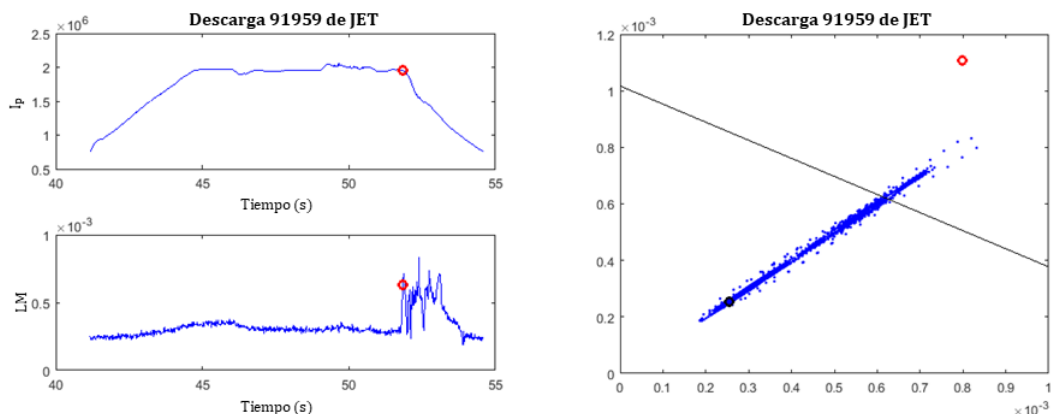
Sin embargo para la descarga 91983, en los instantes previos a la disrupción existen varios puntos donde la segunda componente es menor que la primera  $x_2 < x_1$ ; ya que  $x(t - \tau) > x(t)$ . Por tanto, hay que ir tomando pares de puntos hacia atrás en el dominio del tiempo, hasta conseguir que  $x_2 > x_1$ . Ese hecho sucede cuando se tiene que la

componente  $x_2$  se encuentra en lo alto del “pico” de la gráfica, siendo  $x_1$  el valor de la muestra anterior.

### 6.3.3. Fase de evaluación en JET aplicando el método de los centroides.

Una vez determinados los centroides en la fase de entrenamiento, el siguiente objetivo consiste en validar la eficiencia del método. Para ello, haciendo uso de descargas cuyo comportamiento es conocido, se trata de averiguar si con el método de los centroides se acierta o se falla en la predicción. Igual que en el caso del entrenamiento, serán seleccionadas un conjunto de muestras de cada descarga atendiendo al umbral de intensidad de corriente del plasma  $I_p$ .

Además de obtener aciertos en las predicciones, también pueden darse casos de falsas alarmas, alarmas tardías o alarmas perdidas. Se muestran a continuación varios ejemplos de cada tipo, explicando a continuación los detalles observados en el análisis correspondiente. En la Figura 6.11. se recoge un ejemplo de falsas alarmas.



**Figura 6.11.** Ejemplo de falsas alarmas en JET.

La descarga de la Figura 6.11. muestra todos los síntomas de un comportamiento no disruptivo. A pesar de que se pueden observar ciertas anomalías en el comportamiento de la señal ML, ésta logra recuperarse de las mismas y continua el procesamiento de la descarga. Sin embargo, en la parte derecha de la Figura 6.11., se observa que un conjunto más o menos numeroso de muestras consigue traspasar la frontera y adentrarse en la región disruptiva (región en la que se sitúa el centroide disruptivo, en color rojo).

Aunque tan solo hubiera una muestra que hubiera conseguido traspasar la frontera hacia la región disruptiva, esa descarga pasaría a formar parte de la estadística de falsas alarmas.

En la Figura 6.12. se da un ejemplo de alarma tardía.

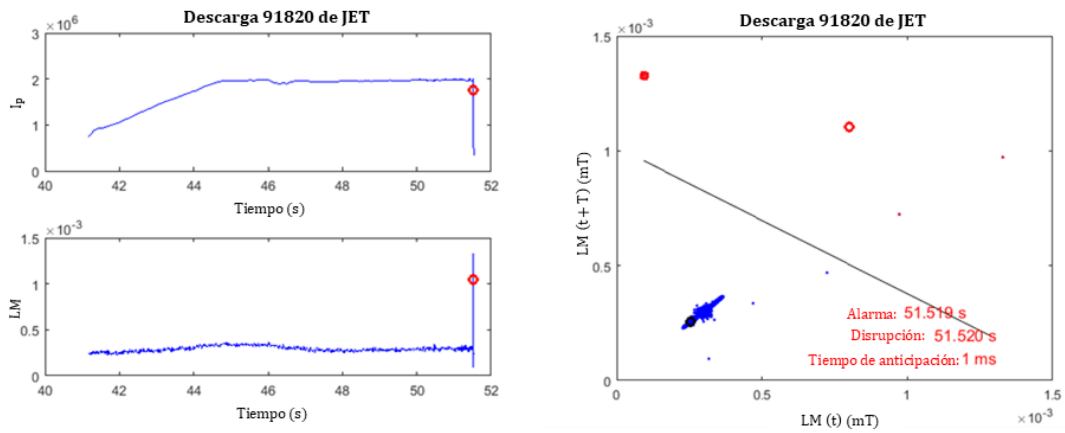


Figura 6.12. Ejemplo de alarma tardía en JET.

En el ejemplo de la Figura 6.12. aparece un comportamiento disruptivo, pero éste no se detecta hasta que es demasiado tarde. El punto rojo grueso de la región disruptiva, representa la primera muestra sobre la que es disparada la alarma. Para que la alarma sea considerada válida en JET, se necesita una anticipación mínima de 10 ms. En la descarga anterior, se predice la disrupción con menor antelación de esos 10 ms necesarios.

Se muestra, en la Figura 6.13., un ejemplo de alarma perdida.

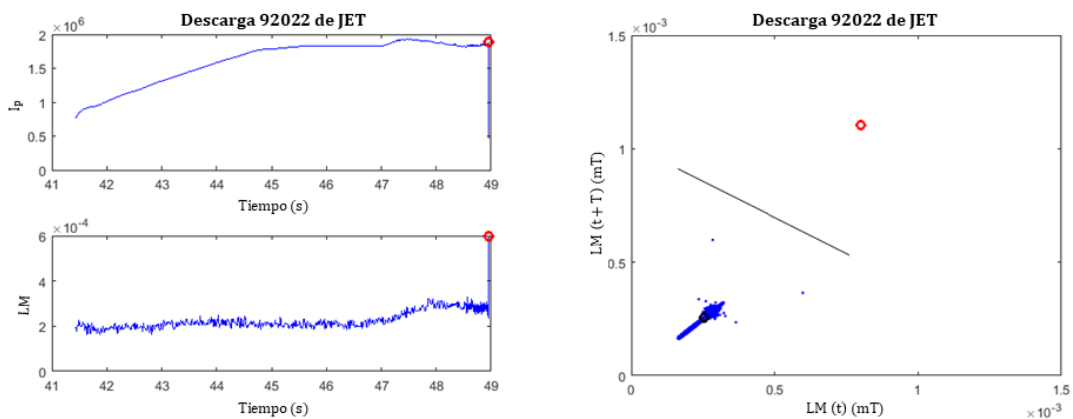


Figura 6.13. Ejemplo de alarma perdida en JET.

Se trata de una descarga que termina siendo disruptiva, pero no llega a detectarse tal comportamiento. Ni siquiera se predice de forma tardía ese comportamiento disruptivo. Teniendo en cuenta la parte derecha de la Figura 6.13., se observa que ninguna

de las muestras de esa descarga llega a traspasar la frontera hacia la región disruptiva. Por lo que, una descarga que es disruptiva en la práctica, el predictor la considera con comportamiento no disruptivo.

## 6.4. Resultados en JET.

Se presentan los resultados atendiendo al análisis de descargas en JET procedentes de las campañas en pared metálica (ILW). Teniendo en cuenta los 81 modelos basados en considerar el desplazamiento del centroide disruptivo, se ha obtenido la clasificación en términos de las tasas de alarma válidas y falsas alarmas junto con la distribución de los tiempos de anticipación. La Figura 6.14. muestra la curva ROC de los 81 modelos diferentes, donde se grafican las tasas de alarma válidas con respecto a las tasas de falsas alarmas.

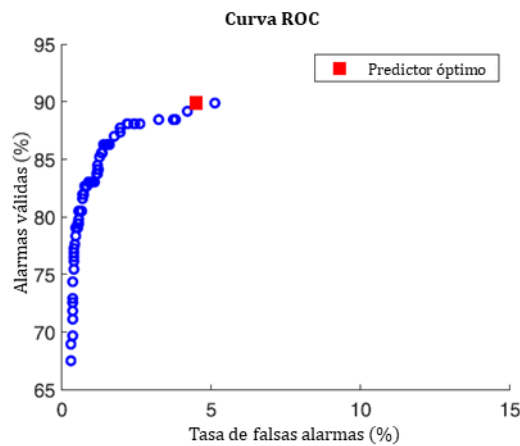


Figura 6.14. Curva ROC para  $ML/I_p$ .

El cuadrado rojo representa el punto más cercano a (0, 100) y determina el mejor predictor. El predictor más cercano a (0, 100) es el predictor óptimo y corresponde al centroide disruptivo ubicado en el punto  $(d_1 - 0.75 \cdot stdx_D, d_2 - stdy_D)$ . La frontera de separación entre comportamientos disruptivo y no disruptivo viene dada por la ecuación de la recta  $X(t) = -0.7441 \cdot X(t - 0.002) + 6.1243 \times 10^{-10}$ , donde  $X(t) = ML(t)/I_p(t)$ , 0.002 s es el período de muestreo, ML está en Tesla e  $I_p$  en A. Por lo tanto, con este modelo y con este periodo de muestreo, se reconoce un comportamiento disruptivo cuando

$$X(t) > -0.7441 \cdot X(t - 0.002) + 6.1243 \times 10^{-10} \quad (9)$$

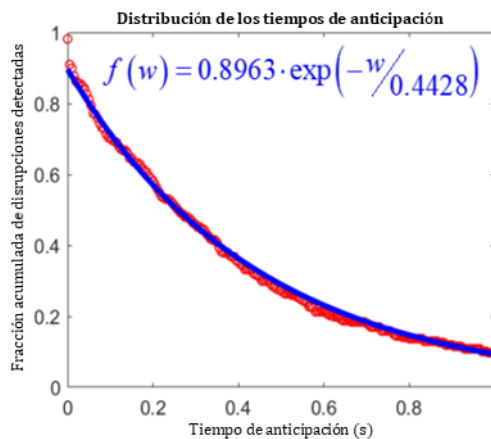
Al aplicar la ecuación (9), que sabe reconocer un comportamiento disruptivo, para el conjunto de datos de prueba de la Tabla 6.3.; se obtienen los resultados que se muestran en la Tabla 6.4.

**Tabla 6.4.** Resultados del predictor seleccionado por la curva ROC.

Acumulación (%)	98
Alarmas válidas (%)	90
Falsas alarmas (%)	4
Tiempo medio de anticipación (ms)	443

El tiempo promedio de anticipación, está determinado por el ajuste de la distribución de tiempos de anticipación a un modelo exponencial.

La Figura 6.15. muestra la distribución de tiempos de anticipación.



**Figura 6.15.** Distribución de los tiempos de anticipación.

La distribución de los tiempos de anticipación ( $w$ ), sigue un modelo exponencial con un tiempo de anticipación promedio de 443 ms. Esta distribución se puede adaptar a un modelo exponencial de la forma  $f(w) = f_0 \cdot e^{(-\frac{w}{T})}$ , donde  $f(w)$  es la fracción de disrupciones detectadas con un tiempo de anticipación  $w$ ,  $f_0$  es la fracción de disrupciones detectadas con tiempos de anticipación positivos y  $T$  es el tiempo de anticipación promedio de las predicciones. Los parámetros del modelo resultante son  $f_0 = 0.8963 \pm 0.0055$  y  $T = 0.4428 \pm 0.0039$  s, donde las estimaciones se han realizado con un 95% de grado de confianza y el factor R-cuadrado del ajuste es 0.9778. El resumen del mejor predictor obtenido por el análisis ROC se muestra en la Tabla 6.4. El resto de predictores

en el análisis ROC se implementan de manera muy similar a la ecuación (9). Todos ellos siguen la expresión:

$$X(t) > A \cdot X(t - 0.002) + B \quad (10)$$

Donde A y B son constantes diferentes para cada predictor, y dependen de las coordenadas de los respectivos centroides. Cabe destacar que una simple desigualdad lineal en dos variables puede predecir con alta confiabilidad (alta tasa de alarmas válidas y baja tasa de falsas alarmas) comportamientos disruptivos en JET. La desigualdad es el resultado de condensar en dos centroides, el carácter disruptivo/no disruptivo de descargas de JET con campañas ILW.

Durante una descarga continua, se obtienen periódicamente puntos en la función y su centroide más cercano determina el estado plasmático. La señal utilizada en la predicción ( $ML/I_p$ ), es una medida común utilizada en JET para reconocer comportamientos disruptivos; pero se ha realizado para un procesamiento particular de datos. La Figura 6.16. muestra diagramas de dispersión de una descarga no disruptiva (Figura 6.16.(a)) y dos predicciones de disrupción exitosas (Figura 6.16. (b) y Figura 6.16. (c)).

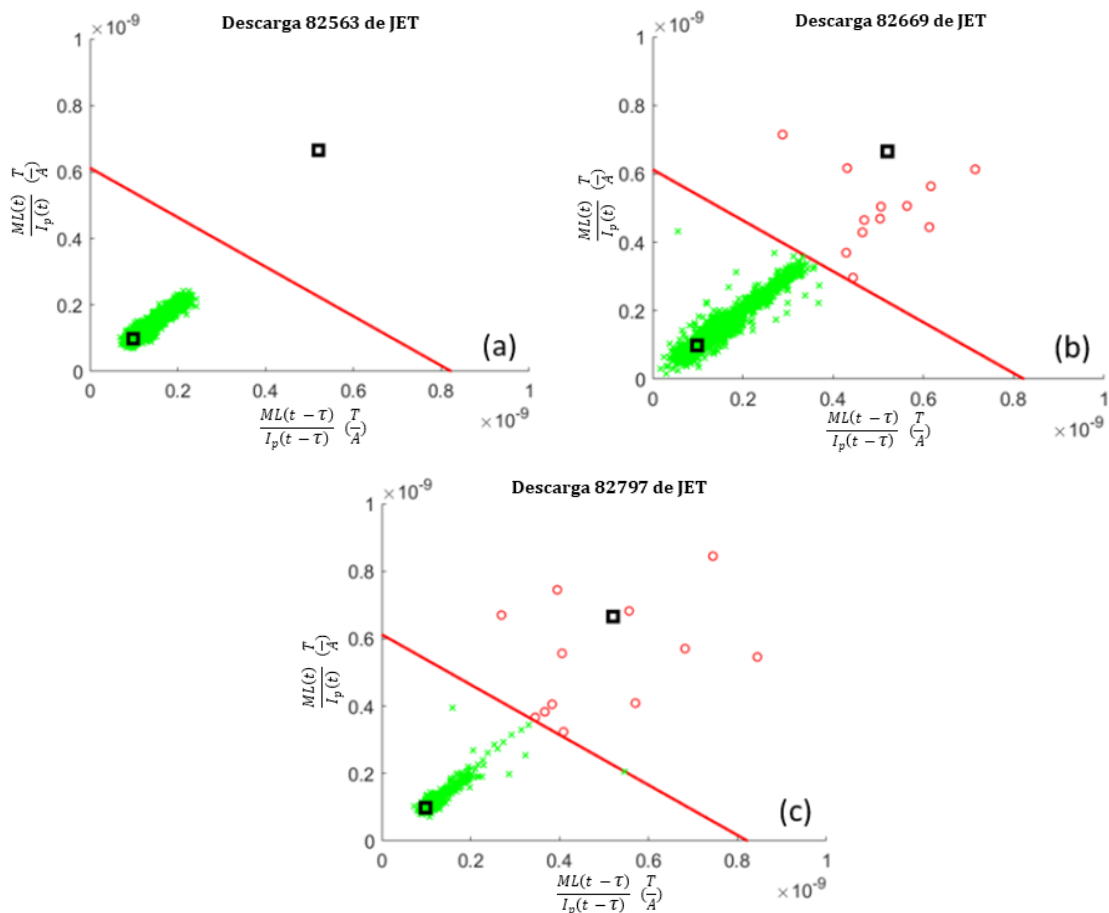


Figura 6.16. Diagramas de dispersión.

Los cuadrados negros son los centroides, las cruces verdes representan comportamientos no disruptivos y se encuentran debajo de la frontera de separación (línea roja). Los círculos rojos son comportamientos disruptivos y aparecen por encima de la frontera de separación. Es importante enfatizar que los puntos no disruptivos están bastante concentrados alrededor de la diagonal del primer cuadrante. Sin embargo, los puntos disruptivos están muy extendidos con respecto a dicha diagonal.

Aplicando un análisis fuera de línea de las descargas disruptivas para campañas ILW en JET, y teniendo en cuenta las alarmas activadas por la ecuación (9); se establece que los puntos disruptivos que se encuentran cerca de la diagonal, permiten obtener tiempos de anticipación mayores que los que están lejos de la diagonal. Es importante tener en cuenta que el término "tiempo de anticipación" es el intervalo real de tiempo entre el instante en el que se lanza la alarma y el instante en el que se produce la disrupción. El hecho empírico de que cuanto más cerca están los puntos de la diagonal, mayores son los tiempos de anticipación, sugiere la hipótesis de que cuando se activa la alarma, el instante a la disrupción (TTD) está relacionado con la distancia a la diagonal  $d_D$ , es decir,  $TTD = f(d_D)$ .

En la sección 9 y en el Apéndice de [Vega, 2020], se dan más detalles acerca de la interpretación física de la distancia desde el espacio de parámetros a la diagonal.

A continuación, se cuantificarán los resultados obtenidos para los distintos análisis de JET.

Se realizó el análisis para las campañas C36-C37, teniendo en cuenta para la fase de entrenamiento 151 descargas disruptivas no intencionadas (descargas 89515 - 90985) y 995 descargas no disruptivas (descargas 89478 – 91000). En la fase de evaluación se utiliza el rango de descargas (91807 – 92504), de las cuales 155 son descargas disruptivas no intencionadas, y 385 son no disruptivas. En la Tabla 6.5. se presentan los resultados según el predictor SPAD y según el predictor basado en la ecuación lineal definida a partir de los centroides, para el conjunto de descargas validadas.

**Tabla 6.5.** Resultados según SPAD y según el método de los centroides C36 - C37.

	Falsas alarmas (%)	Alarmas perdidas (%)	Detecciones tardías (%)	Alarmas válidas (%)	Alarmas prematuras (%)
SPAD	21.30	3.23	1.29	87.09	8.39
LEP	5.46	5,16	3.23	90.97	0.65



A la vista de los resultados, y a pesar de que la tasa de alarmas perdidas y las detecciones tardías son ligeramente superiores teniendo en cuenta el método de los centroides; claramente éste presenta notables ventajas sobre SPAD. La tasa de falsas alarmas se sitúa en torno al 5% exigido por ITER (concretamente 5,46%), frente a la tasa del 21,30% que presenta SPAD. Por otra parte, el método de los centroides presenta una tasa de aciertos del 90,97% frente al 87,09% que presenta SPAD.

También se ofrecen los resultados para la campaña de hidrógeno (C34), interviniendo 16 descargas disruptivas no intencionadas (87640 – 87804) y 68 descargas no disruptivas (87606 – 87805) para la fase de entrenamiento. La fase de test se realizó con 19 descargas disruptivas (87811 – 87878) y con 71 descargas no disruptivas (87807 – 87918). Los resultados se muestran en la Tabla 6.6.

**Tabla 6.6.** Resultados según el método de los centroides para C34.

	Falsas alarmas (%)	Alarmas perdidas (%)	Detecciones tardías (%)	Alarmas válidas (%)	Alarmas prematuras (%)
LEP	0	21,05 (4/19)	0	78,95 (15/19)	0

Aunque no existe presencia de falsas alarmas, detecciones tardías y alarmas prematuras; el 21,05% de tasa de alarmas perdidas y el 78,95% de tasa de aciertos no son lo suficientemente buenas.

Si para C34, en lugar de tener en cuenta los centroides obtenidos a partir de esa misma campaña, se tienen en cuenta los centroides generados a partir de las campañas en pared de carbono; manteniendo el mismo conjunto de test, se obtienen los resultados que ofrece la Tabla 6.7.

**Tabla 6.7.** Resultados según centroides con campaña C34 y con pared de carbono.

	Falsas alarmas (%)	Alarmas perdidas (%)	Detecciones tardías (%)	Alarmas válidas (%)	Alarmas prematuras (%)
LEP (Centroides C34)	0	21,05 (4/19)	0	78,95 (15/19)	0
LEP (Centroides CW)	2,82 (2/71)	15,79 (3/19)	0	84,21 (16/19)	0

Los resultados obtenidos, teniendo en cuenta la generación de los centroides en pared de carbono (CW), mejoran con respecto a los centroides generados a partir de la propia campaña C34. Así, aunque la tasa de falsas alarmas alcanza un 2,82%, disminuyen las alarmas perdidas y aumenta la tasa de aciertos.

Finalmente, se realiza un análisis masivo de datos, interviniendo para el cálculo de los centroides (fase de entrenamiento), las campañas C15 – C20 en pared de carbono

(CW). De dichas campañas, fueron seleccionadas 179 descargas disruptivas no intencionadas (66027 – 73062), y 3041 descargas no disruptivas (65988 – 73127).

Para la fase de evaluación fueron tomadas:

- 1084 descargas disruptivas no intencionadas (73460 – 92500), de las cuales:
  - 278 eran descargas en pared de carbono (CW) procedentes de las campañas C21 – C27.
  - 806 eran descargas en pared metálica (ILW), procedentes de las campañas C28 – C37; excepto la C34 (campaña de hidrógeno).
- 8039 descargas no disruptivas, de las cuales:
  - 4602 eran descargas en pared de carbono (CW) procedentes de las campañas C21 – C27, con números de descarga 73132 - 79853.
  - 3437 eran descargas en pared metálica (ILW), procedentes de las campañas C28 – C37; excepto la C34 (campaña de hidrógeno), con números de descarga 81852 – 92504.

Los resultados se pueden observar en la Tabla 6.8.

**Tabla 6.8.** Resultados según el método de los centroides para C15 - C37.

	Falsas alarmas (%)	Alarmas perdidas (%)	Detecciones tardías (%)	Alarmas válidas (%)	Alarmas prematuras (%)
LEP	5.50	7.47	5.63	82.20	4.70

Tomando el mismo conjunto de entrenamiento anterior (campañas C15 – C20), y teniendo en cuenta para la fase de test:

- 805 descargas disruptivas no intencionadas procedentes de campañas ILW (C28 – C37, excepto C34), con números de descarga 80181 – 92500.
- 3424 descargas no disruptivas procedentes de ILW (C28 – C37, excepto C34), con números de descarga 81852 – 92504.

Introduciendo la aplicación de las barras de error, proporcionando un desplazamiento relativo a las ubicaciones originales de los centroides calculados, y teniendo en cuenta esos desplazamientos relativos para los casos:

- 0 = Se tienen en cuenta las posiciones originales de los centroides.
- -0,25 = Los centroides se ven desplazados  $\frac{1}{4}$  de la longitud de las barras de error, en la dirección decreciente de la diagonal.

- -0,50 = Los centroides se ven desplazados ½ de la longitud de las barras de error, en la misma dirección indicada anteriormente.
- -0,75 = Los centroides se desplazan ¾ de la longitud indicada anteriormente.
- -1 = Los centroides se desplazan hasta el extremo de la diagonal en la parte inferior de las barras de error.

Para ver con mayor detalle, se representa como muestra la Figura 6.17.

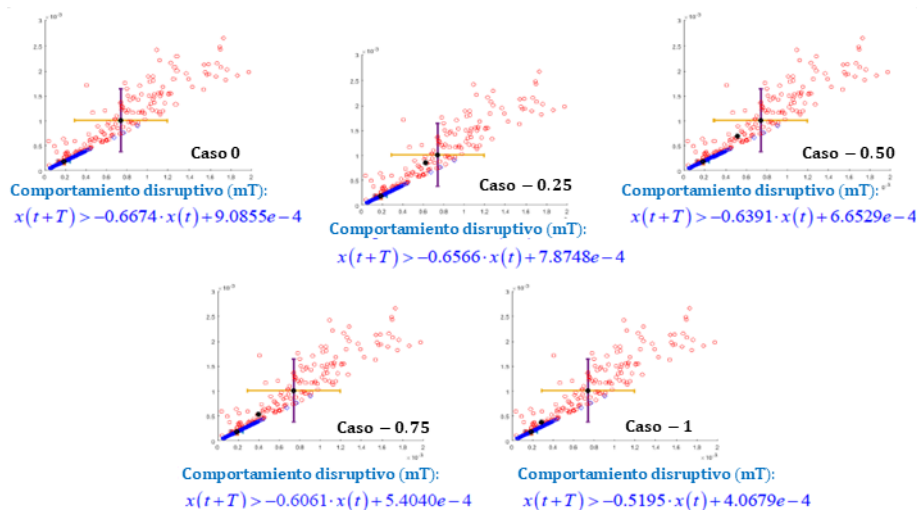


Figura 6.17. Desplazamiento relativo de los centroides obtenidos de C15 – C20.

Los resultados que se obtienen después de aplicar los desplazamientos relativos anteriores a los centroides originales, se muestran en la Tabla 6.9.

Tabla 6.9. Resultados según el método de los centroides con barras de error.

**Alarmas perdidas: tardías + totalmente perdidas**

Caso	Tasa aciertos (%)	Alarmas perdidas (%)	Tardías (%)	T.Perd.(%)	Falsas alarm. (%)
0	89.44	10.56	5.47	5.09	1.75
-0.25	90.81	9.19	5.34	3.85	3.21
-0.50	92.80	7.20	3.73	3.47	8.50
-0.75	94.04	5.96	3.11	2.86	27.19
-1	97.77	2.23	1.61	0.62	70.65

↓  
No detectadas en absoluto  
Detectadas con anticipación < 10 ms

Tal y como se puede observar en la Tabla 6.9., a medida que los centroides se van alejando de su posición original, mejoran todas las tasas; excepto la tasa de falsas alarmas. Ésta última aumenta muy rápido desde el caso -0.50 al caso -1. Por tanto, hay que buscar

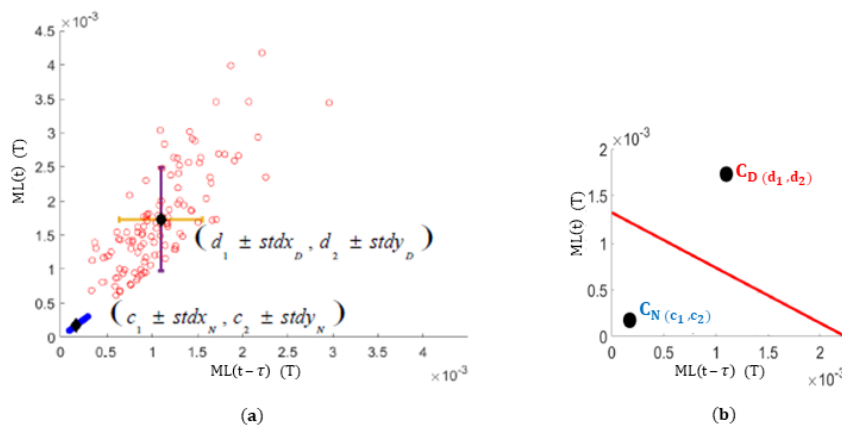
un equilibrio entra la tasa de aciertos y la tasa de falsas alarmas; sin descuidar el resto de los indicadores.

Teniendo en cuenta los datos de entrenamiento y de test de la Tabla 6.10., se realizaron nuevos conjuntos de pruebas. Donde además de tener en cuenta la señal ML, también se tuvo en cuenta la señal ML normalizada según la corriente del plasma ( $ML/I_p$ ).

**Tabla 6.10.** Base de datos utilizada usando campañas ILW.

Tipo/Fase	Nº de descargas	Rango
Disr/Entrenam.	113	80181-82504 SEP 2011-MAR 2012
No disr/Entre.	1397	80176-82550 SEP 2011-MAR 2012
Disr/Test	281	82569-92410 MAR 2012-NOV 2016
No disr/Test	3027	82552-92504 MAR 2012-NOV 2016

Como se ha visto anteriormente, la influencia de las barras de error produce un impacto en los resultados. Las barras de error en el caso de los centroides no disruptivos no tienen ninguna influencia en la estimación de la frontera. Sin embargo, el desplazamiento del centroide disruptivo dentro de la caja cuadrada definida por las barras de error, hace que cambie la pendiente de la recta que define la frontera entre ambas regiones.



**Figura 6.18.** Impacto del desplazamiento de los centroides en la recta que define la frontera.

La Figura 6.18.(a) muestra las posiciones originales de los centroides, incluyendo sus barras de error. La Figura 6.18.(b) muestra la recta que define la frontera de separación entre las regiones disruptiva y no disruptiva. La pendiente y la posición de dicha recta variará dependiendo de los desplazamientos relativos de los centroides, siendo el desplazamiento relativo del centroide disruptivo el que mayor impacto va a tener sobre dicha frontera.

Si en lugar de tener en cuenta las 5 posiciones relativas anteriores para los centroides, se tienen en cuenta nuevas posiciones, tal y como recoge la Figura 6.19.

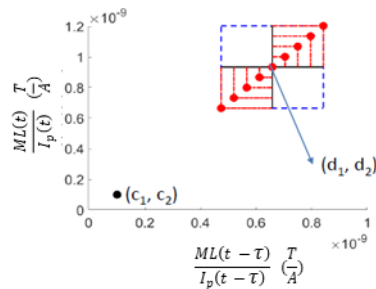


Figura 6.19. Posiciones relativas de los centroides para las nuevas pruebas.

En la Figura 6.19. se están teniendo en cuenta 9 posiciones diferentes para el centroide disruptivo. Dichas posiciones vienen determinadas por los resultados de la Tabla 6.11.

Tabla 6.11. Cálculo de los puntos donde se desplaza el centroide disruptivo.

Etiqu.	Coordenada x	Coordenada y
1	$d_1 + stdx_d$	$d_2 + stdy_d$
0.75	$d_1 + 0.75 \cdot stdx_d$	$d_2 + 0.75 \cdot stdy_d$
0.5	$d_1 + 0.5 \cdot stdx_d$	$d_2 + 0.5 \cdot stdy_d$
0.25	$d_1 + 0.25 \cdot stdx_d$	$d_2 + 0.25 \cdot stdy_d$
0	$d_1$	$d_2$
-0.25	$d_1 - 0.25 \cdot stdx_d$	$d_2 - 0.25 \cdot stdy_d$
-0.5	$d_1 - 0.5 \cdot stdx_d$	$d_2 - 0.5 \cdot stdy_d$
-0.75	$d_1 - 0.75 \cdot stdx_d$	$d_2 - 0.75 \cdot stdy_d$
-1	$d_1 - stdx_d$	$d_2 - stdy_d$

Con las posiciones anteriores para el centroide disruptivo, se generan los siguientes resultados para el conjunto de datos de JET que presenta la Tabla 6.10. Esos resultados se pueden observar en la Tabla 6.12.

Tabla 6.12. Resultados para el conjunto de datos definido en la Tabla 6.10.

Modelo	ML/Ip		ML	
	T.Aciertos (%)	Tasa Falsas Alarmas (%)	T.Aciertos (%)	Tasa Falsas Alarmas (%)
1	95.73	0.30	75.80	0.10
0.75	96.44	0.36	80.43	0.13
0.5	97.51	0.46	86.12	0.13
0.25	97.86	0.66	89.32	0.30
0	98.58	0.99	93.24	0.50
-0.25	99.29	1.35	95.73	0.86
-0.5	99.29	1.95	97.51	1.62
-0.75	99.29	3.73	98.58	2.61
-1	99.29	5.12	99.64	4.33

Los resultados resaltados en rojo en la Tabla 6.12., muestran los mejores resultados de entre los que fueron probados, tanto para la señal  $ML/I_p$  como para la señal ML. Para verificar que son estos, y no otros los mejores resultados, se recurre al uso de las gráficas ROC, que tienen en cuenta la tasa de aciertos frente a la tasa de falsas alarmas; mostrándose en la Figura 6.20.

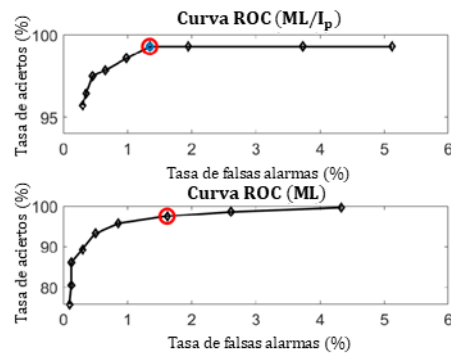


Figura 6.20. Gráficas ROC que muestran de entre todos los resultados, el resultado óptimo.

La interpretación de dichos resultados se traduce en que:

- Considerando la señal  $ML/I_p$ , el resultado óptimo se consigue considerando un desplazamiento relativo del centroide disruptivo igual a  $-0,25$ .
- Teniendo en cuenta la señal ML, el resultado óptimo se consigue con un desplazamiento relativo del centroide disruptivo igual a  $-0,50$ .

A la vista de los resultados anteriores, se puede decir que son excelentes; ya que la tasa de aciertos  $> 95\%$  y la tasa de falsas alarmas  $< 5\%$ .

Ya fue indicado con anterioridad, que los puntos disruptivos que se encuentran cerca de la diagonal permiten obtener tiempos de anticipación mayores que los que están lejos de la diagonal.

Continuando con la optimización del predictor basado en el centroide más próximo, se introduce un tipo de bandas, para filtrar falsas alarmas. Para ello se comenzará con las representaciones de los diagramas de dispersión de distintos comportamientos del plasma, y teniendo en cuenta la evolución temporal de la señal  $ML/I_p$ . En la Figura 6.21. aparecen representados dichos comportamientos.

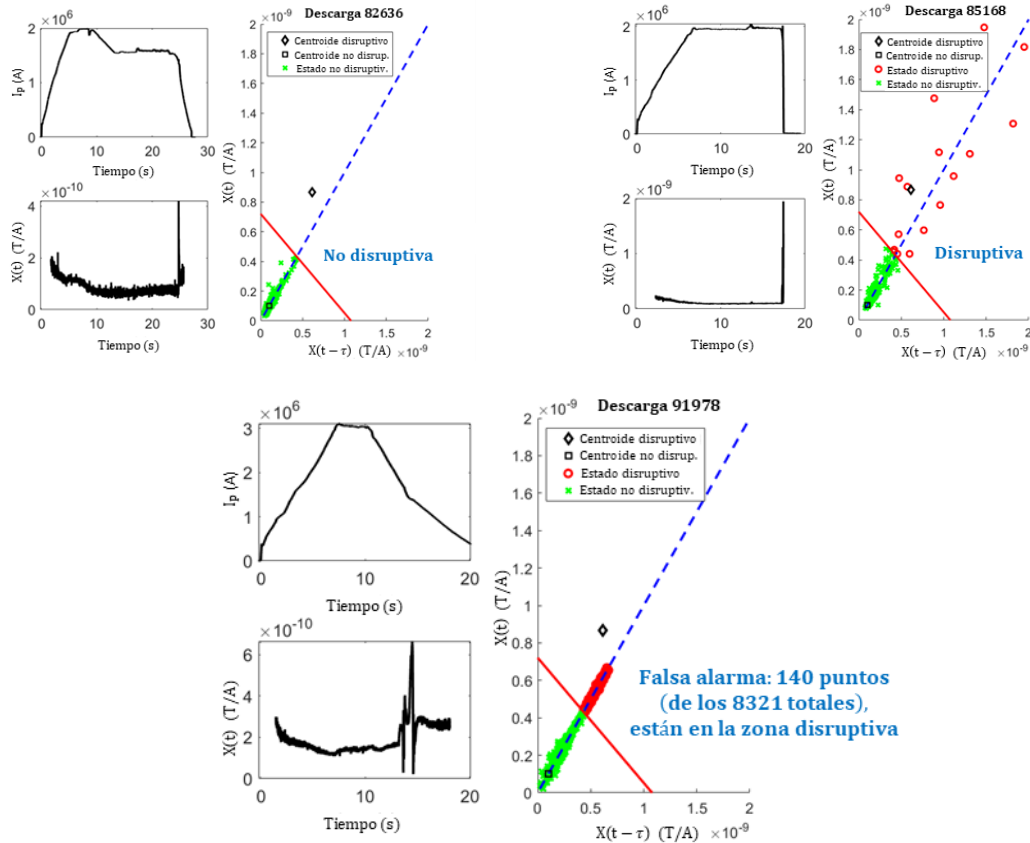


Figura 6.21. Distribución de las muestras en torno a la diagonal.

Las representaciones de la Figura 6.21. sugieren que las muestras en torno a la diagonal, podrían proporcionar algún tipo de información. Por este motivo, se introduce el uso de las bandas, que son rectas paralelas a la diagonal por encima y por debajo de la misma (marcadas en línea discontinua), tal y como ilustra la Figura 6.22.

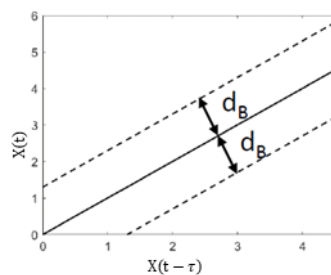


Figura 6.22. Disposición de las bandas en torno a la diagonal.

Algunas de las falsas alarmas pueden estar asociadas a disrupciones menores que el plasma puede recuperar. El resto de las falsas alarmas muestran el mismo patrón, estando bastante concentradas alrededor de la diagonal (están dentro de una banda

alrededor de la diagonal), y donde la diferencia de amplitudes entre las muestras consecutivas no excede un cierto valor:

$$|x(t - \tau) - x(t)| \leq d_B.$$

Si se introduce el filtrado de las bandas para la descarga 86268 (véase Figura 6.23.), se obtiene el siguiente diagrama de dispersión.

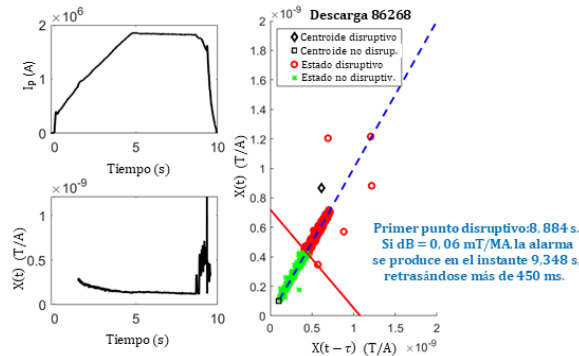


Figura 6.23. Emplazamiento de muestras disruptivas en torno a la diagonal para descarga 86268.

Esto permite retrasar el disparo de la alarma 450 ms más tarde que si no se aplicaran el uso de dichas bandas, lo que conlleva la obtención de mayor información durante el análisis de la descarga. La obtención de la banda óptima permite reducir la cantidad de falsas alarmas. Han sido probados distintos valores:

$d_{B,k} = k \cdot \sigma_D$ , con  $k = 0, 0.01, 0.02, \dots, 0.1$  y siendo  $\sigma_D$  la desviación estándar de la distancia al centroide disruptivo para  $x(t) = x(t - \tau)$ .

Teniendo en cuenta los resultados de la Tabla 6.12., en [Vega, 2018c] se aplicaron dichas bandas, mostrándose sus resultados en la Tabla 6.13.

Tabla 6.13. Aplicación de las bandas a los resultados de la Tabla 6.12.

$$d_{B,k} = k \cdot \sigma_D, k = 0, 0.01, 0.02, \dots, 0.1$$

k	dB,k (mT/MA)	SR (%)	WT>0 (%)	FA (%)	k	dB,k (mT)	SR (%)	WT>0 (%)	FA (%)
0	0	99.29	97.51	1.35	0	0	97.51	96.80	1.62
0.01	0.0021	99.29	97.51	1.32	0.01	0.0049	97.15	96.44	1.55
0.02	0.0041	99.29	97.51	1.32	0.02	0.0097	97.15	96.44	1.49
0.03	0.0062	99.29	97.51	1.32	0.03	0.0146	97.15	96.44	1.42
0.04	0.0083	99.29	97.51	1.29	0.04	0.0195	96.80	96.09	1.35
0.05	0.0103	99.29	97.51	1.29	0.05	0.0244	96.80	96.09	1.22
0.06	0.0124	99.29	97.51	1.16	0.06	0.0292	96.80	96.09	1.16
0.07	0.0145	99.29	97.51	1.16	0.07	0.0341	96.80	96.09	1.12
<b>0.08</b>	<b>0.0165</b>	<b>99.29</b>	<b>97.51</b>	<b>1.12</b>	0.08	0.0390	96.44	95.73	1.06
0.09	0.0186	98.58	97.15	1.12	0.09	0.0438	94.44	95.73	0.89
0.1	0.0207	98.58	97.15	1.02	0.1	0.0487	95.73	95.37	0.83



Los resultados de la Tabla 6.12. y de la Tabla 6.13. arrojan las siguientes conclusiones:

1. Para la señal  $ML/I_p$  (parte izquierda de la Tabla 6.13.), se ha logrado mantener la tasa de aciertos en el 99,29%, reduciendo la tasa de falsas alarmas del 1,35% al 1,12%. Siendo, además, los tiempos de anticipación positivos en el 97,51% de los casos. Los resultados se encuentran resaltados en rojo.
2. Para la señal ML (parte derecha de la Tabla 6.13.), se produce una reducción de las falsas alarmas a medida que se ensancha la banda. Sin embargo, también se va reduciendo la tasa de aciertos; por lo que habría que buscar un equilibrio que haga eficiente el predictor. En cualquier caso, como ITER requiere una tasa de aciertos  $> 95\%$  y una tasa de falsas alarmas  $< 5\%$ , 10 de las 11 variantes cumpliría con dichos requerimientos.

La anchura de esas bandas se podría ampliar, pero, aunque sigue bajando la tasa de falsas alarmas, también se reduce de forma notable la tasa de aciertos. Se puede observar una muestra de ello en la Tabla 6.14.

**Tabla 6.14.** Aplicación de bandas más anchas a los resultados de la Tabla 6.12.

k	dB,k (mT)	SR (%)	WT>0 (%)	FA (%)
0	0	97.51	96.80	1.62
0.1	0.0487	95.73	95.37	0.83
0.2	0.0974	94.66	94.31	0.56
0.3	0.1462	93.59	92.17	0.50
0.4	0.1949	92.88	90.39	0.36
0.5	0.2436	90.75	86.83	0.23
<b>0.6</b>	<b>0.2923</b>	<b>86.48</b>	<b>81.85</b>	<b>0.20</b>
0.7	0.3410	80.07	73.31	0.13
0.8	0.3897	72.24	65.48	0.10
0.9	0.4384	65.48	59.43	0.10
1	0.4872	60.14	54.09	0.07

Para valores grandes de  $k$ , las interrupciones son reconocidas con poco tiempo de antelación, es decir, a medida que las bandas se van alejando de la diagonal, los tiempos de anticipación de la interrupción, se reducen. El tamaño de la banda es proporcional a  $k$  y, por lo tanto, como el ancho de la banda aumenta, habrá menor número de puntos fuera de la banda. Como se puede comprobar en la Tabla 6.14., se reduce el número de falsas alarmas, pero cae de forma notable la tasa de aciertos; por lo que es preciso determinar dónde está el valor límite  $k$  de dichas bandas.

En la Figura 6.24. se muestra el efecto de las bandas dentro de la región disruptiva de la descarga.

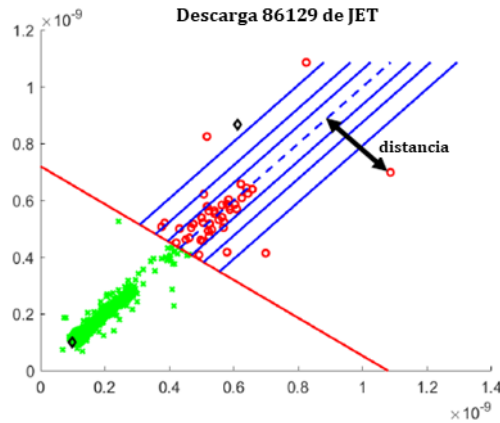


Figura 6.24. Efectos de las bandas en torno a la diagonal para la descarga 86129.

Una muestra tendrá comportamiento disruptivo, si además de encontrarse en la región disruptiva, queda fuera de las bandas que se estén considerando. Como ya se ha mencionado con anterioridad, el ancho de las bandas ha sido un factor empleado para el análisis de las falsas alarmas.

Se analizó el efecto que produce la variación de la anchura de las bandas en torno a la diagonal, considerándose las muestras de la región disruptiva de la descarga. Teniendo en cuenta el tamaño de las bandas mostrado en la Tabla 6.14., el efecto que produce en cuanto a tiempo de anticipación de la disruptción, se presenta en la Figura 6.25.

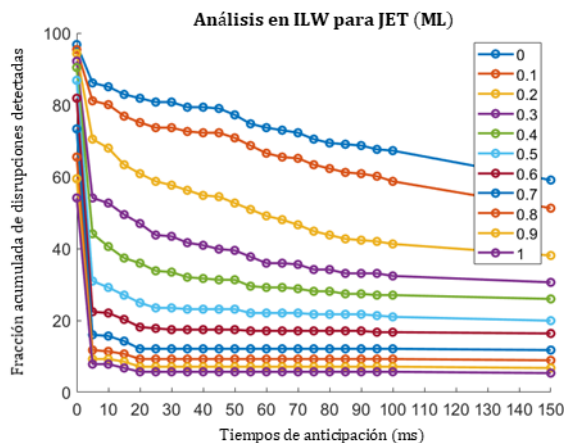


Figura 6.25. Efectos de las bandas con respecto a los tiempos de anticipación.

El tamaño de la banda  $d_{B,0.6} = 0.2923 \text{ mT} = |x(t-0.002) - x(t)|$  parece ser una distancia crítica, desde la cual la disrupción tiene un desarrollo muy rápido; pudiéndose interpretar como la frontera desde la cual la disrupción es inevitable.

Asimismo, se determinan las siguientes reglas que relacionan el tiempo de anticipación de la disrupción con la anchura proporcionada a las bandas, en la región disruptiva de la descarga:

- Si la distancia  $|x(t) - x(t - \tau)| > d_{B,0.6}$ , la disrupción se espera en  $t \leq 5 \text{ ms}$ .
- Si  $d_{B,0.2} < \text{distancia} < d_{B,0.6}$ , la disrupción se espera en  $t \leq 85 \text{ ms}$ .
- Si  $d_{B,0.003} < \text{distancia} < d_{B,0.2}$ , la disrupción se espera en  $t \leq 1 \text{ s}$ .

## 6.5. Predicción del tiempo a la disrupción en cada alarma.

### Antecedentes con la predicción TTD.

La predicción TTD (Time To Disruption) es un problema difícil, sin solución práctica hasta ahora. A pesar de no haber existido muchos intentos para abordar el problema, todos ellos utilizaban para la predicción: una estructura compleja de redes neuronales artificiales, un número limitado de descargas y tipos particulares de disrupciones. La utilización de redes neuronales evita la posibilidad de identificar las razones físicas que causan la disrupción. Mientras que, con la limitación en el número de descargas y la particularización de las mismas, no se pueden generalizar los predictores para una gran cantidad de descargas y de tipos diferentes de disrupciones.

Un primer trabajo relacionado con la predicción TTD [Morabito, 2001], consistió en definir una estructura compleja de redes neuronales. La primera capa de procesamiento se utilizó para descomponer la base de datos original en cuatro subconjuntos. Las salidas de la capa fueron utilizadas para activar cuatro perceptrones multicapa, entrenados exclusivamente en un subconjunto de la base de datos original. La salida del sistema consistía en una capa lineal de un solo nodo, que proporcionaba la estimación de tiempo para la disrupción; así como la activación de una alarma en caso de que procediera. El predictor generó una señal de evolución temporal, activándose la alarma cuando el tiempo hasta la disrupción se encontraba por debajo de cierto umbral. En las pruebas realizadas,

se activó una alarma cuando el TTD se mantuvo en el rango 250 ms – 350 ms para 10 predicciones consecutivas; evitando así las falsas alarmas.

En un segundo trabajo, [Pautasso, 2002] se basó en una red neuronal de dos capas. La primera capa contenía 20 neuronas, y la segunda capa (salida de la red neuronal) proporcionaba el intervalo de tiempo hasta la interrupción. Para filtrar falsas alarmas, se activaba una alarma cuando la salida del predictor era menor o igual a 50 ms durante tres muestras consecutivas. El predictor fue probado con 128 descargas, 28 de las cuales eran disruptivas; obteniéndose tasas de acierto del 79% y 7% de falsas alarmas.

Aplicado a la base de datos de JET, un tercer trabajo de [Cannas, 2004b] también se basó en redes neuronales artificiales. La mejor configuración de red estaba compuesta de nueve entradas, dos capas ocultas con seis y cinco neuronas ocultas respectivamente, y una salida. Para ello, se tuvieron en cuenta sólo señales disponibles en tiempo real. La salida de la red neuronal fue un número real entre 0 y 1; que representaba el riesgo de interrupción que se obtenía cada 20 ms. Se activó una alarma cuando la salida de la red neuronal estaba por encima de cierto umbral. El predictor fue probado con 62 descargas disruptivas y 132 no disruptivas. La salida de la red neuronal fue analizada en el intervalo de tiempo entre 440 ms y 100 ms antes de la interrupción, informándose de una tasa de aciertos del 83,9%.

## **Interpretación física basada en la distancia a la diagonal en el espacio de parámetros para TTD.**

El método de los centroides es un predictor de interrupciones confiable para JET, que se ha desarrollado mediante una sola señal: Mode Lock normalizada a la corriente del plasma. El método comprime en dos puntos individuales (el disruptivo y el no disruptivo) información de descargas ILW para JET. La predicción se basa en el enfoque del centroide más próximo (disruptivo/no disruptivo). El estado del plasma viene determinado por una simple desigualdad lineal en dos variables; donde las variables representan la amplitud de dos muestras consecutivas tomadas cada 2 ms en JET.

Si  $X(t) > 6.1243 * 10^{-10} T * A^{-1}$ , la ecuación (11) se cumple siempre, independientemente de la amplitud  $X(t - 0.002)$ . Además, es importante tener en cuenta que, el método de los centroides es lo suficientemente flexible como para operar en diferentes niveles de riesgo de interrupciones; sin tener que seleccionar umbrales de forma manual.

Con respecto a la distribución de muestras consecutivas en el espacio de parámetros, la propiedad común de los puntos cercanos a la diagonal, es que la diferencia de amplitudes entre muestras consecutivas es realmente pequeña. Debe tenerse en cuenta que la amplitud de la señal ML/Ip está relacionada con el frenado de rotación del plasma. Por lo tanto, una pequeña diferencia entre muestras consecutivas de la señal ML/Ip, puede interpretarse como pequeñas variaciones en la rotación. Estas pequeñas variaciones en la señal ML/Ip no se pueden asociar a comportamientos disruptivos. Sin embargo, dichos comportamientos disruptivos pueden ser identificados cuando los puntos en el espacio de parámetros se hallen suficientemente alejados de la diagonal. Eso significa que hay suficiente diferencia entre las amplitudes de muestras consecutivas. Y, por tanto, cuando existen cambios abruptos en la amplitud de muestras consecutivas, habrá cambios abruptos en la rotación del plasma; estando relacionadas esas variaciones rápidas de la amplitud con comportamientos disruptivos.

Además, teniendo en cuenta la bidimensionalidad del espacio de muestras consecutivas de la señal ML/Ip; la distancia a la diagonal de un punto en ese espacio de parámetros, se halla relacionada con el TTD mediante un modelo Weibull. La estimación del TTD es un logro muy relevante, que ha sido probado con buenos resultados. El aspecto más importante a destacar, es el razonamiento empírico simple para determinar el TTD; en lugar de usar expresiones muy complejas, deducidas de métodos generales como los relacionados con redes neuronales artificiales.

Un análisis fuera de línea de las descargas disruptivas de ILW en JET, y las alarmas activadas según la ecuación (9); y que por efectos prácticos se reescribe como ecuación (11)

$$X(t) > -0.7441 * X(t - 0.002) + 6.1243 * 10^{-10} \quad (11)$$

permiten establecer que, los puntos disruptivos próximos a la diagonal (véase Figura 6.16.), proporcionan mayores tiempos de anticipación a la disrupción, que los puntos que se hallan alejados de dicha diagonal. Tal y como fue comentado en la sección 6.4., el hecho empírico de que cuanto más cerca están los puntos de la diagonal, mayores son los tiempos de anticipación; sugiere la hipótesis de que cuando se active una alarma, el TTD está relacionado con la distancia  $d_D$  a la diagonal. Es decir,  $TTD = f(d_D)$ .

Para ello, se requiere ajustar un primer conjunto de datos. Una vez realizado el ajuste, se necesita un segundo conjunto de datos (diferente al primero), para probar la

hipótesis. Para generar el primer conjunto, fueron consideradas las 113 descargas disruptivas de la Tabla 6.3. El predictor representado por la ecuación (11) se aplicó a cada una de esas 113 descargas, para obtener todos los pares ( $d_D$ , *tiempo de anticipación a la disrupción*); siempre que se trate de una descarga cuyo comportamiento sea disruptivo. Esto significa que, cada descarga puede contribuir posiblemente con varios pares. Es importante tener en cuenta que, el uso de la ecuación (11) para fines de mitigación, sólo necesita el primer reconocimiento de un comportamiento disruptivo para activar la alarma. Sin embargo, la situación para modelar los tiempos de anticipación a la disrupción es diferente; ya que todos los pares posibles pueden contribuir a ajustar los datos a un modelo de distribución específico.

Con respecto al procedimiento de prueba, se genera un segundo conjunto de pares ( $d_D$ , *tiempo de anticipación a la disrupción*). Ese conjunto está compuesto por todos los pares correspondientes a los comportamientos disruptivos que se obtienen con la ecuación (11) para las 277 descargas disruptivas de prueba de la Tabla 6.3.

Entre las diversas alternativas investigadas, la mejor función de densidad de probabilidad para relacionar distancias con TTD, consistió en un modelo de distribución de Weibull [URL, 57] de la forma expresada según la ecuación (12).

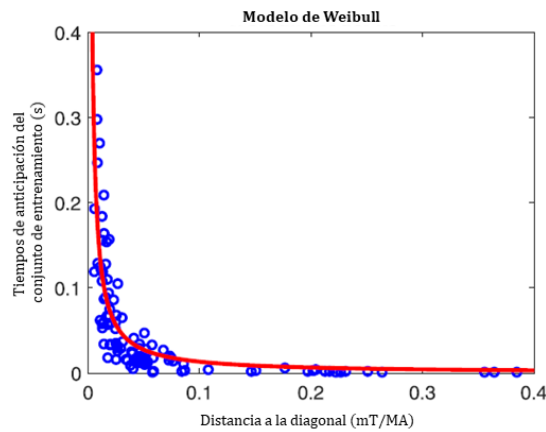
$$TTD(d_D) = \alpha * \beta * d_D^{\beta-1} * \exp(-\alpha * d_D^\beta), \text{ con } d_D > 0 \quad (12)$$

Donde  $TTD(d_D)$  es una estimación del tiempo hasta la disrupción,  $d_D$  es la distancia a la diagonal, siendo  $\alpha$  y  $\beta$  los coeficientes del modelo. Sin embargo, han de tenerse en cuenta algunas consideraciones. La primera de ellas está relacionada con la selección de los pares que componen el conjunto de entrenamiento, para determinar los coeficientes  $\alpha$  y  $\beta$  del modelo. Debido a que el tiempo mínimo en JET para mitigar disrupciones es de 10 ms, el predictor debe ser capaz de hacer estimaciones de TTD mayores de 10 ms. Por lo tanto, sólo se entrenarán parejas cuyo valor para la componente *tiempo de anticipación a la disrupción*  $\geq 10$  ms. La segunda consideración está relacionada con el objetivo de encontrar el mayor tiempo de anticipación, para el cual el modelo de la ecuación (12) proporciona buenos resultados. Para este fin, se han definido 14 intervalos para crear 14 modelos diferentes; cuyos tiempos de anticipación de entrenamiento se encuentran entre los límites temporales de la ecuación (13).

$$[0.01, 0.02 + (k - 1) * 0.01], \text{ con } k = 1, \dots, 14 \quad (13)$$

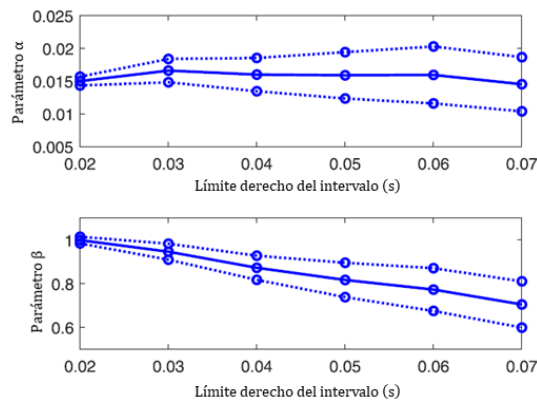
Donde los tiempos están en s. Cabe señalar que la resolución temporal entre intervalos es de 0.01 s. Con esta definición de intervalos, el ancho mínimo es de 0.01 s, y el ancho máximo es de 0.14 s.

La Figura 6.26. muestra como ejemplo, todos los pares con tiempos de anticipación menores de 0.4 s del conjunto de entrenamiento de las 113 descargas disruptivas, y el ajuste a un modelo Weibull.



**Figura 6.26.** Los círculos son los pares (distancia a la diagonal, tiempo de anticipación), que se obtienen al reconocer comportamientos disruptivos con ecuación (11) en el conjunto de datos de 113 descargas disruptivas, que se muestran en la Tabla 6.3. La línea roja continua se ajusta a un modelo de Weibull.

La Figura 6.27. muestra los parámetros  $\alpha$  y  $\beta$  del modelo de la ecuación (12), con límites de confianza del 95% para  $k = 1, \dots, 6$ .



**Figura 6.27.** Variación de los parámetros  $\alpha$  y  $\beta$  del ajuste de Weibull, con límites de confianza del 95% para los casos  $k = 1, \dots, 6$ . El eje horizontal muestra el límite derecho de los intervalos.

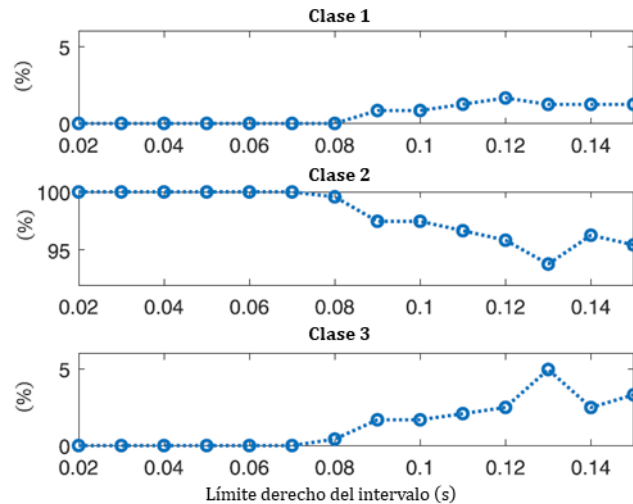
Es importante tener en cuenta que el primer parámetro permanece alrededor de 0.015. Mientras que el segundo disminuye de 1 a 0.7. La ecuación (12) con  $\beta = 1$  se reduce

a un modelo exponencial (que sólo es capaz de explicar el caso para  $k = 1$ ). Sin embargo, los modelos Weibull explican, con las mejores tasas de predicción posibles, los seis casos en los que los tiempos de anticipación y los TTD del conjunto de prueba tienen los mismos valores dentro de intervalos estrechos de tiempo.

Los pares de prueba ( $d_D$ , *tiempo de anticipación a la disrupción*), se generan con la ecuación (11) para todas las descargas de prueba. De este conjunto de pares, se eligen 14 subconjunto; agrupándose según sus tiempos de anticipación a la disrupción. Cada uno de los 14 subconjuntos, incluye los pares dentro de los intervalos establecidos por la ecuación (13). Una vez que se definen los subconjuntos de prueba, la coordenada  $d_D$  de los puntos de prueba que reconoce un comportamiento disruptivo, es utilizada en la ecuación (12) para obtener las estimaciones TTD.

Para evaluar que el modelo se ajusta, se realiza una comparación entre los tiempos de anticipación de la ecuación (11) y los TTD de la ecuación (12). Los TTD son clasificados en tres clases posibles. La primera clase contiene la fracción de descargas de prueba, cuyos TTD están por debajo del límite de los 10 ms. Esto corresponde a descargas en las que el ajuste del modelo produce TTD más pequeños que los tiempos de anticipación reales. La segunda clase incluye la fracción de descargas cuyos TTD se encuentran dentro del intervalo de tiempo de anticipación utilizado para entrenamiento / prueba. Este es el caso ideal. La tercera clase comprende la fracción de descargas cuyos TTD exceden el límite correcto del intervalo de tiempo de anticipación. Esta situación ocurre cuando el modelo de ajuste proporciona mayores TTD que los tiempos de anticipación reales. Por lo tanto, de acuerdo a este criterio de clasificación, un buen modelo produciría tasas muy bajas en las clases 1 y 3; y una alta tasa en la clase 2. La Figura 6.28. muestra las tres tasas de prueba correspondientes a clases 1, 2 y 3 de los 14 modelos.





**Figura 6.28.** Tasas de prueba en las clases respectivas, para cada uno de los 14 modelos. El eje x muestra el límite derecho de los respectivos intervalos de tiempo.

Se puede observar que los modelos de  $k = 1$  a  $k = 6$  (véase ecuación (13)), proporcionan los mejores resultados posibles. Es decir, clase 1: 0%, clase 2: 100%, clase 3: 0%. Esto significa que cualquiera de estos 6 modelos determina buenos predictores de TTD, porque los tiempos de anticipación de prueba y los TTD de prueba muestran los mismos valores en reducidos intervalos de tiempo.

En particular, y de acuerdo con la segunda consideración anterior sobre el mayor tiempo de anticipación que produce mejores resultados; la mejor opción corresponde al modelo con  $k = 6$ . En este caso, los parámetros del modelo de la ecuación (12), con límites de confianza del 95% son  $\alpha = 0.0145 \pm 0.0041$  y  $\beta = 0.7038 \pm 0.1060$ .

Al considerar tres clases, para comparar las predicciones TTD con los tiempos de anticipación reales en las alarmas producidos por el método de los centroides; los mejores resultados posibles (clase 1: 0%, clase 2: 100%, clase 3: 0%) se obtienen en intervalos de tiempo próximos a las interrupciones. El intervalo más grande es [0.01, 0.07] s. Sin embargo, los intervalos de tiempo más largos (hasta 0.15 s), también proporcionan buenos resultados (véase Figura 6.28.). Cabe señalar que la fracción de puntos cuyo TTD predicho y el tiempo de anticipación pertenecen al mismo intervalo (clase 2), está por encima del 95% (excepto en un caso). Del mismo modo, la fracción de pares cuyo TTD predicho es menor que el tiempo de anticipación real (clase 1), está por debajo del 2% en todos los casos. Finalmente, la fracción de pares en la que se predice TTD como mayor que el tiempo de anticipación (clase 3), es siempre inferior al 5%.

Con respecto a los desarrollos metodológicos para clasificación de los tipos de interrupciones, es un tema para futuras investigaciones el abordar el potencial de las

técnicas descritas (método de los centroides y estimación TTD). De hecho, este es un aspecto muy importante para la optimización de estrategias de elusión y mitigación; pudiendo requerir también métricas más sofisticadas.

## 6.6. Análisis preliminar de datos para JT-60U.

Los datos que fueron proporcionados para este dispositivo no estaban clasificados, por lo que hubo que realizar un filtrado previo, para poder operar con ellos. Inicialmente fueron analizadas visualmente las gráficas de las distintas señales proporcionadas, observando la evolución de sus muestras durante la descarga.

En la Figura 6.29. y Figura 6.30. se muestran dos ejemplos de las gráficas facilitadas.

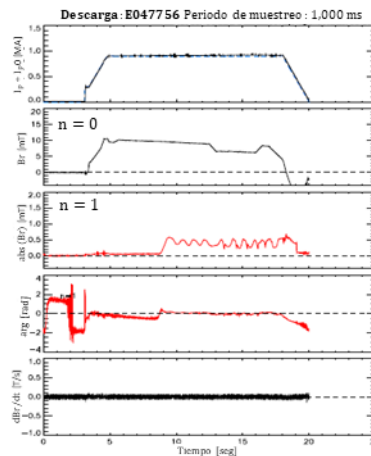


Figura 6.29. Representación de señales para una descarga en JT-60U.

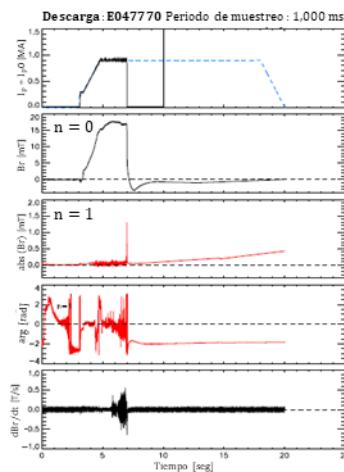
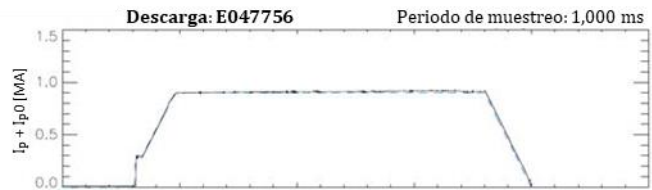


Figura 6.30. Representación de señales para una descarga con distinto comportamiento en JT-60U.

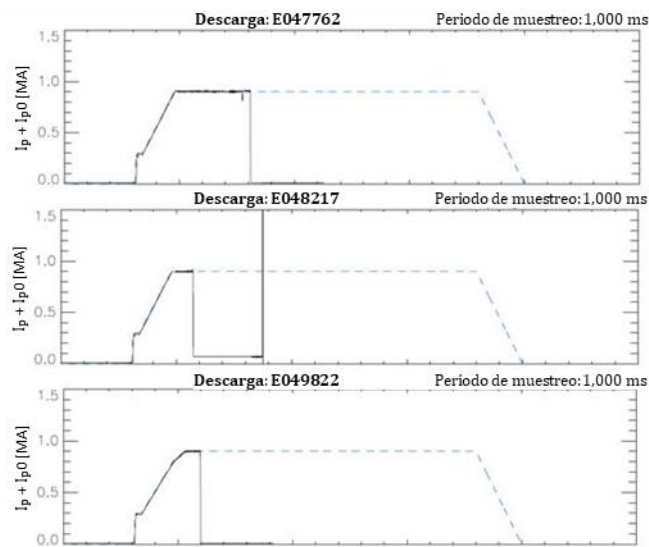
Si la gráfica de la corriente del plasma es similar a la de la Figura 6.31., la descarga se considera no disruptiva.



**Figura 6.31.** Representación de un comportamiento típico no disruptivo en JT-60U.

La intensidad de corriente del plasma asciende hasta llegar a un umbral, se mantiene en el mismo por un intervalo de tiempo, y finalmente desciende progresivamente hasta llegar a anularse.

Si la gráfica de la corriente del plasma fuera similar a uno de los siguientes tipos de la Figura 6.32., se considerarían como descargas disruptivas.



**Figura 6.32.** Representación de varios comportamientos disruptivos en JT-60U.

De esa manera, la desviación entre la corriente real (la proporcionada por las muestras que componen las descargas) y la corriente de referencia (indicada por la línea discontinua que tiene una fase de ascenso, una fase con valor constante y por último una fase de caída hasta hacerse nula); proporciona información relevante para este análisis.

Para completar el análisis, se hizo uso de la tasa de apagado de la corriente, teniendo en cuenta lo siguiente:

- Se obtiene un punto inicial, a partir del cual iría disminuyendo la corriente.

- Se obtiene el punto final, aproximadamente cuando la corriente es de 50 kA.

$$\text{Tasa de apagado de la corriente} = \frac{(y_2 - y_1)}{(x_2 - x_1)}$$

Donde  $(x_1, y_1)$  representa el valor de la corriente  $y_1$ , en el instante inicial  $x_1$ , y  $(x_2, y_2)$  representa el valor de la corriente  $y_2$ , en el instante  $x_2$ ; en el que la corriente desciende aproximadamente a 50 kA.

Con tasas de apagado  $\geq 3,5$  los comportamientos de las descargas resultaban ser disruptivos. Y con tasas de apagado inferiores, se tuvo que recurrir para algunas descargas dudosas, al criterio visual de sus comportamientos gráficos.

Con un número reducido de descargas todavía es posible una evaluación visual de sus comportamientos. No ocurriría lo mismo si se tratara de miles de descargas.

Una vez determinado el comportamiento de las 154 descargas proporcionadas (76 disruptivas y 78 no disruptivas), se destinaron el 40% de cada tipo para el cálculo de los centroides (fase de entrenamiento), y el 60% restante para la fase de evaluación.

Para el dispositivo JT-60U las muestras que forman el espacio de parámetros no se disponen en torno a la diagonal, como ocurría en JET. Suelen concentrarse próximas al punto (0, 0) y guardando una disposición diferente.

## 6.7. Resultados para JT-60U.

La forma mediante la cual se determinan los centroides para este dispositivo, ya fue presentada en el capítulo 5.

Se analizó un conjunto de datos con 154 descargas (76 descargas disruptivas no intencionadas y 78 descargas no disruptivas) en el rango 47756–49826. El predictor lineal ha sido probado bajo dos condiciones de entrenamiento diferentes. El primero utiliza aproximadamente el 40% de descargas disruptivas y no disruptivas para entrenamiento y el resto para prueba. La segunda condición de entrenamiento utiliza un entrenamiento adaptativo. En todos los casos, el predictor activa una alarma cuando dos predicciones consecutivas reconocen un comportamiento disruptivo. Esto se ha establecido para reducir el número de falsas alarmas, sin afectar la tasa de aciertos. Se han considerado cinco parámetros para calificar las pruebas:

**Global Success Rate (GSR):** que representa el número de interrupciones identificadas en el conjunto de prueba.

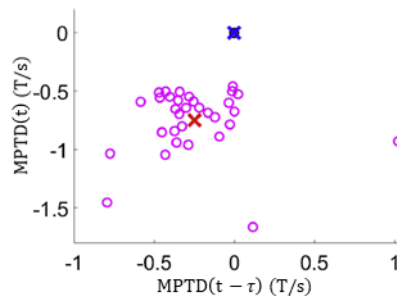
**Success Rate with Positive warning time (SRP):** número de interrupciones identificadas antes de producirse la interrupción, teniendo en cuenta el número total de interrupciones del conjunto de prueba.

**Missed Alarm rate (MA):** número de descargas disruptivas no reconocidas como interrupciones, sobre el número total de descargas disruptivas.

**False Alarm rate (FA):** número de descargas no disruptivas que activaron una alarma, sobre el número total de descargas no disruptivas.

**Average Warning Time (AVGWT):** tiempo de anticipación promedio (en ms) de las alarmas, antes de producirse la interrupción.

En primer lugar, se trata un enfoque no adaptativo, donde los conjuntos de datos de entrenamiento y prueba han sido elegidos al azar, de entre todo el conjunto de descargas disponibles. Para el entrenamiento (cálculo de los centroides), han sido seleccionadas 32 descargas disruptivas y 30 no disruptivas; mientras que para efectuar las pruebas han sido seleccionadas 44 descargas disruptivas y 48 no disruptivas. La Figura 6.33. muestra los centroides individuales de las descargas de entrenamiento representados por círculos, junto con los centroides globales representados por cruces.



**Figura 6.33.** Contribución de las descargas a los centroides en la fase de entrenamiento.

Las cruces azul y roja son, respectivamente, los centroides globales de las descargas no disruptivas y disruptivas. Es importante tener en cuenta que los centroides individuales no disruptivos se encuentran prácticamente en el mismo punto.

La fila 1 de la Tabla 6.15. muestra los resultados obtenidos con el conjunto de datos correspondiente a las descargas anteriores. Sin embargo, para eliminar cualquier posible sesgo, fueron seleccionados 10 conjuntos aleatorios adicionales, para entrenamiento/prueba. Los resultados de las pruebas aleatorias de descargas aparecen de

la fila 2 a la 11, siendo los resultados bastante similares en todos los casos; y representando el resultado global del predictor lineal en la última fila de la Tabla 6.15.

**Tabla 6.15.** Predicciones de JT-60U teniendo en cuenta 11 juegos de prueba aleatorios.

#test	GSR	SRP	MA	FA	AVGWT
1	100	97.7	0	6.2	20
2	97.7	97.7	2.3	4.2	17
3	100	97.7	0	4.2	13
4	95.5	90.9	4.5	4.2	14
5	97.7	95.5	2.3	2.1	14
6	97.7	90.9	2.3	8.3	19
7	100	97.7	0	6.2	18
8	97.7	95.5	2.3	10.4	20
9	100	97.7	0	8.3	18
10	95.5	90.9	4.5	0	17
11	100	97.7	0	4.2	20
<b>Valor medio</b>	<b>98.3</b>	<b>95.4</b>	<b>1.7</b>	<b>5.3</b>	<b>17</b>

Los resultados mostrados en la Tabla 6.15. dependen de los centroides obtenidos a partir de la Figura 6.33. y de la ecuación:

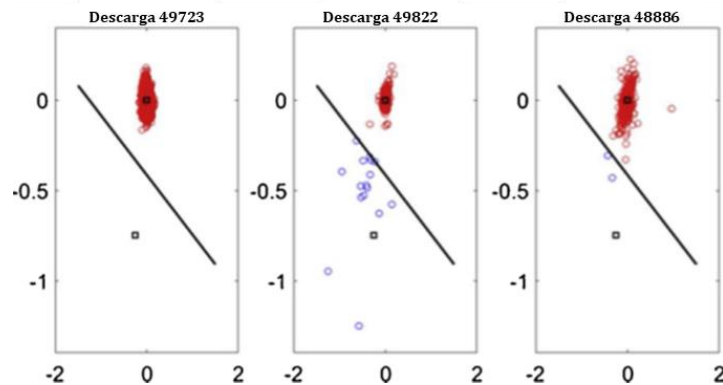
$$x_2 > -\frac{d_1 - c_1}{d_2 - c_2}x_1 + \frac{d_1^2 + d_2^2 - c_1^2 - c_2^2}{2(d_2 - c_2)}$$

Donde  $x_2 < -0.3300 \cdot x_1 - 0.4152$ .

Siendo  $x_2 = MPTD(t)$  y  $x_1 = MPTD(t - \tau)$ .

La señal MPTD (Magnetic Perturbation Time Derivative) está muestreada con un periodo  $\tau = 1$  ms.

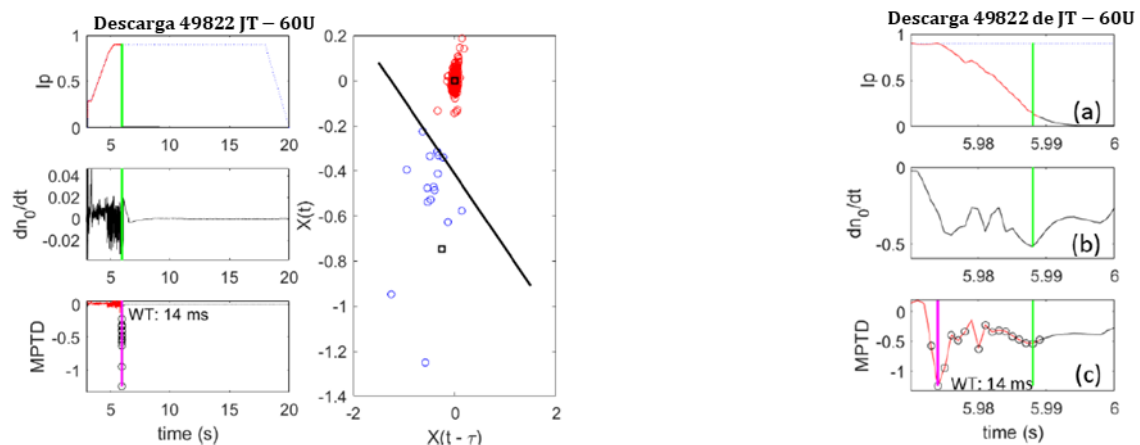
La Figura 6.34. muestra el espacio de parámetros de tres situaciones diferentes: una descarga no disruptiva, un acierto que reconoce un comportamiento disruptivo y una falsa alarma.



**Figura 6.34.** Espacio de parámetros para distintos comportamientos en el JT-60U.

En la Figura 6.34. de izquierda a derecha se muestran los espacios de parámetros correspondientes a una descarga no disruptiva, descarga disruptiva y falsa alarma. Los cuadrados negros representan los centroides globales de cada una de las dos regiones. Es importante destacar que la mayoría de las falsas alarmas corresponden a interrupciones menores. En estos casos el plasma puede recuperarse, pero el predictor activa una alarma, porque al menos dos puntos consecutivos aparecen en la zona disruptiva del espacio de parámetros.

En la Figura 6.35. se representa el comportamiento de una descarga disruptiva en el dispositivo JT-60U; donde, mediante el uso de tres señales, se observan las variaciones entre una muestra y la siguiente. La señal  $I_p$ , al igual que en el caso de JET, representa la intensidad de corriente del plasma. La señal  $dn_0/dt$  es una señal propia del JT-60U, que representa la derivada temporal de la componente  $n = 0$  de la perturbación magnética. Y la señal  $MPTD$  está relacionada con la perturbación magnética.



**Figura 6.35.** Evolución de las señales para una descarga disruptiva en el JT-60U.

En la parte izquierda de la Figura 6.35., se muestra en primer lugar la evolución de la señal  $I_p$ . Se observa a la derecha de la línea vertical en verde, una trayectoria marcada con puntos discontinuos, que debería ser la que siguiera la descarga en caso de tratarse de una descarga no disruptiva. Pero como se puede ver en la imagen ampliada para la señal  $I_p$  mediante la Figura 6.35.(a) de la parte derecha, dicha señal experimenta una reducción brusca hasta que termina anulándose en torno al instante 5,990 s.

La evolución de la señal  $dn_0/dt$  en torno al instante de la interrupción también termina anulándose. Este hecho se observa mejor en la parte izquierda que en la Figura 6.35.(b) de la parte derecha, donde aparece ampliada.

Finalmente, la señal  $MPTD$ , que aparece ampliada en la Figura 6.35.(c) de la parte derecha proporciona la información que se utilizará en el espacio de parámetros; permitiendo predecir la interrupción con una antelación de 14 ms.

En la parte central de la Figura 6.35. se muestra el espacio de parámetros, donde la recta en negro representa la frontera entre las regiones disruptiva y no disruptiva, los cuadrados en negro representan la posición de los centroides, los círculos en rojo constituyen el espacio no disruptivo, y los círculos en azul forman el espacio disruptivo de la descarga del ejemplo.

En la Figura 6.36. se muestra la evolución de una descarga no disruptiva, teniendo en cuenta las mismas señales que en el caso anterior.

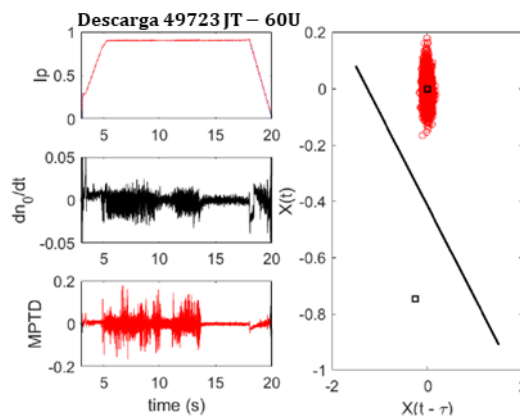


Figura 6.36. Evolución de las señales para una descarga no disruptiva en el JT-60U.

En el caso de la Figura 6.36., la intensidad de corriente del plasma sigue una línea ascendente, se mantiene una vez alcanzado cierto umbral, y termina anulándose mediante una caída progresiva (no abrupta); respondiendo al comportamiento típico de la señal  $I_p$  con descargas no disruptivas. En la representación del espacio de parámetros aparecen todas las muestras que lo componen dentro de la región no disruptiva, no sobrepasando ninguna de las muestras la frontera que divide ambas regiones.

Las falsas alarmas de los tres casos correspondientes a la fila 1 de la Tabla 6.15., vienen representados en la Figura 6.37.



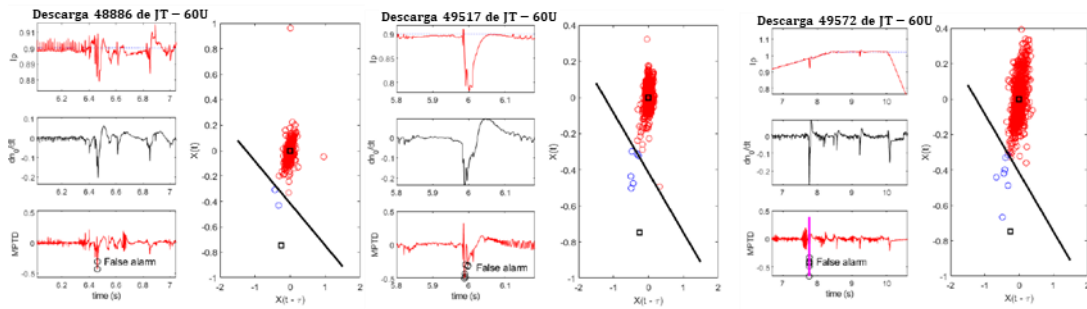


Figura 6.37. Evolución de las señales para descargas con presencia de falsas alarmas en el JT-60U.

La señal  $I_p$  no llega a anularse de forma repentina. Sin embargo, en ciertos instantes se producen cambios abruptos en dicha señal, pero sin llegar a anularla. Esas anomalías detectadas en la señal  $I_p$ , junto con el comportamiento de las señales  $dn_0/dt$  y  $MPTD$ , determinan la pertenencia de dichas descargas a este tipo de falsas alarmas. En la Figura 6.37. se puede observar además que, el espacio de parámetros está formado en casi su totalidad por muestras no disruptivas compactadas. Sin embargo, algunas de dichas muestras consiguen traspasar la frontera, adoptando comportamientos disruptivos.

La Figura 6.38. representa la fracción acumulativa de las disruptivas detectadas para el caso de la primera fila de la Tabla 6.15. Se reconoce el 100% de las disruptivas, pero una de ellas se identifica 2 ms después de la disruptiva (detección tardía).

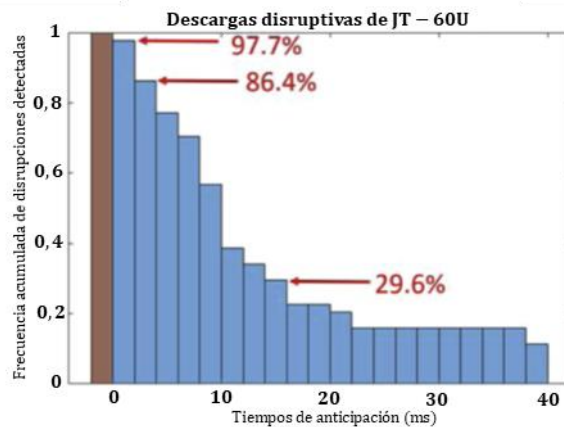


Figura 6.38. Tiempos de anticipación para la fila 1 de la Tabla 6.15. en el JT-60U.

Las barras azules representan tiempos de anticipación positivos y la marrón representa una detección tardía. El ancho de la barra es de 2 ms, obteniéndose un reconocimiento del 29.6%, 86.4% y 97.7% de las disruptivas con un tiempo de

anticipación superior a 14 ms, 2 ms y 0 respectivamente. La fracción acumulativa aumenta principalmente para tiempos de anticipación inferiores a 14 ms, lo que hace que este predictor sea adecuado para fines de mitigación.

También se trató un enfoque adaptativo desde cero. Hasta ahora, no hay teoría de los principios ni un solo modelo de física para explicar la inestabilidad disruptiva [Boozer, 2012]. Por lo tanto, cada Tokamak desarrolla sus propios predictores a partir de su propia base de datos de descargas. En el pasado se intentó crear predictores cruzados, por ejemplo, entre JET y AUG [Windsor, 2005], pero se necesita más trabajo de investigación. Por lo tanto, debido a la falta de bases de datos en nuevos dispositivos como JT-60SA o ITER, los predictores deben crearse desde cero y deben aprender de una forma adaptativa. Ejemplos de predictores de disrupción desde cero pueden ser encontrados en las referencias [Dormido-Canto, 2013], [Vega, 2014], [Murari, 2018], todos ellos aplicados a datos de JET. El objetivo de la predicción desde cero es procesar las descargas en orden cronológico, como sucede en la operación real, y para crear el primer predictor después de tener al menos 1 descarga disruptiva y 1 no disruptiva. El criterio utilizado para reentrenar el predictor, establece que debe llevarse a cabo después de cada alarma perdida. La aplicación de centroides, para predecir desde cero al conjunto de datos de 154 descargas de JT-60U, proporciona los resultados de la Tabla 6.16.

**Tabla 6.16.** Resultados con predicción adaptativa desde cero para JT-60U.

GSR	SRP	MA	FA	AVGWT
97.3	94.7	2.7	19.7	59

El primer predictor utiliza 1 descarga disruptiva y 2 descargas no disruptivas. Solo se pierden 2 disrupciones y, por lo tanto, solo ha sido necesario 2 reentrenamientos. La desigualdad para reconocer comportamientos disruptivos es:  $x_2 < -0.5580 * x_1 - 0.3664$ . Siendo  $x_2 = MPTD(t)$  y  $x_1 = MPTD(t - \tau)$ ; y estando muestreada la señal MPTD a 1 ms.

El enfoque adaptativo parece ser muy sensible a los comportamientos disruptivos, y esta es la razón de la alta tasa de falsas alarmas y alto valor del parámetro AVGWT, en comparación con los valores de Tabla 6.15. Posiblemente, este hecho es una consecuencia del uso de descargas en un amplio rango de más de 2000 descargas (47756 – 49826), habiéndose tomado tan solo 154 descargas de entre esas más de 2000. Para este tipo de

enfoque adaptativo hubiera sido deseable aprender con descargas a corta distancia, es decir, que no hubiera saltos tan considerables como hay desde un número de descarga hasta el siguiente número de descarga del que se tienen datos.

## 6.8. Análisis de datos para DIII-D.

De la misma forma que ocurría con los datos del dispositivo JT-60U, en el caso de DIII-D también ha sido preciso un análisis previo de los datos proporcionados. En la Figura 6.39. se muestra un ejemplo, como resultado de los datos proporcionados.

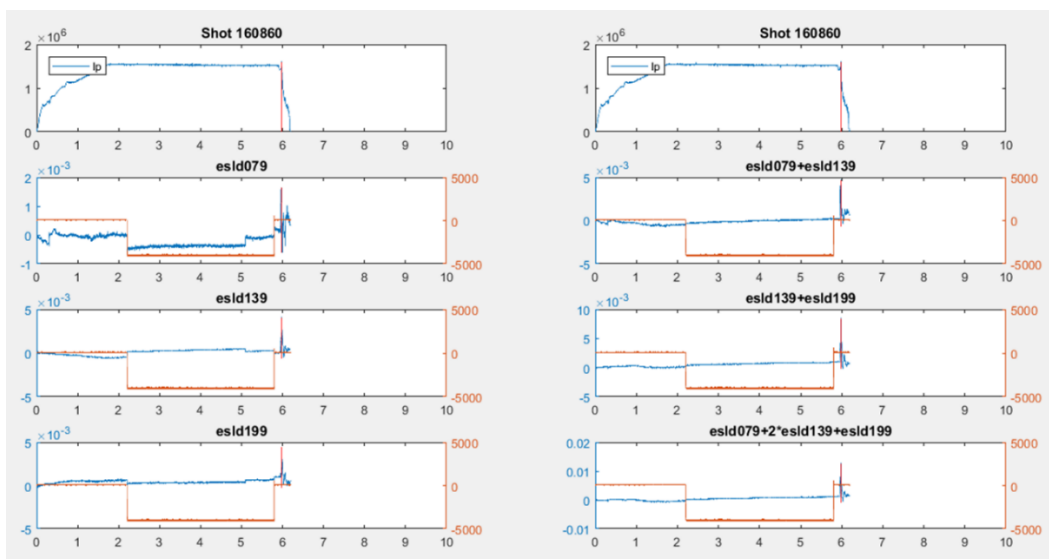


Figura 6.39. Señales proporcionadas para el dispositivo DIII-D.

Para cada descarga fueron proporcionadas diversas señales, entre las cuales se encuentra la intensidad de corriente del plasma  $I_p$ , común también a los otros dos dispositivos analizados con anterioridad. Las señales esld079, esld139 y esld199 son propias del dispositivo DIII-D, siendo señales sin normalizar. Las tres últimas gráficas de la parte derecha de la Figura 6.39., son el resultado de la combinación de esld079, esld139 y esld199; proporcionando datos normalizados acerca del dispositivo.

En una primera inspección visual, se determinaron 299 descargas disruptivas y 697 no disruptivas; tomando como referencia principal la señal esld199. Sin embargo, para determinadas descargas no estaba muy claro cuál era su comportamiento real, por lo que hubo que realizar acciones de filtrado debido al “ruido” que aparecía en alguna de las descargas. Véase el ejemplo de la Figura 6.40, donde desde el instante temporal  $t = 2,208$

s hasta  $t = 5,094$  s aparece una especie de “escalón”, que posteriormente recupera los valores de la señal esld199 precedentes al instante  $t = 2,208$  s.

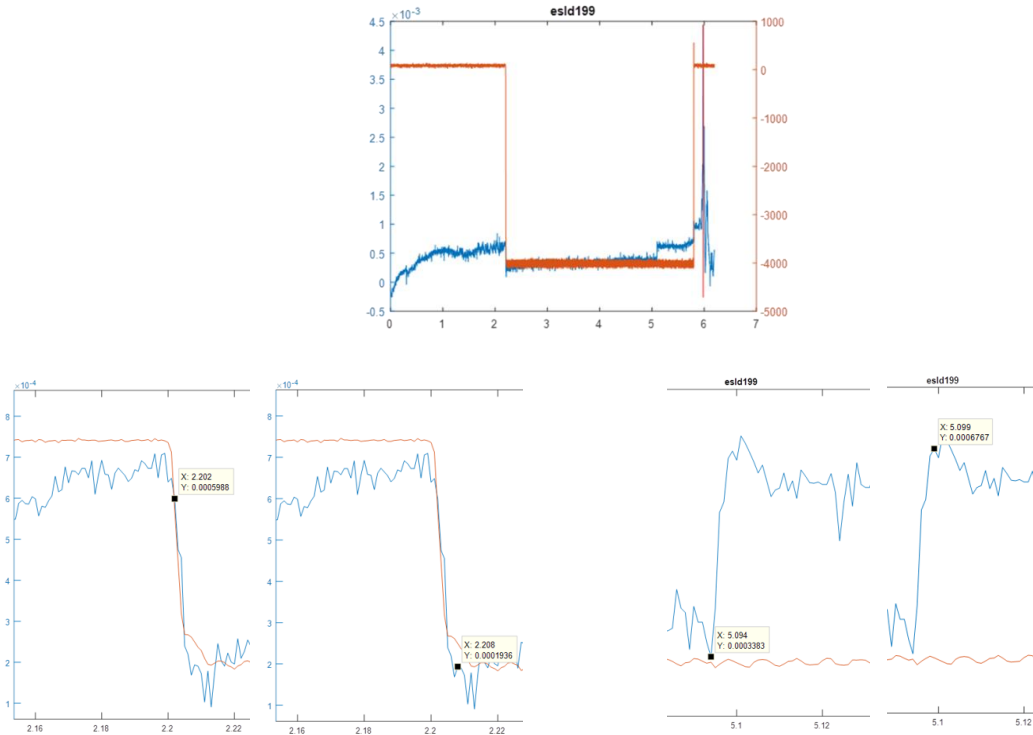


Figura 6.40. Detección con presencia de “ruido” en dispositivo DIII-D.

Además del problema del “ruido”, se observó la doble polaridad en los valores de las señales proporcionadas. Es decir, las muestras tomaban tanto valores positivos como negativos.

Tabla 6.17. Presentación de la doble polaridad en datos DIII-D.

1	2	3	4	5	6	7	8	9	10	11
5.8090	1544720	0.0011	9.5013e-04	-9.7811e-05	39.6729	37.3840	48.0652	40.4358	40.4358	32.0435
5.8100	1.5438e+06	0.0011	9.5721e-04	-1.0180e-04	39.6729	41.9617	41.1987	38.1470	39.6729	57.9834
5.8110	1.5595e+06	0.0011	9.5620e-04	-7.1861e-05	38.9099	41.1987	41.9617	38.9099	40.4358	37.3840
5.8120	1548270	0.0010	9.7239e-04	-3.1938e-05	38.1470	39.6729	41.9617	38.1470	38.1470	36.6211
5.8130	1.5493e+06	9.9023e-04	9.5923e-04	-4.4913e-05	38.9099	43.4875	42.7246	38.1470	41.1987	50.3540
5.8140	1.5489e+06	0.0010	9.7644e-04	-1.9962e-05	37.3840	39.6729	43.4875	40.4358	42.7246	51.8799
5.8150	1547030	9.9817e-04	9.7441e-04	-5.9885e-06	35.8582	38.1470	37.3840	38.9099	41.1987	35.8582
5.8160	1.5498e+06	9.9619e-04	9.6733e-04	5.9884e-06	39.6729	40.4358	45.0134	38.9099	41.9617	42.7246
5.8170	1555350	9.9321e-04	9.6429e-04	1.4971e-05	38.9099	41.9617	42.7246	39.6729	41.1987	41.1987
5.8180	1.5515e+06	9.8526e-04	9.7947e-04	1.8963e-05	38.1470	38.9099	42.7246	38.9099	38.9099	32.0435
5.8190	1.5432e+06	9.4652e-04	9.8453e-04	1.0979e-05	40.4358	43.4875	43.4875	39.6729	41.9617	38.1470
5.8200	1559820	9.9023e-04	9.9263e-04	4.0921e-05	35.0952	41.1987	41.1987	37.3840	41.1987	46.5393
5.8210	1.5503e+06	9.5248e-04	0.0010	4.1919e-05	42.7246	41.1987	48.8281	38.9099	41.9617	2.7105e-09
5.8220	1.5587e+06	9.8228e-04	0.0011	3.3935e-05	34.3323	40.4358	37.3840	38.1470	41.1987	61.0352
5.8230	1551040	0.0010	0.0011	-1.7965e-05	36.6211	41.9617	46.5393	39.6729	40.4358	41.1987
5.8240	1.7637e+06	0.0019	0.0027	-5.2898e-04	36.6211	40.4358	41.9617	36.6211	40.4358	39.6729
5.8250	1223330	0.0019	0.0035	-0.0011	38.1470	41.1987	45.0134	38.1470	40.4358	36.6211
5.8260	9.0332e+05	0.0014	0.0031	-0.0021	38.9099	44.2505	41.9617	37.3840	39.6729	35.8582
5.8270	4.9164e+05	0.0012	0.0026	-0.0023	35.8582	38.9099	41.9617	38.9099	41.1987	38.9099
5.8280	294123	9.9420e-04	0.0021	-0.0021	40.4358	42.7246	42.7246	38.1470	41.1987	37.3840

En la Tabla 6.17., la columna 1 representa el instante temporal en el que se produce la muestra, la columna 2 representa la  $I_p$ , y la columna 5 muestra los valores para la señal esld199. La señal de referencia esld199 en este caso, sería equivalente a la señal ML de JET y a la señal MPTD de JT-60U. Obsérvese que en el instante  $t = 5.8260$  s es el instante en el que por última vez la muestra tiene una  $I_p > 750$  kA. Y si se considera el intervalo temporal desde  $t = 5.8090$  s hasta  $t = 5.8260$  s, en la columna 5 de la Tabla 6.17., se observan tanto valores positivos como negativos.

Debido a la doble polaridad de la señal esld199, se tomó como nueva referencia el valor absoluto de la primera derivada de la señal esld199 ( $abs(diff(esld199))$ ).

Se realizaron pruebas con las condiciones establecidas hasta el momento, pero trabajar en un espacio de parámetros en dos dimensiones, tal y como se había hecho para JET y JT-60U, no permitía obtener un predictor eficiente. Los resultados en cuanto a tasa de aciertos eran aceptables si se restringía el número de descargas disruptivas, pero la tasa de falsas alarmas resultaba demasiado elevada; aun teniendo en cuenta las bandas ligadas a las barras de error de los centroides.

El siguiente escenario consistió en contemplar únicamente los valores de la señal  $abs(diff(esld199))$  no en el plano bidimensional, sino en el espacio unidimensional; como si se tratara de una recta a lo largo del eje X, con valor Y = 0. Para ello, se consideró el planteamiento que describe la Figura 6.41.

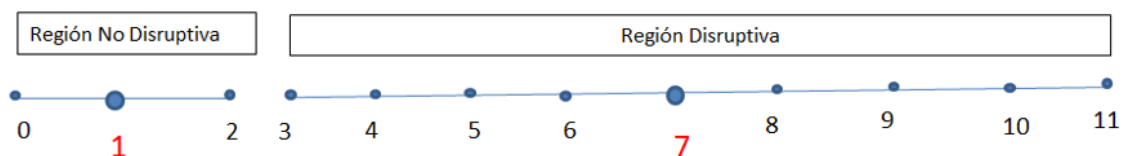


Figura 6.41. Espacio de representación para DIII-D.

Con el nuevo escenario planteado:

- Para la región no disruptiva: Se supone que la dispersión de las descargas que contribuyen al centroide no disruptivo es de 2 unidades (entre  $X = 0$  y  $X = 2$ ). Y que dicho centroide no disruptivo ocupa la posición  $X = 1$ .
- Para la región disruptiva: Se supone que la dispersión de las descargas que contribuyen al centroide disruptivo es de 8 unidades (entre  $X = 3$  y  $X = 11$ ), estando situado dicho centroide disruptivo en la posición  $X = 7$ .

Hasta ahora se ha tenido en cuenta como punto frontera entre las regiones  $(x_{centroide\ no\ disruptivo} + x_{centroide\ disruptivo}) / 2 = (1 + 7) / 2 = 4$ , que representa el punto medio entre los dos centroides. Esto significa que la frontera se halla claramente dentro de la región disruptiva. ¿Qué significado tiene esto? Que en la fase de test y con descargas disruptivas, para las alarmas detectadas con  $X < 4$ ; dicha descarga estaría adoptando un comportamiento no disruptivo. De esa manera, no se estaría tomando en cuenta como acertada, por hallarse en la región no disruptiva.

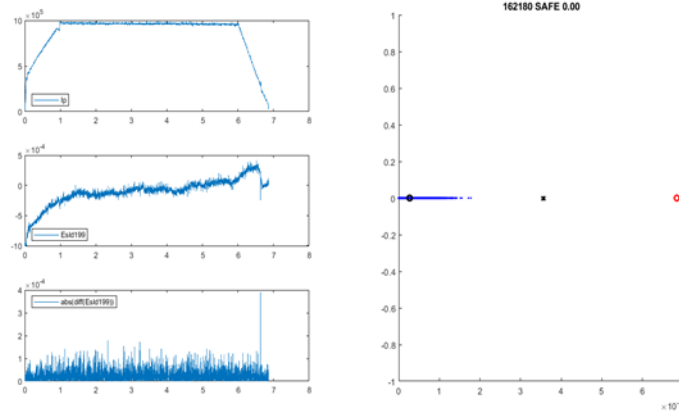
Si se considera que la razón de dispersión entre las regiones no disruptiva y disruptiva es de 1:4, se podría tener en cuenta el punto frontera entre ambas regiones como  $(x_{centroide\ no\ disruptivo} + x_{centroide\ disruptivo}) / Razón\ de\ dispersión = (1 + 7) / 4 = 2$ . De esta manera, al establecer la frontera en  $X = 2$ , en la fase de test de cualquier alarma detectada con  $X > 2$ , sería considerada como acierto si se tratara realmente de una descarga disruptiva.

Fueron implementadas tanto la propuesta anterior, como otra basada en entrenar con las descargas disruptivas que en la fase de evaluación (test) se habían detectado antes del instante de la disrupción (el instante teórico); pero no se consiguió mayor eficiencia con el predictor.

Después de los procesos previos de filtrado de datos procedentes del dispositivo DIII-D, el número de descargas no disruptivas que inicialmente se estimó en 697; se redujo a 365. Y las 299 descargas disruptivas estimadas inicialmente, después de su filtrado fueron reducidas a 156.

## 6.9. Resultados para DIII-D.

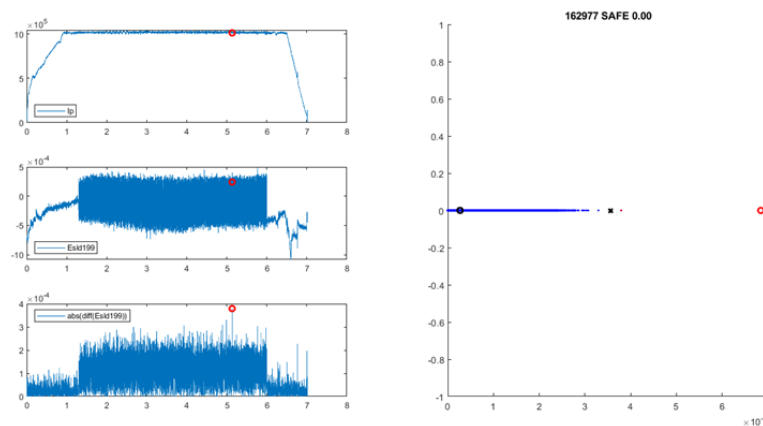
Existen distintos tipos de situaciones que pueden darse, y que se muestran a continuación. En la Figura 6.42. se observa un comportamiento típico no disruptivo.



**Figura 6.42.** Descarga no disruptiva en DIII-D, con todas sus muestras situadas en la región no disruptiva.

En la parte izquierda de la Figura 6.42. y de arriba abajo, aparecen representadas: la intensidad de corriente del plasma, la amplitud (esld199), y el valor absoluto de la derivada de la amplitud. La parte derecha de la Figura 6.42. representa las posiciones de ambos centroides, la frontera entre las regiones no disruptiva y disruptiva, y las posiciones que ocupan las muestras de la descarga que se está evaluando. El círculo negro representa la posición del centroide no disruptivo, el círculo en rojo representa la posición del centroide disruptivo, el símbolo “x” en negro representa la frontera entre ambas regiones, y los puntos en azul representan las muestras que constituyen el espacio de parámetros para la descarga que está siendo evaluada. Claramente se aprecia que ninguna de las muestras traspasa la frontera hacia la región disruptiva, por lo que eso unido al comportamiento que ofrece la señal  $I_p$  para esa descarga, valida el hecho de que su comportamiento sea no disruptivo.

En la Figura 6.43. se muestra la detección de una falsa alarma.



**Figura 6.43.** Descarga con la presencia de una falsa alarma en DIII-D.

La descarga ofrece un comportamiento teóricamente no disruptivo, y se presenta una falsa alarma; cuyo comportamiento si es disruptivo. La presencia de esa falsa alarma se puede observar como un ‘punto en rojo’, situado a la derecha de la frontera en la parte derecha de la Figura 6.43. Obsérvese además en las gráficas de la izquierda que, a pesar de esa alarma, la señal consigue recuperarse y terminar la descarga con normalidad.

En la Figura 6.44. se muestra un ejemplo de descarga disruptiva.

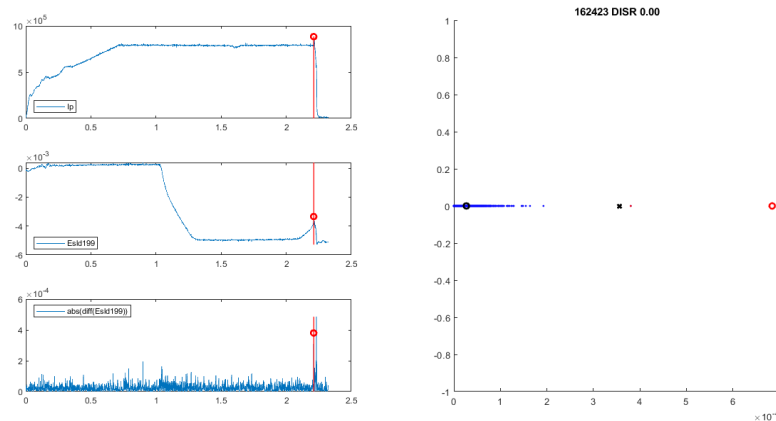


Figura 6.44. Descarga que presenta comportamiento disruptivo en DIII-D.

En la parte izquierda de la Figura 6.44., se observa la caída repentina de la intensidad de corriente del plasma, terminando la descarga de forma abrupta. Ampliando esas gráficas según Figura 6.45.

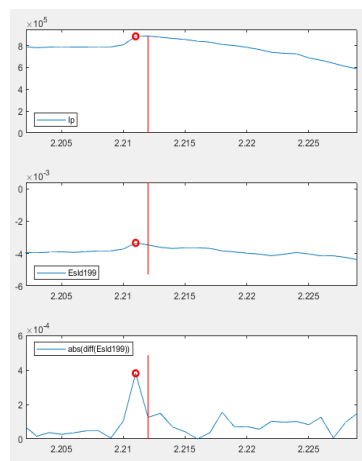
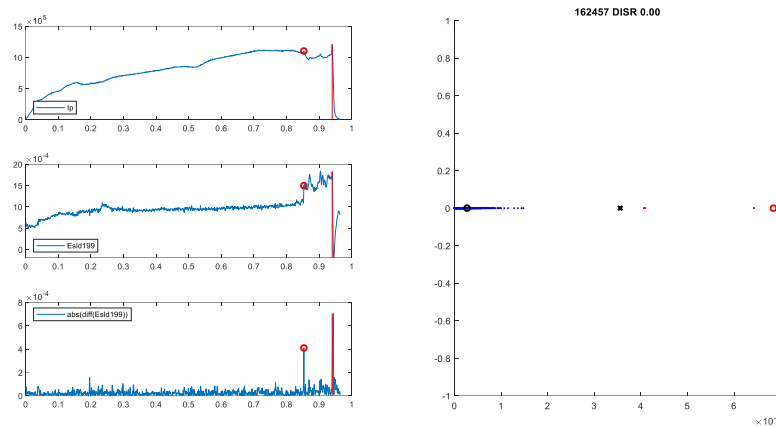


Figura 6.45. Ampliación de las señales que intervienen para la descarga 162423 de DIII-D.

Se observa que, la línea vertical en rojo representa el instante en el que tiene lugar la disruptión, y el círculo en rojo representa el instante en el que es detectada la alarma (por encontrar una muestra disruptiva).

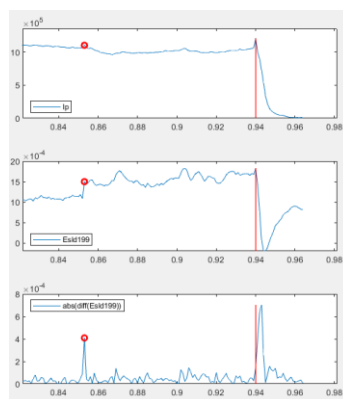


La Figura 6.46. muestra la presencia de una alarma prematura, donde la alarma es detectada antes del instante de la disrupción; por lo que se podría tratar de una alarma válida, siempre que la anticipación a la disrupción sea suficiente como para poder realizar acciones de elusión o mitigación.



**Figura 6.46.** Descarga que presenta comportamiento disruptivo, con presencia de alarma prematura.

Ampliando la mitad izquierda de la Figura 6.46., se obtiene la Figura 6.47.



**Figura 6.47.** Ampliación de las señales que intervienen para la descarga 162457 de DIII-D.

La alarma es detectada con una antelación de casi 100 ms antes del instante de la disrupción; tratándose de una alarma prematura. Aunque para JET se consideraba una alarma como prematura, si ésta era detectada con 1 s o más de antelación con respecto al instante de la disrupción. Se observa, además, que la descarga logra recuperarse después de dicha alarma. Con un tiempo de anticipación tan amplio, se podrían emprender acciones de elusión para evitar que esa descarga llegue a perderse.

La suma de alarmas válidas y de alarmas prematuras contribuye a la tasa de aciertos.

En la Figura 6.48. se observa la presencia de una alarma tardía.

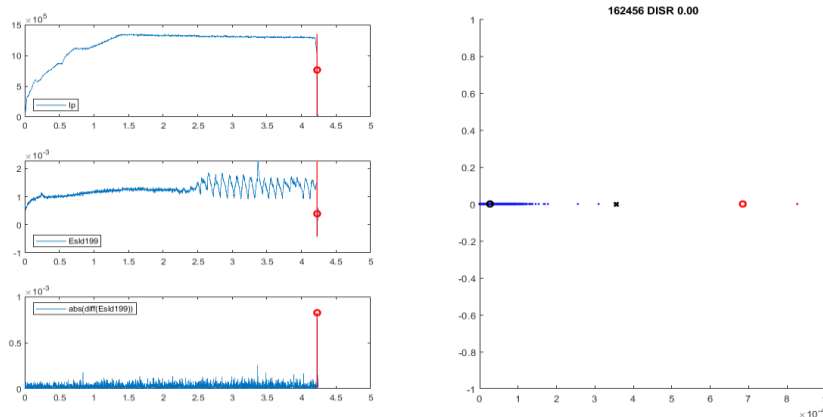


Figura 6.48. Descarga que presenta comportamiento disruptivo, con presencia de alarma tardía en DIII-D.

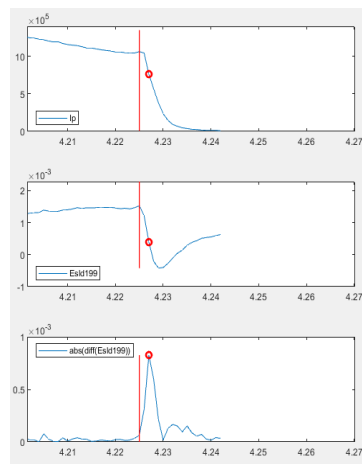


Figura 6.49. Ampliación de las señales que intervienen para la descarga 162456 de DIII-D.

En las imágenes ampliadas de la Figura 6.49., la alarma es detectada con posterioridad al instante de la disrupción, con lo cual no sirve de nada su detección; tratándose de una alarma tardía. Aquellas alarmas que sean detectadas durante el instante de la disrupción, o en instantes previos y muy próximos a dicha disrupción; también serán alarmas tardías, ya que no existiría tiempo suficiente de reacción para realizar las correspondientes tareas de elusión y/o mitigación.

Se realizaron un número considerable de pruebas, para validar la eficiencia de los predictores obtenidos durante la fase de entrenamiento; y donde también fueron consideradas innumerables variantes para la obtención de los centroides.

En la Figura 6.50. se muestran las gráficas ROC, teniendo en cuenta cuatro versiones diferentes empleadas para la obtención de los centroides.

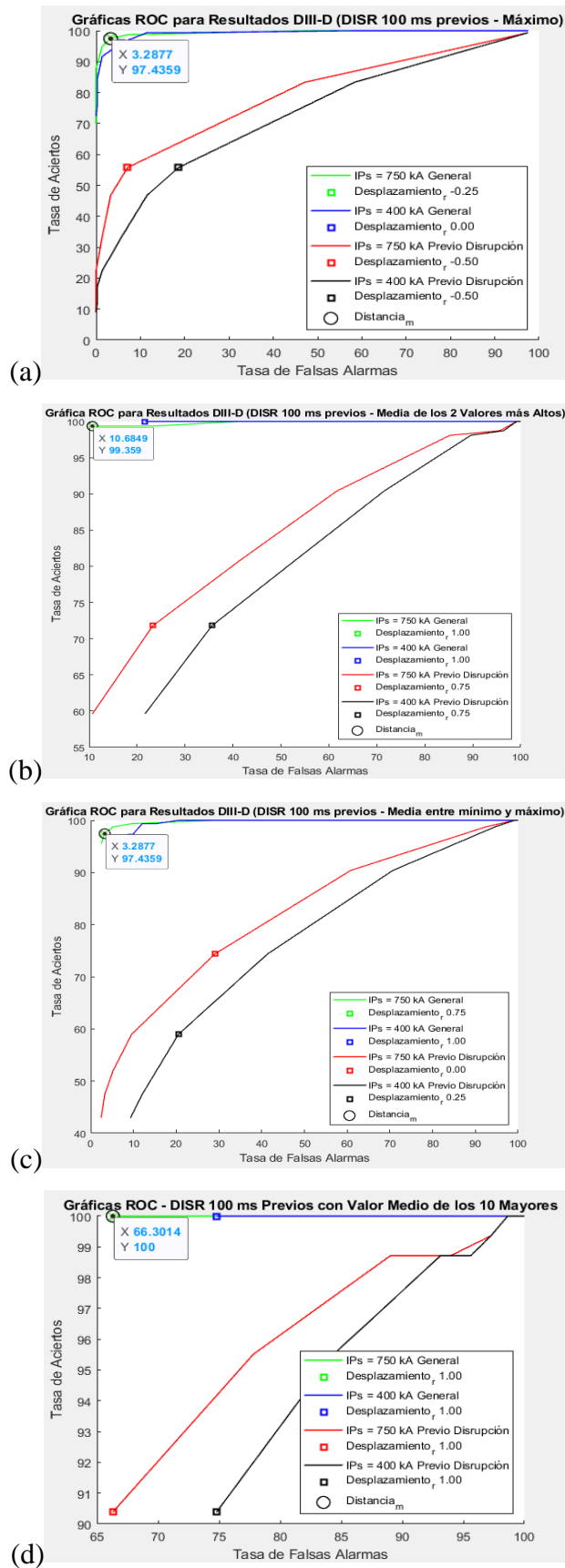


Figura 6.50. Resultados para cuatro versiones diferentes en la obtención de los centroides en DIID-D.

Las gráficas de la Figura 6.50. muestran la tasa de aciertos en relación a la tasa de falsas alarmas. Analizando cada una de las gráficas anteriores, se puede decir lo siguiente:

1. En la primera gráfica (Figura 6.50.(a)): se tienen en cuenta para el cálculo del centroide disruptivo los 100 ms previos a la interrupción, y de las muestras comprendidas en ese intervalo temporal, se selecciona el valor máximo de la señal  $\text{abs}(\text{diff}(\text{esld199}))$ . Los mejores resultados se consiguen trabajando con umbrales de corriente por encima de los 750 kA, y teniendo en cuenta un desplazamiento relativo de los centroides de -0,25 en relación a las barras de error. Los resultados obtenidos son: 97,44% de tasa de aciertos y 3,29% de falsas alarmas.
2. En la segunda gráfica (Figura 6.50.(b)): se tienen en cuenta para el cálculo del centroide disruptivo los 100 ms previos a la interrupción, y de las muestras comprendidas en ese intervalo temporal, se selecciona el valor medio de los dos valores más altos de la señal  $\text{abs}(\text{diff}(\text{esld199}))$ . Los mejores resultados se consiguen trabajando con umbrales de corriente por encima de los 750 kA, y teniendo en cuenta un desplazamiento relativo de los centroides de +1,00 en relación a las barras de error. Los resultados obtenidos son: 99,36% de tasa de aciertos y 10,68% de falsas alarmas.
3. En la tercera gráfica (Figura 6.50.(c)): se tienen en cuenta para el cálculo del centroide disruptivo los 100 ms previos a la interrupción, y de las muestras comprendidas en ese intervalo temporal, se selecciona el valor medio de los valores mínimo y máximo de la señal  $\text{abs}(\text{diff}(\text{esld199}))$ . Los mejores resultados se consiguen trabajando con umbrales de corriente por encima de los 750 kA, y teniendo en cuenta un desplazamiento relativo de los centroides de +0,75 en relación a las barras de error. Los resultados obtenidos son: 97,44% de tasa de aciertos y 3,29% de falsas alarmas, igual que en el primer caso descrito.
4. En la cuarta gráfica (Figura 6.50.(d)): se tienen en cuenta para el cálculo del centroide disruptivo los 100 ms previos a la interrupción, y de las muestras comprendidas en ese intervalo temporal, se selecciona el valor medio de los 10 valores más altos de la señal  $\text{abs}(\text{diff}(\text{esld199}))$ . Los mejores resultados se consiguen trabajando con umbrales de corriente por encima de los 750 kA, y teniendo en cuenta un desplazamiento relativo de los centroides de +1,00 en relación a las barras de error. Los resultados obtenidos son: 100,00% de tasa de aciertos y 66,30% de falsas alarmas.

Los resultados anteriores 1. y 3. son muy buenos, sin embargo, no pueden considerarse válidos; ya que un elevado número de alarmas es detectado de forma tardía. Es decir, se están contabilizando como aciertos, los resultados de la evaluación en descargas disruptivas; cuando las alarmas de disrupción, ni siquiera se han detectado con antelación con respecto al instante en el que se produce la disrupción.

Por tanto, cuando se consideran todas las alarmas (sea cual sea el instante de presentación), alta tasa de aciertos y baja tasa de falsas alarmas. Sin embargo, considerando sólo las alarmas efectivas (las que se presentan antes de producirse la disrupción); la tasa de aciertos disminuye drásticamente y aumenta ligeramente la tasa de falsas alarmas.

El análisis de curvas **ROC** (receiver operating characteristic curve) constituye un método estadístico para determinar la exactitud diagnóstica de los test.

Se realizó un último intento para evaluar la eficiencia del predictor empleando **regresión logística**. La regresión logística convierte la variable respuesta, que es cualitativa, en una probabilidad, que es cuantitativa. En la Figura 6.51. se muestra el proceso seguido mediante Matlab.

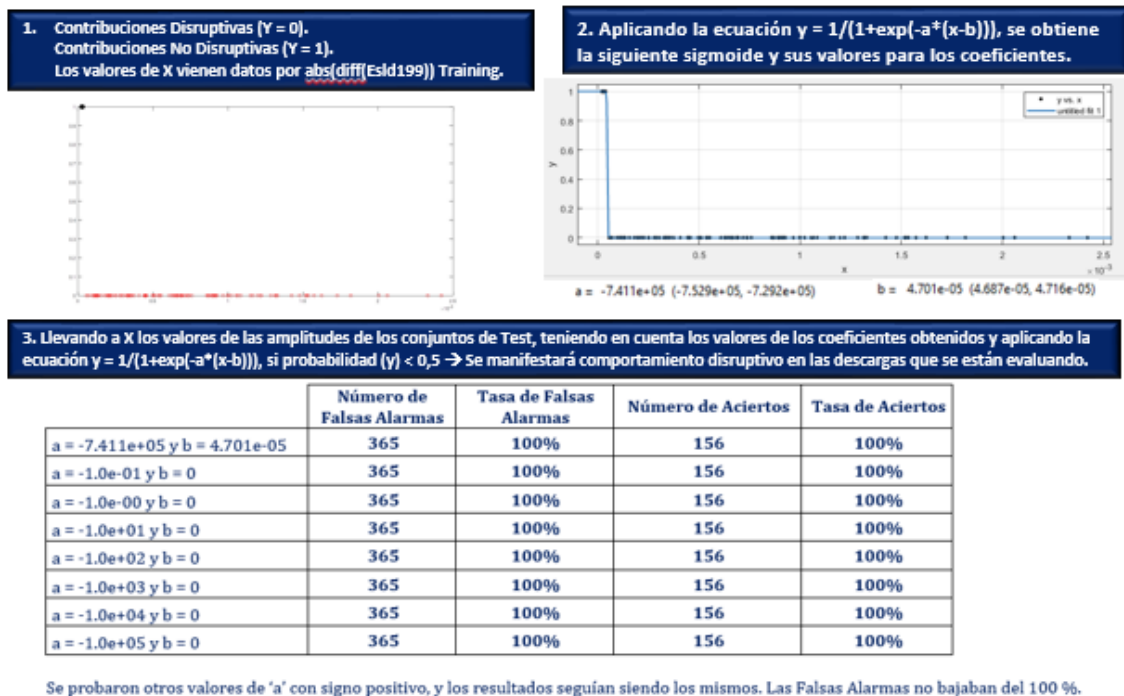


Figura 6.51. Regresión logística aplicada a DIII-D.

La regresión logística aplicada a los datos disponibles, obtiene resultados excelentes en cuanto a tasa de aciertos (alarmas encontradas en descargas evaluadas con

comportamiento disruptivo); pero la tasa de falsas alarmas obtenida es inadmisibles (100 % de falsas alarmas, cuando debería tender al 0,00 %).

Por tanto, para el dispositivo DIII-D, no ha resultado posible encontrar un predictor de interrupciones que sea efectivo, teniendo en cuenta los planteamientos adoptados. Este hecho pone de manifiesto, que la metodología implementada es dependiente de las señales que se tengan disponibles. En este sentido, se podría hablar de buscar soluciones *ad hoc*, para los distintos dispositivos de fusión termonuclear.

Tal y como fue comentado en uno de los capítulos anteriores de la tesis, los resultados que se presentan en este capítulo, han sido publicados en varias referencias de las que el doctorando es coautor: [Vega, 2017a], [Vega, 2017b], [Vega, 2017c], [Vega, 2018a], [Vega, 2018b], [Vega, 2018c], [Vega, 2019] y [Vega, 2020].

# CONCLUSIONES Y LÍNEAS FUTURAS DE INVESTIGACIÓN

## 7.1. Conclusiones.

El trabajo de investigación realizado en esta tesis se centra en el campo de la predicción de disrupciones en los dispositivos de fusión termonuclear de tipo Tokamak. Con anterioridad a detallar las principales conclusiones a las que ha dado lugar el presente trabajo de investigación, cabe mencionar que los resultados obtenidos están dentro de una línea de investigación muy consolidada de mis directores de tesis. En este sentido las aportaciones realizadas toman una mayor importancia, desde el punto de vista de la mejora de resultados de la investigación precedente en el campo de la predicción de disrupciones.

Con el fin de poner en contexto las principales aportaciones de esta tesis, es necesario repasar la evolución sufrida por los predictores de disrupción en la comunidad internacional de fusión a lo largo de los últimos 20 años. Empíricamente, desde los primeros dispositivos experimentales de fusión, se ha sabido de la relación directa entre actividad MHD y producción de disrupciones. La amplitud de la señal “Mode Lock” se ha utilizado siempre como precursor de eventos disruptivos, aunque su fiabilidad ha estado lejos del 100%. A principios del presente siglo se empezaron a desarrollar

predictores basados en aprendizaje automático. Si bien alguno de estos sistemas alcanzó en JET tasas de acierto del 98% y de falsas alarmas del 1%, de cara a ITER, estos predictores han mostrado tres problemas importantes. En primer lugar, aunque son capaces de predecir el fenómeno disruptivo, no proporcionan información acerca del evento físico que desencadena la disrupción. En segundo lugar, estos predictores no han podido extrapolarse entre diferentes Tokamaks, ya que el proceso de aprendizaje está fuertemente influenciado por la interacción plasma-pared y por la física de cada dispositivo. En tercer lugar, requieren de conjuntos de entrenamiento muy grandes, lo que los hace inviables en ITER para tener predictores fiables con un reducido número de descargas.

El tercero de los problemas mencionado se abordó mediante la denominada “aproximación partiendo de cero”. Esta aproximación construye predictores adaptativos a partir de una sola descarga disruptiva y muy pocas descargas no disruptivas. A medida que hace falta añadir conocimiento, se van reentrenando los predictores. Esta aproximación ha dado muy buenos resultados en JET. Sin embargo, aunque es posible predecir disrupciones, los predictores no dan información acerca de las causas físicas que las originan.

Posteriormente apareció una nueva familia de predictores, cuyo principal objetivo es no necesitar entrenamiento de pasadas descargas. Puesto que lo habitual es que el plasma evolucione de manera no disruptiva, los sistemas aprenden cómo es una evolución no disruptiva a medida que avanza la descarga. La detección de anomalías en esta evolución permite reconocer la proximidad de un evento disruptivo. Este tipo de sistemas fue también probado con éxito en JET. La principal desventaja de la implementación específica es la producción de falsas alarmas (superiores al 15%).

Con estos antecedentes se comenzó una investigación centrada en obtener un predictor adaptativo, lo más simple posible y cuyas predicciones fuesen interpretables desde el punto de vista de la física. Esta investigación fue lo que se propuso como trabajo de esta tesis. El predictor es el llamado “método de los centroides”. Este predictor ha demostrado su fiabilidad en tres máquinas diferentes: JET, JT-60U y DIII-D.

El método de los centroides utiliza una única señal, la clásica señal del “Mode Lock”, pero las predicciones no se basan en la simple amplitud, sino en la relación existente entre las amplitudes de muestras consecutivas tomadas periódicamente. La señal del “Mode Lock” crece cuando disminuye la velocidad de rotación de modos MHD del



plasma. Si la rotación llega a anularse, la amplitud del modo MHD crece de manera muy rápida, dando lugar a una disrupción. Por tanto, si la rotación del plasma se mantiene, las amplitudes del “Mode Lock” entre muestras consecutivas son muy similares. Esto significa que en el espacio bidimensional formado por muestras consecutivas del “Mode Lock”, los puntos aparecen agrupados alrededor de la bisectriz de dicho espacio. Cuando la diferencia entre muestras consecutivas es grande, la rotación del plasma ha sufrido un cambio brusco, lo que puede ser precursor de una disrupción. Este cambio súbito puede obedecer o a un frenado violento o a una aceleración rápida en la rotación del plasma. Este razonamiento muestra la relación entre el predictor y la interpretación física de las predicciones.

Por otro lado, es importante resaltar que en el espacio de muestras consecutivas que se ha considerado, la frontera de separación entre los comportamientos disruptivos y no disruptivos es la más simple posible: una línea recta. A partir de esta línea recta se explica claramente, que predictores basados simplemente en la amplitud del “Mode Lock” no sean fiables. Una amplitud en un instante determinado representa un comportamiento disruptivo dependiendo de la amplitud de la muestra anterior. Es decir, depende linealmente de la diferencia entre muestras consecutivas. No obstante, existe un umbral de amplitud por encima del cual el comportamiento es disruptivo, independientemente de la amplitud de la muestra anterior. Este umbral viene dado por la ordenada en el origen de la frontera de separación.

Como se ha descrito a lo largo de esta memoria, el método de los centroides se ha aplicado a tres Tokamaks diferentes con buenos resultados. Es importante enfatizar que en todos ellos se ha utilizado una sola señal: el “Mode Lock”. Aunque se habla de forma genérica de la señal de “Mode Lock”, lo cierto es que el diagnóstico que mide la señal en cada Tokamak es diferente. En términos generales, la señal de “Mode Lock” mide variaciones de campo magnético. Estas variaciones se miden con bobinas magnéticas, que en cada dispositivo tienen diferente geometría y una electrónica *ad hoc*. Esto se traduce en que las señales obtenidas en cada Tokamak no son exactamente equivalentes: las señales tienen diferentes amplitudes y polaridades (por ejemplo, la señal de JET es unipolar positiva mientras que las de las otras dos máquinas es bipolar y no necesariamente simétrica). Por consiguiente, la aplicación del método de los centroides a cada Tokamak ha requerido una optimización detallada para cada uno de ellos en esta tesis.

Los predictores basados en el método de los centroides comparados con cualesquiera otros, se manifiestan como los predictores más simples, sencillos y eficientes (en términos de tasas de acierto y de falsas alarmas), además de proporcionar una interpretación física directa. Pero adicionalmente a su capacidad para reconocer eventos disruptivos, se les ha encontrado una utilidad importantísima como predictores del tiempo a la disrupción. Conocer el tiempo a la disrupción sería fundamental en los dispositivos actuales, no solo para maximizar el tiempo de experimento con los beneficios que esto conlleva, sino también para elegir entre mitigación o elusión y seleccionar una técnica adecuada. La predicción del tiempo a la disrupción solamente se ha probado hasta ahora en JET (que conviene recordar que a día de hoy es el Tokamak más importante del mundo). Estos primeros resultados son muy satisfactorios y estimulan a optimizar el método y a ser aplicado en otros dispositivos.

## 7.2. Líneas futuras de investigación.

Esta tesis se ha enmarcado en el flujo principal de investigación sobre predicción de disrupciones, en el ámbito internacional de la fusión nuclear. Los resultados de esta tesis refuerzan y amplían el conocimiento previo en el campo. También ayudan a clarificar cuáles deben de ser las líneas inmediatas de investigación, para caracterizar la naturaleza del evento disruptivo. El objetivo es ayudar en su reconocimiento y control, para evitar potenciales riesgos en los dispositivos de siguiente generación, tales como ITER y DEMO.

Las necesarias líneas de investigación pueden resumirse en cuatro puntos:

- **Predicción del tiempo a la disrupción.** Esta necesidad se ha comentado ampliamente a lo largo de la memoria. Se ha enfatizado su relevancia en cuanto a permitir decidir en tiempo real si tomar acciones de elusión o mitigación y dentro de la elección, escoger el método más adecuado. Un punto de partida para esta línea puede ser el trabajo realizado en esta tesis sobre la predicción del tiempo a la disrupción, utilizando el método de los centroides [Vega, 2020].
- **Predictores cruzados.** Hasta ahora, todos los predictores se entrenan, validan y operan en el mismo dispositivo. En esta tesis se ha utilizado el método de los centroides en tres dispositivos diferentes y cada uno ha necesitado su propio entrenamiento. Pensando en ITER (y también en DEMO), se requiere el desarrollo

de predictores en máquinas más pequeñas (por ejemplo, JET o JT-60SA) que puedan utilizarse en dispositivos más grandes. Pueden mencionarse dos intentos en este sentido en el que el predictor se entrenaba con datos de ASDEX Upgrade y se probó con datos de JET [Windsor, 2005], [Rattá, 2018]. Los resultados muestran que todavía es necesaria mucha investigación en este campo.

- **Clasificación de los tipos de disrupción.** Aunque de forma general se habla de disrupciones, lo cierto es que existen varios tipos de ellas en función de los eventos físicos que las originan: bloqueo de modos MHD, límite de densidad, colapso radiactivo, eventos de desplazamiento vertical y otros [de Vries, 2011]. Una correcta clasificación de los tipos de disrupción mediante técnicas de aprendizaje automático permitiría escoger, en tiempo real, una medida de elusión o mitigación adecuada. Hasta la fecha hay muy pocos trabajos relativos a la clasificación de disrupciones. Un ejemplo es [Murari, 2009].
- **Creación de modelos físicos de disrupciones.** Las técnicas de aprendizaje automático han sido ampliamente utilizadas para la predicción de disrupciones. Estas técnicas funcionan muy bien desde el punto de vista de la ingeniería. Esto significa que se puede reconocer la presencia de disrupciones con cierta antelación, para que el sistema de control del Tokamak obre en consecuencia. Sin embargo, la causa física de la disrupción es desconocida. Debido a la complejidad de las interacciones altamente no lineales que aparecen en los plasmas de fusión, es muy difícil estudiar el fenómeno disruptivo desde primeros principios. Por esta razón, se necesitan generar modelos físicos orientados a datos a partir de técnicas de minería de datos. Existe un trabajo muy reciente [Vega, 2020b], que propone métodos automáticos en diferido para la detección de anomalías en señales de plasma. Se trata de relacionar estas anomalías mediante métodos de inteligencia artificial y poder llegar a crear modelos con sentido físico. Estos modelos ayudarían a la prevención de disrupciones, puesto que en tiempo real podría detectarse la proximidad de disrupciones conociendo su origen físico.



## A

- [Acero, 2014] Acero A., Vega J., Dormido-Canto S., Guinaldo M., Murari A. and JET-EFDA Contributors, 2014, “Assessment of probabilistic Venn machines as real-time disruption predictors from scratch: application to JET with a view on ITER”, Conf. Record of the 19th IEEE Real-Time Conf. (Nara, Japan, 26–30 May 2014), <http://rt2014.rcnp.osaka-u.ac.jp/AbstractsRT2014.pdf>.
- [Alves, 2011] D. Alves, R. Vitelli, L. Zaccarian, L. Zabeo, A. Neto, and F. Sartori *et al.*, “The new error field correction coil controller system in the joint european torus tokamak”, *Fus. Eng. Des.*, vol. 86, pp. 1024–1038, 2011.
- [Alves, 2012] D. Alves *et al.*, “Vessel thermal map real-time system for the JET tokamak”, special topics—accelerators and beams, *Phys. Rev.*, vol. 15, 2012.

## B

- [Barlow, 2001] Barlow I. *et al* 2001, “The error field correction coils on the JET machine”, *Fusion Eng. Des.* 58–59 189–93.
- [Betancourt, 2005] G.A. Betancourt, “Las Máquinas de Soporte Vectorial (SVMs)”, *Scientia et Technica* Año XI, No 27, Abril 2005. UTP. ISSN 0122-1701, Universidad Tecnológica de Pereira.
- [Bishop] C.M. Bishop, “Neural Networks for Pattern Recognition”, Clarendon, Oxford.
- [Bishop, 1994] C.M. Bishop, “Novelty detection and neural network validation”, in: *Proceedings of the IEE Conference on Vision and Image Processing*, 1994, pp. 217–222.
- [Bishop, 1998] Bishop C.M., Svens’ en M. and Williams C.K.I., 1998, “The generative topographic mapping”, *Neural Comput.* 10 215–34.
- [Bishop, 2004] Bishop C.M., 2004, “Neural Networks for Pattern Recognition”, (Oxford: Oxford University Press).
- [Boozer, 2012] A. H. Boozer, “Theory of tokamak disruptions,” *Phys. Plasmas*, vol. 19, p. 25, Mar. 2012.
- [Burges, 1998] C. Burges, “A tutorial on support vector machines for pattern recognition”, *Data Mining Knowl. Disc.*, vol. 2, pp. 121–167, 1998.

## C

- [Cannas, 2002] B. Cannas, R.S. Delogu, A. Fanni, P. Sonato, M.K. Zedda, JET-EFDA contributors, “On-line prediction and mitigation of disruption in ASDEX Upgrade”, *Nucl. Fusion* 37 (2002) 100–108.

- [Cannas, 2004a] B. Cannas, A. Fanni, G. Sias, P. Sonato, M.K. Zedda, JET-EFDA contributors, “Neural approaches to disruption prediction at JET”, in: Proceedings of the 31st EPS Conference on Plasma Physics, 28G, London, UK, 2004, p. 1.167.
- [Cannas, 2004b] Cannas B. et al, 2004, “Disruptions forecasting at JET using neural networks”, Nucl. Fusion 44 68–76.
- [Cannas, 2005] B. Cannas, A. Fanni, P. Sonato, M.K. Zedda, JET-EFDA contributors to the EFDA-JETwork programme, “Novelty Detection for on-line disruption prediction systems”, in: Proceedings of the 32nd EPS Conference on Controlled Fusion and Plasma Physics, Tarragona, Spain, June 27–July 1, 2005.
- [Cannas, 2006a] B. Cannas, F. Cau, A. Fanni, P. Sonato, M.K. Zedda, JET-EFDA contributors, “Automatic disruption classification at JET: comparison of different pattern recognition techniques”, Nucl. Fusion 46 (2006) 699–708.
- [Cannas, 2006b] Cannas B. et al, 2006, “Disruption prediction at ASDEX Upgrade using neural network”, 33th EPS Conf. on Controlled Fusion and Plasma Physics, (Rome, Italy, 19–23 June 2006) vol 30I (ECA) P-2.143, [http://epsppd.epfl.ch/Roma/pdf/P2\\_143.pdf](http://epsppd.epfl.ch/Roma/pdf/P2_143.pdf).
- [Cannas, 2007a] B. Cannas, R.S. Delogu, A. Fanni, P. Sonato, M.K. Zedda, JET-EFDA contributors, “Support vector machines for disruption prediction and novelty detection at JET”, Fusion Engineering and Design 82 (2007) 1124–1130.
- [Cannas, 2007b] Cannas B. et al, 2007, “A prediction tool for real-time application in the disruption protection system at JET”, Nucl.Fusion 47 1559–69.
- [Cannas, 2010] B. Cannas, A. Fanni, G. Pautasso, G. Sias and P. Sonato, “An adaptive real-time disruption predictor for ASDEX Upgrade”, Nucl. Fusion 50 (2010) 075004 (12pp), [doi:10.1088/0029-5515/50/7/075004](https://doi.org/10.1088/0029-5515/50/7/075004).
- [Carmona, 2014] E. J. Carmona Suárez, “Tutorial sobre Máquinas de Vectores Soporte (SVM)”, Dpto. de Inteligencia Artificial, ETS de Ingeniería Informática, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain, 11 de Julio de 2014.
- [Chang, 2001] C. C. Chang, C. J. Lin, “LIBSVM a library for support vector machines”, Software available at: <http://www.csie.ntu.edu.tw>, 2001.
- [Cherkassky, 1998] V. Cherkassky, F. Mulier, “Learning from data”, Hoboken, NJ, USA: John Wiley & Sons, Inc., 1998.
- [Cherkassky, 2007] Cherkassky V. and Mulier F., 2007, “Learning from Data 2nd edn”, (New York: Wiley).
- [Cios, 2007] Cios K.J., Pedrycz W., Swiniarski R.W. and Kurgan L.A., 2007, “Data Mining. A Knowledge Discovery Approach”, (Berlin: Springer).
- [Cristianini, 2000] N. Cristianini, Shawe-Taylor, “An Introduction to Support Vector Machines”, Cambridge University Press, 2000.

## D

- [Dashevskiy, 2008] Dashevskiy M. and Luo Z., 2008, “Reliable probabilistic classification and its application to internet traffic”, ICIC 2008 (Shanghai, China) (Lecture Notes in

Computer Science vol 5226), ed D.-S. Huang et al (Heidelberg: Springer) pp 380–8.

[Demuth, 1998] H. Demuth, M. Beale, “Neural Network Toolbox User’s Guide”, Versión 3.0, 1998.

[Dormido-Canto, 2013] S. Dormido-Canto, J. Vega, J.M. Ramírez, A. Murari, R. Moreno, J.M. López, A. Pereira and JET-EFDA Contributors, 2013, “Development of an efficient real-time disruption predictor from scratch on JET and implications for ITER”, Nucl. Fusion 53 (2013) 113001 (8pp) doi:10.1088/0029-5515/53/11/113001.

[Duda, 2001] Duda R.O., Hart P.E. and Stork D.G., 2001, “Pattern Classification 2nd edn”, (New York: Wiley-Interscience).

## E

[Esposito, 2009] Esposito B. et al, 2009, “Disruption control on FTU and ASDEX upgrade with ECRH”, Nucl. Fusion 49 065014.

[Esquembri, 2018] S. Esquembri , Student Member, IEEE, J. Vega, A. Murari, M. Ruiz, Senior Member, IEEE, E. Barrera , Member, IEEE, S. Dormido-Canto, R. Felton, M. Tsalas, D. Valcarcel, and JET Contributors, “Real-Time Implementation in JET of the SPAD Disruption Predictor Using MARTE”, IEEE Transactions on Nuclear Science, 65 832-42, February 2018.

## F

[Felton, 1999] R. Felton et al., “Real-time plasma control at JET using an ATM network”, in Proc. 11th IEEE NPSS Real Time Conf., Santa Fe, NM, USA, June 1999, pp. 175–181.

## G

[Giannone, 2004a] Giannone et al, 2004, “Regime Identification in ASDEX Upgrade”, 31th EPS Conf. on Plasma Physics, (London 28 June–2 July 2004), [http://epsppd.epfl.ch/London/pdf/P4\\_131.pdf](http://epsppd.epfl.ch/London/pdf/P4_131.pdf).

[Giannone, 2004b] Giannone L. et al, 2004, “Regime identification in ASDEX Upgrade”, Plasma Phys. Control. Fusion 46 835–56.

## H

[Hender, 2007] Hender T.C. et al, 2007, “MHD stability, operational limits and disruptions”, Nucl. Fusion 47 S128–202.

[Hertz, 1991] Hertz J.A., Krogh A.S. and Palmer R.G., 1991, “Introduction to the Theory of Neural Computation”, Redwood City, CA: Addison-Wesley.

- [Hollmann, 2015] E. M. Hollmann, P. B. Aleynikov, T. Fülöp, D. A. Humphreys, V. A. Izzo, M. Lehnen, V. E. Lukash, G. Papp, G. Pautasso, F. Saint-Laurent, and J. A. Snipes, “Status of research toward the ITER disruption mitigation system”, *Environmental Science Physics of Plasmas* **22**, 021802 (2015), <https://doi.org/10.1063/1.4901251>

## I

- [ITER, 1999] ITER Physics Expert Group on Disruptions et al, 1999, “MHD stability, operational limits and disruptions”, *Nucl. Fusion* **39** 2137–664.

## J

- [Japkowicz, 2002] Japkowicz N. and Stephen S., 2002, “The class imbalance problem: a systematic study”, *Intell. Data Anal.* **6** 429–49.

## K

- [Kharecha, 2013] A. Kharecha, James E. Hansen, “Prevented Mortality and Greenhouse Gas Emissions from Historical and Projected Nuclear Power”, *Environmental Science & Technology* **47** (2013) es3051197 (7 pp).  
<https://pubs.acs.org/doi/pdf/10.1021/es3051197>
- [Kohonen, 1982] T. Kohonen, “Self-organized formation of topologically correct feature maps”, *Biol. Cybern.* **43** (1982) 59–69.
- [Kohonen, 1998] Kohonen T., 1998, “The self-organizing map *Neurocomputing* **21** 1–6”.

## L

- [Lambrou, 2012] Lambrou A., Papadopoulos H., Nouretdinov I. and Gammerman A., 2012, “Reliable probability estimates based on support vector machines for large multiclass datasets”, *AIAI 2012 Workshops, (Halkidiki, Greece), IFIP AICT vol 382, (Berlin: Springer) pp 182–91.*
- [López, 2012] J.M. López, J. Vega, D. Alves, S. Dormido-Canto, A. Murari, J.M. Ramírez, et al., 2012, “Implementation of the disruption predictor APODIS in JET real-time network using the MARTe framework”, *Proc. of the 18th IEEE-NPSS Real Time Conf.*, June 11–15, Berkeley (CA), USA, doi:10.1109/RTC.2012.6418168.
- [López, 2014] J. M. López, J. Vega, D. Alves, *Member, IEEE*, S. Dormido-Canto, A. Murari, J. M. Ramírez, R. Felton, M. Ruiz, and G. de Arcas, “Implementation of the Disruption Predictor APODIS in JET’s Real-Time Network Using the MARTe Framework”, *IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 61, NO. 2, APRIL 2014.*



## M

- [Markou, 2003a] M. Markou, S. Singh, “Novelty detection: a review. Part 1. Statistical Approaches”, *Signal Process.* 83 (2003) 2481–2497.
- [Markou, 2003b] M. Markou, S. Singh, “Novelty detection: a review. Part 2. Neural network-based approaches”, *Signal Process.* 83 (2003) 2499–2521.
- [Milani, 1998] F. Milani, PhD thesis, University of Aston, Birmingham, December 1998.
- [Morabito, 2001] Morabito F.C., Versaci M., Pautasso G., Tichmann C. and ASDEX Upgrade Team, 2001, “Fuzzy-neural approaches to the prediction of disruptions in ASDEX Upgrade”, *Nucl. Fusion* 41 1715–23.
- [Murari, 2008] Murari A. et al, 2008, “Prototype of an adaptive disruption predictor for JET based on fuzzy logic and regression trees”, *Nucl. Fusion* 48 68–76.
- [Murari, 2009] A. Murari, J. Vega, G. A. Rattá, G. Vagliasindi, M. F. Johnson, S. H. Hong and JET-EFDA Contributors. “Unbiased and non-supervised learning methods for disruption prediction at JET”. *Nuclear Fusion*. 49 (2009) 055028 (11pp).
- [Murari, 2018] A. Murari, M. Lungaroni, E. Peluso, P. Gaudio, J. Vega, S. Dormido-Canto, et al., *Nucl. Fusion* 58 (2018) 05600216pp.

## N

- [Neto, 2010] A.C. Neto, F. Sartori, F. Piccolo, R. Vitelli, G. de Tommasi, L. Zabeo, et al., “MARTE: a multiplatform real-time framework”, *IEEE Transactions on Nuclear Science* 57 (April (2)) (2010) 479–486.
- [Neto, 2012] A. Neto *et al.*, “Exploitation of modularity in the JET tokamak vertical stabilization system”, *Contr. Eng. Practice*, vol. 20, no. 9, pp. 846–856, Sep. 2012.
- [Nouretdinov, 2012] Nouretdinov I. et al, 2012, “Multiprobabilistic Venn predictors with logistic regression”, *AIAI 2012 Workshops (Halkidiki, Greece), IFIP AICT vol 382* (Berlin: Springer) pp 224–33.
- [NuclFus39, 1999] Chapter 3 of: ITER Physics Expert Group on Disruptions et al., *Nucl. Fus.* 39 (1999) 2251.

## P

- [Papadopoulos, 2013] Papadopoulos H., 2013, “Reliable probabilistic classification with neural networks”, *Neurocomputing* 107 59–68.
- [Pautasso, 1998a] G. Pautasso et al., *EPS Conference on Controlled Fusion and Plasma Physics*, Prague, 1998.

- [Pautasso, 1998b] Pautasso G. et al, 1998, “Causes, precursor and mechanisms of disruptions in ASDEX Upgrade”, ICPP & 25th EPS Conf. on Control Fusion and Plasma Physics (Praha, Czech Republic, 29 June–3 July 1988) vol 22C p 520, [https://fusion.gat.com/conferences/meetings/eps98/web/98icpp\\_w/b169pr.pdf](https://fusion.gat.com/conferences/meetings/eps98/web/98icpp_w/b169pr.pdf).
- [Pautasso, 2001] G. Pautasso, S. Egorov, Ch. Tichmann, J.C. Fuchs, A. Herrmann, M. Maraschek, F. Mast, V. Mertens, I. Perchermeier, C.G. Windsor, T. Zehetbauer, ASDEX Upgrade Team, “Prediction and mitigation of disruptions in ASDEX Upgrade”, Fusion Association, Culham Science Center, OX14 3DB, UK, *Journal of Nuclear Materials* 290-293 (2001) 1045-1051 (7pp).
- [Pautasso, 2002] Pautasso G. et al, 2002, “On-line prediction and mitigation of disruption in ASDEX Upgrade”, *Nucl. Fusion* 42 100–08.
- [Pautasso, 2009] Pautasso G. et al, 2009, “Disruptions studies in ASDEX Upgrade in view of ITER”, *Plasma Phys. Control Fusion* 51 124056.
- [Polikar, 2001] Polikar R., 2001, “Learn++: an incremental learning algorithm for supervised neural networks”, *IEEE Trans. on Systems Man Cybern.—Part C: Appl. Rev.* 31 497–508.
- ## R
- [Rattá, 2008] Rattá G.A., Vega J., Murari A., Johnson M. and JET-EFDA Contributors, “Feature extraction for improved disruption prediction analysis at JET”, *Review of Scientific Instruments*, 79 (2008) 10F328.
- [Rattá, 2010] G.A. Rattá, J. Vega, A. Murari, G. Vagliasindi, M.F. Johnson, P.C. de Vries and JET EFDA Contributors, “An advanced disruption predictor for JET tested in a simulated real-time environment”, *Nucl. Fusion* 50 (2010) 025005 (10pp) [doi:10.1088/0029-5515/50/2/025005](https://doi.org/10.1088/0029-5515/50/2/025005) .
- [Rattá, 2012] G.A. Rattá, J. Vega, A. Murari, JET-EFDA Contributors, *Fusion Engineering and Design*, 2012, <http://dx.doi.org/10.1016/j.fusengdes.2012.07.002>.
- [Rattá, 2014] Rattá G.A., Vega J., Murari A. and JET EFDA Contributors, 2014, “Simulation and real-time replacement of missing plasma signals for disruption prediction: an implementation with APODIS Plasma”, *Phys. Control. Fusion* 56 114004.
- [Rattá, 2018] G. A. Rattá, J. Vega, A. Murari. “Viability Assessment of a Cross-Tokamak AUG-JET Disruption Predictor”. *Fusion Science and Technology*, 74 (2018), 13-22. <https://doi.org/10.1080/15361055.2017.1390390> .
- [Reux, 2013] Reux C. et al (JET-EFDA Contributors), 2013, “Use of the disruption mitigation valve in closed loop for routine protection at JET”, *Fusion Eng. Des.* 88 1101–4.
- [Ruiz, 2011] M. Ruiz, J. Vega, G. Arcas, G. Ratta, E. Barrera, A. Murari, J. M. Lopez, and R. Melendez, “Real-time plasma disruptions detection in JET implemented with the

ITMS platform using FPGA based IDAQ”, *IEEE Trans. Nucl. Sci.*, vol. 58, no. 4, pt. 1, pp. 1576–1581, Aug., 2011.

## S

- [Santagiustina, 1995] Santagiustina A. et al 1995, “Studies of tearing mode control in JET Proc. 22nd EPS Conf. on Controlled Fusion and Plasma Physics”, (Bournemouth, 3–7 July 1995) (European Conference Abstracts), 19c IV p 461 (EPS, Geneva).
- [Saunders, 1999] Saunders C., Gammerman A. and Vovk V., 1999, “Transduction with confidence and credibility”, Proc. 16th Int. Joint Conf. on Artificial Intelligence (Bled, Slovenia), vol 2 (San Francisco, CA: Morgan Kaufmann) pp 722–6.
- [Schuller, 1995] Schuller F.C., 1995, “Disruptions in tokamak”, *Plasma Phys. Control. Fusion* 37 135–62.
- [Stephen, 2011] A. V. Stephen *et al.*, “Centralised coordinated control to protect the JET ITER-like wall”, presented at the 13th Int. Conf. Accelerator and Large Experimental Physics control Systems, 2011.

## T

- [Tamprasert, 2002] Tanprasert T. and Kripruksawan T., 2002, “An approach to control aging rate of neural networks under adaptation to gradually changing context”, ICONIP’02 9th Int. Conf. on Neural Information, (Singapore, 18–22 November 2002), vol 1 pp 174–8, <http://www3.ntu.edu.sg/Iconip02/>

## U

- [URL, 1] La Energía de Fusión Nuclear. INSTITUTO DE LA INGENIERÍA DE ESPAÑA.  
<https://www.iies.es/events/la-energia-de-fusion-nuclear>
- [URL, 2] La fusión nuclear como solución a las necesidades energéticas. MAPFRE Global Risks.  
<https://www.mapfreglobalrisks.com/gerencia-riesgos-seguros/articulos/la-fusion-nuclear-como-solucion-a-las-necesidades-energeticas/>
- [URL, 3] ¿Qué es la fusión nuclear? Taylor Hopkinson  
<https://www.taylorhopkinson.com/fusion-nuclear-el-futuro-de-la-energia/>
- [URL, 4] ITER y la energía de fusión. Joaquín Sánchez (CIEMAT). LYCHNOS. Cuadernos de la Fundación General CSIC Edición Digital.

- [http://www.fgcsic.es/lychnos/es\\_ES/articulos/iter\\_y\\_la\\_energia\\_de\\_fusion](http://www.fgcsic.es/lychnos/es_ES/articulos/iter_y_la_energia_de_fusion)
- [URL, 5] Fusión: la energía del siglo XXI. Laboratorio Nacional de Fusión (CIEMAT)  
[http://www-fusion.ciemat.es/New\\_fusion/es/Fusion/basica.shtml](http://www-fusion.ciemat.es/New_fusion/es/Fusion/basica.shtml)
- [URL, 6] Energía de fusión. Laboratorio Nacional de Fusión  
<http://www.fusion.ciemat.es/energia-de-fusion/>
- [URL, 7] El proceso de la fusión. Laboratorio Nacional de Fusión  
<http://www.fusion.ciemat.es/energia-de-fusion/el-proceso-de-la-fusion/>
- [URL, 8] Hito en el desarrollo de fusión nuclear, la energía del futuro. Ciencia BBC Mundo, 10 de octubre de 2013  
[https://www.bbc.com/mundo/noticias/2013/10/131010\\_ciencia\\_hito\\_energia\\_nuclear\\_fusion\\_nif\\_ch](https://www.bbc.com/mundo/noticias/2013/10/131010_ciencia_hito_energia_nuclear_fusion_nif_ch)
- [URL, 9] El proyecto de fusión nuclear ITER. Foro Nuclear  
<https://www.foronuclear.org/es/el-experto-te-cuenta/122980-el-proyecto-de-fusion-nuclear-iter>
- [URL, 10] ¿Qué es un reactor de fusión nuclear? Foro Nuclear  
<https://www.foronuclear.org/es/100957-faqas-sobre-energia/capitulo-8/115744-112-ique-es-un-reactor-de-fusion-nuclear>
- [URL, 11] Fusión nuclear. Energía nuclear  
<https://energia-nuclear.net/que-es-la-energia-nuclear/fusion-nuclear>
- [URL, 12] Nuclear fusión. Consejo de Seguridad Nuclear  
<https://www.csn.es/en/fusion-nuclear>
- [URL, 13] Nuclear energy in Spain. Consejo de Seguridad Nuclear  
<https://www.csn.es/en/la-energia-nuclear-en-espana>
- [URL, 14] Fusión nuclear. Norvento ENERXIA. Alberto Loarte. ITER y combustibles empleados  
<https://www.norvento.com/fusion-nuclear/>
- [URL, 15] ITER avanza, conseguir la energía de las estrellas en la Tierra está más cerca. Juan Ranchal. Publicado el 25 de julio de 2019.  
<https://www.muycomputer.com/2019/07/25/iter/>
- [URL, 16] ¿Es la fusión nuclear la energía del futuro? THE CONVERSATION.  
<http://theconversation.com/es-la-fusion-nuclear-la-energia-del-futuro-109969>
- [URL, 17] La Energía de Fusión Nuclear. Instituto de la Ingeniería de España  
<https://www.iies.es/single-post/2019/04/23/La-Energía-de-Fusión-Nuclear>
- [URL, 18] Fusión. OIEA (Organismo Internacional de Energía Atómica)  
<https://www.iaea.org/es/temas/fusion>
- [URL, 19] ¿Qué es y cómo funciona la fusión nuclear, la energía limpia? Twenergy. 25 agosto 2019  
<https://twenergy.com/energia/energia-nuclear/fusion-nuclear-energia-limpia-e-inagotable-1921/>
- [URL, 20] Adiós, combustibles fósiles: en 2050 Europa dependerá de las renovables y la energía nuclear.  
<https://www.xataka.com/energia/adios-combustibles-fosiles-2050-europa-dependera-renovables-energia-nuclear>

- [URL, 21] Los combustibles fósiles frente a la energía nuclear.  
[http://recursostic.educacion.es/eda/web/eda2008/profesores\\_newton/practicas\\_newton/p3/Eda2008%20Newton/fernandez\\_munim/textos/6no\\_renovables/no\\_renovables.htm](http://recursostic.educacion.es/eda/web/eda2008/profesores_newton/practicas_newton/p3/Eda2008%20Newton/fernandez_munim/textos/6no_renovables/no_renovables.htm)
- [URL, 22] Combustible nuclear. Energía Nuclear  
<https://energia-nuclear.net/funcionamiento-central-nuclear/combustible-nuclear>
- [URL, 23] Ventajas e inconvenientes de las energías no renovables.  
<https://www.aelca.es/es/nuestro-blog/energia-no-renovable/>
- [URL, 24] LA INFORMACIÓN. Ciencia y tecnología. Los científicos sí tienen un plan para sustituir las energías fósiles por la fusión nuclear. JOSÉ GARCÍA. Viernes, 03 febrero 2017, 00:00.  
[https://www.lainformacion.com/tecnologia/fusion-nuclear-energia-renovable\\_0\\_996201958.html](https://www.lainformacion.com/tecnologia/fusion-nuclear-energia-renovable_0_996201958.html)
- [URL, 25] Mejoramos el retorno de la inversión en I + D. Kim (Knowledge Innovation Market)  
<https://kinglobal.com/es/futta-broker/>
- [URL, 26] Fusión Nuclear en 2030, la energía de las estrellas en la Tierra. Juan Ranchal. Publicado el 31 de enero de 2017.  
<https://www.muycomputer.com/2017/01/31/fusion-nuclear-2030/>
- [URL, 27] Energía fósil (combustibles fósiles).  
<https://www.oiltanking.com/es/publicaciones/glosario/detalles/term/energia-fosil.html>
- [URL, 28] Energía y Fusión. ¿Realmente necesitamos la fusión?  
[http://200.23.188.74/sites/ciencia/volumen3/ciencia3/135/html/sec\\_9.html](http://200.23.188.74/sites/ciencia/volumen3/ciencia3/135/html/sec_9.html)
- [URL, 29] Energía. Muy Interesante  
<https://www.muyinteresante.es/ciencia/articulo/20-preguntas-sobre-energia-nuclear-551555262577>
- [URL, 30] Acuerdo de París contra el cambio climático.  
<https://www.xataka.com/ecologia-y-naturaleza/acuerdo-historico-en-la-cumbre-de-paris-contra-el-cambio-climatico>
- [URL, 31] La contaminación del aire causa más muertes al año que el tabaco. Muy Interesante  
<https://www.muyinteresante.es/salud/articulo/la-contaminacion-del-aire-causa-mas-muertes-al-ano-que-el-tabaco-421552386516>
- [URL, 32] Avances en fusión nuclear, energía limpia e ilimitada. Juan Ranchal. Publicado el 13 de febrero de 2014.  
<https://www.muycomputer.com/2014/02/13/avances-en-fusion-nuclear/>
- [URL, 33] Conferencia de las Naciones Unidas sobre el Cambio Climático, diciembre de 2019. 02 Dic, 2019 - 13 Dic, 2019.  
<https://unfccc.int/es/cop25>
- [URL, 34] Ministerio para la Transición Ecológica. Gobierno de España.  
<https://www.miteco.gob.es/es/cop25/>
- [URL, 35] Todo lo que debes saber sobre la COP25 Chile-Madrid 2019.  
[https://www.cnnchile.com/cop25/que-es-la-cop25-preguntas-y-respuestas\\_20191129/](https://www.cnnchile.com/cop25/que-es-la-cop25-preguntas-y-respuestas_20191129/)
- [URL, 36] IFEMA, sede de la Cumbre Mundial del Clima COP25.

- <https://www.ifema.es/cop25>
- [URL, 37] Cambio climático: cinco cosas que debes saber sobre la COP25.  
<https://news.un.org/es/story/2019/11/1465981>
- [URL, 38] ¿Qué es la COP25 y para qué sirve?  
<https://es.greenpeace.org/es/noticias/que-es-la-cop25-y-para-que-sirve/>
- [URL, 39] COP25: una cumbre de poca altura.  
<https://www.lavanguardia.com/vida/20191214/472206288094/cop25-cumbre-clima-poca-altura.html>
- [URL, 40] Redes Neuronales, Aprendizaje Supervisado y No Supervisado.  
<http://redesneuronal.blogspot.com.es/>
- [URL, 41] Implications of Disruption Mitigation for the ITER Vacuum System.  
[https://lss.fnal.gov/conf/C070709.13/DM\\_Baylor.pdf](https://lss.fnal.gov/conf/C070709.13/DM_Baylor.pdf)
- [URL, 42] Energía Nuclear. Fusión Nuclear.  
<http://energia-nuclear.net/que-es-la-energia-nuclear/fusion-nuclear>
- [URL, 43] ITER.  
<https://www.iter.org/>
- [URL, 44] Concepto del reactor de fusión TOKAMAK del Instituto Kurchatov, Moscú.  
[http://www.lapizarradeyuri.com/wp-content/uploads/2015/03/15\\_tokamak\\_iter.jpg](http://www.lapizarradeyuri.com/wp-content/uploads/2015/03/15_tokamak_iter.jpg)
- [URL, 45] UK Atomic Energy Authority. CCFE.  
<http://www.ccf.ac.uk/CCFE.aspx>
- [URL, 46] Reactor TOKAMAK. Revista Sputnik.  
<https://sputnik87.wordpress.com/tag/reactor-tokamak/>
- [URL, 47] Foro Nuclear. ¿Qué diferencia hay entre fisión y fusión nuclear?  
<https://www.foronuclear.org/es/el-experto-te-cuenta/122516-que-diferencia-hay-entre-fision-y-fusion-nuclear>
- [URL, 48] ITER. JT-60SA.  
<https://www.iter.org/newsline/-/3273>
- [URL, 49] JT-60SA. Relation to ITER.  
[http://www.jt60sa.org/b/index\\_nav\\_1.htm?n1/ITER.htm](http://www.jt60sa.org/b/index_nav_1.htm?n1/ITER.htm)
- [URL, 50] General Atomics. TOKAMAK Operations and Engineering.  
<http://www.ga.com/tokamak-operations-and-engineering>
- [URL, 51] General Atomics. Plasma Control.  
<http://www.ga.com/plasma-control>
- [URL, 52] ASDEX Upgrade.  
<https://www.ipp.mpg.de/16195/asdex>
- [URL, 53] Wavelets. Proyecto Sistemas Informáticos.  
[http://eprints.ucm.es/16705/1/SI\\_Elena\\_Rupérez\\_Cerezo\\_Jun12.pdf](http://eprints.ucm.es/16705/1/SI_Elena_Rupérez_Cerezo_Jun12.pdf)
- [URL, 54] Distancia de Mahalanobis.  
[https://www.uv.es/ceaces/multivari/cluster/d\\_mahalanobis.htm](https://www.uv.es/ceaces/multivari/cluster/d_mahalanobis.htm)
- [URL, 55] Distancia Euclídea.

<http://bioinformatica.upf.edu/2004/projectes04/3.1.2/distancias.html>

- [URL, 56] Criterio de Lawson.  
[http://www-fusion.ciemat.es/New\\_fusion/es/Fusion/confinamiento.shtml](http://www-fusion.ciemat.es/New_fusion/es/Fusion/confinamiento.shtml)
- [URL, 57] Modelo de distribución de Weibull.  
<http://www.ub.edu/stat/GrupsInnovacio/Statmedia/demo/Temas/Capitulo4/B0C4m1t9.htm>

## V

- [Vannucci, 1999] A. Vannucci et al., Nucl. Fus. 39(1999) 255.
- [Vapnik, 1998] V. Vapnik, “Statistical Learning Theory”, Wiley-Interscience, New York, 1998.
- [Vega, 2013] Jesús Vega, Sebastián Dormido-Canto, Juan M. López, Andrea Murari, Jesús M. Ramírez, Raúl Moreno, Mariano Ruiz, Diogo Alves, Robert Felton, JET-EFDA Contributors, “Results of the JET real-time disruption predictor in the ITER-like wall campaigns”, Fusion Engineering and Design 88 (2013) 1228–1231.
- [Vega, 2014] J. Vega, A. Murari, S. Dormido-Canto, R. Moreno, A. Pereira, A. Acero and JET-EFDA Contributors, “Adaptive high learning rate probabilistic disruption predictors from scratch for the next generation of tokamaks”, Nucl. Fusion 54 (2014) 123001 (17pp) doi:10.1088/0029-5515/54/12/123001.
- [Vega, 2015a] Vega J., Moreno R., Pereira A., Dormido-Canto S., Murari A. and JET Contributors, 2015, “Advanced disruption predictor based on the locked mode signal: application to JET”, 1st EPS Conference on Plasma Diagnostics (14–17 April 2015, Frascati, Italy) 02/03, (<http://www.ecpd2015.enea.it/index.php/scientificprogramme>).
- [Vega, 2015b] Vega J., Murari A., Dormido-Canto S., Moreno R., Pereira A., Rattá G.A. and JET Contributors, 2015, “Disruption precursor detection: combining the time and frequency domains”, Proc. of the 26th Symp. on Fusion Engineering (SOFE 2015), (Austin, TX, 31 May–4 June 2015), (<http://ece-events.unm.edu/ppcsofe2015/>).
- [Vega, 2016a] J. Vega, A. Murari, S. Dormido-Canto, R. Moreno, A. Pereira, S. Esquembri and JET, “Real-time anomaly detection for disruption prediction: the JET case”, Fusion Energy Conference 2014, Saint Petersburg, Russia EUROFUSION WPJET1-PR(16) 14855.
- [Vega, 2017a] J. Vega, S. Dormido-Canto, F. Hernández, A. Murari, T. Cruz, G. A. Rattá and JET Contributors, “Disruption prediction: linear equation in two variables as boundary and comparison of predictors”, T17-14, July 11<sup>th</sup>, 2017.
- [Vega, 2017b] J. Vega, S. Dormido-Canto, F. Hernández, A. Murari, T. Cruz, G. A. Rattá and JET Contributors, “Strong reduction of false alarms with the centroid method and the mode lock”, T17-14, September 13<sup>th</sup>, 2017.

- [Vega, 2017c] J. Vega, S. Dormido-Canto, F. Hernández, A. Murari, T. Cruz, G. A. Rattá and JET Contributors, “Physics based linear equation in two variables as boundary of disruptive and non-disruptive regions in JET”, TFM, September 14<sup>th</sup>, 2017.
- [Vega, 2018a] J. Vega, F. Hernández, S. Dormido-Canto, “Disruption predictor for JT-60SA: first approach”, WPSA 5<sup>th</sup> PPM, Barcelona, 5-7 Feb 2018.
- [Vega, 2018b] J. Vega, F. Hernández, S. Dormido-Canto, A. Isayama, E. Joffrin, G. Matsunaga, T. Suzuki, “Assessment of linear disruption predictors using JT-60U data”, (Contribution to the SOFT Symposium), RCM-7, Naka, 4-8 Jun 2018.
- [Vega, 2018c] J. Vega, S. Dormido-Canto, A. Murari, F. Hernández, P. Lomas, T. Cruz, D. Gadariya, G. A. Rattá and JET Contributors, “A linear equation based on signal increments to predict disruptive behaviours and the time to disruption in JET”, EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK, TFM, June 28<sup>th</sup>, 2018.
- [Vega, 2019] J. Vega, F. Hernández, S. Dormido-Canto, A. Isayama, E. Joffrin, G. Matsunaga, T. Suzuki, “Assessment of linear disruption predictors using JT-60U data”, Fusion Engineering and Design, 146 (2019) 1291-1294, <https://doi.org/10.1016/j.fusengdes.2019.02.061> .
- [Vega, 2020] J. Vega, A. Murari, S. Dormido-Canto, F. Hernández, T. Cruz, D. Gadariya, G. Rattá and JET Contributors, “A linear equation based on signal increments to predict disruptive behaviours and the time to disruption on JET”, International Atomic Energy Agency, Nucl. Fusion 60 (2020) 026001 (13pp), <https://doi.org/10.1088/1741-4326/ab5880> .
- [Vega, 2020b] J. Vega, R. Castro, S. Dormido-Canto, G. A. Rattá, M. Ruiz. “Automatic recognition of plasma relevant events: Implications for ITER”. Fusion Engineering and Design 156 (2020). <https://doi.org/10.1016/j.fusenddes.2020.111638> .
- [Vesanto, 1999] J. Vesanto, “SOM-based data visualization method”, Intelligent Data Analysis 3 (1999) 111–126.
- [Vovk, 1999] Vovk V., Gammerman A. and Saunders C., 1999, “Machine learning applications of algorithmic randomness”, Proc. 16th Int. Conf. on Machine Learning (Bled, Slovenia), vol 2 (San Francisco, CA: Morgan Kaufmann) pp 444–53.
- [Vovk, 2005] Vovk V., Gammerman A. and Shafer G., 2005, “Algorithmic Learning in a Random World”, (Berlin: Springer).
- [de Vries, 2009] de Vries P.C., Johnson M.F. and Segui I. 2009, “Statistical analysis of disruptions in JET”, Nucl. Fusion 49 055011.
- [de Vries, 2011] P. C. de Vries, M. F. Johnson, B. Alper, P. Buratti, T. C. Hender, H. R. Koslowski, V. Riccardo and JET-EFDA Contributors. “Survey of disruption causes at JET”. Nucl. Fusion 51 (2011) 053018 (12pp). doi:10.1088/0029-5515/51/5/053018.
- [de Vries, 2012] De Vries P.C. et al, 2012, “The impact of the ITER-like wall at JET on disruptions”, Plasma Phys. Control. Fusion 54 124032–40.
- [de Vries, 2016] De Vries P.C. et al (Snipes and the COMPASS Team, the ASDEX Upgrade Team and JET Contributors), 2016, “Scaling of the MHD perturbation amplitude required to trigger a disruption and predictions for ITER”, Nucl. Fusion 56 026007.



## W

- [Weiss, 2003] Weiss G.M. and Provost F., 2003, "Learning when training data are costly: the effect of class distribution on tree induction", *J. Artif. Intell. Res.* 19 315–54.
- [Wesson, 1989] Wesson J.A. et al, 1989, "Disruptions in JET", *Nucl. Fusion* 29 641–66.
- [Wesson, 2004] Wesson J. 2004 *Tokamak* 4th edn (Oxford: Clarendon).
- [Windsor] C.G. Windsor, N.M. Ershow, A.M. Popov, T.N. Todd, "Disruption prediction in COMPASS-D using a neural network", unpublished report.
- [Windsor, 2005] C.G.Windsor, G. Pautasso, C. Tichmann, R.J. Buttery, T.C. Hender, JET EFDA Contributors and the ASDEX Upgrade Team. "A cross-tokamak neural network disruption predictor for the JET and ASDEX Upgrade tokamaks". *Nucl. Fusion* 45 (2005) 337–350. doi:10.1088/0029-5515/45/5/004.
- [Wroblewki, 1997] D. Wroblewki et al., *Nucl. Fus.* 37 (1997) 725.

## Y

- [Yoshino, 2003] R. Yoshino, "Neural-net disruption predictor in JT-60U", *Nucl. Fusion* 43 (2003) 1771–1786.
- [Yoshino, 2005] Yoshino R., 2005, "Neural-net predictor for beta limit disruptions in JT-60U", *Nucl. Fusion* 45 1232–46.

