

**DETECCIÓN, CLASIFICACIÓN Y LOCALIZACIÓN
DE EVENTOS DE CALIDAD DE ENERGÍA
UTILIZANDO TÉCNICAS AVANZADAS DE
PROCESAMIENTO DE SEÑALES E INTELIGENCIA
ARTIFICIAL**

por

David Marcelo de Yong

**Tesis presentada para la obtención del grado de Doctor en
Ciencias de la Ingeniería**

**FACULTAD DE INGENIERÍA
UNIVERSIDAD NACIONAL DE RÍO CUARTO**

Febrero del 2016

**DETECCIÓN, CLASIFICACIÓN Y LOCALIZACIÓN
DE EVENTOS DE CALIDAD DE ENERGÍA
UTILIZANDO TÉCNICAS AVANZADAS DE
PROCESAMIENTO DE SEÑALES E INTELIGENCIA
ARTIFICIAL**

por

David Marcelo De Yong

Dr. Fernando H. Magnago

Director

Dr. Claudio A. Reineri

Co-Director

Aprobada por:

Dr. Enrique Mombello

Jurado

Dr. Jorge C. Vaschetti

Jurado

Juan M. Fontana

Jurado

Mercedes Carnero

Directora de la Secretaría de Posgraduación de la FI-UNRC

Gabriela Maldonado

Directora de la Secretaría de Posgraduación y Cooperación Internacional - UNRC

Febrero del 2016

Un camino de mil millas comienza con un paso.

Benjamin Franklin.

A la memoria de mi padre y Fernando, por siempre presentes en mi corazón.

A mi patria, mi orgullo.

A mi amada esposa.

A mi madre y hermanos.

AGRADECIMIENTOS

En primer lugar quiero agradecer a mi amada y compañera de la vida, Paola Perotto Ghi. Sin su infinito amor, su inagotable paciencia y su apoyo diario me hubiera resultado imposible avanzar por este hermoso camino. Su optimismo ilumina cada uno de mis días.

Mi reconocimiento también para mi madre, cuya fortaleza inquebrantable resulta un ejemplo para mí. Ella me aconsejó desde joven a luchar por mis sueños y a ella le debo haber llegado hasta aquí.

Quiero compartir este importante momento con mis hermanos, Alejandro, Diego, Mauricio, Mariano, Matías y María Belén. Ellos son parte de lo que soy y lo que quiero ser.

A Pamela, Ivana, Sabrina, Micaela y Sol, mis queridas cuñadas.

A mi mentor, el Dr. Fernando Magnago, director de la presente tesis y compañero de muchas horas de trabajo. Su invaluable y desinteresado apoyo durante estos cinco años, su ejemplo y sus consejos sin dudas me hicieron crecer como profesional y como persona. A Fernando le debo la posibilidad de hoy encontrarme en esta instancia.

Al Dr. Claudio Reineri, co-director del presente trabajo de tesis. Sus recomendaciones resultó un aporte importante para culminar esta tesis.

Mi más sincero reconocimiento al personal docente y al personal administrativo de la Facultad de Ingeniería de la Universidad Nacional de Río Cuarto cuyo trabajo diario, sin dudas, hicieron y hacen posible la educación de calidad de la que sentimos orgullo.

Quiero agradecer a mi país que me brindó una educación de grado y postgrado. Ruego poder devolver a mi patria todo lo recibido. Lucharé para que otros, al igual que yo, el día de mañana, puedan sentirse realizados haciendo lo que más les gusta. En este sentido, quiero reconocer también el apoyo económico brindando por CONICET.

Quiero también agradecer a todos los integrantes del Grupo de Análisis de Sistemas Eléctricos de Potencia (GASEP), en especial al Dr. Luis Aromataris. De sus consejos he aprendido mucho en todo este tiempo.

Al Dr. Enrique Mombello, al Dr. Jorge Vaschetti y al Dr. Juan Manuel Fontana por su labor como miembros jurados de la comisión evaluadora de la presente tesis. Su importante aporte indudablemente concluye en una mejora de este trabajo.

A mis amigos y colegas de la Facultad de Ingeniería de Fabián Rinaudo, Juan Alemany, Carlos Massei, Ricardo Lima, Eduardo Toledo, Jorge Federico Aguirre, Santiago Lovera, Rodrigo Prat, Julián Durigutti y Diego Gagliesi.

Mi eterno reconocimiento a todos los que de una u otra manera han contribuido con esta tarea que encaré con entusiasmo, responsabilidad y pasión.

TABLA DE CONTENIDOS

AGRADECIMIENTOS	X
TABLA DE CONTENIDOS	XII
INDICE DE TABLAS	XIX
INDICE DE FIGURAS	XX
Introducción	26
I Situación Energética Nacional	28
II Incorporación de energías renovables y políticas de eficiencia energética	29
III Convergencia hacia las Redes Inteligentes de Energía Eléctrica	30
IV Calidad de Energía – Descripción de la problemática	31
V Calidad de Energía - Interés de la comunidad científica	32
VI Objetivos	34
VII Organización de la tesis	36
VIII Tabla de acrónimos	38
Capítulo 1: Clasificación de Perturbaciones y Sistemas de Monitoreo	40
1.1 <i>Clasificación de Eventos de Calidad de Energía</i>	42
1.1.1 Clasificación de acuerdo a la duración	42
1.1.2 Clasificación de acuerdo a la IEC	42
1.1.3 Clasificación de acuerdo a la IEEE	43
1.1.3.1 Transitorios	45
1.1.3.2 Variaciones de corta duración	46
1.1.3.3 Variaciones de larga duración	50
1.1.3.4 Desbalance de Tensión	50
1.1.3.5 Distorsión de la forma de onda	50
1.1.3.6 Fluctuaciones de tensión	53
1.1.3.7 Variaciones de la frecuencia fundamenta	54
1.1.4 Modelos empleados para la simulación de los eventos de calidad de energía	54
1.2 <i>Sistemas de Monitoreo de Calidad de Energía</i>	56
1.2.1 Necesidades de Monitoreo	56
1.2.2 Reseña Histórica	57
1.2.3 Clasificación de Sistemas de Monitoreo	57
1.2.4 Sistemas de Monitoreo de Calidad de Energía	58

1.3	<i>Proceso de Análisis de Datos</i>	61
1.3.1	Procesamiento de Señales.....	62
1.3.2	Detección.....	64
1.3.3	Extracción de Parámetros Característicos.....	65
1.3.4	Clasificación de Eventos de Calidad de Energía.....	66
1.3.5	Localización de Fallas en un Sistema Eléctrico.....	68
1.3.6	Caracterización y Presentación de Resultados.....	72
1.4	<i>Conclusiones</i>	73
	Primera Parte: Conceptos Teóricos	74
	Capítulo 2: Transformada Wavelet	76
2.1	<i>Introducción</i>	78
2.2	<i>Transformada Wavelet</i>	79
2.2.1	Conceptos Previos.....	79
2.2.1.1	Clasificación de Señales.....	79
2.2.1.2	Transformada de Fourier.....	80
2.2.1.3	Transformada de Fourier dependiente del tiempo.....	81
2.2.2	La Transformada Wavelet.....	84
2.2.2.1	Wavelets Madre.....	84
2.2.2.2	Expansión y Traslación de una Wavelet Madre.....	87
2.2.2.3	La Transformada Wavelet de Tiempo Continuo.....	88
2.2.2.4	La Transformada Wavelet Discreta – DWT.....	89
2.2.2.5	Interpretación de la DWT como un banco de Filtros.....	90
2.2.2.6	Teorema de Parseval.....	93
2.3	<i>La Transformada Wavelet para la Detección de Perturbaciones</i>	94
2.4	<i>Conclusiones</i>	96
	Capítulo 3: Máquinas de Vectores de Soporte	98
3.1	<i>Introducción</i>	100
3.2	<i>Teoría del Aprendizaje Estadístico</i>	102
3.2.1	Aprendizaje Supervisado.....	102
3.2.2	Limitaciones en el Aprendizaje.....	103
3.2.3	Proceso de Aprendizaje para las Máquinas de Vectores de Soporte.....	103
3.3	<i>Teoría de la Optimización Matemática</i>	104

3.3.1	Conceptos.....	104
3.3.2	Formulación del Problema.....	104
3.3.3	Definiciones Importantes.....	105
3.3.3.1	Programación Lineal.....	105
3.3.3.2	Programación Cuadrática.....	105
3.3.3.3	Función Cóncava y Convexa.....	105
3.3.3.4	Transformación de las Restricciones.....	105
3.3.4	Teoría de Lagrange.....	106
3.3.5	Dualidad.....	108
3.4	<i>Máquinas de Vectores de Soporte</i>	109
3.4.1	Clasificador de Vectores de Soporte.....	109
3.4.2	Clasificador de Máximo Margen.....	109
3.4.2.1	Clasificador para datos linealmente separables.....	110
3.4.2.2	Límites de decisión.....	111
3.4.2.3	Optimización mediante multiplicadores de Lagrange.....	115
3.4.2.4	Clasificación para datos no linealmente separables.....	118
3.4.3	Transformación de sub espacio usando funciones kernel.....	120
3.4.3.1	Funciones kernel.....	122
3.4.3.2	Incorporación de las funciones kernel al clasificador SVM.....	123
3.4.4	Máquinas de Vectores de Soporte multiclase.....	124
3.4.4.1	Uno contra el Resto (One vs Rest)	124
3.4.4.2	Uno contra Uno (One vs One)	125
3.4.4.3	Árbol de jerarquía de clasificadores.....	125
3.5	<i>Aplicación de Máquinas de Vectores de Soporte en estudios eléctricos</i>	127
3.5.1	Clasificación de perturbaciones.....	127
3.5.2	Localización de fallas.....	129
3.6	<i>Conclusiones</i>	131
Segunda Parte: Desarrollo de la Propuesta		132
Capítulo 4: Desarrollo de un clasificador basado en Inteligencia Artificial		134
4.1	<i>Introducción</i>	136
4.1.1	Objetivos del capítulo.....	137
4.2	<i>Algoritmo de Diseño y Entrenamiento del Clasificador</i>	138

4.2.1	Procesamiento de Señales.....	139
4.2.1.1	Técnicas de pre procesamiento de señales.....	139
4.2.1.2	Técnicas de procesamiento de señales.....	140
4.2.2	Extracción de parámetros característicos.....	140
4.2.2.1	Cálculo de parámetros característicos.....	140
4.2.2.2	Normalización de parámetros característicos.....	141
4.2.3	Minería de datos.....	141
4.2.3.1	Generación de Etiquetas.....	142
4.2.3.2	Filtrado Heurístico.....	142
4.2.3.3	Algoritmo de Búsqueda Exhaustiva.....	144
4.2.4	Optimización de los parámetros del clasificador.....	145
4.2.5	Entrenamiento de las etapas del clasificador.....	148
4.3	<i>Algoritmo de Clasificación.....</i>	149
4.3.1	Procesamiento de Señales.....	149
4.3.2	Detección de Eventos.....	150
4.3.3	Extracción de parámetros característicos.....	150
4.3.4	Clasificación.....	150
4.4	<i>Diagramas de Flujo.....</i>	151
4.5	<i>Conclusiones.....</i>	154
	Capítulo 5: Software Educativo para el Análisis de Eventos de Calidad de Energía.....	156
5.1	<i>Introducción.....</i>	158
5.2	<i>Software para el Análisis de Eventos de Calidad de Energía.....</i>	160
5.2.1	Acceso a la herramienta.....	160
5.2.2	Generación de señales de Entrenamiento y Prueba.....	162
5.2.3	Extracción de parámetros característicos de las señales de entrenamiento.....	164
5.2.4	Entrenamiento y Prueba del clasificador.....	165
5.2.4.1	Red Neuronal Backpropagation.....	165
5.2.4.2	Máquinas de Vectores de Soporte.....	167
5.2.5	Incorporación de señales reales.....	168
5.2.6	Prueba del clasificador.....	168
5.3	<i>Conclusión.....</i>	170
	Capítulo 6: Clasificación de Perturbaciones Complejas.....	172

6.1	<i>Introducción</i>	174
6.1.1	Objetivos.....	179
6.1.2	Metodología.....	179
6.2	<i>Entrenamiento y Diseño del Sistema Clasificador</i>	180
6.2.1	Arquitectura del Clasificador.....	180
6.2.2	Configuración del set de Entrenamiento.....	181
6.2.3	Resultados: Diseño y Entrenamiento del clasificador.....	182
6.2.3.1	Procesamiento de Señales.....	182
6.2.3.2	Extracción de Parámetros Característicos.....	188
6.2.3.3	Minería de Datos.....	190
6.2.3.4	Optimización de parámetros del clasificador.....	197
6.2.4	Entrenamiento.....	198
6.2.4.1	Resultados: Prueba de rendimiento.....	198
6.2.4.2	Prueba 1- Perturbaciones simples.....	199
6.2.4.3	Prueba 2 – Perturbaciones complejas.....	199
6.2.4.4	Comparación de resultados.....	202
6.3	<i>Conclusiones</i>	202
	Capítulo 7: Localización de Fallas	204
7.1	<i>Introducción</i>	206
7.1.1	Objetivos.....	208
7.1.2	Metodología.....	209
7.2	<i>Algoritmo para la Localización e Identificación del Tipo de Falla</i>	210
7.2.1	Modelo del Sistema Eléctrico de Potencia.....	210
7.2.2	Generación del set de Entrenamiento y Prueba.....	212
7.2.3	Diseño y Entrenamiento del Clasificador.....	213
7.2.3.1	Procesamiento de Señales.....	214
7.2.3.2	Extracción de Parámetros Característicos.....	215
7.2.3.3	Minería de Datos.....	216
7.2.3.4	Optimización de parámetros del clasificador.....	219
7.2.4	Clasificador.....	220
7.2.4.1	Procesamiento de Señales.....	220
7.2.4.2	Detección de Fallas.....	220

7.2.4.3	Clasificación.....	220
7.2.4.4	Limitaciones Observadas.....	221
7.3	<i>Conclusiones</i>	222
Capítulo 8: Conclusión, Aportes y Trabajos Futuros		224
8.1	<i>Resumen del contenido de los capítulos de la tesis</i>	226
8.2	<i>Resumen de las motivaciones</i>	230
8.3	<i>Aportes realizados</i>	232
8.3.3.1	Desarrollo del Software de Monitoreo.....	232
8.3.3.2	Metodología para el diseño de un clasificador basado en técnicas de AI.....	232
8.3.3.3	Implementación de un clasificador de perturbaciones simples y complejas.....	233
8.3.3.4	Implementación de un localizador de fallas en un sistema de distribución.....	233
8.4	<i>Herramientas utilizadas</i>	236
8.5	<i>Trabajos futuros</i>	237
8.6	<i>Publicaciones realizadas</i>	239
8.6.1	Revistas Indexadas.....	239
8.6.2	Conferencias.....	240
8.6.3	Trabajos de tesis.....	240
Referencias		241

INDICE DE TABLAS

Capítulo 1 - Clasificación de Perturbaciones y Sistemas de Monitoreo

Tabla 1. 1 - Principales fenómenos que causan perturbaciones electromagnéticas.....	43
Tabla 1. 2 - Categorías y Características de fenómenos electromagnéticos en un SEP.....	44
Tabla 1. 3 - Modelo de perturbaciones.	55

Capítulo 4 - Algoritmo de Entrenamiento y Prueba de un clasificador basado en Máquinas de Vectores de Soporte

Tabla 4. 1 - Selección de parámetros característicos.....	145
Tabla 4. 2 - Algoritmo de búsqueda tipo grilla.....	147

Capítulo 6 - Clasificación de perturbaciones complejas

Tabla 6. 1 - Set de perturbaciones simples.....	182
Tabla 6. 2 - Set de perturbaciones complejas.....	182
Tabla 6. 3 - Parámetros característicos.....	188
Tabla 6. 4 - . Asignación de etiquetas para perturbaciones simples.....	191
Tabla 6. 5 - Asignación de etiquetas para perturbaciones complejas.....	191
Tabla 6. 6 - Resultados del filtrado heurístico de parámetros característicos.....	191
Tabla 6. 7 - Combinación de máxima precisión - Huecos.....	193
Tabla 6. 8 - Combinación de máxima precisión – Sobretensión.....	194
Tabla 6. 9 - Combinación de máxima precisión - Armónicos.....	195
Tabla 6. 10 - Combinación de máxima precisión – Interrupción.....	196
Tabla 6. 11 - Combinación de máxima precisión – Fluctuaciones.....	197
Tabla 6. 12 - Resultado del algoritmo de búsqueda exhaustiva.....	197
Tabla 6. 13 - Resultados del algoritmo de búsqueda de grilla.....	198
Tabla 6. 14 - Matriz de confusión - Perturbaciones simples.....	199
Tabla 6. 15 - Matriz de confusión - Perturbaciones complejas.....	200
Tabla 6. 16 - Comparación de resultados – Perturbaciones complejas.....	201

Capítulo 7 - Localización de Fallas

Tabla 7. 1 - Parámetros de las líneas.....	212
Tabla 7. 2 - Longitud de las líneas y potencia de las cargas.....	212
Tabla 7. 3 - Tipos de falla.....	213
Tabla 7. 4 - Localización de la falla.....	213
Tabla 7. 5 - Parámetros característicos.....	216
Tabla 7. 6 - Resultado de filtrado heurístico.....	217
Tabla 7. 7 - Resultado del algoritmo de búsqueda exhaustiva - Tipo de falla.....	217
Tabla 7. 8 - Resultado del algoritmo de búsqueda exhaustiva - Localización de falla.....	219
Tabla 7. 9 - Resultados del algoritmo de búsqueda tipo grilla.....	219

INDICE DE FIGURAS

Introducción

Figura I. 1 - Escenario futuro de las energías renovables propuesto por CADER.....	29
Figura I. 2 - Publicaciones por año en el área de Localización de Fallas.....	33
Figura I. 3 - Publicaciones por año en el área de Calidad de Energía.....	34

Capítulo 1 - Clasificación de Perturbaciones y Sistemas de Monitoreo

Figura 1. 1 - Transitorio Oscilatorio de alta frecuencia.....	46
Figura 1. 2 - Transitorio Oscilatorio de baja frecuencia.....	46
Figura 1. 3 - Interrupción ideal.....	47
Figura 1. 4 - Variación de la tensión instantánea y valor rms ante una interrupción.....	48
Figura 1. 5 - Tensión instantánea y perfil rms de una forma de onda ante hueco de tensión.....	49
Figura 1. 6 - Tensión instantánea y perfil rms de una forma de onda ante una sobretensión.....	49
Figura 1. 7 - Tensión instantánea y espectro de una señal con conteniendo armónicos.....	51
Figura 1. 8 - Tensión instantánea de un sistema ante una muesca.....	52
Figura 1. 9 - Tensión instantánea del sistema contaminada con ruido.....	53
Figura 1. 10 - Tensión instantánea del sistema ante fluctuaciones.....	54
Figura 1. 11 - Sistema de monitoreo.....	59
Figura 1. 12 - Análisis de las muestras.....	61
Figura 1. 13 - Taxonomía de las técnicas de procesamiento de señales.....	62
Figura 1. 14 - Taxonomía de los clasificadores basados en inteligencia artificial.....	66
Figura 1. 15 - Clasificación de sistemas de localización de fallas.....	70

Capítulo 2 - Transformada Wavelet

Figura 2. 1 - Señal estacionaria.....	79
Figura 2. 2 - Señal no estacionaria.....	79
Figura 2. 3 - Señal estacionaria.....	81
Figura 2. 4 - Señal no estacionaria.....	81
Figura 2. 5 - Transformación de dominios de la STFT.....	83
Figura 2. 6 - a) Wavelet Haar b) Wavelet Sombrero Mexicano.....	85
Figura 2. 7 - Wavelet Gaussiana Compleja b) Wavelet Shannon Compleja.....	85
Figura 2. 8 - a) Wavelet Meyer b) Wavelet Morlet.....	86
Figura 2. 9 - Wavelet madre Daubechie 4.....	86
Figura 2. 10 - Traslación de la wavelet madre.....	87
Figura 2. 11 - Escalamiento de la wavelet madre.....	87
Figura 2. 12 - Plano escala - tiempo de la transformada wavelet.....	89
Figura 2. 13 - Aproximación mediante banco de filtros.....	90
Figura 2. 14 - Banco de filtros y sub muestreo.....	91
Figura 2. 15 - Coeficientes A y D de una señal senoidal más ruido blanco.....	91
Figura 2. 16 - Árbol de descomposición.....	92
Figura 2. 17 - Árbol de descomposición para una perturbación tipo muesca.....	92

Capítulo 3 - Máquina de Vectores de Soporte

Figura 3. 1 - Fronteras de decisión.....	110
Figura 3. 2 - Caso linealmente separable.....	110
Figura 3. 3 - Límite de decisión.....	111
Figura 3. 4 - Separación entre los hiperplanos H_0 y H_1	113
Figura 3. 5 - Vectores de soporte.....	115

Figura 3. 6 - Hiperplano para el caso no linealmente separable.	120
Figura 3. 7 - Transformación del espacio de características usando funciones Kernel.	122
Figura 3. 8 - Arbol de decisión binaria.	126

Capítulo 4 - Algoritmo de Entrenamiento y Prueba de un clasificador basado en Máquinas de Vectores de Soporte

Figura 4. 1 - Sistema de clasificación.	136
Figura 4. 2 - Algoritmo de diseño y entrenamiento.	138
Figura 4. 3 - Minería de datos.	142
Figura 4. 4 - Algoritmo de clasificación.	149
Figura 4. 5 - Diagrama de flujo del algoritmo de entrenamiento del clasificador.	152
Figura 4. 6 - Algoritmo de clasificación.	153

Capítulo 5 - Software para el Análisis de Eventos de Calidad de Energía

Figura 5. 1 - Acceso a la herramienta.	161
Figura 5. 2 - Panel principal.	161
Figura 5. 3 - Generador de señales de prueba.	162
Figura 5. 4 - Generador automático de perturbaciones simples.	163
Figura 5. 5 - Generador automático de perturbaciones complejas.	163
Figura 5. 6 - Panel de generación del vector característico.	164
Figura 5. 7 - Clasificador basado en una Red Neuronal Backpropagation.	166
Figura 5. 8 - Clasificador basado en Máquinas de Vectores de Soporte.	167
Figura 5. 9 - Hueco de tensión adquirido de la red.	168
Figura 5. 10 - Localización y clasificación de una señal real.	169

Capítulo 6 - Clasificación de Perturbaciones Complejas

Figura 6. 1 - Armónicos y hueco.	174
Figura 6. 2 - Armónicos e interrupción.	175
Figura 6. 3 - Fluctuaciones e interrupción.	175
Figura 6. 4 - Hueco y sobre tensión.	176
Figura 6. 5 - Armónicos, interrupción y sobre tensión.	176
Figura 6. 6 - Arquitectura uno contra el resto.	181
Figura 6. 7 - Algoritmo de diseño y entrenamiento.	182
Figura 6. 8 - Perfil rms.	184
Figura 6. 9 - Perturbación compleja - Armónicos y sobre tensión.	185
Figura 6. 10 - Coeficientes de detalle ($D_1 - D_4$).	186
Figura 6. 11 - Coeficientes de detalle ($D_5 - D_8$).	187
Figura 6. 12 - Coeficiente de detalle D_9 y de aproximación A_9	187
Figura 6. 13 - Resultados de la normalización.	190
Figura 6. 14 - Resultado del algoritmo de búsqueda exhaustiva - SVM Huecos.	192
Figura 6. 15 - Resultado del algoritmo de búsqueda exhaustiva - SVM sobre tensión.	193
Figura 6. 16 - Resultado del algoritmo de búsqueda exhaustiva - Armónicos.	194
Figura 6. 17 - Resultado del algoritmo de búsqueda exhaustiva - Interrupción.	195
Figura 6. 18 - Resultado del algoritmo de búsqueda exhaustiva - Fluctuaciones.	196
Figura 6. 19 - Algoritmo de clasificación.	198

Capítulo 7 - Localización de Fallas

Figura 7. 1 - Diseño, Entrenamiento y Prueba del sistema de localización de fallas.....	210
Figura 7. 2 - Sistema eléctrico de potencia de entrenamiento y prueba.....	211
Figura 7. 3 - Falla en línea 3 - Medición en barra 1.	214
Figura 7. 4 - Resultado del algoritmo de búsqueda exhaustiva – Tipo de falla.	218
Figura 7. 5 - Resultado del algoritmo de búsqueda exhaustiva – Localización de falla.	218
Figura 7. 6 - Perfil rms de tensión.	220

RESUMEN

La convergencia de la red eléctrica tradicional con los sistemas de generación distribuida, las fuentes de energía renovable y la incorporación de dispositivos electrónicos no lineales, generan una gran cantidad de fenómenos electromagnéticos que provocan un deterioro en la calidad de la energía consumida por el usuario.

La desregulación de los mercados y el hecho que los usuarios residenciales e industriales se hayan vuelto más exigentes en lo referente a la calidad de la energía que consumen genera marcadas presiones a todos los actores del mercado para brindar una energía de mayor calidad.

Por este motivo resulta imprescindible que las empresas eléctricas cuenten con herramientas que permitan el monitoreo del estado de la red en tiempo real para gestionar adecuadamente tareas de mantenimiento preventivo y reactivo de la red y, de este modo, mejorar los índices que miden la calidad de la energía.

En el futuro, las empresas prestadoras de energía eléctrica se verán obligadas a incurrir en importantes gastos para implementar dispositivos de medición, redes de comunicación para centralizar la información y algoritmos de procesamiento que permitan extraer información importante de los datos colectados.

Sin importar los objetivos técnicos involucrados en la implementación de un sistema de monitoreo el principal motivo que fundamenta su implementación radica en la necesidad, de mitigar las pérdidas económicas generadas por las perturbaciones en suministro eléctrico.

Un algoritmo de clasificación de perturbaciones y localización de fallas tiene por objetivo principal encontrar la fuente de la perturbación para establecer diagnósticos de la salud de la red y, en caso de ser necesario, rápidamente restablecer el servicio.

En este contexto, en la presente tesis se presentan algoritmos innovadores de procesamiento de mediciones obtenidas en un sistema eléctrico con el objetivo de detectar, caracterizar, clasificar y localizar perturbaciones en un sistema eléctrico de potencia, empleando la menor cantidad posible de información.

La mayoría de los algoritmos propuestos por la comunidad científica se enfocan a detectar y clasificar una perturbación, dentro de un conjunto de muestras, evitando abordar casos que comúnmente se presentan en la realidad tal como las perturbaciones complejas.

Por otro lado, la mayoría de las publicaciones relativas a la localización de fallas se enfocan en sistemas eléctricos de transmisión siendo muy escasos los trabajos que abordan la problemática desde el punto de vista de un sistema de distribución.

En la presente tesis se presenta el desarrollo de diversos algoritmos cuyo objetivo radica en la detección y clasificación de perturbaciones simples y complejas así como también en la localización de fallas en sistemas eléctricos de potencia.

Los desarrollos realizados están fundamentados en avanzadas técnicas de procesamiento de señales, diversas estrategias para la extracción de parámetros característicos, métodos de minería de datos para la selección de parámetros característicos y técnicas innovadoras de *Inteligencia Artificial* para la clasificación y localización de eventos de calidad de energía.

Los métodos desarrollados presentan destacados resultados en comparación con las publicaciones que abordan la problemática propuesta en la presente tesis. Por este motivo, los algoritmos desarrollados se enfocarán en la detección y clasificación de perturbaciones complejas y en la localización de fallas en sistemas eléctricos de distribución.

ABSTRACT

The combination of the traditional power grid with distributed generation systems, renewable energy sources and incorporating non-linear, electronic devices generate several electromagnetic phenomena that cause a deterioration in the parameters that measure the quality of energy consumed by the users.

On the other hand, residential and industrial users have become more demanding regarding the quality of the energy they consume.

For this reason, it is essential that utilities have tools to monitor the status of the electrical network in real time to manage properly different tasks such as preventive and reactive maintenance to improve thus the characteristic indices that measure the power quality.

In the future, electrical utilities will be forced to incur significant costs to implement measurement devices, communication networks to centralize information, and processing algorithms, which are designed to extract relevant information from the collected data.

An algorithm dedicated to classifying power quality events and localize faults has as the primary task to find the source of the disturbance as fast as possible to pursue the service restoration rapidly.

In this context, this thesis presents a novel algorithm to detect, characterize, classify and locate disturbances in a power system, using the fewest possible measurement devices.

Most of the algorithms proposed by the scientific community are focused on detecting and classifying a disturbance avoiding addressing some particular situation present in real life scenarios known as complex disturbances.

Moreover, most publications related to fault location are focused on transmission systems and not on radial distribution systems. This thesis presents various algorithms developed to detect and classify simple and complex disturbances and to locate faults in radial distribution systems.

The proposed developments are based on advanced signal processing, different strategies for extracting characteristic parameters, and innovative artificial intelligence techniques to classify the various types of disturbances and their geographical location within the grid.

The developed methods have outstanding results compared to publications that address the proposed problem in this thesis.

Introducción

I. Situación Energética Nacional.

En los últimos cinco años el sector energético ha se ha vuelto un área de incesantes estudios debido a la creciente demanda, la situación económica del país, el contexto energético mundial y las políticas llevadas a cabo por el gobierno nacional [1].

El balance energético nacional en el último tiempo se ha visto afectado debido a las siguientes situaciones:

- El país ha necesitado aumentar las importaciones de combustibles, que alcanzó un máximo en el año 2013, llegando a 13.000 millones de dólares.
- La caída del precio internacional del barril de petróleo, observada desde mediados del 2014, tuvo un impacto en dos direcciones, por un lado afecta negativamente los planes de extracción no convencional por parte de YPF pero, por otro lado, redundó en una reducción de la presión de las importaciones energéticas debido a la disminución de los costos del gas natural y combustibles líquidos que el país necesita adquirir.
- Los cortes de energía eléctrica producidos en Capital Federal que, en muchos casos, se prolongan por varios días están generando manifestaciones de rechazo de los clientes que derivó, por parte del ministro de Ambiente y Desarrollo Sustentable, en la declaración de la emergencia energética nacional el 14 de diciembre del 2015. Estos hechos indican, por un lado una creciente concientización de la población acerca de los problemas energéticos que se deben afrontar y, por otro, desnuda indudables falencias del sistema energético nacional en lo relativo a generación, transporte y distribución de electricidad.
- Continuó la expansión del parque de generación térmico y se incorporó al mercado eléctrico una nueva central nuclear.
- En cuanto a la visión futura, la mayoría de los analistas de la coyuntura energética nacional concluyen que será necesario al menos duplicar la potencia instalada con miras al 2035 llegando, en algunos casos, a la conclusión que puede ser necesario incluso triplicarla. Los niveles de inversión necesarios para lograr estos objetivos son extremadamente altos y será difícil llevarlos a cabo en los tiempos requeridos.
- Las exigencias ambientales también juegan un rol fundamental a la hora de delinear políticas energéticas.

Según los analistas, para poder afrontar la problemática mencionada, la Argentina deberá incorporar, en el corto plazo, fuentes de energía renovable en proporciones considerables y alentar políticas de eficiencia energética que resultan en el ahorro de la energía consumida.

II. Incorporación de energías renovables y políticas de eficiencia energética.

La mayoría de los expertos afirman que incorporar energías renovables a la matriz energética es un factor de vital importancia para alcanzar la potencia que los usuarios demandarán en el futuro. Por ejemplo, un escenario futuro planteado por la Cámara Argentina de Energías Renovables (CADER), como se puede ver en la Figura I. 1, supone un incremento en la potencia instalada de energías renovables hasta alcanzar una participación de un 25% de la matriz energética para el año 2030.

En la gráfica también se observa un fuerte crecimiento de la energía solar y eólica producto de la disponibilidad de estos recursos en el país, su bajísimo impacto ambiental así como el alto potencial de reducción de costos y mejora tecnológica en el corto y mediano plazo.

En este sentido, con un amplio consenso parlamentario, y a casi una década de su precursora la ley 26.190, el Congreso Nacional Argentino aprobó la ley N 27.191 (B.O. 21.10.15). “Régimen de Fomento Nacional para el Uso de Fuentes Renovables de Energía Destinada a la Producción de la Energía Eléctrica”.

Su precursora proponía un escenario a 2016 con un 8% de la matriz energética cubierto por fuentes renovables. Sin embargo, en la actualidad, en vistas que no se alcanzó cubrir ni el 1% de la energía generada la nueva ley postergó esta meta para el 2017.

Asimismo, en una segunda etapa se fijó como objetivo lograr una participación de las energías renovables del 25% para el año 2030 en la generación eléctrica nacional.

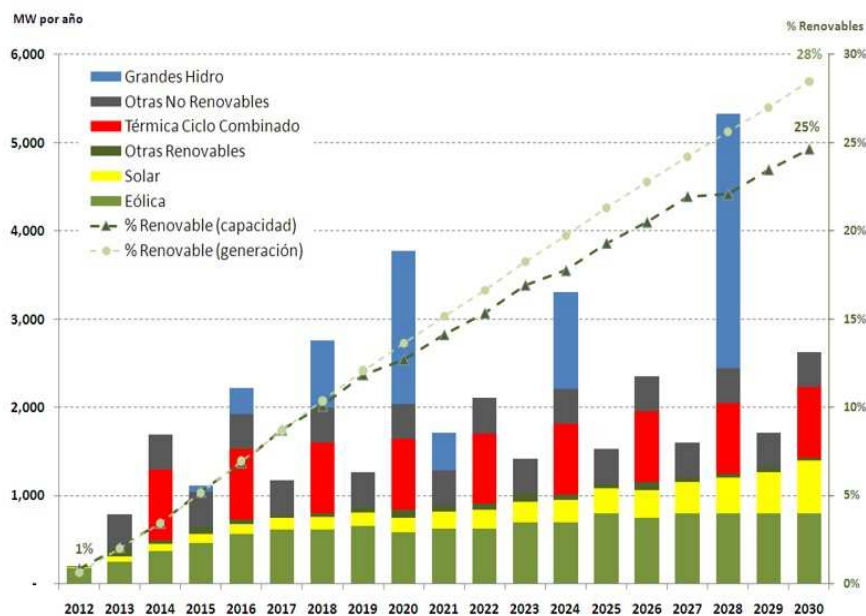


Figura I. 1 - Escenario futuro de las energías renovables propuesto por CADER.

Una segunda alternativa para mejorar la capacidad de afrontar la demanda futura es aplicar políticas de *eficiencia energética*.

La *eficiencia energética* se entiende a toda iniciativa para reducir el consumo de energía de cualquier dispositivo, sistema o infraestructura. Es por ello que a estas iniciativas comúnmente se las conoce como una fuente equivalente de energía o fuente oculta de combustible (en inglés, *hidden fuel*).

De acuerdo a la *Agencia Internacional de Energía (IEA)* la escala de las inversiones mundiales para los desarrollos relacionados con la eficiencia energética, conjuntamente con su contribución al en la demanda energética han alcanzado un nivel que las equipara con aquellas realizadas para el desarrollo de energías fósiles y energías renovables [2].

El ahorro energético obtenido por los países que adoptaron políticas de eficiencia energética son tan masivos que han llevado a re plantear la definición primera que comparaba al mercado de eficiencia con una fuente oculta de combustible erigiéndola al estatus de la fuente más importante de energía del mundo .El impacto acumulativo de la implementación de políticas de eficiencia energética es enorme. Según la IEA, desde el 2005 al 2010, gracias a estas prácticas, la energía ahorrada sus once miembros, solamente en petróleo, equivalen a 420 billones de dólares.

De no ser por las medidas de eficiencia energética implementada en los países que adhirieron a estas políticas, en los últimos años se hubieran consumido, y pagado, cerca de dos tercios más de lo que actualmente utilizan. Para tomar dimensión del ahorro logrado, los países que vienen adoptado estas medidas, desde 1974 a la fecha, evitaron quemar 1.7 billones de toneladas equivalentes petróleo, mayor que la demanda combinada de *Estados Unidos y Alemania* en 2012.

Por lo tanto, debido a las políticas de las principales potencias en el área (estándares energéticos, acceso a financiamiento privado y estatal y obligaciones reglamentadas a las prestadoras), la creciente presión de organismos ecológicos, precio del combustible y el aumento de la demanda, el área de la eficiencia energética continuará siendo un área de sumo interés para las economías del mundo.

III. Convergencia hacia las Redes Inteligentes de Energía Eléctrica.

La incorporación de las fuentes de energía renovable y la necesidad de aplicar políticas de eficiencia energética hace que las redes eléctricas deban evolucionar hacia lo que se conoce como redes inteligentes de energía eléctrica (*Smart Grids*).

Si bien no existe una definición precisa del concepto *Smart Grids*, el departamento de energía eléctrica de los Estados Unidos (DOE) sugiere que una red inteligente debe ser capaz de soportar un flujo bidireccional de energía e información, cuya infraestructura de comunicación puede ser empleada para obtener abundante información del estado de los equipos de la red e incluso coleccionar datos sobre las preferencias de consumo del usuario.

Muchos países se encuentran en una etapa de despliegue de infraestructura de soporte para las redes eléctricas del futuro. Se estima que para el año 2020, por ejemplo, cerca de 800 millones de estos dispositivos de medición inteligente (*AMI, Advances Metering*

Infraestructure) serán instalados a nivel global, siendo China el país que más rápidamente aplicará estas tecnologías instalado entre 380 y 500 millones de unidades [3].

La convergencia hacia las redes eléctricas inteligentes proveerá de innumerables beneficios a los actores del mercado energético, sin embargo, también representará un desafío importante ya que sin dudas existirá un impacto en las estrategias destinadas a mantener calidad de energía.

La convergencia hacia redes inteligencias tendrá incidencia en la calidad de energía proporcionada al usuario ya que:

- Las maniobras realizadas para incorporar fuentes de energía renovable y generación distribuida a la red eléctrica generarán alteraciones en su normal operación. Las fuentes de energía renovables, al ser de naturaleza impredecible sin dudas obligarán a las compañías a adoptar mayores prevenciones para sostener una calidad energética aceptable.
- La cantidad de nodos de medición interconectados a la red crecerá notoriamente. Esto permitirá a las compañías implementar tecnologías que les permitan tener mayor visibilidad del estado de la infraestructura eléctrica.
- Los algoritmos de detección, clasificación y localización de perturbaciones deberán adaptarse a las nuevas topologías de la red permitiendo el procesamiento de enormes cantidades de información.

IV. Calidad de Energía – Descripción de la problemática.

En la comunidad científica no existe una unicidad de criterios a la hora de dar una definición de lo que el término *calidad de energía* abarca.

Calidad de energía es un concepto que puede diferir de acuerdo al actor que lo formule (cliente, empresa prestadora, ente gubernamental). La definición que más se aproxima a los estudios presentados en el presente trabajo de tesis es aquella que se enfoca en la perspectiva del usuario:

Calidad de energía se refiere a un concepto de alimentación y puesta a tierra de equipamiento sensible en términos si los mismos son adecuados para el funcionamiento de dicho equipamiento [4].

Para la *Comisión Internacional de Electrotecnia (IEC)* calidad de energía se define como: *Las características de la electricidad en un determinado punto del sistema eléctrico.*

La primera de las definiciones aquí presentadas se enfoca en la energía medida en los equipos sensibles a perturbaciones electromagnéticas. Esta definición está más vinculada a la percepción de la calidad de energía por parte del usuario.

La segunda se basa en la medición de perturbaciones en diferentes nodos del sistema eléctrico. En este caso, la definición se orienta a la perspectiva de un operador de red.

Para el desarrollo que en esta tesis se presenta la, definición más apropiada es aquella que se enfoca en los eventos de calidad de energía.

Se entiende por eventos de calidad de energía, tanto en las formas de onda de tensión como las de corriente, a toda desviación en los parámetros que gobiernan a las estas señales tomando como referencia a una señal sinusoidal ideal [5].

Una señal de tensión ideal se entiende como una forma de onda sinusoidal cuya frecuencia y magnitud son constantes y se corresponden con sus valores nominales. De la misma manera, una forma de onda ideal de corriente también se considera como una señal sinusoidal cuyos parámetros de frecuencia y magnitud se corresponden con sus valores nominales, son constantes y, adicionalmente, la frecuencia y fase de la corriente se deben corresponder con la frecuencia y fase de la señal de tensión. Cualquier desviación tanto de corriente o de tensión de la forma de onda ideal se conoce como *evento de calidad de energía*.

V. Calidad de Energía - Interés de la comunidad científica.

En los últimos años la comunidad científica involucrada en el estudio de los sistemas eléctricos de potencia ha demostrado un creciente interés en la investigación de los fenómenos de calidad de energía.

Si bien existen muchas razones que justifican este fenómeno, aquí solo se detallan algunas de las más importantes.

Las exigencias de los usuarios en lo relativo a la calidad de la energía eléctrica hacen que los esfuerzos de las operadoras de red se deban enfocar, con mayor compromiso, en detectar las causas que lo producen y amortiguar sus efectos.

Los problemas de calidad de energía se pueden presentar como: (i) distorsiones en las formas de onda de tensión y corriente o (ii) interrupción temporal en la provisión del servicio. Ambos casos provocan pérdidas económicas tanto para el usuario como para las empresas prestadoras.

Los procesos productivos, en la mayoría de las industrias, presentan una dependencia vital de los equipamientos alimentados eléctricamente. Dichos equipamientos, a su vez, se han vuelto día a día más sensibles a las perturbaciones electromagnéticas. Para agravar la situación, las industrias se han vuelto menos tolerantes a la detención de sus líneas de producción por lo que el interés de las compañías en contar con una energía libre de perturbaciones se ha incrementado drásticamente en las últimas décadas.

A pesar de ser más sensibles a los eventos de calidad de energía, estos equipos son uno de los principales generadores de perturbaciones en la forma de onda de la corriente. El incremento en el empleo de convertidores electrónicos de energía eléctrica agrava el problema de las distorsiones generando un amplio espectro de perturbaciones en un SEP.

La desregulación de los mercados de la energía eléctrica ha generado un incremento en la necesidad de contar con indicadores de calidad de energía. Los clientes demandan mucha más información acerca de la calidad de energía que lo que solían hacer tiempo atrás.

La incorporación de equipos de consumo eficiente se ha transformado en una fuente importante de problemas de calidad de energía. Dispositivos de velocidad ajustable y lámparas de bajo consumo son ejemplos de equipos que introducen distorsiones en las formas de onda de la corriente y, a su vez, son dispositivos altamente sensibles a dichas perturbaciones.

La incorporación de la generación distribuida y fuentes energía renovable a la red eléctrica, una realidad cada día más palpable, ha creado también nuevos desafíos en lo relativo a la calidad de la energía [6].

Esta problemática ha provocado un creciente interés de los investigadores para desarrollar herramientas que permitan mejorar la calidad de energía y la localización de fallas.

En las Figura I. 2 y Figura I. 3 donde se muestra la cantidad de artículos publicados en la materia de acuerdo a la base de datos *Scopus* [7]. Estas gráficas se obtuvieron mediante la consulta empleando las palabras clave: “*Fault Location*” y “*Power Quality*”.

Como se puede observar en ambas gráficas el crecimiento de los artículos en ambas temáticas presenta una creciente tendencia en los últimos 25 años. Es de esperar, a su vez, que esta tendencia continúe en alza producto de la migración hacia una red inteligente y la convergencia con las fuentes de energía renovable y distribuida.

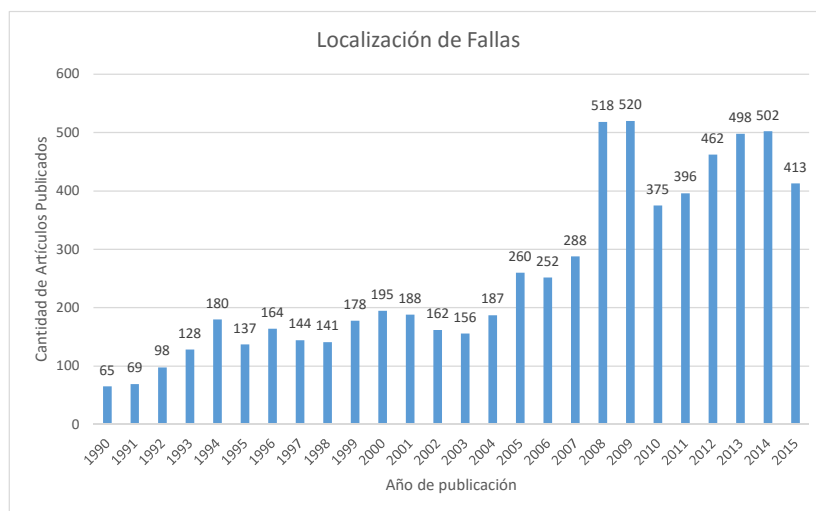


Figura I. 2 - Publicaciones por año en el área de Localización de Fallas.

A partir de lo anterior, una cosa está clara: el desarrollo aplicado a la optimización del uso de recursos renovables, la eficiencia energética y la calidad de energía continuarán jugando un papel clave en la política energética de cualquier país.

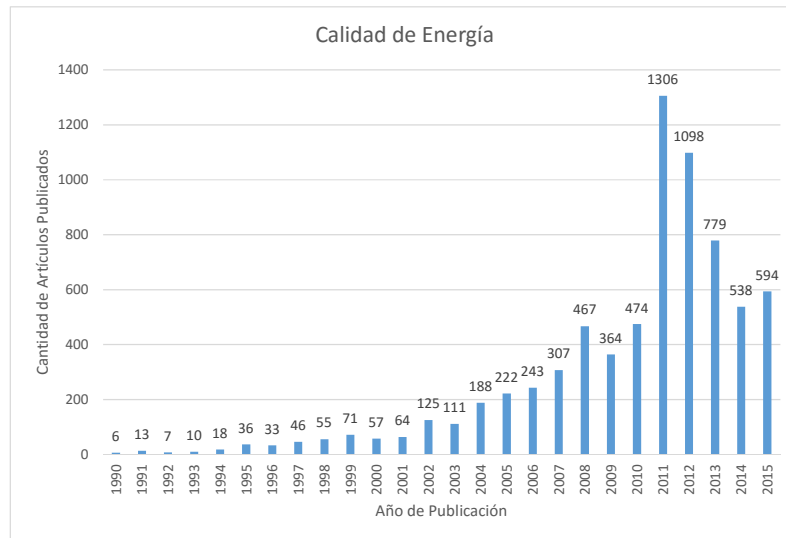


Figura I. 3 - Publicaciones por año en el área de Calidad de Energía.

Para mejorar la calidad de la energía se ha desarrollado una normativa que estandariza los tipos de anomalías que se pueden presentar en un SEP. En la sección 1.1.3 del *capítulo 1* se presenta un resumen de los estándares que definen los tipos de perturbaciones según los parámetros característicos de las formas de onda y las posibles causas que pueden provocar cada tipo de evento en una red eléctrica.

Con el objetivo de disminuir el impacto que los eventos de calidad de energía tienen en un sistema eléctrico de potencia se han desarrollado herramientas para el monitoreo de la calidad de energía. Estas herramientas permiten la identificación, localización y clasificación de un evento en tiempo real. En la sección 1.2 del *capítulo 1* se presentan los aspectos más importantes relacionados a un sistema de monitoreo de eventos de calidad de energía.

VI. Objetivos.

El objetivo general de la tesis consiste en el desarrollo y prueba de algoritmos que utilizan técnicas innovadoras para la detección, clasificación y localización de perturbaciones en Sistemas Eléctricos de Potencia (SEP) para generar mejoras en los parámetros que miden la calidad de la energía eléctrica.

Resultado del análisis de la bibliografía y publicaciones científicas mostró que los sistemas de procesamiento de perturbaciones recientemente adoptaron a las *Máquinas de Vectores de Soporte (SVM, por sus siglas en inglés)* como técnica de clasificación de patrones, las publicaciones rara vez utilizan señales medidas de sistemas reales para el entrenamiento y prueba de los algoritmos y el estudio de *perturbaciones complejas* es un área poco estudiada. Del mismo modo se detecta que la mayoría de las publicaciones analiza los eventos de calidad de energía en un nodo en particular de la red eléctrica descartando la influencia de esta perturbación en los nodos vecinos. Los algoritmos usualmente detectan y clasifican una perturbación electromagnética pero no brindan información del nodo de la red eléctrica

donde se estima que se encuentra el motivo que genera la perturbación. En concordancia con estas ideas se plantean los siguientes objetivos parciales:

- Desarrollar algoritmos avanzados para el procesamiento de señales con el objetivo de permitir la extracción y selección de parámetros característicos que describan las formas de onda adquiridas teniendo en cuenta los siguientes métodos:
 - Transformadas Wavelet (*WT por sus siglas en inglés*).
 - Cálculo del perfil cuadrático medio.
 - Técnicas Estadísticas: Cálculo de la media, mediana, desviación estándar, energía, entre otros.
 - Técnicas de minería de datos para la selección de los parámetros característicos que mejor describen a cada evento en particular.
- Desarrollar algoritmos avanzados para la clasificación de perturbaciones utilizando metodologías de Inteligencia Artificial, tales como, *Máquinas de Vectores de Soporte* (SVM por sus siglas en inglés). El algoritmo desarrollado deberá:
 - Ser fácil de implementar, de bajo costo computacional y lo suficientemente flexible para permitir una rápida incorporación de nuevos tipos de perturbaciones
 - Permitir la clasificación de perturbaciones complejas (perturbaciones que se presentan cuando más de un evento de calidad de energía ocurre de manera simultánea) sin la necesidad de un entrenamiento específico.
- Desarrollar algoritmos avanzados para la identificación y localización de fallas en un Sistema Eléctrico de Potencia de distribución, de topología radial, utilizando técnicas de Inteligencia Artificial, tales como, *Máquinas de Vectores de Soporte* (SVM por sus siglas en inglés). El desarrollo propuesto deberá:
 - Permitir la identificación del tipo de fallas, enfocándose principalmente en las fallas de baja impedancia conocidas como fallas de una fase a tierra, dos fases a tierra, tres fases a tierra y fallas fase a fase.
 - Permitir la identificación de la línea donde la falla se produce.
 - Utilizar la menor cantidad de nodos de medición para la identificación del tipo y localización de una falla.
 - Ser preciso, rápido y lo suficientemente flexible para permitir una rápida adaptación a nuevas condiciones y topologías de la red eléctrica.

VII. Organización de la tesis.

La presente tesis está organizada en dos partes.

La primer parte aborda los estándares involucrados en el monitoreo de un SEP así como los fundamentos teóricos y matemáticos de diferentes métodos de procesamiento de señales y técnicas de clasificación de patrones que luego son utilizados en los desarrollos propuestos. En la segunda parte se presentan los desarrollos destinados a la detección, clasificación y localización de eventos de calidad de energía.

Primera parte – Fundamentación teórica

Para establecer una normativa que caracterice los tipos de anomalías que se pueden presentar en un SEP y las estrategias de monitoreo que se deben llevar a cabo para las realizar las mediciones, diferentes organismos han desarrollado estándares dentro de los cuales, el que se adoptará en este trabajo, es el “*IEEE Std 1159-2009*”. El *capítulo 1* de esta tesis presenta los aspectos más destacados del mencionado estándar.

Tal vez los componentes más importantes de un sistema de monitoreo sean sus algoritmos específicamente diseñados para detectar, clasificar y localizar perturbaciones provenientes de mediciones de la red eléctrica. Estos algoritmos son desarrollados como una combinación de métodos de disparo, códigos de decisiones basados en reglas, técnicas de procesamiento de señales y máquinas de aprendizaje lineal, las que, mancomunadamente, permiten recabar información de una red eléctrica y afrontar con mayor conocimiento las causas que generan una disminución de la calidad de energía.

Dentro de las técnicas de procesamiento de señales la más ampliamente difundida y adoptada en la comunidad científica avocada al estudio de los fenómenos eléctricos es la *Transformada Wavelet (WT, por sus siglas en inglés)*. Esta herramienta permite un análisis simultáneo en el dominio del tiempo y de la frecuencia, lo que facilita la tarea de la detección y localización en el tiempo de fenómenos electromagnéticos. El *capítulo 2* abordará conceptualmente la *WT* y se presentará una reseña de los trabajos más importantes relacionados a la utilización de ésta técnica en el área de calidad de energía y localización de fallas.

En el *capítulo 3* se presenta una técnica de máquinas de aprendizaje que se fundamenta en dos bases sólidas: *La teoría de aprendizaje estadístico* y *la optimización matemática*. Este método, denominado *Máquinas de Vectores de Soporte (SVM, por sus siglas en inglés)* presenta una serie de ventajas frente a otros métodos de *Inteligencia Artificial (AI, por sus siglas en inglés)* adoptados por los investigadores. En el *capítulo 3* el lector podrá encontrar un breve recorrido por las publicaciones científicas que adoptaron a los *SVM* como método de clasificación de perturbaciones y la localización de fallas en un SEP.

Segunda parte – Desarrollo y prueba de un sistema de monitoreo

En la segunda parte del trabajo de tesis se presentan los algoritmos desarrollados para la detección y clasificación de eventos de calidad de energía y los resultados obtenidos. Los

métodos de procesamiento de señales utilizados en el desarrollo mostrado en esta segunda parte de la tesis se sustentan en el marco teórico y matemático presentado en los capítulos que conforman la primera parte de la presente tesis.

En el *capítulo 4* se presenta el desarrollo de una metodología general para el diseño, entrenamiento y prueba de un sistema de clasificación basado *Inteligencia Artificial*, enfocado en aquellos métodos que utilizan técnicas de aprendizaje supervisado. La metodología aquí presentada será posteriormente utilizada en esta tesis, en el *capítulo 6*, para el entrenamiento de un sistema de clasificación de perturbaciones eléctricas. Del mismo modo la metodología propuesta se emplea, en el *capítulo 7*, para el desarrollo de un sistema de clasificación y localización de fallas en un sistema eléctrico de potencia.

En el *capítulo 5* se presenta el desarrollo de una interfaz gráfica que permite, de manera manual, la configuración de dos clasificadores, uno basado en una red neuronal backpropagation y un segundo basado en un clasificador *SVM One vs Rest*.

En el *capítulo 6* se presenta el diseño de un clasificador multiclase basado en *SVM* organizados en una arquitectura *One vs Rest* optimizado para la clasificación de perturbaciones simples y complejas. Los resultados obtenidos son comparados con otras publicaciones similares.

De la misma manera, en el *capítulo 7*, se propone un algoritmo para detectar, clasificar y localizar fallas en un sistema eléctrico de distribución, de topología radial. El algoritmo propuesto, basado en la *WT* y *SVM*, será diseñado y entrenado de acuerdo a la metodología propuesta en el *capítulo 4*.

Finalmente, en el *Capítulo 8* se presentan las conclusiones, aportes, trabajos futuros y se listan las publicaciones realizadas en el marco de las investigaciones aquí presentadas.

VIII. Tabla de acrónimos

Acrónimo	Nombre en inglés	Nombre en español
ACO	Ant Colony Optimization	Optimización Colonia de Hormigas
AI	Artificial Intelligence	Inteligencia Artificial
ANN	Artificial Neural Network	Red Neuronal Artificial
ART	Adaptive Resonance Theory	Teoría de la Resonancia Adaptiva
CFT	Continuous Fourier Transform	Transformada de Fourier de tiempo continuo
CHT	Cross Hilbert Transform	Transformada Cross Hilbert
CWT	Continuous Wavelet Transform	Transformada Wavelet de Tiempo Continuo
CHHT	Cross Hilbert-Huang Transform	Transformada Cross Hilbert-Huang
DFT	Discrete Fourier Transform	Transformada Discreta de Fourier
DST	Discrete Stockwell Transform	Transformada de Stockwell Discreta
DWT	Discrete Wavelet Transform	Transformada Wavelet Discreta
ELM	Extreme Learning Machine	Máquina de Aprendizaje Extremo
EEMD	Ensemble Empirical Mode Decomposition	Descomposición Empírica de Modos por Conjuntos
ES	Expert System	Sistema Experto
FFT	Fast Fourier Transform	Transformada Rápida de Fourier
FKNN	Fuzzy k-nearest neighbour	Algoritmo difuso de k-ésimo vecino más cercano
FL	Fuzzy Logic	Lógica Difusa
FT	Fourier Transform	Transformada de Fourier
GA	Genetic Algorithm	Algoritmo Genético
GT	Gabor Transform	Transformada Gabor
GWT	Gabor-Wigner Transform	Transformada Gabor-Wigner
HMM	Hidden Markov Model	Modelo Markoviano Oculto
HT	Hilbert Transform	Transformada Hilbert
KF	Kalman Filter	Filtro de Kalman
LVQ	Learning Vector Quantization	Vector de Aprendizaje Cuántico
MLP	Multi Layer Perceptron	Perceptrón Multi Capa
PNN	Probabilistic Neural Network	Red Neuronal Probabilística
PSO	Particle Swarm Optimization	Optimización de Enjambre de Partículas
RBF	Radial Basic Function	Función de Base Radial
SANN	Self Adapting Neural Network	Red Neuronal Auto Adaptiva
SOM	Self Organizing Map	Mapas Auto Organizados
STFT	Short Time Fourier Transform	Transformada Fourier de Tiempo Corto
SVM	Support Vector Machine	Máquinas de Vectores de Soporte
OvA	One versus All	Uno contra todos
OvO	One versus One	Uno contra Uno
OvN	One versus Normal	Uno contra Normal
OvR	One versus Rest	Uno contra Resto
URONN	Univariate Random Optimized Neural Network	Red Neuronal Optimizada Invariable
WPT	Wavelet Packet Transform	Paquete de Transformada Wavelet
WT	Wavelet Transform	Transformada Wavelet

Capítulo 1:

Clasificación de Perturbaciones y Sistemas de Monitoreo

1.1 CLASIFICACION DE EVENTOS DE CALIDAD DE ENERGIA.

Los principales motivos por los cuales es recomendable clasificar los eventos en diferentes categorías se debe a la necesidad de aplicar las técnicas adecuadas para obtener mediciones, resolver problemas, extraer parámetros característicos o establecer métodos de clasificación que dependen del tipo o clase de evento en cuestión.

Existen diversos criterios para la clasificación de eventos de calidad de energía de acuerdo a la variación o los parámetros que difieren de la forma de onda ideal. En este trabajo se presenta un criterio de división de acuerdo a la duración de la perturbación y luego se abordan los criterios establecidos por la *IEC* y la *IEE*.

1.1.1 Clasificación de acuerdo a la duración.

Una importante división dentro de los tipos de perturbaciones es aquella que se realiza teniendo en cuenta la duración del evento. Existen fenómenos de carácter estacionario (o cuasi estacionario) que persisten por períodos de tiempo mucho mayores a los períodos de análisis. En tal caso, la perturbación suele llamarse *variación*. Un ejemplo de una variación es una alteración en la frecuencia fundamental de la forma de onda de tensión.

El segundo tipo de variaciones, conocidas como *perturbaciones repentinas*, son eventos que comienzan y finalizan dentro de un período corto de tiempo que usualmente se encuentra dentro del tiempo de análisis (también llamado *ventana*). Una caída de tensión por debajo de un umbral predeterminado, durante un tiempo reducido es un ejemplo de *perturbación repentina* también conocida como *perturbación de corta duración*.

Esta clasificación es importante a la hora de definir las técnicas de procesamiento de señales que permiten la extracción de parámetros característicos para la detección, clasificación y caracterización de perturbaciones en un SEP. Un claro ejemplo de esto es el método empleado para la ejecución de los algoritmos de detección y clasificación.

1.1.2 Clasificación de acuerdo a la IEC.

En la *IEC 61000-2-5 1995* [8] los eventos electromagnéticos se clasifican de acuerdo a lo mostrado en la Tabla 1. 1.

Tabla 1. 1 - Principales fenómenos que causan perturbaciones electromagnéticas.

Grupo	Ejemplos
Fenómenos de baja Frecuencia – Conducidos	Armónicos, Interarmónicos
	Señales de sistema (Power Line Carrier)
	Fluctuaciones de Voltaje
	Huecos e Interrupciones de Voltaje
	Desbalance de Voltaje
	Variaciones en la frecuencia
	Voltajes inducidos de baja frecuencia
	Tensiones de continua en redes de corriente alterna.
Fenómenos de baja frecuencia – Radiados	Campos Magnéticos
	Campos Eléctricos
Fenómenos de alta frecuencia – Conducidos	Señales de continua inducidas (Corriente o Voltaje)
	Transitorios unidireccionales
	Transitorios oscilatorios
Fenómenos de alta frecuencia – Radiados	Campos Magnéticos
	Campos Eléctricos
	Campos Electromagnéticos
	Ondas continuas
	Transitorios
Fenómenos de descarga electro estática (ESD)	_____
Pulsos Electromagnéticos de origen Nuclear (NEMP)	_____

1.1.3 Clasificación de acuerdo a la IEEE.

En la Tabla 1. 2 se detalla lo propuesto por el estándar *1159-2009 de la IEEE* [9]. Esta normativa trata sobre las prácticas recomendadas para el monitoreo de calidad de energía eléctrica.

El estándar propone una clasificación en siete grandes grupos de acuerdo parámetros como la duración de la perturbación, su contenido espectral y su magnitud típica.

En las siguientes secciones se describirán las características distintivas de cada una de las categorías. Debido a que el presente trabajo se enfoca en el procesamiento de señales para la detección y clasificación de perturbaciones aquí no se profundizará en las causas que generan cada perturbación ni en el estudio de las consecuencias que estos eventos traen aparejados en un sistema eléctrico de potencia.

Tabla 1. 2 - Categorías y Características de fenómenos electromagnéticos en un SEP.

Categoría	Contenido espectral típico	Duración típica	Magnitud de tensión típica
1.0 Transitorios			
1.1 Impulsivo			
1.1.1 nanosegundos	5 nseg (levantamiento)	< 50 nanoseg	
1.1.2 microsegundos	1 microseg (levantamiento)	50 nanoseg - 1 miliseg	
1.1.3 milisegundos	0.1 miliseg (levantamiento)	> 1 miliseg	
1.2 Oscilatorio			
1.2.1 Baja frecuencia	< 5 KHz	0.3 - 50 miliseg	0 - 4 pu*
1.2.2 Media frecuencia	5 - 500 KHz	20 microseg	0 - 8 pu
1.2.3 Alta frecuencia	0.5 - 5 MHz	5 microseg	0 - 4 pu
2.0 Variaciones de corta duración			
2.1 Instantáneo			
2.1.1 Sag		0.5 - 30 ciclos	0.1 - 0.9 pu
2.1.2 Swell		0.5 - 30 ciclos	1.1 - 1.8 pu
2.2 Momentáneo			
2.2.1 Interrupción		0.5 ciclos - 3 seg	< 0.1 pu
2.2.2 Sag		30 ciclos - 3 seg	0.1 - 0.9 pu
2.2.3 Swell		30 ciclos - 3 seg	1.1 - 1.4 pu
2.3 Temporánea			
2.3.1 Interrupción		3 seg - 1 min	< 0.1 pu
2.3.2 Sag		3 seg - 1 min	0.1 - 0.9 pu
2.3.3 Swell		3 seg - 1 min	1.1 - 1.2 pu
3.0 Variaciones de larga duración			
3.1 Interrupción sostenida		> 1 min	0.0 pu
3.2 Subtensiones		> 1 min	0.8 - 0.9 pu
3.3 Sobretensiones		> 1 min	1.1 - 1.2 pu
4.0 Desbalance de tensión		estado estacionario	0.5 - 2 %
5.0 Distorsión de la forma de onda			
5.1 Offset de continua		estado estacionario	0 - 0.1 %
5.2 Armónicos	Orden de armónico : 0 - 100	estado estacionario	0 - 20 %
5.3 Interarmónicos	0 - 6 KHz	estado estacionario	0 - 2 %
5.4 Muestras		estado estacionario	

5.5 Ruido	Banda ancha	estado estacionario	0 - 1 %
6.0 Fluctuaciones de tensión	< 25 Hz	intermitente	0.1 - 7 %
7.0 Variaciones de la frecuencia fundamental		< 10 seg	

*Nota La cantidad pu se refiere a una cantidad por unidad (per unit) que es adimensional. La cantidad 1.0 pu se corresponde al 100% del valor nominal.

1.1.3.1 Transitorios.

- Transitorios impulsivos (En inglés *Impulsive Transient*): Un transitorio impulsivo es un cambio súbito y unidireccional (positivo o negativo) en la condición de estado estable de la tensión, la corriente o ambos y de frecuencia diferente a la frecuencia del sistema eléctrico de potencia. Un transitorio impulsivo usualmente es caracterizado por su tiempo de subida y el tiempo de decaimiento. Este tipo de fenómeno puede producirse por una descarga eléctrica o por fuentes de iluminación.
- Transitorios oscilatorios (En inglés *Oscillatory Transient*): Un transitorio oscilatorio es un cambio repentino y bidireccional de la condición de estado estacionario de la tensión, corriente o ambos que incluyen valores positivos y negativos. Según la Tabla 1. 2 los transitorios oscilatorios se pueden clasificar en como de alta, media y baja frecuencia dependiendo del contenido espectral del evento. Los transitorios oscilatorios con una componente de frecuencia primaria mayor a los 500 [KHz] y una duración típica medida en micro segundos (Varios ciclos de la frecuencia fundamental) son considerados como transitorios oscilatorios de alta frecuencia. Estos eventos se producen mayoritariamente por algún evento de conmutación o ser el resultado de una respuesta oscilatoria de un sistema ante una excitación del tipo transitorio impulsivo.

Un transitorio cuyo componente de frecuencia primaria se encuentre entre los 5 y los 500 [KHz] y cuya duración sea de decenas de microsegundos (varios ciclos de la frecuencia fundamental) se considera como un transitorio oscilatorio de media frecuencia. La energización de un arreglo back to back de capacitores resulta en corrientes oscilatorias cuya frecuencia supera la decena de [KHz].

Un transitorio con un componente de frecuencia primaria que no supera los 5 [KHz] y cuya duración va de los 0.3 a los 50 [ms] se considera como un transitorio oscilatorio de baja frecuencia. Esta categoría de eventos se encuentran en los sistemas de sub transmisión y distribución y son causados por diversos tipos de eventos, principalmente por la energización de un banco de capacitores.

La Figura 1. 1 muestra un transitorio oscilatorio de alta frecuencia y la Figura 1. 2 muestra un transitorio de baja frecuencia.

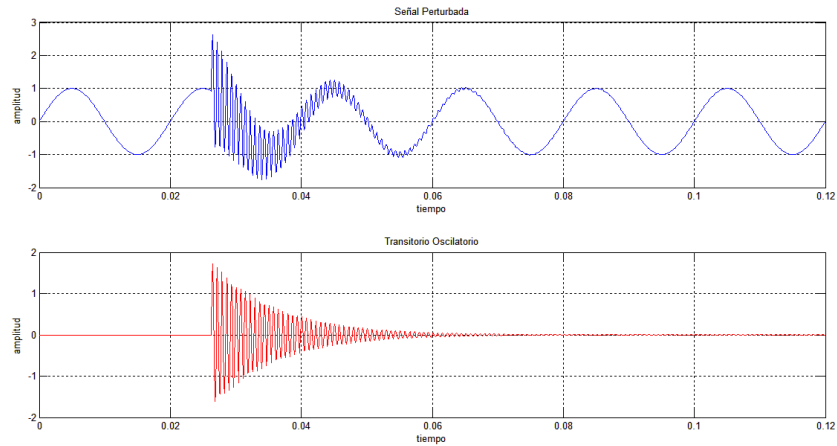


Figura 1. 1 - Transitorio Oscilatorio de alta frecuencia.

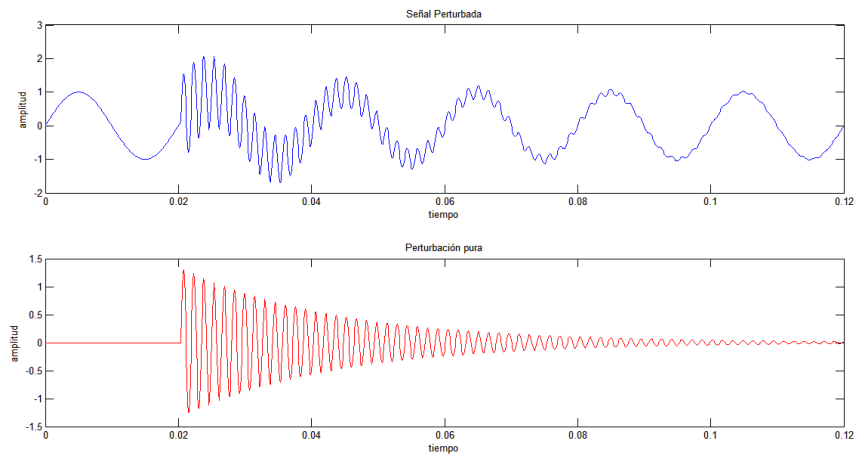


Figura 1. 2 - Transitorio Oscilatorio de baja frecuencia.

1.1.3.2 Variaciones de corta duración.

El estándar divide a las variaciones de corta duración en tres categorías: *Huecos* (en inglés *Sag**), *Subas* (en inglés *Swell*) e *Interrupción* (en inglés *Outage*). Cada una de estas categorías se subdividen de acuerdo a la duración como: *Instantáneas*, *Momentáneas* o *Temporarias*. La mayoría de las variaciones de corta duración son causadas por condiciones de falla o la alimentación de grandes cargas que requieren altas corrientes de arranque. Otro motivo se puede encontrar en la pérdida intermitente de conexión con el cableado de alimentación.

*Nota: Sag es la terminología empleada por el estándar propuesto por la IEEE. Un término equivalente comúnmente empleado por la IEC es “Voltage Dip”

- Interrupciones momentáneas y temporales (en inglés *Outage*): Una interrupción ocurre cuando el voltaje de alimentación o la corriente de carga disminuye a un valor

inferior al $0.1 [p.u.]$ por un período de tiempo que no excede un minuto. Las interrupciones pueden ser el resultado de fallas en los sistemas eléctricos, falla en los equipos y funcionamiento inadecuado de los sistemas de control. El tiempo de inicio para la medición de una interrupción tiene lugar cuando la magnitud del voltaje cae por debajo del 10% del valor nominal.

Las interrupciones generadas por una falla generalmente se ven precedidas por un hueco de tensión cuya duración está relacionada con el tiempo de reacción de los dispositivos de protección. La duración de la interrupción dependerá mayoritariamente de la capacidad de recierre de los dispositivos de protección. La Figura 1. 3 muestra una interrupción ideal ya que los tiempos de decaimiento y de subida son nulos.

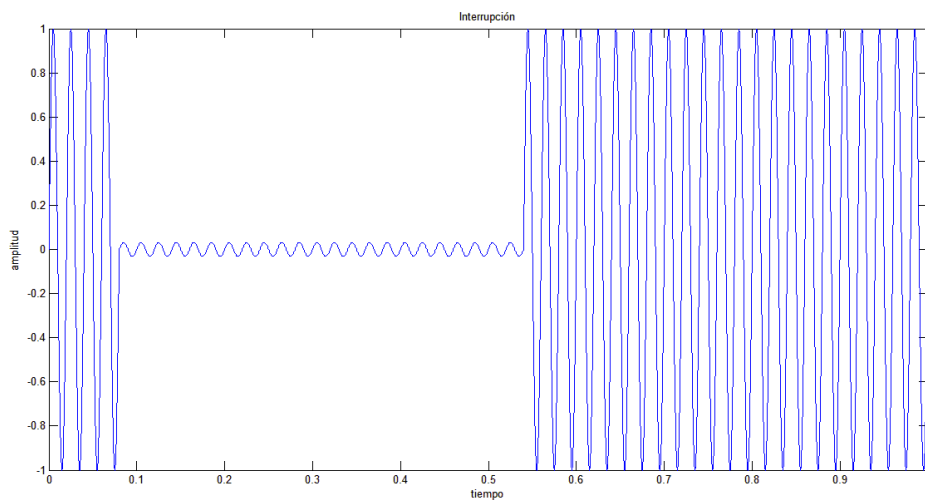


Figura 1. 3 - Interrupción ideal.

La tensión (o su valor rms) generalmente no cae a cero en forma instantánea ante una operación de apertura de un interruptor, debido a las cargas inductivas y/o motores conectados a la red. Esto se produce por ejemplo cuando hay conectada una significativa cantidad de motores de inducción en un SEP. Estos efectos pueden observarse en la Figura 1. 4 (a) y (b).

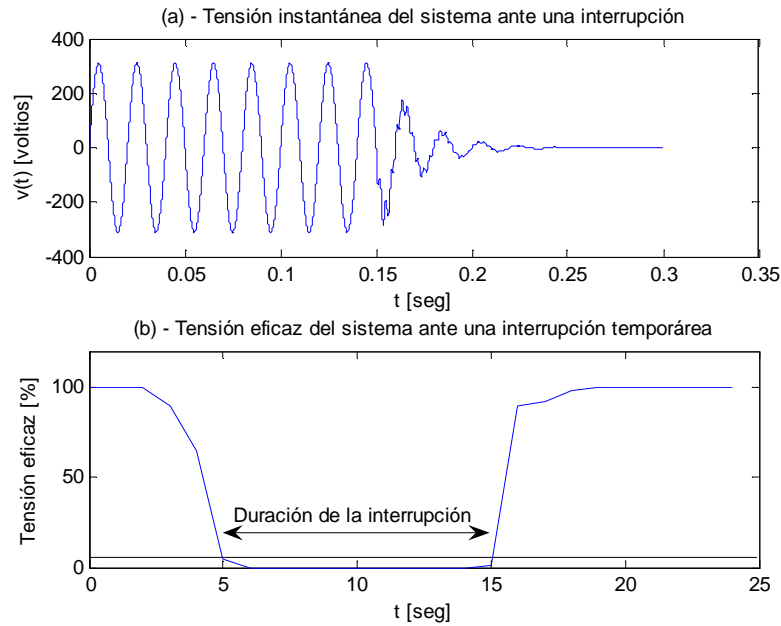


Figura 1. 4 - Variación de la tensión instantánea y valor rms ante una interrupción.

- Huecos de tensión (en inglés *Voltage Sag*): Un hueco de tensión es una caída en el valor medio del voltaje entre $0.1 [p.u.]$ y $0.9 [p.u.]$ por un tiempo que va desde los 0.5 ciclos hasta un minuto. La terminología utilizada para describir un hueco de tensión suele ser algo ambigua. Un hueco del 20% puede significar como resultado de una caída del voltaje de $0.8 [p.u.]$ o de $0.2 [p.u.]$ dependiendo si la terminología empleada hace referencia al voltaje resultante o a la caída de voltaje. El estándar propone utilizar el voltaje resultante para ser empleado como terminología común. En este caso, un hueco de tensión del 80% significa un perturbación que resulta en una tensión de $0.8 [p.u.]$.

Los huecos de tensión se dan como resultado de una falla en el sistema, de la conmutación de grandes cargas o el encendido de motores grandes.

La duración de un hueco de tensión ha sido definida desde los $2 [ms]$ (aproximadamente un octavo de un ciclo) a varios minutos en diferentes publicaciones. Caídas de tensión cuya duración no supera la mitad de un ciclo no puede ser caracterizada efectivamente por un cambio en el valor eficaz de la frecuencia fundamental. Por otro lado, caídas de tensión que duran más de un minuto son clasificadas como eventos de larga duración.

La Figura 1. 5 (a) muestra un Hueco de tensión ideal ya que las transiciones ocurren en un tiempo nulo mientras que en la Figura 1. 5 (b) se puede ver el valor eficaz instantáneo de la señal.

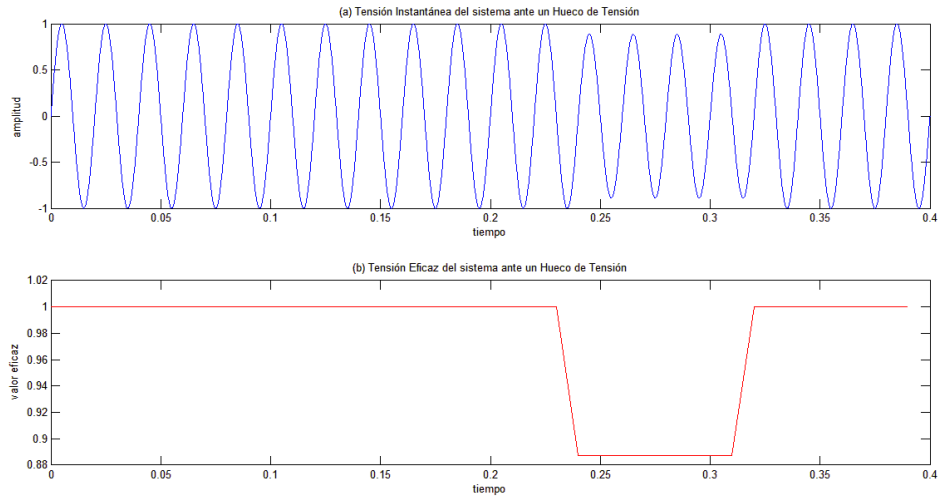


Figura 1. 5 - Tensión instantánea y perfil rms de una forma de onda ante hueco de tensión.

- **Sobretensión** (en inglés *Voltage Swell*): Una sobretensión es un incremento en el valor rms de la señal de voltaje por encima del valor $1.1 [p.u.]$ por un tiempo que puede ir desde los 0.5 ciclos hasta 1 minuto. Las magnitudes típicas de sobre tensión van desde 1.1 a $1.2 [p.u.]$.

Al igual que los huecos de tensión, las sobretensiones suele estar asociadas con condiciones de fallas del sistema aunque este tipo de perturbaciones son mucho menos comunes. Las sobretensiones suelen también ser causadas por la conmutación de grandes cargas, divisores de cargas o grandes bancos de capacitores.

La Figura 1. 6 (a) muestra una sobretensión ideal ya que los tiempos de elevación y decaimiento del evento ocurren en tiempos idealmente nulos. La Figura 1. 6 (b) ilustra el valor eficaz instantáneo de la señal.

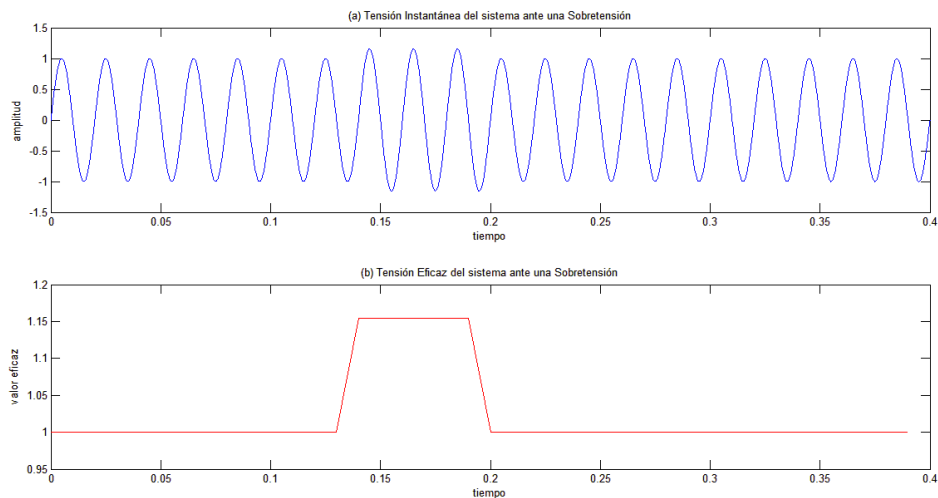


Figura 1. 6 - Tensión instantánea y perfil rms de una forma de onda ante una sobretensión.

La severidad de una sobretensión durante una condición de falla es función de la localización de la falla, la impedancia del sistema y del método de puesta a tierra del sistema.

1.1.3.3 Variaciones de larga duración.

Las variaciones de larga duración se corresponden con desviaciones del valor eficaz de una señal por un tiempo superior al minuto. En la Tabla 1. 2 se puede ver que las variaciones de larga duración pueden tratarse tanto de sobretensiones como subtensiones y generalmente son causadas por variaciones en el sistema y operaciones de conmutación. A continuación se abordan los distintos tipos de variaciones de larga duración.

- Sobretensión (En inglés *Overvoltaje*): Una sobre tensión es un incremento del valor eficaz de una tensión por un valor superior a $1.1 [p.u.]$ cuya duración supera el minuto. Los valores típicos de la sobretensión se encuentran entre $1.1 [p.u.]$ y $1.2 [p.u.]$. Las sobre tensiones pueden ser el resultado de la conmutación de cargas, la conmutación de bancos de capacitores o la configuración incorrecta del tap de un transformador.
- Subtensión (En Inglés *Undervoltaje*): Una caída prolongada en la tensión es un decremento del valor eficaz de tensión por un valor inferior a $0.9 [p.u.]$ por un tiempo superior a un minuto. Los valores típicos de las caídas de tensión se encuentran entre $0.8 [p.u.]$ y $0.9 [p.u.]$. Algunas de las causas de una subtensión pueden ser: la conexión de grandes cargas, la desconexión de bancos de capacitores o la sobrecarga de los circuitos.
- Interrupciones sostenidas (En inglés *Sustained Interruptions*): La caída de tensión a un valor menor al 10% del valor nominal y por un período de tiempo que supera el minuto se considera como una interrupción sostenida. Interrupciones que superan el minuto usualmente son consideradas como permanentes ya que requieren de la intervención manual para restablecer las condiciones normales de operación.

1.1.3.4 Desbalance de Tensión.

Se define como desbalance de un sistema trifásico a la relación entre la componente de la secuencia negativa respecto de la positiva, expresada en porcentaje. Esta definición puede aplicarse tanto para el voltaje como para la corriente. El desbalance típico de un SEP se encuentra por debajo del 3% en medidos en tensión.

1.1.3.5 Distorsión de la forma de onda.

Los fenómenos clasificados como perturbaciones de distorsión de forma de onda, de acuerdo a lo detallado en la Tabla 1. 2, se clasifican en: *Componente de Continua*, *Armónicos*, *Interarmónicos*, *Muecas*, *Ruido*. Cada una de estas categorías se explican brevemente a continuación.

- Corrimiento de continua (en inglés *DC offset*): La presencia de componentes de continua en un sistema de corriente y tensión alterna se denomina como corrimiento

de continua. Este fenómeno ocurre como resultado de perturbaciones geomagnéticas o debido a la presencia de rectificadores de media onda en el Sistema Eléctrico de Potencia. Este tipo de perturbaciones pueden generar pérdidas debido al incremento de la saturación en los transformadores y su calentamiento asociado.

- Armónicos (en inglés *Harmonics*): Los armónicos han existido desde que se adoptó a la corriente alterna como medio de transmisión de la energía eléctrica. Sin embargo este efecto se ha hecho más significativo en estos últimos tiempos debido al incremento de los dispositivos no lineales. Un dispositivo no lineal produce corrientes no sinusoidales aun cuando la tensión sea sinusoidal. En el caso ideal, las fuentes de tensión son sinusoidales. Las corrientes distorsionadas originadas en las distintas cargas no lineales del sistema producen en las impedancias del sistema, consideradas lineales, caídas de tensión distorsionadas. Estas caídas de tensión, sumadas a la fuente ideal sinusoidal, reproducen en las barras de carga tensiones que finalmente son distorsionadas.

Los dispositivos cuyo funcionamiento se basa en electrónica de potencia son los mayores contribuyentes de armónicos en los SEP. La forma de caracterizar la distorsión armónica es mediante un gráfico que contenga la magnitud y fase de cada componente armónica del sistema. En un típico gráfico del espectro de frecuencias de una señal con contenido armónico se puede apreciar una fuerte presencia de energía en las frecuencias correspondientes a múltiplos enteros de la frecuencia fundamental del sistema. Existe un parámetro, denominado *Distorsión Total Armónica* (por sus siglas en inglés *THD*), que es comúnmente empleado para cuantificar la distorsión armónica presente en la red.

La Figura 1. 7 a) Muestra la tensión instantánea de un sistema contaminado con segundo, tercer cuarto y quinto armónico. En la Figura 1. 7 b) se puede ver el resultado de la *Transformada de Fourier* (*FT*, por sus siglas en inglés) de la señal contaminada con armónicos.

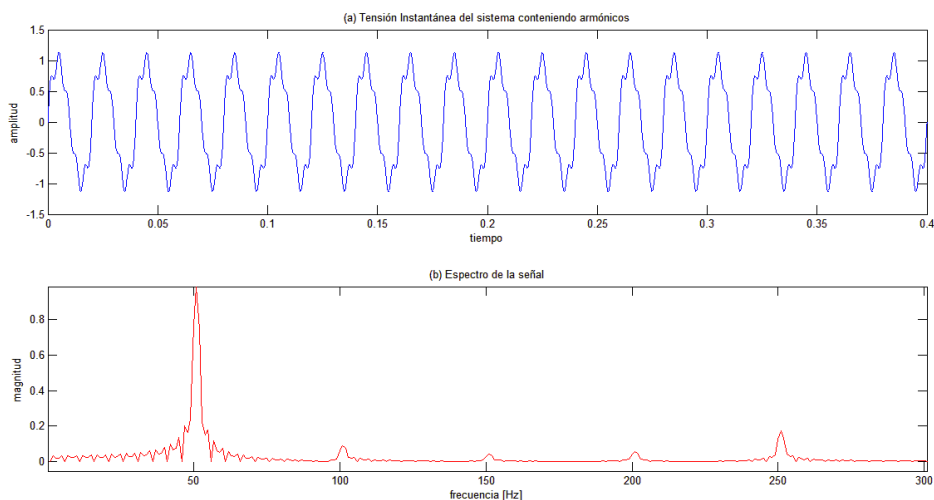


Figura 1. 7 - Tensión instantánea y espectro de una señal con conteniendo armónicos.

- **Interarmónicos (En inglés Interharmonics):** Se denomina interarmónicos a aquellos componentes del espectro de frecuencia de la señal perturbada que no se corresponden con múltiplos enteros de la frecuencia fundamental del sistema eléctrico. Las principales causas de los interarmónicos son los convertidores estáticos de frecuencia, los ciclos convertidores, los hornos de inducción y los dispositivos de arco, especialmente aquellos cuyo control no se encuentra apropiadamente sincronizado con la frecuencia del sistema.
- **Muecas (En inglés Notch):** Las muecas son una perturbación periódica provocada por la operación de dispositivos que emplean la electrónica de potencia cuando la corriente se conmuta de una fase a otra. Este fenómeno representa un caso especial ya que es periódico, por lo que podría interpretarse como un fenómeno de estado estacionario, y a su vez presenta componentes de alta frecuencia. Esto hace difícil enmarcarlo como un transitorio de corta duración o una distorsión de forma de onda. Los convertidores trifásicos (inversores) son los mayores generadores de este tipo de perturbación en un sistema de potencia.

La Figura 1. 8 muestra una señal contaminada con una perturbación del tipo muesca. En esta figura se puede apreciar la periodicidad de la perturbación.

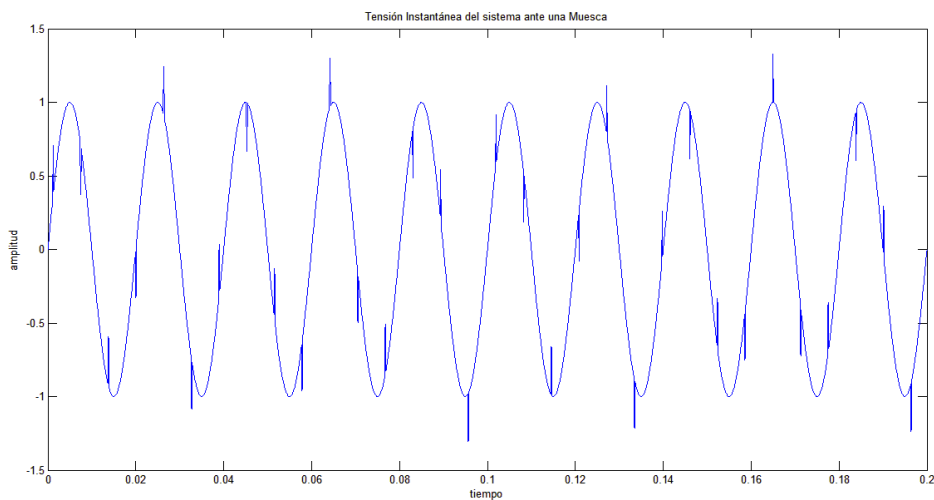


Figura 1. 8 - Tensión instantánea de un sistema ante una muesca.

- **Ruido (En inglés Noise):** Un ruido es una señal eléctrica no deseada, con una amplia composición espectral, típicamente con valores máximos de 200 [KHz], que se agrega a la señal de tensión o corriente de manera aditiva. Básicamente, se define ruido a toda alteración no deseada que no puede ser clasificada como distorsión armónica o perturbaciones del tipo transitorios. El ruido puede ser generado por equipos de electrónica de potencia, circuitos de control, equipos que utilizan arcos voltaicos, cargas con rectificadores de estado sólido y fuentes de energía conmutada.

En la Figura 1. 9 se puede ver una señal de tensión de un sistema eléctrico de potencia contaminada con una perturbación del tipo ruido.

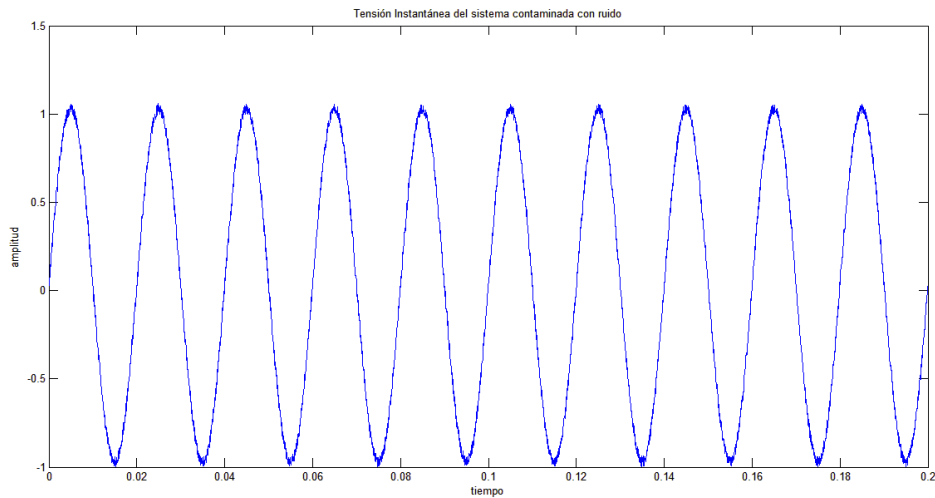


Figura 1. 9 - Tensión instantánea del sistema contaminada con ruido.

1.1.3.6 Fluctuaciones de tensión.

Las fluctuaciones de tensión son variaciones sistemáticas de la envolvente de la señal del voltaje o una serie de cambios aleatorios en la magnitud del voltaje, magnitud que normalmente no excede el rango comprendido entre los $0.95 [p.u.]$ y los $1.05 [p.u.]$.

Cualquier carga que tenga importantes variaciones cíclicas, especialmente su componente reactiva, puede producir fluctuaciones de voltaje. Usualmente suele emplearse el término *flicker* de manera errónea para denominar a las *fluctuaciones* de tensión. Mientras que las primeras se refieren a un fenómeno electromagnético, el término *flicker* está relacionado al resultado no deseado de estas variaciones percibidas por el ojo humano como una oscilación en la intensidad lumínica de una lámpara.

Los hornos de arco eléctrico son los generadores más comunes de este tipo de fenómeno.

Las fluctuaciones de tensión suelen aparecer como una modulación en la magnitud de la forma de onda fundamental (similar a lo que ocurre en la modulación AM de una señal de radio). Se puede definir la magnitud de la fluctuación como la magnitud rms de la señal modulante. Esto puede fácilmente obtenerse mediante la demodulación de la envolvente a través de la remoción de la frecuencia fundamental y mediante la medición de la magnitud de la componente moduladora.

La Figura 1. 10 muestra la tensión instantánea de un sistema eléctrico de potencia ante la presencia de fluctuaciones en su magnitud.

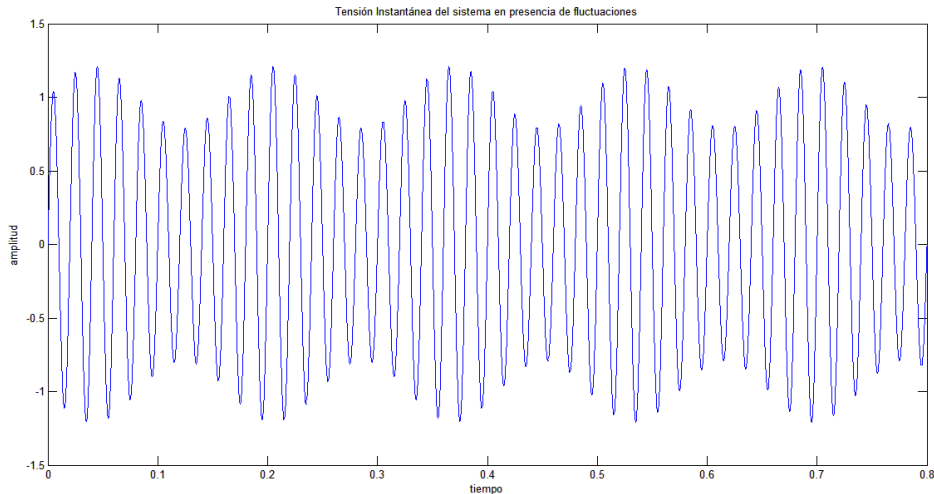


Figura 1. 10 - Tensión instantánea del sistema ante fluctuaciones.

1.1.3.7 Variaciones de la frecuencia fundamental.

Este tipo de fenómeno se explica como una desviación de la frecuencia nominal del sistema eléctrico de potencia (normalmente $50 [Hz]$ o $60 [Hz]$). La frecuencia fundamental en estado estacionario está directamente relacionada con la velocidad de rotación de los generadores del sistema. En cualquier instante de tiempo, la frecuencia fundamental depende del balance entre la carga y la capacidad disponible de generación. Cuando este balance se altera, pueden ocurrir pequeños cambios en la frecuencia. El cambio de la frecuencia y su duración son variables que dependen de la carga y de la capacidad de respuesta del sistema de generación ante estas variaciones en la carga.

1.1.4 Modelos empleados para la simulación de los eventos de calidad de energía.

En Tabla 1. 3 se resumen los modelos matemáticos empleados en este trabajo para la simulación de las formas de onda perturbadas. Estos modelos fueron propuestos por Panigrahi y Pandi (2009) [10].

Tabla 1. 3 - Modelo de perturbaciones.

<i>Perturbaciones</i>	<i>Modelo</i>	<i>Variación de los parámetros</i>
<i>Hueco de tensión</i>	$x(t) = A\{1-\alpha [u(t-t_1) - u(t-t_2)]\}\sin(\omega t)$ $t_2 > t_1 \text{ y } u(t) = \begin{cases} 1 & \geq 0 \\ 0 & \leq 0 \end{cases}$	$0.05 \leq \alpha \leq 0.85$ $T \leq t_2 - t_1 \leq 7T$
<i>Sobre tensión</i>	$x(t) = A\{1+\alpha [u(t-t_1) - u(t-t_2)]\}\sin(\omega t)$ $t_2 > t_1 \text{ y } u(t) = \begin{cases} 1 & \geq 0 \\ 0 & \leq 0 \end{cases}$	$0.05 \leq \alpha \leq 0.8$ $T \leq t_2 - t_1 \leq 7T$
<i>Armónicos</i>	$x(t) = A[\alpha_1 \sin(\omega t) + \alpha_2 \sin(2\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t)]$	$\alpha_1 = 1$ $\alpha_2 - \alpha_7 = 0.0 - 0.3$
<i>Fluctuaciones</i>	$x(t) = A\{1 + \alpha_f [\sin(\beta_f \omega t)]\}\sin(\omega t)$	$0.01 \leq \alpha_f \leq 0.25$ $2[\text{Hz}] \leq f_f \leq 8[\text{Hz}]$
<i>Transitorio Oscilatorio</i>	$x(t) = A\{\sin(\omega t) + be^{-\gamma(t-t_1)}\}\sin(\omega_{tr}(t-t_1))$	$-2 \leq b \leq 2$ $50 \leq \gamma \leq 100$ $500[\text{Hz}] \leq f_{tr} \leq 1500[\text{Hz}]$
<i>Interrupción</i>	$x(t) = A\{1-\alpha [u(t-t_1) - u(t-t_2)]\}\sin(\omega t)$ $t_2 > t_1 \text{ y } u(t) = \begin{cases} 1 & \geq 0 \\ 0 & \leq 0 \end{cases}$	$0.88 \leq \alpha \leq 0.85$ $T \leq t_2 - t_1 \leq 7T$
<i>Armónicos y Hueco</i>	$A\{1-\alpha [u(t-t_1) - u(t-t_2)]\}[\alpha_1 \sin(\omega t) + \alpha_2 \sin(2\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t)]$	$0.05 \leq \alpha \leq 0.85$ $T \leq t_2 - t_1 \leq 7T$ $\alpha_1 = 1$ $\alpha_2 - \alpha_7 = 0.0 - 0.3$
<i>Armónicos y Sobretenión</i>	$A\{1-\alpha [u(t-t_1) - u(t-t_2)]\}[\alpha_1 \sin(\omega t) + \alpha_2 \sin(2\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t)]$	$0.05 \leq \alpha \leq 0.8$ $T \leq t_2 - t_1 \leq 7T$ $\alpha_1 = 1$ $\alpha_2 - \alpha_7 = 0.0 - 0.3$

1.2 SISTEMA DE MONITOREO DE EVENTOS DE CALIDAD DE ENERGIA

1.2.1 Necesidades de monitoreo

En la mayoría de las disciplinas tecnológicas el primer paso para resolver cualquier problema radica en la detección de la fuente que lo origina. El monitoreo de un SEP permite la recolección de información de las características del estado de la red desde el punto de vista del comportamiento de las variables más importantes que la gobiernan en condiciones de operación normal y/o transitoria [5]. Las mencionadas variables se corresponden a los valores de amplitud, frecuencia y fase de las formas de onda de tensión y corriente. Los datos recolectados son procesados con el objetivo de obtener información acerca de los tipos de perturbaciones que se presentan en el SEP, las posibles causas que la generan y las acciones que se deben tomar para mitigarlas.

Según las definiciones del estándar *IEEE Std 1159-2009 Prácticas recomendadas para el monitoreo de calidad de energía* [9] el monitoreo de un sistema eléctrico de potencia es una herramienta imprescindible para detectar, clasificar, caracterizar y localizar los fenómenos electromagnéticos.

No existe un único y común objetivo para el monitoreo de las variables que gobiernan un SEP. En algunos casos, el monitoreo se realiza para diagnosticar incompatibilidades entre las fuentes de energía y la carga. En otros casos el monitoreo es empleado para evaluar las condiciones eléctricas de una ubicación particular con el fin de refinar técnicas de modelado para desarrollar normas y referencias de calidad de energía. La elección de las técnicas, los equipamientos de medición, los métodos de recolección de información, los disparadores, los umbrales y las técnicas de análisis de datos a emplear se deben realizar teniendo en cuenta el objetivo del sistema de monitoreo. Por ejemplo, si un sistema eléctrico presenta reiteradas salidas de servicio, el propósito del monitoreo de la red puede ser el de detectar las perturbaciones que generan condiciones de operación no previstas en equipamientos sensibles. En este caso, el equipamiento y técnicas de medición deben ser seleccionados de modo que permitan la detección de las condiciones que provocan que los equipos funcionen fuera de los rangos de tolerancia exigidos por el o los fabricante/s.

Sin importar los objetivos técnicos involucrados en la implementación de un sistema de monitoreo el principal motivo que fundamenta su implementación radica en las *pérdidas económicas generadas por los eventos de calidad de energía*. Los efectos en el equipamiento y en los procesos de una industria pueden incluir la mala operación, el daño del equipamiento y la detención de una línea productiva.

Todos estos efectos tienen como consecuencia importantes pérdidas económicas tanto para el usuario como para la empresa proveedora de energía eléctrica. El recambio de un equipo dañado, las pérdidas económicas que se generan por mantener una línea productiva fuera de operación, la incorrecta medición de la energía consumida, el detrimento de la vida útil de las instalaciones, etc., son algunas de las pérdidas económicas producidas por eventos de calidad de energía que un usuario y la empresa proveedora deben afrontar.

1.2.2 Reseña Histórica

En la década de 1920 *General Electric* desarrolló lo que denominó una grabadora de arco. Este artefacto estaba conformado por un motor que movía una tira de papel desde un tambor hacia otro. Un par de electrodos generaban un arco a través del papel el que dejaba unas marcas proporcionales a las mediciones realizadas. Este dispositivo era capaz de grabar en el papel información cualitativa de la amplitud de una señal así como el tiempo y la fecha en que el evento era registrado. Un técnico capacitado debía ser capaz de interpretar las marcas en el papel y en sus habilidades se basaba la inteligencia del sistema de monitoreo.

En los comienzos de la década de 1960 *Martzloff* aplicó un arreglo inteligente con una cámara fotográfica y un osciloscopio analógico de alta persistencia para registrar eventos de calidad de energía. En este caso lo que se modificó es la forma en que las mediciones eran registradas pero los datos necesariamente debían ser interpretados visualmente por un operario por lo que la inteligencia de este dispositivo también recaía en la pericia de los técnicos que operaban el artefacto.

En los 70 aparecieron los equipos de monitoreo de eventos de calidad de energía que proporcionaban indicadores basados en texto, caracterizando una perturbación de acuerdo a su tipo y duración. Estos equipos gradualmente comenzaron a incorporar herramientas para la visualización gráfica y resúmenes estadísticos de los eventos capturados.

A mediados de 1980 los instrumentos comenzaron a contar con la capacidad de muestreo en el orden de los mega Hertz [MHz] para proveer información detallada de los transitorios de altas frecuencias.

Desde la década del 2000 a la actualidad los dispositivos de monitoreo incorporan procesadores digitales cada día con mayor potencia de cálculo. Además la capacidad de almacenamiento crece continuamente, la integración con una computadora ya está estandarizada y pueden encontrarse en el mercado cualquier interfaz de comunicación para incorporar los dispositivos de monitoreo a los distintos tipos de red. Estos adelantos se realizaron con el objetivo de mejorar la interoperabilidad de los equipos y fabricantes mediante la estandarización de interfaces, base de datos e instrumentación para permitir mayor flexibilidad y reducir costos.

1.2.3 Clasificación de sistemas de monitoreo

Una tipificación comúnmente empleada para los sistemas de monitoreo es la división entre equipos portables y equipos fijamente instalados. La implementación de cada uno de ellos depende de cada necesidad en particular.

- Equipos portables: Para solucionar problemas en instalaciones edilicias, industrias o eventos puntuales en sistemas de distribución se suelen emplear dispositivos de medición portable ya que las necesidades de almacenamiento de información para ciertas tareas no son determinantes.

- Equipos de monitoreo permanente: Estos equipos son empleados para la gestión de grandes sistemas eléctricos de potencia donde se necesita realizar la recolección de información de rendimiento y confiabilidad de las instalaciones eléctricas para proveer funcionalidades de alarmas ante la aparición de eventos de calidad de energía. Para este tipo de soluciones, el sistema de monitoreo suele ser integrado con los sistemas de control de procesos para correlacionar información de los eventos de calidad de energía y determinar su impacto en los procesos productivos.

De acuerdo a la complejidad y a las funcionalidades proporcionadas por los sistemas de monitoreo se puede realizar la siguiente clasificación:

- Indicadores de Evento: Los indicadores de eventos son los más simples y baratos de los dispositivos de monitoreo de perturbaciones. Estos dispositivos son capaces de clasificar varios tipos de eventos mostrando su ocurrencia a través de luces indicadoras, gráficos de barra, alarmas audibles o una combinación de estas tres funcionalidades. Este tipo de dispositivos generalmente no es capaz de almacenar los tiempos de inicio y duración del evento.
- Osciloscopios: Los osciloscopios digitales pueden almacenar formas de onda de voltaje y corriente. Algunos osciloscopios digitales permiten el cálculo de parámetros básicos como el valor pico, el valor medio, el valor eficaz, entre otros. Las formas de onda se almacenan manteniendo un sincronismo temporal entre las señales de tensión y corriente. Generalmente los osciloscopios son diseñados para que el usuario intervenga después de realizar una medición y no son capaces de capturar y almacenar repetitivamente eventos de calidad de energía de manera automática.
- Sistemas de monitoreo de calidad de energía: Un sistema de monitoreo de eventos de calidad de energía está conformado por una serie de dispositivos permanentemente instalados que cumplen con la función de adquisición y procesamiento de información.

Debido a que el presente trabajo de tesis se enfoca en los sistemas de monitoreo (principalmente en la etapa de procesamiento de información) este tipo de arreglo será abordado más profundamente en la siguiente sección.

1.2.4 Sistemas de monitoreo de calidad de energía

La Figura 1. 11 muestra un diagrama funcional de un sistema de monitoreo de calidad de energía.

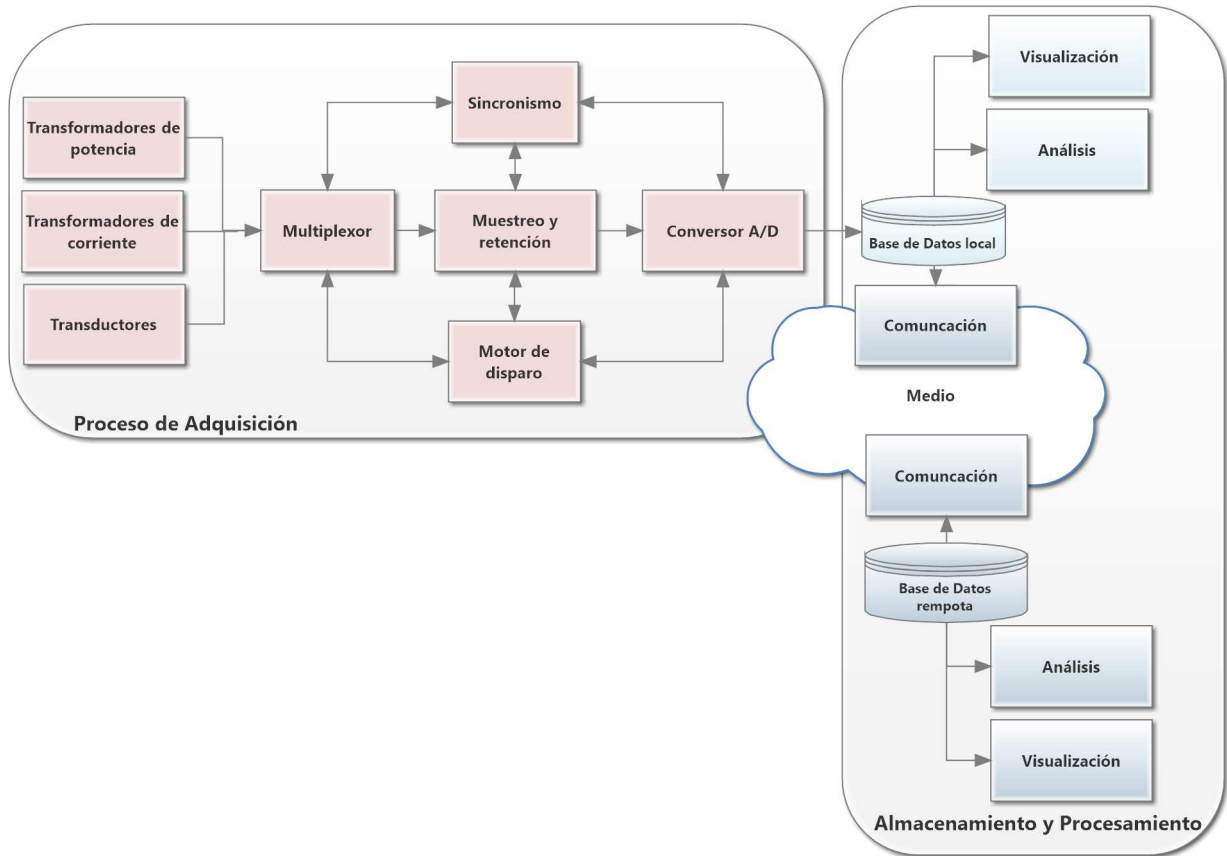


Figura 1. 11 - Sistema de monitoreo.

Según lo que se muestra en la Figura 1. 11 un sistema de monitoreo está conformado por dos bloques de procesos:

- **Procesos de adquisición de datos:** Aunque el presente trabajo se enfoca en el estudio de diferentes técnicas de procesamiento de señales es importante destacar que un sistema de monitoreo involucra equipamiento de adquisición de datos cuyas prestaciones son determinantes para el procesamiento de la información. Esta etapa del sistema de monitoreo está conformada por transformadores de potencia, transformadores de corriente, transductores, muestreadores, conversores Analógico/Digital, dispositivos de sincronismo, multiplexores y complejos algoritmos de disparo.
- **Procesos de análisis, almacenamiento y visualización de los datos:** Esta etapa de un sistema de monitoreo está conformada por el hardware y software de almacenamiento, procesamiento y visualización de la información relacionada con los eventos de calidad de energía en un sistema eléctrico de potencia. La Figura 1. 11 muestra que las actividades relacionadas con el análisis y visualización de la información se puede realizar localmente o de manera remota empleando redes de comunicaciones. El proceso de *Análisis de la información* obtiene la información de la base de datos generada por el *proceso de adquisición*

de datos, extrae y calcula parámetros relacionados a los eventos de calidad de energía (tiempos de ocurrencia, duración, magnitud, energía, etc.) y prepara la información resultante del análisis para ser almacenada y/o graficada.

Debido a que este trabajo de tesis se enfoca principalmente en el análisis y comparación de diferentes técnicas de procesamiento de datos para la detección y clasificación de eventos de calidad de energía, en la sección 0, se abordará con más detalle el proceso *Análisis* mostrado en la Figura 1. 11.

1.3 PROCESO DE ANALISIS DE LOS DATOS

Los datos son información sin contexto. La adquisición de los datos es una materia importante para un sistema de monitoreo, pero más importante aún es transformar esos datos en información útil para el usuario y presentarla de manera precisa y amigable. La mayoría de los instrumentos de monitoreo colectan información de manera continua con altas velocidades de muestreo. Si no existiera una herramienta para extraer información de los datos colectados los sistemas de almacenamiento rápidamente colapsarían.

El proceso de *Análisis* de los datos obtenidos del muestreo de las formas de onda cumple la importante función de extraer información útil de los datos colectados por el proceso de adquisición.

El proceso de *Análisis* está conformado por las etapas que se muestran en la Figura 1. 12:

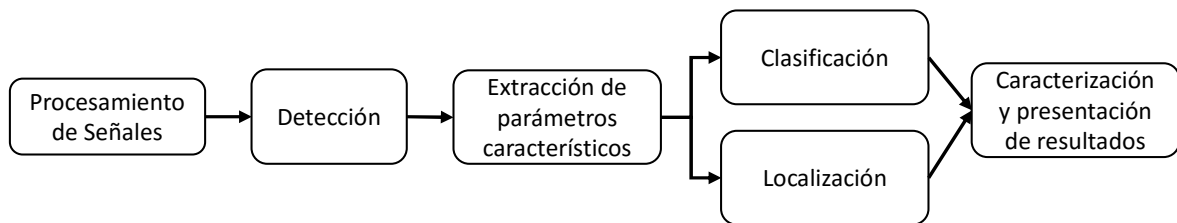


Figura 1. 12 - Análisis de las muestras.

En las siguientes sub secciones se describe las principales funcionalidades de cada módulo.

1.3.1 Procesamiento de Señales.

El módulo de procesamiento de señales es uno de los más importantes dentro del sistema de análisis de perturbaciones ya que su elección definirá aspectos como:

- La vulnerabilidad del clasificador a la presencia de ruido en las muestras.
- La capacidad del módulo de *extracción de parámetros característicos* para reducir la dimensión de los datos a ser procesados por las etapas posteriores.
- Los procedimientos para la detección de perturbaciones estacionarias y de corta duración.
- Los parámetros característicos que podrán ser calculados.
- El costo computacional y los tiempos de procesamiento.
- La habilidad del clasificador para reconocer perturbaciones a través de los parámetros característicos.

Varias técnicas de procesamiento han sido propuestas por la comunidad científica dentro de las cuales se destaca la *Transformada de Fourier (FT)*, la *Transformada rápida de Fourier (FFT)*, la *Transformada de Fourier de tiempo corto (STFT)*, la *Transformada de Stockwell*

(*ST*), *Transformada de Hilbert (HT)*, *Transformada Wavelet (WT)*, *Paquete de Transformadas Wavelet (WPT)*, *Filtrado de Kalman (KF)*, *Transformada de Gabor (GT)*, *Transformada de Sinusoides amortiguadas (DST)*, cálculo del valor cuadrático medio rms.

En Figura 1. 13 la se muestra la taxonomía de las técnicas de procesamiento de señales más implementadas por la comunidad científica para el procesamiento de señales [11].

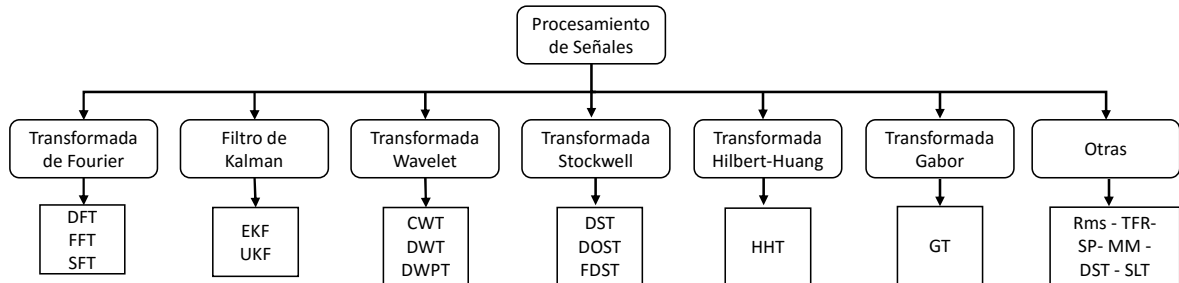


Figura 1. 13 - Taxonomía de las técnicas de procesamiento de señales.

A continuación se realizará un recorrido de las publicaciones más importantes relacionadas con la detección de anomalías en un sistema eléctrico de potencia.

La *FT* ha sido históricamente empleada para el análisis del contenido espectral de una señal de naturaleza estacionaria debido a que esta técnica solo es capaz de detectar la existencia de ciertos componentes espectrales sin brindar ninguna información del tiempo en que ese componente aparece.

Roscoe; Burt, y McDonald (2009) proponen el uso de la independientemente *FT* para la medición de perturbaciones estacionarias del tipo armónicos [12].

Por otro lado, Zhang, Geng y Yuan (2001) emplean la *Transformada ventaneada de Fourier* para disminuir el error introducido por la *FFT* en la medición de armónicos [13].

Sin embargo, las perturbaciones electromagnéticas son usualmente de naturaleza no estacionaria lo que hace que la *FT* (y sus derivadas, la *Transformada de Fourier Discreta (DFT)* y la *FFT* y la *Transformada de Fourier Continua (CFT)* sean inapropiadas para el análisis de este tipo de eventos.

La *FT* dependiente del tiempo y la *Transformada de Fourier de tiempo (STFT, por su siglas en inglés)* corto son adaptaciones derivadas de la transformada original para dotarla, en cierto modo, de una capacidad de resolución temporal.

Heydt, Fjeld, Liu y Pierce (1999) justifican el uso de la *STFT* para la detección de eventos de calidad de energía de corta duración [14].

En [15] Gu y Bollen (2000) proponen la aplicación de la *STFT* para el análisis de perturbaciones no estacionarias mediante la descomposición de señales variantes en el tiempo en componentes que residen en un plano tiempo y la frecuencia.

Como se explicará con más detalle en el *capítulo 2*, una vez que se selecciona el tamaño y forma de la ventana, la resolución de la STFT queda completamente definida. Esto representa una desventaja notoria frente a otras técnicas de procesamiento de señales.

El *KF* es una herramienta de procesamiento de señales muy empleada que se vale un modelo de estado para la estimación de la amplitud, ángulo, fase y frecuencia de cada armónico por medio de una descomposición de su envolvente en elementos constantes y elementos fluctuantes.

En [16] los autores proponen un método que combina la *Transformada Wavelet de tiempo discreto (DWT)* y un *KF* con un *Sistema Experto Difuso (FES, por sus siglas en inglés)* para la identificación y clasificación de perturbaciones.

Caujolle, Petit, Fleury y Berthet (2010) emplean el *KF* [17] para la detección de perturbaciones en sistemas eléctricos de potencia.

La *DWT* modificada, comúnmente denominada *Transformada Stockwell* o simplemente *Transformada S (ST)*, fue propuesta por Stockwell; Mansinha y Lowe (1996) como una alternativa a la *DWT* [18]. Al igual que la *WT*, esta transformación permite una gran representación simultánea en el dominio del tiempo y la frecuencia.

Posteriormente Dash, Panigrahi y Panda (2003) emplearon esta técnica como una herramienta potente para la detección, localización y clasificación de problemas relacionados con la calidad de energía [19]. Siguiendo la misma línea de investigación Fengzhan y Rengang (2007) demostraron que la *ST* es un método rápido, eficiente, escalable y muy poco sensible al ruido [20].

En [21] Dash (2004) propone el uso de la *ST* en combinación con el *KF* para detección y clasificación de perturbaciones de corta duración.

Tal vez una de las mayores críticas que suele hacerse a la *ST* radica en sus altos costos de cómputo. En [22] Biswall (2013) proponen utilizar una modificación a la *ST* que reduce considerablemente la carga computacional. Esta técnica es conocida como *fast Discrete S Transform (FDST)*.

Erişti, (2014) [23] propuso un clasificador que emplea la *ST* para la extracción de parámetros característicos en combinación con Extreme Learning Machine (*ELM*) para la clasificación de perturbaciones.

En [24] Kumar (2015) presenta un algoritmo que utiliza la *ST* para la extracción de parámetros característicos para la clasificación de perturbaciones complejas.

Una recientemente nueva y prometedora técnica de procesamiento de señales para el análisis de señales no estacionarias se conoce como *Transformada Hilber-Huang (HHT)*. Esta metodología emplea una combinación de un procedimiento conocido como *Descomposición en Modos Empíricos (EMD)* y la *HT* [25]. Biswal (2014) [26] emplea la transformada *HHT* para la extracción de parámetros característicos de perturbaciones electromagnéticas.

La *Transformada Gabor (GT)* es otra transformada destacada empleada para una estimación fasorial precisa.

Soo-Hwan Cho, Gilsoo Jang y Sae-Hyuk Kwon (2010) emplean la *Transformada Gabor-Wigner (GWT)* para la detección de fenómenos estacionarios y de corta duración [27].

Los autores Kawady y Elkalashy [28] (2014) proponen la utilización de la *GT* para la detección de arcos de tensión.

Conjuntamente con la *ST*, la técnica de procesamiento de señales más extensamente aceptada por la comunidad científica es la *Transformada Wavelet*. Existen muchas razones que hace que la *WT* ocupe un lugar de privilegio en la comunidad científica desde hace muchas décadas. Se trata de una técnica que tiene una buena resolución tiempo-frecuencia, permite descomponer la señales en diferentes coeficientes y no es costosa computacionalmente.

En el *capítulo 2* de este trabajo se profundizará en el estudio de la *WT* como método para la detección de perturbaciones en un sistema eléctrico de potencia y la extracción de parámetros que resultan de utilidad para la etapa de clasificación. También se realizará una más detallada revisión bibliográfica de los trabajos que adoptan esta transformada como método de procesamiento de señales.

1.3.2 Detección.

Los eventos de calidad de energía se presentan de manera aleatoria y no existen métodos para predecir su ocurrencia.

Debido a la naturaleza azarosa de las perturbaciones un sistema de monitoreo debe continuamente muestrear y almacenar información de las formas de onda aunque estas no contengan perturbación alguna.

Para que la base de datos del sistema de monitoreo no se vea inundada de información no relevante es preciso realizar el procesamiento de las muestras para descartar la información que represente al sistema en un estado de operación normal y almacenar solo aquellos datos que representan a una perturbación.

Un sistema de monitoreo contará, por lo tanto, con un módulo denominado módulo o etapa de *detección* cuya función principal es la de identificar la presencia de una perturbación dentro de un conjunto de datos almacenados en una base de datos.

Los primeros sistemas de monitoreo se basaban métodos de disparo soportados por umbrales rígidos y pre configurados. Un umbral rígido era difícil de configurar ya que si el mismo era establecido con un valor muy alto una gran cantidad de eventos de calidad de energía ocurrían sin ser detectados por el sistema de monitoreo. Por el contrario si el umbral se configuraba muy bajo se detectaba una enorme cantidad de eventos de calidad de energía y los usuarios de los sistemas debían lidiar con grandes cantidades de información para realizar la discriminación entre los distintos tipos de perturbaciones por medio de la inspección

visual. Esta primera generación de sistemas de monitoreo fue reemplazada por sistemas con umbrales de detección variables en el tiempo y con sistemas de clasificación automáticos.

Un algoritmo de detección consiste, principalmente, en el monitoreo de ciertos parámetros que se presentan solo cuando un evento ocurre. Por este motivo, el método de detección está estrechamente vinculado con la técnica de procesamiento de señales seleccionada.

La exploración del coeficiente d_l de la *WT*, el análisis del perfil rms de una señal, el estudio de la energía contenida en los coeficientes de la *FT*, son algunas de las técnicas más populares.

El alcance de esta tesis no abarca el estudio exhaustivo de las técnicas de detección de perturbaciones, sin embargo, a modo de referencia, se presentan algunos de los trabajos que emplean la *WT* como método de detección.

En los trabajos de Santoso (1996) [29], Robertson (1996) [30] y Pillay (1996) [31] proponen, por primera vez, emplear el análisis de los coeficientes de la *WT* para la detección de perturbaciones no estacionarias.

En Weiming (2006) [32] propone utilizar los *Paquetes de Transformadas Wavelets (WPT)* para la detección y clasificación de perturbaciones.

1.3.3 Extracción de parámetros característicos.

La performance de un sistema de análisis de perturbaciones depende, en gran medida, del vector de parámetros característicos empleado para describir los eventos que se desean clasificar.

Los parámetros estadísticos extraídos se calculan con el objetivo de reducir el tamaño o dimensión de los datos para obtener parámetros distintivos de las perturbaciones. De esta manera, una señal senoidal pura, compuesta por ejemplo de 4000 muestras, puede ser descripta por solo tres parámetros: la amplitud, frecuencia y fase.

En señales no determinísticas (como es el caso de las perturbaciones) los parámetros empleados para describir un evento no pueden obtenerse de manera tan sencilla, sin embargo la idea principal de la extracción de parámetros característicos es idéntica: Describir una señal, de la manera más precisa posible, empleando la menor cantidad posible de parámetros.

Los estadísticos más ampliamente empleados en la órbita de calidad de energía son: *energía (E)*, *entropía (H)*, *valor mínimo*, *valor máximo*, *desviación estándar (DSt)*, *media*, *mediana*, *rms*, o una combinación de estos parámetros [33].

Nuevamente, los estadísticos empleados para describir una perturbación dependerán en gran medida de la técnica de procesamiento de señales adoptada.

Así, si se emplea la *FT* en la etapa de procesamiento de señales, un buen descriptor puede ser la energía contenida en una determinada frecuencia. Del mismo modo, la *WT*, permite extraer parámetros de cada uno de sus coeficientes de aproximación y detalles.

A pesar que no existe una discusión amplia de cuáles son los parámetros característicos que mejor describen a una perturbación, escasas investigaciones se han realizado para optimizar técnicas para la selección óptima de parámetros característicos [11].

Uno de los principales aportes de este trabajo de tesis será implementar técnicas que permitan optimizar la extracción de los parámetros característicos en función de cada una de las clases presentes en el universo de perturbaciones.

1.3.4 Clasificación de eventos de calidad de energía

La etapa de detección es la encargada de descubrir la presencia de una anomalía en una señal de tensión o corriente pero es el proceso de clasificación el que realiza el diagnóstico acerca del tipo de evento presente en las muestras.

El proceso de clasificación, por lo tanto, relaciona un conjunto de parámetros característicos de entrada con un conjunto de patrones de perturbaciones conocido.

Una revisión recientemente publicada [11] presenta una taxonomía de los clasificadores que utilizan técnicas de Inteligencia Artificial. Esta clasificación se presenta en la Figura 1. 14.

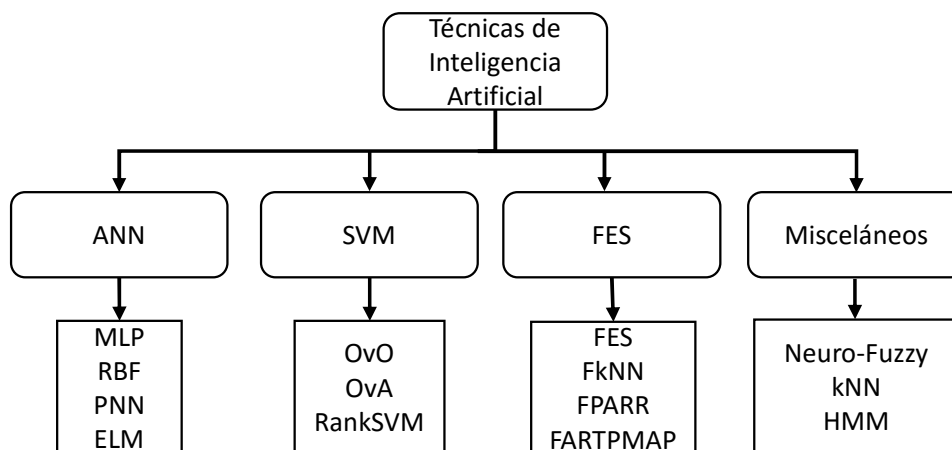


Figura 1. 14 - Taxonomía de los clasificadores basados en inteligencia artificial.

A continuación se realizará una breve reseña de las técnicas empleadas por la comunidad científica para la clasificación de perturbaciones en un sistema eléctrico de potencia.

El trabajo de Gosh and Lubkeman (1995) proponen una comparación donde se concluye que la red *Time Delay Neural Network (TDNN)* es mucho más precisa que la red *Feedforward Neural Network (FFNN)* para la clasificación de perturbaciones [34].

Surya Santoso (2000) fue uno de los primeros en utilizar un esquema combinado basado en la *WT* para la extracción de parámetros característicos y *Redes Neuronales Artificiales (ANN, por sus siglas en inglés)* para la clasificación de perturbaciones. A diferencia del trabajo presentado por Gosh and Lubkeman, Santoso propuso un esquema donde los parámetros característicos se obtienen en el dominio de la *WT* y se utiliza un esquema de múltiples *ANN*,

por sus siglas en i del tipo *Learning Vector Quantization (LVQ por sus siglas en inglés)* para la clasificación de patrones. Su trabajo fue publicado en dos artículos, el primero enfocado en los aspectos teóricos [35] y el segundo apuntó a una implementación práctica de los algoritmos [36].

Una de las principales falencias de un sistema basado en *ANN* radica en su escasa flexibilidad para adaptarse a nuevas perturbaciones. Por este motivo Wijayakulasooriya, Putrus y Minns presentaron un clasificador que utiliza un topo de redes *ANN* Auto Organizadas denominado *Self Adapting Neural Networks (SANN)* (2002) [37].

Monedero, Leon, Roper, Garcia, Elena y Montano (2007) desarrollaron un sistema de monitoreo basado en la *WT* y una arquitectura de cuatro redes del tipo *Perceptron Multi Layer Neural Network* para la clasificación de eventos de calidad de energía en tiempo real. El trabajo aborda metodologías y técnicas de bajo consumo computacional para lograr una aplicación en tiempo real [38].

El trabajo de Reaz, Choong, Sulaiman Mohd-Yasin y Kamada (2007) presenta un clasificador que utiliza *ANN* aleatoriamente optimizadas conocidas como *Univariate Random Optimized Neural Network (URONN)* en combinación con la *WT* y tecnologías basadas en *Lógica Difusa (FL, por sus siglas en inglés)* [39].

Otro tipo de red neuronal denominado *Proabilistic Neural Network (PNN)* se emplea en un sistema de clasificación desarrollado por Mishra, Bhende y Panigrahi (2008). Este tipo de red neuronal se implementa utilizando un modelo probabilístico conocido como proceso Bayesiano. Las ventajas de este tipo de red neuronal es que no requiere procesos de aprendizaje y no necesita una configuración previa de sus pesos y, si el número de muestras es lo suficientemente grande, la red neuronal converge a un clasificador bayesiano [40].

La publicación presentada por Cheng (2012) donde se emplea una red neuronal adaptiva línea comúnmente conocida como *Adaptive Lineal Neural Network (ADALINE)*. El empleo de este tipo de red neuronal se ve justificado por su mecanismo de ajuste online, su fácil implementación y sus características propicias para la cancelación de ruido [41].

Finalmente, es interesante tener en cuenta una configuración rápida y precisa conocida como *Extreme Learning Machine (ELM)*. En un trabajo [23] publicado en 2014, Erişti logra excelentes resultados en cuanto a tiempo y precisión alcanzada.

Jaehak, Powers, Grady y Bhatt (2002) propusieron un método de clasificación que combina un sistema basado en reglas, la *WT* y el *Modelo oculto de Markov* [42].

Una combinación de técnicas basadas en *Redes Neuronales (ANN, por sus siglas en inglés)* y *FL* fue discutido en el trabajo de Jiansheng, Negnevitsky y. Nguyen (2002) [43].

En el trabajo de Reaz, Choong, Sulaiman (2007) se presenta un esquema que emplea la *WT* para la extracción de parámetros característicos de la señal perturbada, las *ANN* para la clasificación de patrones y un post-procesamiento basado en *FL* para mejorar la precisión del proceso de clasificación. [39].

Una clase importante de algoritmos de I se conoce como algoritmos de *FL*. Esa técnica extiende la lógica binaria clásica mediante el razonamiento bajo incertidumbre. Este tipo de lógica toma dos valores aleatorios pero contextualizados entre sí y toma decisiones basándose en ese contexto. Las aplicaciones particulares de la *AI* donde se aplican estos conceptos se conocen como *Sistemas Expertos Difusos (FES)*.

Dash (2000) [44] propone una técnica basada en la *FT* para la extracción de la amplitud y fase de cada componente fundamental y un *Fuzzy Expert System* para el reconocimiento de cada tipo de perturbación.

Liao (2004) [45] propone la implementación de un *FES* para la toma de decisiones basándose en parámetros característicos extraídos de la *FT* y la *WT*.

Un sistema híbrido para la detección de la operación en isla en un entorno de generación distribuida fue propuesto por Dash (2012) [46].

Existen varios métodos misceláneos para la clasificación de perturbaciones.

En su trabajo Panigrahi y Pandi (2009) presentan un sistema automático de clasificación que emplea un algoritmo genético denominado *fuzzi k-nearest neighbour* para seleccionar las características que mejor caracterizan a un conjunto de perturbaciones.

Jaehak propone el empleo de un método que combina un algoritmo basado en reglas con modelos markovianos, más precisamente un modelo conocido como *Hidden Markov Model* para la clasificación de perturbaciones [42].

En los últimos años la comunidad científica ha mostrado creciente interés en la aplicación de las *Máquinas de Vectores de Soporte (Support Vector Machine)* para la clasificación de eventos de calidad de energía.

En el *capítulo 3* de este trabajo se abordará con mayor profundidad la fundamentación teórica de esta técnica así como su empleo en un sistema de monitoreo y localización de fallas.

1.3.5 Localización de fallas en un sistema eléctrico.

En las últimas décadas se ha venido produciendo un rápido crecimiento de las redes eléctricas alrededor del mundo. La desregulación de los mercados energéticos ha incrementado la presión ejercida por parte de los usuarios de contar con una ininterrumpida y confiable fuente de energía eléctrica.

Uno de los factores más importantes que atentan contra la calidad la energía eléctrica percibida por el usuario son las *fallas en el suministro*. Cualquier flujo de corriente anormal en un sistema eléctrico de potencia se considera una *falla*. Estas fallas no pueden ser completamente evitables ya que muchas de ellas dependen de impredecibles causas naturales o errores humanos.

Las fallas pueden ser de varios tipos como, por ejemplo, transitorias, persistentes, simétricas o asimétricas [47]. A su vez, una falla puede ocurrir en un sistema eléctrico de transmisión

o distribución, cuya topología varía radicalmente de uno a otro. Por lo tanto no existe un procedimiento universal para la detección, clasificación y localización de una falla.

Una cosa es cierta, la precisión con que la falla sea localizada será de gran importancia a la hora de despachar los operarios al sitio para realizar las reparaciones pertinentes y así restablecer el servicio.

En este sentido, el Departamento de Energía de los Estados Unidos presentó un documento, conocido como “*Grid 2030 – A national vision for electricity’s second 100 years*” [48] destinado a presentar tendencias de la evolución y necesidades de una red eléctrica a mediano y largo plazo. Este documento hace especial referencia en la en la necesidad de mejorar los parámetros de confiabilidad de una red de distribución.

Por lo tanto, la habilidad para localizar fallas de manera confiable y con alto grado de precisión, será un importante aporte para la mejora de los índices que reflejan la calidad de energía percibida por los usuarios [49].

Un sistema de localización de fallas es un conjunto de software y hardware que actúan con el objetivo de detectar cualquier clase de flujo de corriente anormal en el sistema eléctrico, identificar el tipo de evento que lo genera y luego, lo más precisamente posible, localizar la posición de la falla en el sistema eléctrico de potencia [5].

Tradicionalmente, en una subestación, unidades terminales remotas (*Remote Terminal Units RTU*) adquieren periódicamente mediciones analógicas y digitales (voltajes y corrientes, flujo de ramas, frecuencia, estado de los seccionadores, posición de los tap de los transformadores, etc). Estas mediciones son colectadas a través de un sistema de control y adquisición de datos (SCADA) las cuales son enviadas a los sistemas de administración de energía (EMS) cada dos a diez segundos.

Sin embargo, con el rápido avance de la tecnología, otros dispositivos inteligentes comenzaron a reemplazar a las antiguas RTU. Estos dispositivos se conocen como IED (*en inglés Intelligent Electronic Devices*). Los IED, muy parecidos a una pequeña computadora, están dotados con la capacidad de adquirir y almacenar una importante cantidad de información con una periodicidad que dependerá de la función a la que están destinados.

Los operadores de las redes del futuro contarán con mayor cantidad de información la que, en caso de ser apropiadamente procesada, podrá acarrear beneficios a la hora de tratar de predecir, monitorear y analizar eventos en el sistema eléctrico.

Esta tendencia pone de manifiesto la necesidad de adaptar los algoritmos de localización de fallas para que sean capaces de procesar mayor cantidad de información y, teniendo en cuenta los parámetros de la red, su topología y el tipo de datos adquiridos, ser capaces de localizar eventos de manera rápida y precisa.

Como ya se mencionó, no existe un método universal para la localización de fallas. La Figura 1. 15 muestra una clasificación de estos sistemas.

De acuerdo al tipo de datos procesados, un localizador de fallas se puede fundamentar en mediciones *fasoriales* o mediciones en el *dominio del tiempo*.

Luego, una segunda clasificación se realiza de acuerdo a la cantidad de nodos de medición empleados para obtener los resultados.

Los métodos que emplean un único nodo de medición se conocen como *single-end methods*. Estos algoritmos suelen ser menos precisos que los métodos que emplean múltiples mediciones, también conocidos como *double-end methods*. Sin embargo, los métodos que emplean un único nodo de medición necesitan una menor cantidad de *IED*, los requerimientos de comunicación son inferiores y no requieren de tecnología adicional para sincronizar las muestras o los fasores.

En un sistema de transmisión, la longitud de las líneas puede alcanzar desde la decena hasta cientos de kilómetros. Por lo tanto, identificar el sitio donde la falla se localiza, para este tipo de redes, suele ser de gran importancia. Los métodos *double-end* suelen ser mayormente empleados en los sistemas de transmisión ya que permiten mayor precisión para localizar la falla y, para un administrador de un sistema de transmisión, los costos de despliegue de múltiples *IED* suelen ser más justificables en comparación con los presupuestos manejados por un administrador de un sistema de distribución.

En un sistema de distribución, que no presente cableado subterráneo y la longitud de sus líneas no sea excesiva, identificar la línea, rama o zona donde la falla se produce, sin la necesidad de incorporar un medidor en cada extremo de rama es la solución que mejor se adapta.

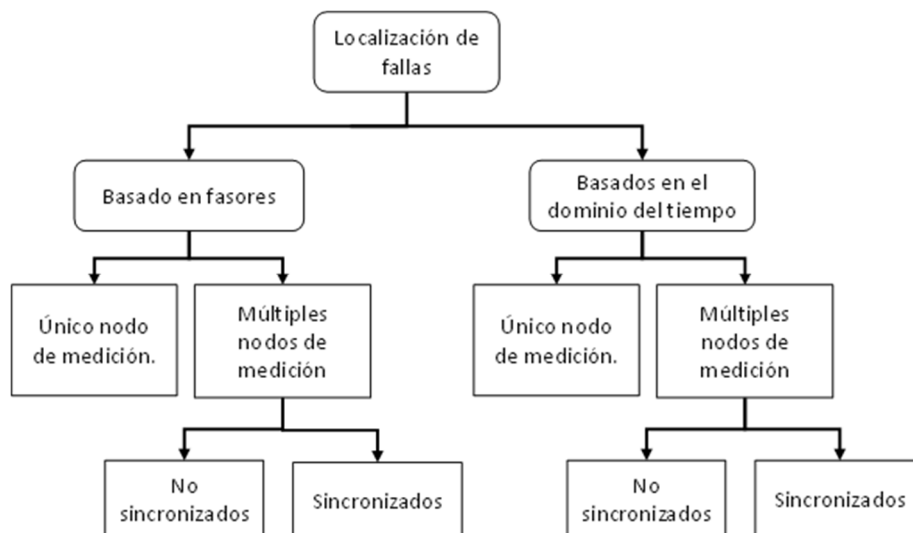


Figura 1. 15 - Clasificación de sistemas de localización de fallas.

La mayoría de los algoritmos de localización de fallas desarrollados por la comunidad científica se enfocaron en la red de transmisión [50, 51]. Estos desarrollos no pueden ser directamente adaptados a la red de distribución ya que esta última presenta una topología radial, cuenta con una cantidad limitada de dispositivos de medición, no siempre se tiene

información centralizada acerca del estado de seccionadores e interruptores en todas las ramas que la conforman, las longitudes de las líneas son menores y las mediciones en diferentes momentos del día pueden cambiar debido a la naturaleza variante de la carga y el cambiante número de generadores conectados al sistema.

Por estos motivos los algoritmos de localización de fallas para un sistema de distribución difieren de los desarrollados para un sistema de transmisión.

De acuerdo a la revisión realizada por Kezunovich (2011) [49], los algoritmos de localización de falla en un sistema eléctrico de distribución pueden clasificarse en:

1. Basados en la impedancia aparente [52, 53].
2. Análisis directo del circuito trifásico [54, 55].
3. Componentes sobre impresas [56].
4. Ondas viajeras [57-59].
5. Inteligencia artificial (AI) [60-69].

Un sistema de localización de fallas basado en técnicas de *AI* procesa las muestras adquiridas en un determinado nodo de la red eléctrica con el objetivo de clasificar el fenómeno de acuerdo a una serie de clases pre-establecidas e identificar la zona geográfica donde el evento se produce.

En este sentido, un sistema de localización de fallas está conformado por las siguientes etapas:

- Clasificador del tipo de falla: Como su nombre lo indica, este clasificador será el responsable de identificar si la falla es del tipo fase a tierra, fase a fase, dos fases a tierra, o tres fases a tierra.
- Localizador de la falla: El módulo de localización tiene por objetivo distinguir la línea y fase del sistema eléctrico de potencia donde más probablemente la falla se encuentra de acuerdo a un conjunto de mediciones de campo.

Cuando una falla se produce en una línea de un sistema eléctrico, las mediciones de tensión y corriente de las líneas o barras vecinas presentan perturbaciones generalmente del tipo hueco de tensión y/o transitorios oscilatorios.

El método para la localización de fallas propuesto en esta tesis se fundamenta en el hecho que cuando una falla ocurre, las mediciones de tensión y corriente obtenidas en nodos estratégicos de la red presentarán características particulares según la ubicación geográfica donde la falla se produce. En este sentido, si una falla se produce en cercanías del punto de medición los datos adquiridos presentarán determinadas características que serán marcadamente diferentes a aquellos obtenidos para un escenario en donde la falla se genera en nodos más alejados del punto de medición.

1.3.6 Caracterización y presentación de resultados.

La etapa de caracterización consiste en la generación de parámetros que proporcionen una medida del estado de las ondas de tensión y corriente en un sistema eléctrico de potencia. Esta información puede ser empleada por los fabricantes para diseñar los componentes de red eléctrica, por los entes reguladores para establecer indicadores de calidad de energía, por los usuarios para realizar reclamos o bien por las empresas proveedoras de energía para efectivizar penalizaciones. El módulo de caracterización es el encargado de calcular estos indicadores basándose en las señales perturbadas.

El estudio las técnicas de caracterización de los eventos de calidad de energía se encuentran fuera del alcance de los objetivos del presente trabajo razón por la cual esta problemática no se abordará más profundamente.

1.4 CONCLUSIONES

Los conceptos que aquí se abordaron demuestran que el monitoreo de eventos de calidad de energía en un sistema eléctrico de potencia representa una herramienta fundamental para disminuir de pérdidas económicas para las empresas proveedoras de energía eléctrica y para sus clientes.

Debido a la amplia variedad de fenómenos electromagnéticos que se pueden encontrar en un sistema eléctrico de potencia es necesario adoptar un estándar que defina los tipos de perturbaciones así como también las causas que generan las perturbaciones, los instrumentos empleados medirlas y las prácticas recomendadas para limitar sus efectos. La lectura de este capítulo permitió al lector conocer los aspectos más importantes del estándar más reciente que la *IEEE* recomienda.

En el capítulo se presentó el modelo esquemático de un sistema de monitoreo convencional, destacando la etapa de adquisición de datos, la etapa de análisis, la etapa de procesamiento, almacenamiento y visualización de la información y la etapa de comunicación (para aquellos sistemas de monitoreo distribuido en grandes áreas geográficas). En este trabajo se explicaron los aspectos básicos de cada etapa realizando un enfoque más profundo en el estudio de diversas técnicas para el procesamiento de la información obtenida por la etapa de adquisición de datos.

En la última parte de este capítulo se presentó una revisión bibliográfica de la metodología utilizada por la comunidad científica para el procesamiento de las señales perturbadas y diferentes técnicas de *AI* para la detección, clasificación y localización de perturbaciones en un sistema eléctrico de potencia.

Primera Parte
Conceptos
Teóricos

Capítulo 2:

Transformada Wavelet

2.1 INTRODUCCION

En el análisis de señales existe un gran número de herramientas que se han ido desarrollando con el paso de los años, entre las que se destaca, sin lugar a dudas, la *FT*. Sin embargo, no hace más de 20 años, se han desarrollado nuevas herramientas que permiten analizar señales desde otra perspectiva, surgiendo sobre todo de la necesidad de poder analizar señales que no se comportan de manera estacionaria o que presentan cambios bruscos en intervalos de tiempo muy pequeños.

Como se presentó en el *capítulo 1* del presente trabajo de tesis existe un gran número de perturbaciones eléctricas que poseen la característica de ser no estacionarias.

Generalmente los artículos o libros que abordan la *Transformada Wavelet (WT)* lo hacen desde el punto de vista matemático, el cual es ciertamente complejo.

El objetivo de este capítulo es presentar los conceptos más importantes relacionados con la *WT*, y las características que hacen que esta herramienta sea la adecuada para la detección de fenómenos electromagnéticos en un sistema eléctrico de potencia. Los conceptos aquí adquiridos servirán al lector para comprender la función que la *Transformada Wavelet* cumple en un sistema de monitoreo.

En la primera sección de este capítulo se presenta la necesidad de contar con una herramienta de análisis de señales que permitan extraer información en el dominio del tiempo y la frecuencia de manera simultánea. Se explicarán las herramientas de procesamiento de señales tradicionales como la *FT* y la *STFT* para demostrar que estos métodos presentan falencias a la hora de procesar señales no estacionarias.

En la segunda sección del capítulo se aborda la *WT* desde un punto de vista práctico. Se presentará la *Transformada Wavelet Continua (CWT)*, la *Transformada Wavelet Discreta (DWT)* y la representación alternativa de la *DWT* como un banco de filtros seguidos por un proceso de diezmado.

Finalmente se presentará una revisión de los trabajos más destacados que utilizan la *WT* para el análisis de eventos de calidad de energía.

2.2 TRANSFORMADA WAVELET

2.2.1 Conceptos previos

2.2.1.1 Clasificación de señales

La primera clasificación que suele hacerse a una señal se hace de acuerdo a su comportamiento en el tiempo. Las señales pueden ser *estacionarias* o *no estacionarias*.

Se dice que una señal es *estacionaria* o *invariante en el tiempo* cuando la forma de onda es definida por sus parámetros característicos que no varían en el tiempo. Un ejemplo de señal estacionaria puede ser una función senoidal ya que la misma queda definida por un conjunto de parámetros (Amplitud, frecuencia y fase) que no se modifican en el tiempo.

Por otro lado, las señales *no estacionarias* son aquellas en las que sus parámetros característicos varían con el tiempo.

La Figura 2. 1 muestra un ejemplo de una señal estacionaria continua. La Figura 2. 2 muestra la representación de una señal no estacionaria continua:

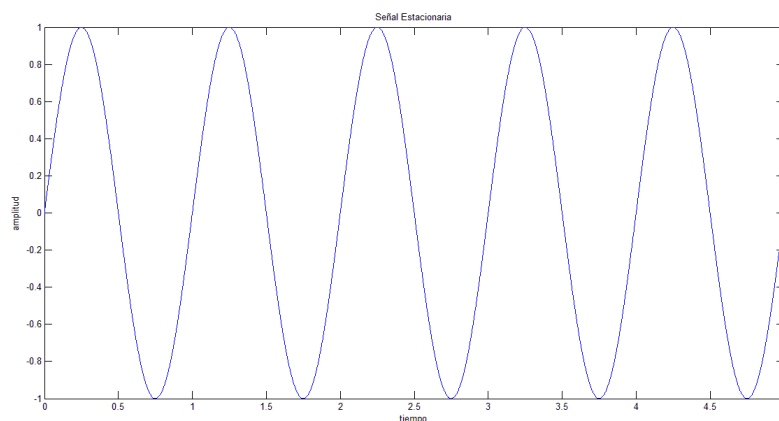


Figura 2. 1 - Señal estacionaria.

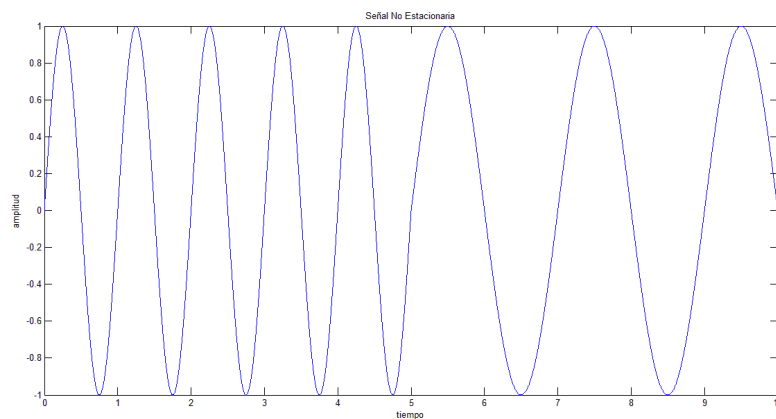


Figura 2. 2 - Señal no estacionaria.

2.2.1.2 Transformada de Fourier.

El estudio clásico del procesamiento de señales ha enfocado sus esfuerzos principalmente en el estudio de sistemas y señales de carácter estacionario. En este sentido la *FT* ha sido la herramienta de análisis más con mayor difusión para el procesamiento de señales.

En términos generales se puede definir a la *FT* como una herramienta matemática empleada para representar una señal mediante una combinación lineal de exponenciales complejas infinitesimalmente separadas en frecuencia.

Sea $f(t)$ una función no periódica cuya variable real es t , la *FT* de esta señal viene dada por:

$$\mathfrak{F}\{f(t)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-j\omega t} \delta t \quad (2.1)$$

donde ω es una variable que representa la frecuencia y se mide en radianes por segundo [rad/seg]. Usualmente la frecuencia también puede expresarse en ciclos por segundo y su símbolo representativo es la letra f . La relación entre ω y f viene dada por:

$$f = \frac{2\pi}{\omega} \left[\frac{\text{ciclos}}{\text{segundos}} \right] \quad (2.2)$$

Para obtener una señal a través de su *FT* se debe aplicar la transformada inversa de Fourier. Esta transformada viene dada por la siguiente ecuación:

$$\mathfrak{F}^{-1}\{F(\omega)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega)e^{j\omega t} \delta \omega \quad (2.3)$$

La siguiente ecuación se corresponde a la Transformada de Fourier para señales de tiempo discreto (*DTFT por sus siglas en inglés*)

$$\mathfrak{F}\{x[n]\} = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \quad (2.4)$$

La *TF*, al tratarse de una herramienta ideada para el análisis de señales estacionarias, presenta limitaciones para el análisis de fenómenos de corta duración también conocidos como fenómenos transitorios. Para aclarar este concepto es conveniente analizar los siguientes ejemplos:

La Figura 2.3 a) muestra una señal senoidal en el dominio del tiempo y la Figura 2.3 b) su *FT*. Al tratarse de una señal estacionaria, los parámetros característicos (Amplitud, frecuencia y fase) de la señal senoidal no varían en el tiempo de análisis. En este caso el espectro de frecuencias obtenido por la *FT* se considera como una representación alternativa de la señal ya que, partiendo de la información del diagrama espectral, la señal puede ser completamente reconstruida en el dominio del tiempo.

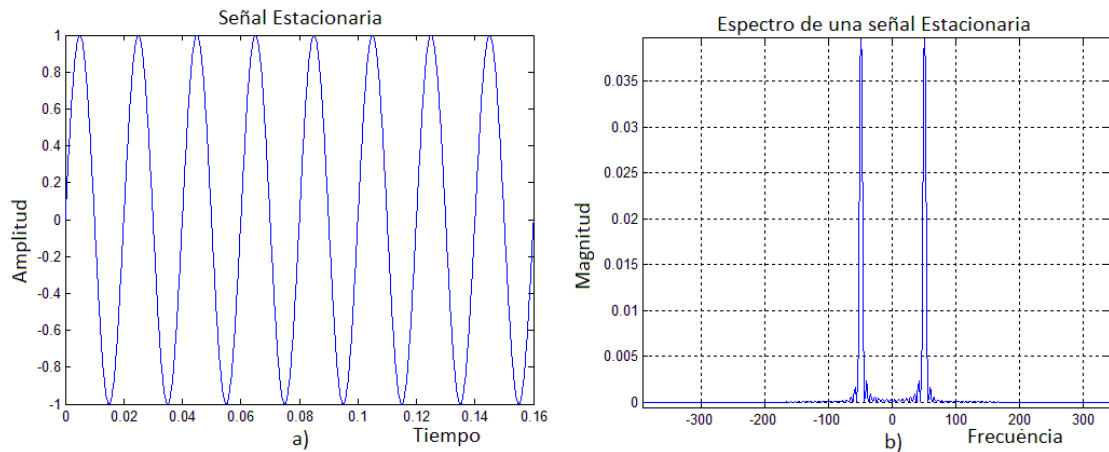


Figura 2. 3 - Señal estacionaria.

Para el segundo ejemplo se propone la señal mostrada en la Figura 2. 4 a). Esta señal es de carácter no estacionario ya que su frecuencia, originalmente de $50 [Hz]$, se desplaza hacia una frecuencia de oscilación de $300 [Hz]$ en un determinado instante de tiempo.

En el diagrama espectral de la señal, mostrado en la Figura 2. 4 b), se pueden identificar los dos pulsos de frecuencia ubicados en los $50 [Hz]$ y $300 [Hz]$. A primera vista se podría concluir que la FT representa completamente a la señal en el dominio del tiempo. Sin embargo, analizando más profundamente el ejemplo propuesto se puede notar que *el diagrama de frecuencias no presenta información del tiempo en que se produce la transición en la frecuencia de la señal.*

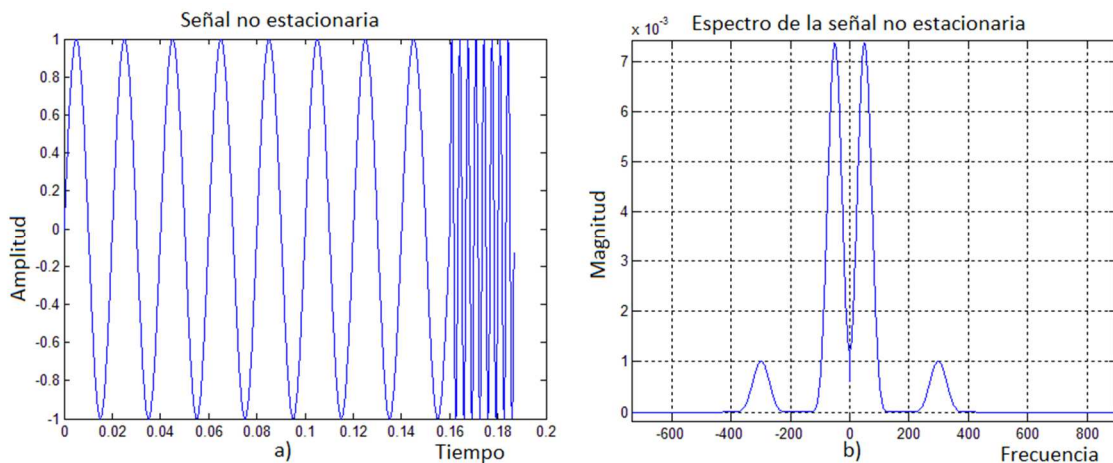


Figura 2. 4 - Señal no estacionaria.

Mediante este ejemplo se ilustra una importante limitación en la FT para representar una señal que contenga transiciones en el dominio del tiempo.

En el *capítulo 1* de este trabajo explicó que las perturbaciones en un sistema eléctrico de potencia pueden ser del tipo estacionarias (perturbaciones armónicas o fluctuaciones) o no

estacionarias (huecos, interrupciones, sobre tensión temporal, fenómenos transitorios, muescas, entre otras).

El análisis de Fourier, a pesar de ser una herramienta de gran utilidad para el procesamiento de perturbaciones estacionarias, resulta a la vista *insuficiente* para el procesamiento de perturbaciones de carácter transitorio. Por este motivo resulta claro que se necesita una herramienta matemática que permita el análisis de una señal en el dominio del tiempo y la frecuencia de manera simultánea.

Es importante aclarar que los diagramas en frecuencia mostrados en las gráficas de la Figura 2. 3 y Figura 2. 4 no se corresponden con el diagrama espectral real ya que:

- Las señales sinusoidales empleadas en el ejemplo no son señales de tiempo continuo. Por el contrario, estas señales se conforman mediante muestras de una señal senoidal graficadas como señales de tiempo continuo.
- El método obtenido para obtener el espectro de la señal utiliza técnicas de enventanado. Cualquier método que emplea una porción de la señal para obtener el espectro introduce distorsión en el resultado obtenido. Por este motivo el espectro de una señal senoidal, la cual idealmente debería corresponderse a un impulso en el dominio de la frecuencia, se muestra como un pulso que abarca más de una única componente espectral.

2.2.1.3 Transformada de Fourier dependiente del tiempo

Para salvar las limitaciones que la *FT* presenta para el análisis de señales no estacionarias *Denis Gabor (1946)* desarrolló una modificación en el método para restringir el análisis a un determinado intervalo de tiempo. Para ello se divide la señal en diferentes secciones en el tiempo y se considera que la señal es periódica dentro de cada uno de dichos intervalos. Al tratar a cada porción de la señal como si fuera periódica se puede aplicar la Transformada de Fourier tradicional y se presenta un espectro de frecuencias por cada una de las secciones en las que la señal original fue dividida. Una forma de presentar los resultados es mediante una grilla bidimensional cuyos ejes son las frecuencias y el tiempo. En esta gráfica se conoce como *espectrograma*.

Esta herramienta se denomina *Transformada de Fourier dependiente del tiempo* o (*STFT por sus siglas en inglés*) y puede ser expresada mediante la siguiente ecuación:

$$\text{STFT}(\tau, f) = \int_{-\infty}^{\infty} x(t)w(t - \tau)e^{-j\omega t} \delta t \quad (2. 5)$$

donde w se corresponde a la función empleada para el ventaneo cuya forma y longitud determinan la resolución en el dominio del tiempo y frecuencia y τ es una variable empleada para el desplazamiento de la ventana a lo largo de la señal.

Una expresión equivalente para las señales en tiempo discreto puede ser expresada como:

$$\text{STFT}[m, n] = \sum_{n=-\infty}^{\infty} x[n]w[n - m]e^{-j\omega n} \quad (2.6)$$

donde w nuevamente se corresponde a la ventana.

m es la variable empleada para trasladar la ventana a lo largo de la señal.

Debido a que una misma ventana se emplea para el análisis de la señal indistintamente de la frecuencia la resolución del método es la misma en todas las ubicaciones del plano tiempo frecuencia.

En la Figura 2. 5 se puede ver como la *STFT* mapea una señal en el dominio del tiempo al dominio tiempo-frecuencia.

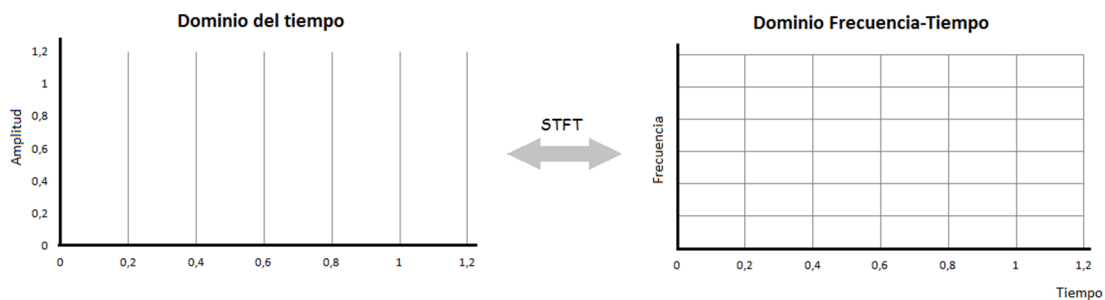


Figura 2. 5 - Transformación de dominios de la STFT.

En este momento se presenta una relación de compromiso entre la resolución en el dominio tiempo y en el dominio de la frecuencia. Si se hace el ancho de la ventana lo suficientemente grande se obtendrán resultados con una buena resolución en frecuencia pero una pobre resolución en el tiempo. Para ventanas muy grandes, la *STFT* mucho a la *TF* tradicional. Por otro lado, hacer la ventana muy estrecha en el tiempo brindará una buena resolución en el tiempo en detrimento de la información utilizada para obtener la señal en el dominio de la frecuencia. Dada la forma de una ventana es imposible ajustar la resolución del método en el dominio del tiempo sin obtener el no deseado deterioro en la resolución de los resultados en el dominio de la frecuencia. Esta definición es usualmente conocida como *el principio de incertidumbre de Heisenberg*.

Una segunda limitación de este método radica en la carga computacional producto del ventaneo solapado entre las secciones de la señal a analizar.

El análisis de perturbaciones en un SEP, las señales de tensión o corriente pueden presentar componentes en frecuencia que van desde la corriente continua a componentes de muy alta frecuencia. Esta variedad de perturbaciones hace que una técnica que utiliza una resolución fija sea inadecuada.

2.2.2 La Transformada Wavelet

La *Transformada Wavelet (WT)*, herramienta matemática desarrollada a mediados de los años '80, representa un método eficiente para el análisis local de señales no estacionarias y de rápida transitoriedad.

La clásica *FT* es una herramienta que transforma una señal del dominio del tiempo al dominio de la frecuencia por medio de una representación que emplea exponenciales complejas infinitesimalmente separadas en frecuencia. Cada coeficiente de la *FT* se puede interpretar como una proyección de la señal en cada una de las bases ortonormales formadas por exponenciales complejas.

La *WT*, por el contrario, en lugar de utilizar exponenciales complejas de duración y energía infinita, emplea un grupo de funciones base conocida como onditas o, por su traducción al inglés, *wavelets*. La *WT*, por lo tanto, se entiende como una herramienta que realiza una proyección de la señal en distintas bases ortonormales dadas por funciones especiales denominadas wavelets.

Las familias de onditas, representadas simbólicamente como $\psi(t)$, son un grupo especial de señales cuya principal característica es su flexibilidad para adaptarse a cualquier tipo de evento de corta duración.

En este capítulo se detallarán los aspectos matemáticos más relevantes de la *WT* en tiempo continuo para luego hacer extensivo el análisis a señales en tiempo discreto.

2.2.2.1 Wavelets Madre

Las wavelets son familias de funciones que se encuentran definidas en el espacio o tiempo. Cada componente de una familia de wavelets se genera a partir de una función denominada *wavelet madre* a través de las variables de escala (a) y la variable de traslación en el tiempo (b).

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2.7)$$

Las wavelets madre son un tipo especial de señales que deben cumplir con un las siguientes condiciones:

- 1) Una wavelet debe tener energía finita:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 \delta t < \infty \quad (2.8)$$

- 2) La función debe cumplir con el criterio de la constante de admisibilidad:

$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(f)|^2}{f} \delta f < \infty \quad (2.9)$$

donde f es la frecuencia y la condición implica que la wavelet debe tener una componente de frecuencia $\psi(f = 0) = 0$ o, lo que resulta equivalente, debe tener valor medio nulo.

3) Para wavelets complejas la $FT \psi(f)$ debe ser real y desvanecida para frecuencias negativas.

Algunas de wavelet madre más comúnmente empleadas pueden verse en la Figura 2. 6, Figura 2. 7 y Figura 2. 8.

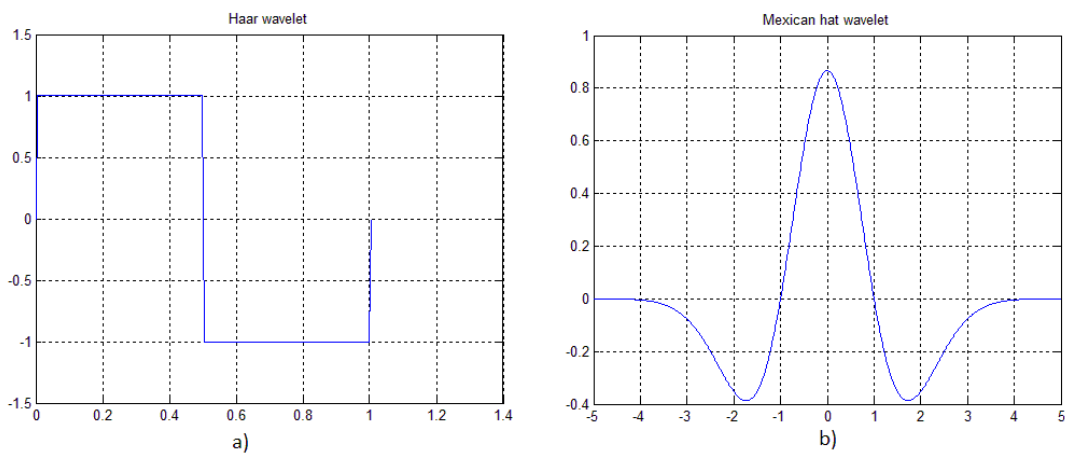


Figura 2. 6 - a) Wavelet Haar b) Wavelet Sombrero Mexicano.

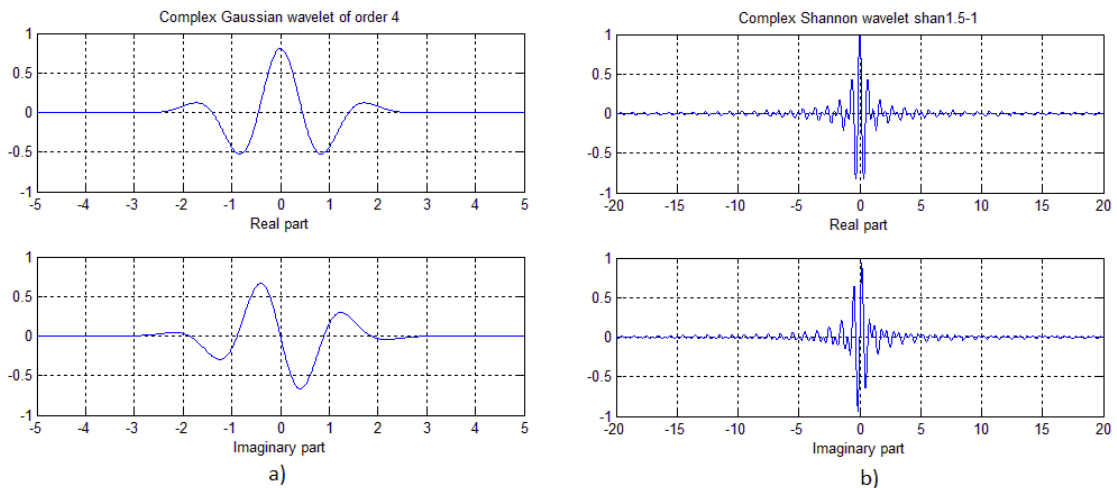


Figura 2. 7 - Wavelet Gaussiana Compleja b) Wavelet Shannon Compleja.

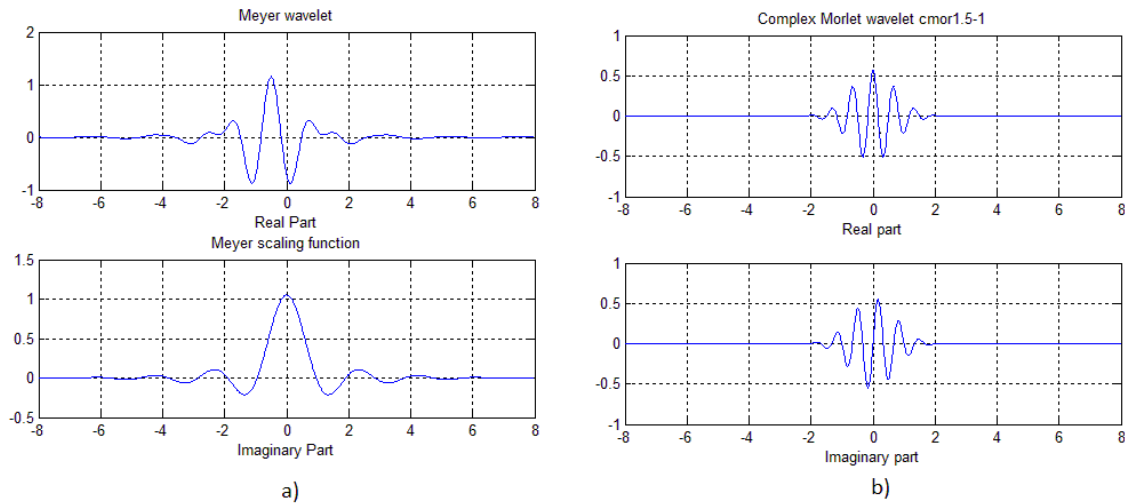


Figura 2. 8 - a) Wavelet Meyer b) Wavelet Morlet.

La selección de la wavelet madre más apropiada depende de la aplicación y no existe una regla general para realizar esta tarea. Por ejemplo, Srivastava, Siddharth y. Singh presentaron un trabajo donde se concluye que la wavelet madre Gaussiana Compleja es la más apropiada para el *análisis de perturbaciones del tipo armónico* [70].

El aporte realizado por Vega, Duarte, Ordonez y Kagan compara diferentes tipos de wavelets madre para la detección de las perturbaciones del tipo hueco, sobre tensión, transitorios, fluctuaciones y armónicos. Este trabajo concluye que la mejor wavelet madre es la Biorotogonal 3.9 [71].

Sin embargo, la wavelet madre más comúnmente empleada para la detección de eventos de calidad de energía es la Daubechie número 4 que se muestra en la Figura 2. 9:

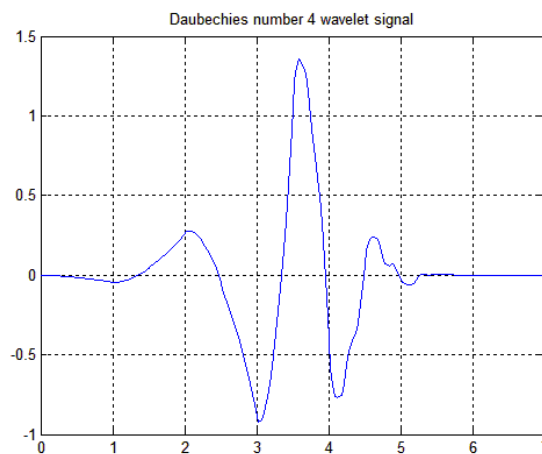


Figura 2. 9 - Wavelet madre Daubechie 4.

2.2.2.2 Expansión y traslación de una wavelet madre.

De acuerdo a lo explicado en la sección 2.2.2.1 una familia de wavelet se genera a partir de la traslación y escalamiento en el tiempo de la función wavelet madre.

La Figura 2. 10 muestra el efecto de la traslación en el tiempo de una wavelet madre del tipo Morlet. El término traslación se relaciona con la localización de la ventana a medida que ésta se desplaza a lo largo de la señal. Obviamente este término corresponde a la información del tiempo en el dominio transformado.

La Figura 2. 11 se ilustra el efecto de compresión o dilatación de la wavelet madre producto de la operación denominada escalamiento. En términos de frecuencia, las bajas frecuencias (altas escalas) corresponden a una información global que comúnmente abarca toda la señal, mientras que las altas frecuencias (escalas bajas) corresponden a una información detallada de una característica oculta en la señal, que se produce en un tiempo relativamente pequeño.

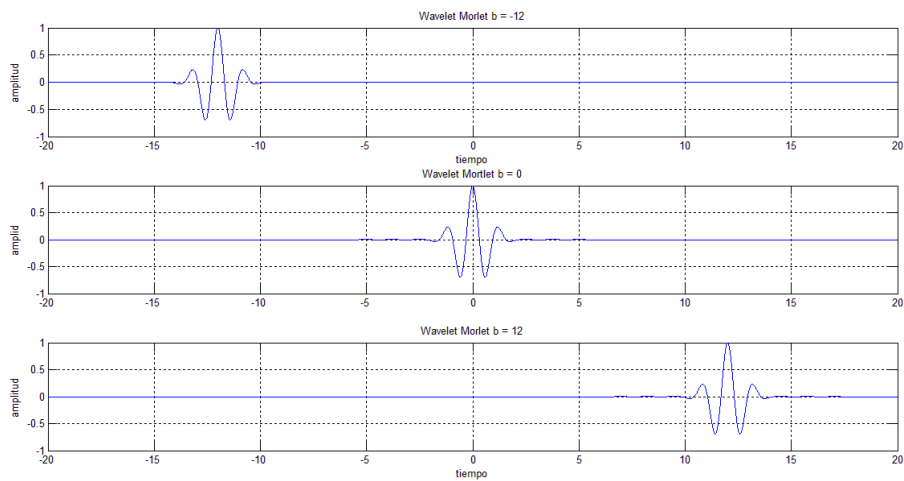


Figura 2. 10 - Traslación de la wavelet madre.

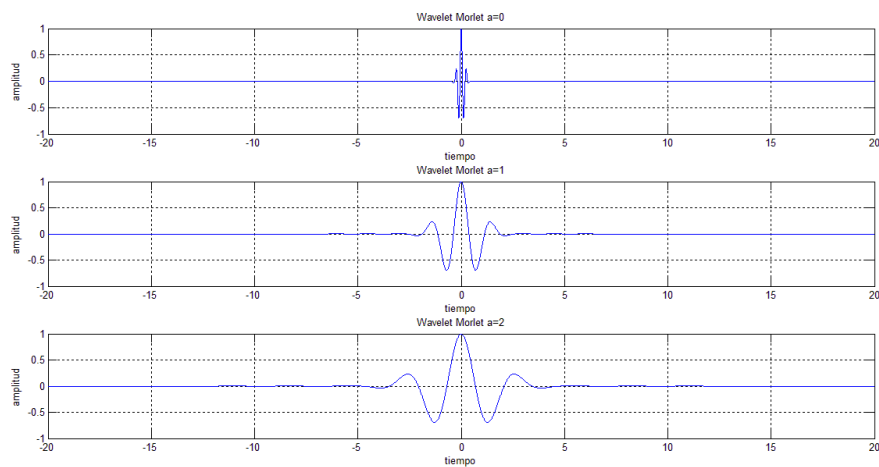


Figura 2. 11 - Escalamiento de la wavelet madre.

2.2.2.3 La transformada Wavelet de tiempo continuo.

La WT de tiempo continuo se define como:

$$CWT(a, b; \psi) = a^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) \delta t \quad (2.10)$$

La ecuación correspondiente a la transformada wavelet continua (CWT) se puede interpretar de manera similar a la STFT. En ambos casos existe una función que multiplica a la señal y en ambos casos esta función se traslada a lo largo del tiempo. La función wavelet madre $\psi(t)$ hace de veces de la función empleada para el ventaneo de la señal $w(t)$ en la STFT. La principal diferencia entre ambas transformadas radican en que, para la STFT, el ancho de la función empleada para el ventaneo permanece constante, mientras que, para la CWT, la wavelet madre es comprimida o expandida por un factor a .

El proceso del cálculo de la WT continua es abordado de manera práctica por Miiti, Openheim y Pogy [72] y se puede resumir de la siguiente manera:

1. Se selecciona una wavelet madre.
2. Se calcula el coeficiente C que indica el grado de correlación entre la wavelet y la primera porción de la señal. Mientras más alto sea el coeficiente mayor será la similitud de la sección de la señal con la wavelet. Observar que el grado de similitud está íntimamente ligado a la elección de la wavelet madre.
3. Se traslada la wavelet mediante el parámetro b y se repite el *paso 2* hasta cubrir toda la señal a analizar. En este punto se obtienen distintos coeficientes de correlación para un único valor de la escala a y para distintos valores de traslación b .
4. Se modifica la escala de la wavelet y se repiten los pasos 2 y 3.
5. Se repiten los pasos 2, 3 y 4 para todas las escalas.

Cuando se han finalizado los pasos se obtiene un conjunto de coeficientes de correlación vinculados a distintos valores de escala a y traslación b .

Para realizar una representación que sea fácil de interpretar se puede realizar un gráfico en donde el eje x se corresponda a la posición de la wavelet con respecto a la señal, en el eje y se represente el valor de la escala y una escala de colores represente la magnitud del parámetro C para cada par de valores (a, b) .

Es importante establecer que existe una correspondencia entre las escalas y la frecuencia de la señal a analizar:

Para bajas escalas (valores chicos del parámetro a) la wavelet madre se comprime, lo cual la hace adecuada para el análisis de los detalles y cambios rápidos de la señal. Los cambios

rápidos se corresponden a altas frecuencias. En resumen, bajas escalas tienen una correspondencia directa a altas frecuencias.

Para escalas altas (valores grandes del parámetro a) la wavelet madre se ensancha, lo cual la hace apropiada para la detección de fenómenos lentos. Fenómenos lentos se corresponden a frecuencias bajas. Por lo tanto, escalas altas tienen una relación directa con bajas frecuencias.

Una representación gráfica tiempo-escala se muestra en la siguiente Figura 2. 12.

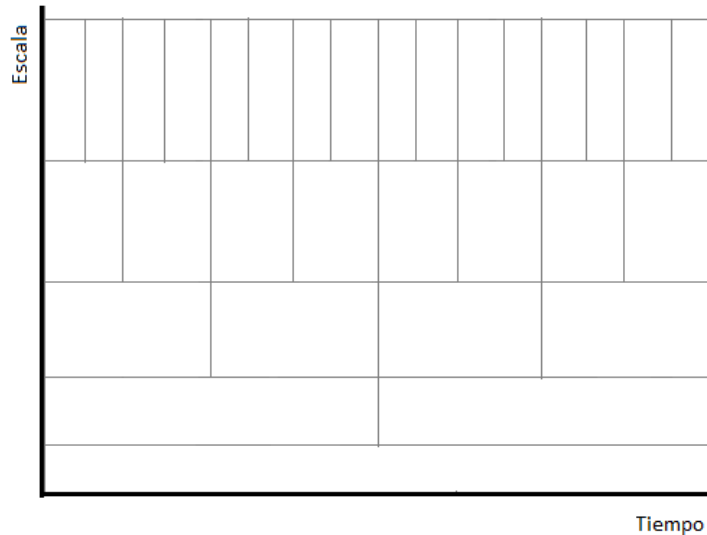


Figura 2. 12 - Plano escala - tiempo de la transformada wavelet.

El área de cada división es fija y viene determinada por la función de la wavelet madre. Al comparar el plano Escala - Tiempo de la *CWT* con el plano Frecuencia - Tiempo de la *STFT* se concluye que la *WT* es un método multi-resolución ya que el método permite ajustar la resolución para detectar los detalles (altas frecuencias) de una señal.

2.2.2.4 Transformada Wavelet Discreta (*DWT*)

El cálculo de la *CWT* se trata de un procedimiento bastante complejo que involucra la resolución de complicadas ecuaciones analíticas (si la transformada se resuelve de manera analítica) o el procesamiento de una gran cantidad de datos (si la transformada es resultado a través de métodos numéricos).

En la *DWT* los parámetros de escalamiento a y traslación b pueden asumir un número discreto de valores lo que facilita en gran medida los cálculos.

Se ha demostrado que si se seleccionan los valores de escalamiento y traslación empleando una potencia de dos (también llamada escala diádica) los resultados de los análisis resultan mucho más eficientes.

La función continua $\psi_{a,b}(t)$ se transforma en la wavelet discreta $\psi_{j,k}(n)$ dada por la siguiente ecuación:

$$\psi_{j,k}(n) = a^{-\frac{j}{2}}\psi(a^{-j}n - kb) \quad (2.11)$$

Los parámetros a y b se definen empleando la escala diádica donde j representa el nivel de descomposición y k el factor de traslación tal como se muestra en la siguiente ecuación:

$$a = 2^j; b = k2^j \quad a > 0; b > 1 \quad (j; k) \in Z \quad (2.12)$$

La *CWT* viene dada por:

$$DWT(j, k) = a^{-\frac{1}{2}} \int f[n]\psi(a^{-j}n - kb)dn \quad (2.13)$$

2.2.2.5 Interpretación de la DWT como un banco de filtros

Una forma eficiente para representar la *DWT* fue presentada en 1998 por *Mallat* [73].

Para muchas señales, el contenido de bajas frecuencias representa la parte más importante de la señal. El contenido de frecuencias, por otro lado, se considera como un el componente que representa matiz o detalle.

Esta distinción es lo que en la *WT* se conoce como aproximación y detalle. La aproximación se corresponde con altas escalas o bajas frecuencias mientras que el detalle se corresponde con bajas escalas o altas frecuencias. Un banco de filtros que cumple con la función de diferenciar estas componentes puede visualizarse en la Figura 2. 13.

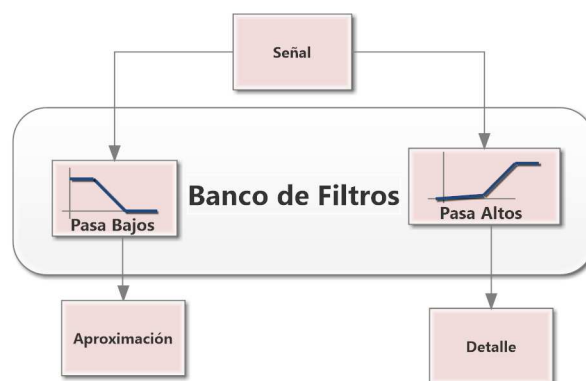


Figura 2. 13 - Aproximación mediante banco de filtros.

La señal ingresa a la etapa de bancos de filtros y, como resultado, dos señales emergen. Es importante notar que esta operación multiplica considerablemente la cantidad de información que debe procesarse. Si la señal cuenta con 1000 muestras, la salida de cada filtro estará formada por esa misma cantidad de muestras. El proceso de filtrado duplicó la cantidad de información a procesar.

Por este motivo es preciso incorporar una técnica denominada sub-muestreo. El sub-muestreo básicamente consiste en eliminar muestras que se consideran redundantes. Si se elimina una de cada dos muestras se dice que el sub-muestreo es de orden $M=2$.

El procedimiento de filtrado y sub-muestreo mostrado en la Figura 2. 14 genera los coeficientes de aproximación y detalle de la *DWT*.

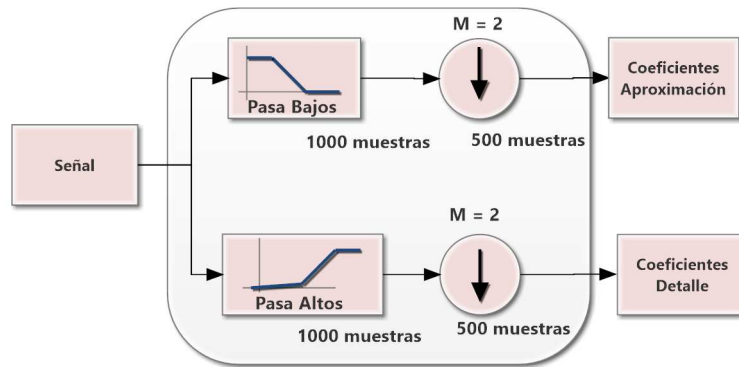


Figura 2. 14 - Banco de filtros y sub muestreo.

Como ejemplo se introduce al banco de filtros una señal senoidal corrupta con ruido blanco gaussiano. La Figura 2. 15 muestra los coeficientes de aproximación y detalle.

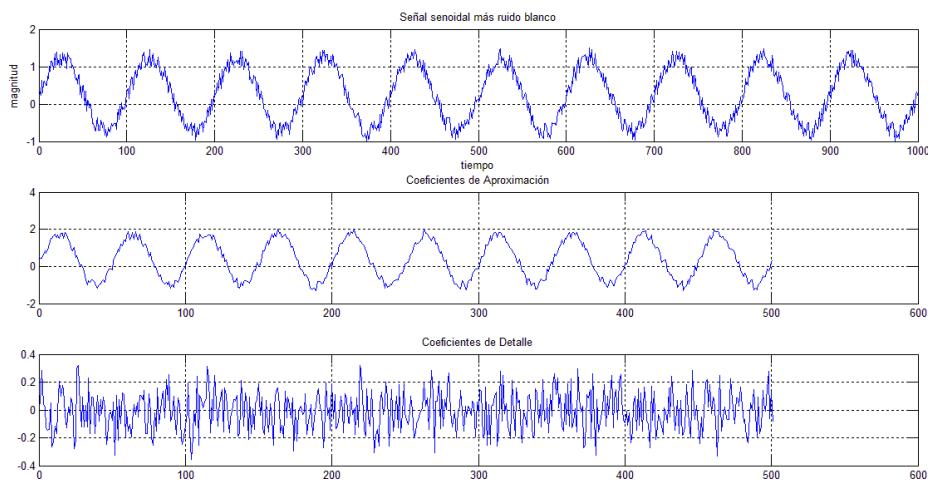


Figura 2. 15 - Coeficientes A y D de una señal senoidal más ruido blanco.

Este proceso puede iterarse, obteniendo sucesivas aproximaciones de forma tal que la señal sea descompuesta en varios componentes de menor resolución. Al esquema mostrado en Figura 2. 16 se lo conoce como *árbol de descomposición*.

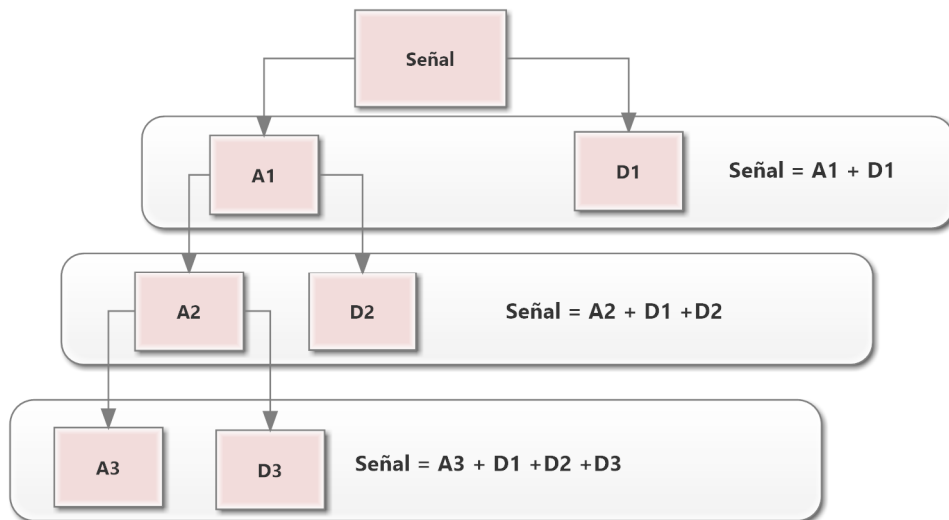


Figura 2. 16 - Árbol de descomposición.

El árbol de descomposición representa una herramienta fundamental para el análisis de perturbaciones en sistemas eléctricos de potencia. En la Figura 2. 17 se muestra un árbol de descomposición para una perturbación del tipo *Muesca*. Los coeficientes de detalle brindan información importante relacionada con la periodicidad y duración de los eventos. Estos coeficientes muestran información relacionada con los cambios rápidos en la señal original. De la misma forma, los coeficientes de aproximación contienen información de la señal de baja frecuencia.

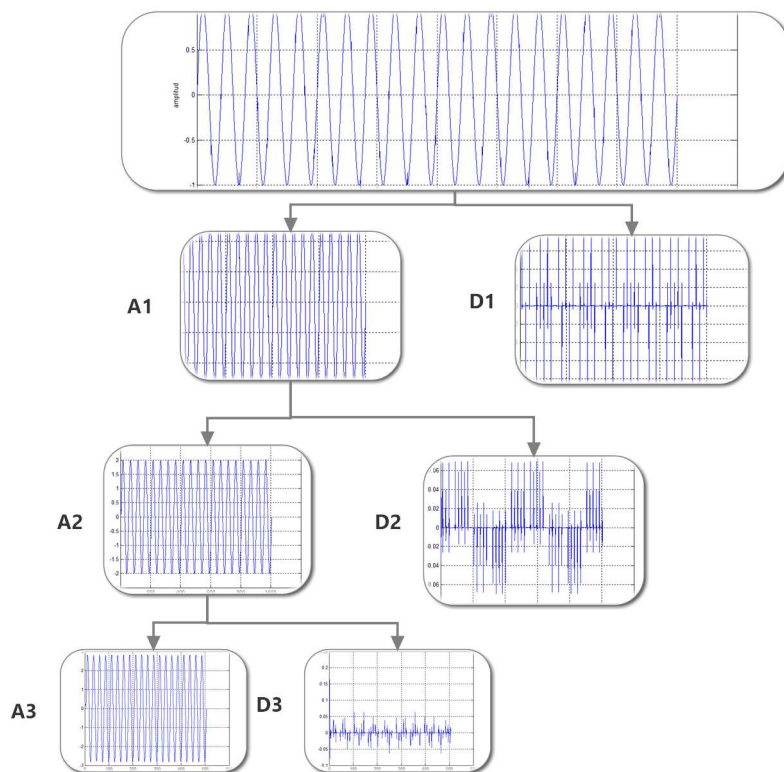


Figura 2. 17 - Árbol de descomposición para una perturbación tipo muesca.

2.2.2.6 Teorema de Parseval.

Este teorema establece una relación entre la energía total de la señal $x(t)$ y la energía asociada a cada uno de los coeficientes de la descomposición.

Esta relación puede expresarse a través de la siguiente ecuación:

$$\int |x(t)|^2 = \sum_{l=-\infty}^{\infty} |c[l]|^2 + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |d_j[k]|^2 \quad (2.14)$$

La energía total de la señal está dividida en tiempo, a través de k y en escala o frecuencia, por medio de j . Esta es la partición tiempo-frecuencia que efectúa la *DWT*.

Daubechies [74] mostró que es posible para las funciones de escalado y funciones wavelet, constituir un paquete compacto que tome valores distintos de cero solamente en una región de duración finita de forma de localizar en el tiempo el contenido espectral.

El sistema wavelet permite contar con características similares a la *STFT*, pero con la diferencia que aquí es posible tener ortogonalidad y buena resolución tiempo frecuencia, característica imposible de lograr en la *STFT*.

2.3 LA TRANSFORMADA WAVELET PARA LA DETECCION DE PERTURBACIONES

De todos los métodos para la detección de eventos de calidad de energía y la extracción de parámetros característicos, la técnica más extensamente empleada se basa en la *WT*. En este sentido, uno de los aportes más importantes es el trabajo presentado por Santos, Powers, Grady y Hofmann, (1996) [29]. Este artículo es uno de los precursores en lo referente al empleo de la *WT* para la detección de perturbaciones en un sistema eléctrico de potencia.

En su trabajo, Possion (1999), concluye que la *CWT* es una buena herramienta para la detección de fenómenos transitorios rápidos con buena precisión en frecuencia aunque no presenta buenas prestaciones para el análisis de fenómenos estacionarios. Este trabajo compara la *CWT* con otros métodos como un algoritmo de análisis multi resolución (*MRA*), la *Transformada de Fourier (FT)* clásica y la *Transformada Cuadrática (QT)* [75] aludiendo las ventajas y desventajas de uno de ellos.

Ese mismo año Gaouda, Salama, Sultan y Chikhani (1999) presentaron un importante aporte donde se fundamenta el uso de una técnica descomposición multiresolución basada en la *DWT* para el análisis de eventos transitorios [76].

Un segundo trabajo fue presentado por estos autores en donde se aplica la transformada multiresolución para el análisis de eventos de corta duración (2000) [77].

Liao (2004) [45] propone una técnica que combina la *FFT* y la *WT* para la extracción de parámetros característicos.

Weiming, S. Xuelei, L. Jingbo and Z. Zhiheng (2006) presentaron un método que combina la *STFT* y el *WPT* para la extracción de parámetros característicos de una señal perturbada con el objetivo de reducir la dimensión de los datos manejados por un sistema de clasificación.

En los trabajos propuestos por Jahehak (2002) [42], Liyan (2008) [78] y Liu también se emplea *WPT* para la detección y extracción de parámetros característicos de perturbaciones electromagnéticas [79].

Además de la utilización de las *WT* para la detección, localización y extracción de parámetros característicos de un evento de calidad de energía, esta herramienta también ha demostrado ser un poderoso método de compresión de datos. En este sentido Santos (1997) presentó un trabajo donde se emplea una técnica de umbralización para descartar información y reducir de este modo el tamaño del archivo que contiene la información de la señal perturbada en una tasa que va desde un tercio a un sexto de su tamaño original [80].

Zhengyou (2011) propone, en su trabajo [81], la utilización de la *WT* para el cálculo de parámetros característicos a través de la obtención de la entropía de los coeficientes de detalle.

Una combinación de *DWT* y una *Wavelet Network* es empleada en por Masoum (2010) [82] para la detección y clasificación de un gran número de perturbaciones.

En numerosos trabajos se ha demostrado que el rendimiento de la *WT* sufre una degradación importante frente al ruido aditivo. Para mejorar la capacidad los sistemas de monitoreo basados en la *WT* Hong-Tzer and L. Chiung-Chou presentaron un esquema de cancelación de ruido [83].

2.4 CONCLUSIONES

En este capítulo se abordaron algunos conceptos básicos de la *Transformada de Fourier* y la *Transformada de Fourier dependiente del tiempo* para el análisis de señales estacionarias y no estacionarias demostrando las limitaciones y desventajas que ambos métodos presentan frente a la *WT*.

Luego fueron presentadas algunas nociones teóricas y analíticas relativas a la *WT* y se demostraron sus ventajas frente a los métodos de procesamiento de señales tradicionales. Dentro de los temas abordados en este capítulo se encuentra la *Transformada Wavelet Continua (CWT)*, la *Transformada Wavelet Discreta (DWT)* y una útil interpretación de la *WT* discreta representada por un banco de filtros.

Se demostró que la *WT*, al tratarse de un método de análisis multi resolución, permite la localización temporal precisa, con una aceptable resolución en frecuencias, de un evento de calidad de energía en un sistema eléctrico de potencia.

Finalmente se presentan los aportes más relevantes de distintos autores que utilizan la *WT* para detectar y clasificar eventos de calidad de energía.

Las claras ventajas que la *WT* presentadas frente a distintos métodos de procesamiento de señales justifican su elección para el desarrollo del software de detección y clasificación de eventos de calidad de energía presentado en los *capítulo 5, 6 y 7* del presente trabajo de tesis.

Capítulo 3:

Máquina de Vectores de

Soporte

3.1. INTRODUCCION.

En la actualidad los sistemas de procesamiento se encuentran difundidos alrededor de todo el mundo y representan una de las herramientas más formidables con las que ha contado la humanidad ya que permiten implementar fácilmente algoritmos para resolver multitud de problemas y cálculos que antes resultaban engorrosos y complicados. Sin embargo debido a que la mayoría de los sistemas de procesamiento emplean una programación rígida la siguiente limitación se presenta:

Los sistemas de procesamiento tradicional son capaces de resolver solo aquellos problemas que admiten tratamiento algorítmico pero no pueden resolver problemas que resultan naturalmente sencillos para el cerebro humano como es el caso, por ejemplo, de la clasificación de objetos por rasgos comunes [84].

Esta limitación impulsó a muchos científicos a estudiar las capacidades del cerebro humano como fuente de ideas para el desarrollo de una nueva generación de máquinas lo que dio lugar a una reciente y prominente rama en la ciencia de la computación denominada *AI*. Así, la *AI* es un intento por descubrir y describir aspectos de la inteligencia humana que pueden ser simulados mediante máquina.

Durante mucho tiempo, las *ANN* fueron tenidas en cuenta como el método principal de *AI* para el reconocimiento de patrones y para la aproximación de funciones difíciles de reproducir mediante técnicas heurísticas.

Estas herramientas son ampliamente utilizadas debido a su estructura de procesamiento en paralelo y su capacidad de aprender y, por lo tanto, de generalizar el conocimiento adquirido [84]. Los investigadores emplearon las *ANN* para la clasificación de eventos de calidad de energía debido a su capacidad de resolver problemas donde los conjuntos son *linealmente no separables*, su habilidad de relacionar elementos de entrada con elementos de salida, su poder de adaptarse a diferentes problemáticas, su tolerancia a fallos y su uniformidad de análisis y diseño.

Sin embargo, las *ANN* presentan una serie de problemas que pusieron en duda sus capacidades. Los algoritmos empleados para el entrenamiento de una *ANN* generalmente necesitan de un número grande de muestras para lograr una precisión deseada lo que provoca que, en implementaciones reales, se deba contar con un número grande de mediciones para entrenar un algoritmo de este tipo. Por otro lado; el proceso de entrenamiento generalmente toma mucho tiempo y puede finalizar de manera no esperada debido a un atascamiento en un mínimo local. Por último, la no existencia de reglas teóricas para definir la arquitectura de una *ANN* por lo que la misma debe ser definida de manera empírica.

Los mencionados problemas que las *ANN* presentan impulsaron a los investigadores a estudiar algunas técnicas de aprendizaje lineal superadoras para resolver problemas de reconocimiento de patrones y aproximación de funciones. Dentro de la familia de algoritmos de *AI* existe una técnica conocida como Máquinas de Soporte de Vectores (*SVM por sus*

sigmas en inglés) que, por sus características intrínsecas, resuelven muchas de las desventajas que presentan las *ANN*.

Las *Maquinas de Vectores de Soporte* es una técnica de máquinas de aprendizaje que se fundamenta en dos bases sólidas: *La teoría de aprendizaje estadístico* y *la optimización matemática*. Esta técnica ha venido siendo aplicada con éxito en una gran variedad de problemas, que van desde reconocimiento de caracteres hasta reconocimiento de objetos basado en sus patrones de apariencia [85].

Algunos de los aspectos que fundamentan el éxito de esta técnica frente a otros métodos de aprendizaje lineal (como las *ANN*) están relacionados con el hecho que los *SVM* exhiben una gran habilidad de generalización en muchos problemas de la vida real, la técnica se encuentra bien fundamentada matemáticamente, el proceso de entrenamiento elimina el posible atascamiento en un mínimo local y el modelo presenta pocos parámetros para ajustar.

Debido a que la matemática involucrada en este tipo de máquinas de aprendizaje lineal es compleja este capítulo se enfocará en proporcionar una introducción a los conceptos esenciales acerca de las *Maquinas de Vectores de Soporte*.

En este contexto la primera parte del capítulo introducirá al lector en las teorías de aprendizaje y la optimización matemática. Estos conceptos serán utilizados en la segunda sección del capítulo para la explicación conceptual de un clasificador binario basado en *Máquinas de Vectores de Soporte*.

La tercera parte del capítulo detalla las diferentes estrategias para la confección de un clasificador multiclase basado en clasificadores *SVM* binarios.

Finalmente, en la última sección del capítulo, se resume el estado del arte de la aplicación de esta técnica en la clasificación de eventos de calidad de energía en un sistema eléctrico de potencia y la localización de fallas.

3.2. TEORIA DE APRENDIZAJE ESTADISTICO.

Existe una gran variedad de problemas que no pueden ser resueltos por técnicas de programación clásica ya que no siempre se puede formular un modelo matemático que responda a un problema de la vida real. Un ejemplo comúnmente citado es el problema de reconocimiento de patrones en la escritura a mano alzada. Para resolver este tipo de problemas las *máquinas de aprendizaje* se asoman como la mejor solución.

El proceso de aprendizaje se conoce como *la habilidad de una entidad inteligente para adecuarse a nuevos entornos, para resolver nuevos problemas o, en base a la experiencia, solucionar estos problemas de manera más eficiente* [86].

3.2.1. Aprendizaje supervisado.

Una alternativa para resolver el problema del reconocimiento de texto es realizar un programa que permita relacionar la entrada con la salida por medio de ejemplos, de la misma manera que un ser humano aprende a reconocer las letras por medio de la inspección de los patrones dibujados en un papel. La aproximación que utiliza ejemplos para sintetizar un programa se conoce como metodología de aprendizaje y, cuando se utilizan varios ejemplos para generalizar una relación de entrada y salida, el entrenamiento se denomina *aprendizaje supervisado*.

El entrenamiento supervisado se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo que determina la respuesta que el modelo debería entregar ante una determinada entrada. El supervisor controla la salida del modelo y, en caso que la salida no coincida con la respuesta esperada, dicho proceso modifica parámetros del algoritmo para que la respuesta se aproxime a la deseada.

Si un problema de aprendizaje supervisado tiene como salida una variable binaria se conoce como *clasificación binaria*. Por el contrario, un problema de aprendizaje con un espacio de salida compuesto por más de dos elementos (reconocimiento de patrones de escritura) se conoce como *clasificación multiclase*.

Un problema del tipo *regresión* se da cuando el espacio de la salida está compuesto por una función real.

Un problema fundamental de este tipo de aprendizaje es que, incluso cuando es posible encontrar una hipótesis consistente que relacione de manera correcta un dato de entrada con la salida deseada, puede que esta hipótesis no sea válida para clasificar correctamente datos que no han sido representados en el set de entrenamiento. La habilidad de un sistema para clasificar correctamente una clase de datos no presente en el set de entrenamiento se conoce como *generalización* y esta es la propiedad más importante que se debe potenciar en un algoritmo de clasificación.

Para que una máquina de aprendizaje logre una buena habilidad de generalización se debe realizar un correcto entrenamiento. Si el set de entrenamiento está compuesto por un número muy grande de componentes es posible que la máquina memorice la relación en lugar de

aprenderla. En este caso se dice que se ha *sobre entrenado* la máquina de aprendizaje y su capacidad de generalización se ve reducida. Una máquina sobre entrenada responderá incorrectamente ante una entrada que difiera (aunque en proporciones ínfimas) a cualquier entrada usada para el entrenamiento.

3.2.2. Limitaciones en el aprendizaje.

El método de aprendizaje tiene 4 limitaciones que son importantes destacar:

- Dependiendo de las funciones de entrenamiento empleadas, los algoritmos de aprendizaje pueden atascarse en un mínimo local.
- El tamaño de la hipótesis de salida puede volverse muy grande e impreciso.
- Si hay únicamente un número limitado de ejemplos demasiado ricos en información un clasificador podrá sobre entrenarse y la habilidad de generalización se verá afectada.
- Usualmente los algoritmos de entrenamiento son controlados por un gran número de parámetros que complejizan en gran manera al sistema y lo hacen difícil de utilizar.

3.2.3. Proceso de Aprendizaje para las Maquinas de Vectores de Soporte.

Los SVM son sistemas de aprendizaje que utilizan espacios de hipótesis de funciones lineales en un espacio característico de mayor dimensión, entrenados con un algoritmo que se basa en la teoría de optimización matemática y que implementan un sesgo de aprendizaje proveniente de la teoría de aprendizaje estadístico.

En las futuras secciones se aborda, de manera genérica, la teoría de optimización matemática, para luego describir el proceso de aprendizaje para las *Máquinas de Vectores de Soporte*. En lo sucesivo, se utilizará el acrónimo *SVM* (*derivadas de Support Vector Machine*) para mencionar a este tipo de inteligencia artificial

3.3. TEORIA DE LA OPTIMIZACION MATEMATICA.

3.3.1. Conceptos.

El entrenamiento de un *SVM* involucra la resolución de un problema de optimización matemática. Por este motivo, con el objetivo de entender el funcionamiento de la etapa de aprendizaje del algoritmo de *SVM* es importante un conocimiento previo sobre la *teoría de optimización matemática*.

La teoría de la optimización es la rama de las matemáticas relacionada con la tarea de hallar la mejor solución dentro de un conjunto de posibles respuestas a un problema y desarrollar algoritmos que encuentren dichas soluciones de manera eficiente.

En el caso más simple, un problema de optimización consiste en maximizar o minimizar una función real eligiendo sistemáticamente valores de entrada (tomados de un conjunto permitido) y computando el valor de la función. De forma general, la optimización incluye el descubrimiento de los "mejores valores" de alguna función objetivo dado un dominio definido, incluyendo una variedad de diferentes tipos de funciones objetivo y diferentes tipos de dominios [87].

En esta sección solo se describen los casos en los que la función principal a optimizar es una función cuadrática convexa y todas sus restricciones son lineales. Esta clase de problemas es la que se utiliza para el entrenamiento de los *SVM*.

3.3.2. Formulación del problema.

El objetivo de resolver problema de optimización matemática es el de *encontrar el valor máximo o el valor mínimo de una función sujeto a diferentes tipos de restricciones*. La definición formal es:

Dada las funciones $f, g_i \ i = 1, 2, \dots, k$ y $h_i \ i = 1, 2, \dots, m$ definidas en el dominio $\Omega \subseteq \mathbb{R}^n$

$$\text{minimizar } f(w) \quad w \in \mathbb{R}^n \quad (3.1)$$

Sujeto a

$$g_i(w) \leq 0 \quad i = 1, 2, \dots, k \quad (3.2)$$

$$h_i(w) = 0 \quad i = 1, 2, \dots, m \quad (3.3)$$

Donde $f(w)$ se denomina *función objetivo* y las restantes funciones son llamadas respectivamente *restricciones de desigualdad* (g_i) y *restricciones de igualdad* (h_i).

El valor óptimo de la función objetivo se conoce como valor del problema de optimización.

La región del dominio donde la función objetivo es definida y donde todas las restricciones son satisfechas se llama *región de factibilidad* y se denota por:

$$R = \{w \in \Omega, g(w) \leq 0, h(w) = 0\} \quad (3.4)$$

Una solución al problema de optimización se corresponde con el punto $w^* \in \mathbb{R}$ de tal forma que no exista otro punto $w \in \mathbb{R}$ tal que $f(w) < f(w^*)$. Este punto es conocido como *mínimo global*.

Un punto $w^* \in \Omega$ se llama *mínimo local* de $f(w)$ si $\exists \xi > 0$ tal que:

$$f(w) \geq f(w^*) \forall w \in \Omega \text{ tal que } \|w - w^*\| \leq \xi \quad (3.5)$$

3.3.3. Definiciones Importantes.

El entrenamiento de un *SVM* involucra un problema de optimización matemática cuyas restricciones serán limitadas a *funciones lineales* y la función objetivo será una función *convexa cuadrática*. Antes de abordar el método de entrenamiento de un *SVM* es preciso realizar las siguientes definiciones:

3.3.3.1 Programación lineal.

Un problema de optimización matemática cuya función objetivo y las restricciones son funciones lineales es llamado problema de *programación lineal*.

3.3.3.2 Programación cuadrática.

Un problema de optimización matemática cuya función objetivo es una función cuadrática y las restricciones son funciones lineales es llamado *problema de optimización cuadrática*.

3.3.3.3 Función cóncava y convexa.

Una función real f definida en un intervalo (o en cualquier conjunto convexo C de algún espacio vectorial) se dice que es cóncava, si para dos puntos x e y cualesquiera definidas en su dominio C , y para cualquier t en $[0, 1]$, se cumple:

$$f(tx + (1 - t)y) \geq tf(x) + (1 - t)f(y) \quad (3.6)$$

Además, $f(x)$ es cóncavo en $[a, b]$ si y sólo si la función $-f(x)$ es convexa en $[a, b]$.

Si una función $f(x)$ es *convexa* para todo x entonces si existe un valor x^* para el cual la función $f(x^*)$ asume su valor mínimo entonces se puede asegurar que ese valor se corresponde con un *mínimo global*. [88]. Esta última definición presenta un concepto realmente importante ya que asegura que un problema de optimización cuya función es convexa no puede sufrir un atascamiento en un mínimo local. Al ser una función convexa, el mínimo encontrado se corresponde siempre a un mínimo global.

3.3.3.4 Transformación de las restricciones.

Existe una estrategia para transformar las restricciones de desigualdad en restricciones de igualdad. El método consiste en incorporar a la función una variable ξ denominada *variable de relajación* de la siguiente manera:

$$g_i(w) \leq 0 \Leftrightarrow g_i(w) + \xi_i = 0 \quad \xi_i \geq 0 \quad (3.7)$$

3.3.4. Teoría de Lagrange.

El propósito de la teoría de Lagrange radica en caracterizar la solución de un problema de optimización matemática donde, inicialmente se propondrá, que no existen restricciones de desigualdad.

Antes de abordar la teoría de Lagrange es preciso realizar la primera definición:

Definición 1: Teorema de Fermat.

Una condición necesaria para que w^* sea un mínimo de $f(w)$, $f \in C^1$ es que la derivada de la función en ese punto sea nula.

$$\frac{\partial f(w^*)}{\partial w} = 0 \quad (3.8)$$

Una función f es de clase C^l si sus derivadas primera y segunda existen y son continuas.

La teoría de Lagrange se fundamenta en dos conceptos básicos: Los multiplicador es de Lagrange y las funciones lagrangianas.

Definición 2: Multiplicador de Lagrange y función lagrangiana.

Dado un problema de optimización matemática, con función objetivo $f(w)$ y restricciones de igualdad $h_i(w) = 0$ para $i = 1, 2, \dots, m$ se define la función lagrangiana como:

$$\mathcal{L}(w, \beta) = f(w) + \sum_{i=1}^m \beta_i h_i(w) \quad (3.9)$$

donde β_i se conoce como *multiplicador lagrangiano*.

Si un punto w^* es un mínimo local para un problema de optimización con únicamente condiciones de igualdad es posible que $\frac{\partial f(w^*)}{\partial w} \neq 0$. Esto quiere decir que es posible moverse en alguna dirección de w , violando algunas de las restricciones, y encontrar que $f(w)$ disminuye su valor.

Para respetar todas las restricciones de igualdad h_i es preciso moverse de manera perpendicular a $\frac{\partial h_i(w^*)}{\partial w}$. De esta manera, para respetar todas las restricciones nos debemos mover de manera perpendicular al sub espacio V creado por:

$$\left\{ \frac{\partial h_i(w)}{\partial w} \quad i = 1, 2, \dots, m \right\} \quad (3.10)$$

Si todas las $\frac{\partial h_i(w^*)}{\partial w}$ son linealmente independientes, ningún movimiento legal va a cambiar el valor de la función objetivo, sin importar donde $\frac{\partial h_i(w^*)}{\partial w}$ se encuentre en el sub espacio V . En otras palabras, cuando existe un valor de β_i tal que:

$$\frac{\partial f(w)}{\partial w} + \sum_{i=1}^m \beta_i h_i(w^*) = 0 \quad (3.11)$$

Esta observación fundamenta siguiente definición.

Definición 3: Multiplicador de Lagrange y función lagrangiana.

Una condición necesaria que un punto w^* debe cumplir para ser un mínimo de $f(w)$ sujeto a $h_i(w)=0, i=1,2,\dots,m$ con $f, h_i \in C^1$ es:

$$\frac{\partial \mathcal{L}(w^*, \beta^*)}{\partial w} = 0 \quad (3.12)$$

$$\frac{\partial \mathcal{L}(w^*, \beta^*)}{\partial \beta} = 0 \quad (3.13)$$

para algunos valores de β^* . Las condiciones arriba mencionadas son también suficientes debido a que $\mathcal{L}(w, \beta^*)$ es también una función convexa de w .

Las tres definiciones presentadas hasta el momento abren el camino para una definición más completa, que contempla no solo las restricciones de igualdad sino que también contempla las restricciones de desigualdad.

Definición 4: Teorema de Kuhn Tucker.

Dado un problema de optimización matemática cuya función objetivo es una función convexa $\Omega \subseteq \mathbb{R}$

$$\text{minimizar } f(w) \quad w \in \Omega \quad (3.14)$$

Sujeta a:

$$g_i(w) \leq 0 \quad w \in \Omega \quad (3.15)$$

$$h_i(w) = 0 \quad w \in \Omega \quad (3.16)$$

con $f \in C^1$ convexa y g_i, h_i son restricciones de desigualdad e igualdad respectivamente. Las condiciones necesarias y suficientes para que un punto w^* sea un valor óptimo es la existencia de un α^* y β^* tal que:

$$\frac{\partial \mathcal{L}(w^*, \alpha^*, \beta^*)}{\partial w} = 0 \quad (3.17)$$

$$\frac{\partial \mathcal{L}(w^*, \alpha^*, \beta^*)}{\partial \beta} = 0 \quad (3.18)$$

$$\alpha^*_i g_i(w^*) = 0 \quad i = 1, \dots, k \quad (3.19)$$

$$g_i(w^*) \leq 0 \quad i = 1, \dots, k \quad (3.20)$$

$$\alpha_i^* \geq 0 \quad i = 1, \dots, k \quad (3.21)$$

La tercera condición se conoce como condición complementaria de *Kurush-Kuhn-Tucker* (*KKT*). Esta condición implica que para restricciones activas $\alpha^* \geq 0$ y para restricciones inactivas $\alpha^* = 0$.

Una forma de interpretar los resultados presentados anteriormente implica que la solución pueda encontrarse en una o dos posibles posiciones con respecto a la restricción de desigualdad o bien en el interior de la región de factibilidad, con la restricción inactiva, o bien en el límite de la región de factibilidad cuando la restricción se encuentra activa. En el primer caso las condiciones de optimalidad para tales restricciones están dadas por el *teorema de Fermat* y, por lo tanto α_i debe ser cero. Para el segundo caso se utiliza el *teorema de Lagrange* con un valor no cero de α_i . Entonces las condiciones *KKT* implican si una restricción es activa, significa que $g_i(w^*) = 0$, o que el multiplicador satisface que $\alpha_i = 0$. Eso se representa en la ecuación $\alpha^*_i g_i(w^*) = 0$.

3.3.5. Dualidad.

La resolución de problemas de optimización convexa empleando el tratamiento de Lagrange puede realizarse por medio de una descripción alternativa dual que hace que sea más fácil la resolución el problema principal ya que la resolución de las restricciones de desigualdad suele ser una tarea compleja. El problema dual se obtiene por medio de la introducción de multiplicadores lagrangianos, también llamados *variables duales*.

Los métodos duales se basan en la idea que las *variables duales* son la incógnita principal del problema de optimización. Se puede transformar un problema primario en un problema dual fijando a cero la derivada del lagrangiano con respecto a las variables primarias y substituyendo las relaciones obtenidas en el lagrangiano removiendo de esta manera la dependencia de las variables primarias. Esto se corresponde a computar la siguiente función:

$$\theta(\alpha, \beta) = \min_{w \in \Omega} \mathcal{L}(w, \alpha, \beta) \quad (3.22)$$

La función resultante contiene únicamente variables duales y debe ser maximizada bajo restricciones simples.

3.4. MAQUINAS DE VECTORES DE SOPORTE.

Las Maquinas de Vectores de Soporte (*SVM*) pueden ser aplicadas en problemas de reconocimiento de patrones, problemas de regresión, extracción de características o detección de novedades. Como el objetivo del presente trabajo de tesis es el de estudiar algoritmos de reconocimiento y clasificación de patrones aquí solo se abordarán los aspectos relacionados a este tipo de problemas.

Una característica importante de este método es que el algoritmo es entrenado empleando representaciones duales y ralas de las restricciones lo que resulta en algoritmos extremadamente eficientes. La eficiencia es una de las características fundamentales que distingue a un clasificador basado en *SVM* frente a otras máquinas lineales como las *ANN* por ejemplo.

Otra característica importante de las aproximaciones que se basan en un *SVM* radica en el hecho que el problema siempre involucra problemas de *optimización matemática convexa* lo que elimina la posibilidad de la aparición de mínimos locales. Los mínimos locales representan un gran problema en el entrenamiento de otros algoritmos de clasificación.

3.4.1. Clasificador de vectores de soporte.

El objetivo de un clasificador que emplea vectores de soporte es el de proveer una forma computacional y eficiente de aprendizaje que realice una óptima separación de dos clases por medio de un hiperplano ubicado en un espacio característico de mayor dimensión.

En caso de ser necesario, un *SVM* primero mapea todos los puntos de entrada a un espacio característico de mayor dimensión. Por ejemplo, si los datos de entrada están representados en \mathbb{R}^2 los *SVM* mapean estos puntos a una dimensión en \mathbb{R}^3 para luego encontrar un hiperplano óptimo que los separe. Existen varios métodos para encontrar el hiperplano óptimo, dentro de los cuales se puede destacar la estrategia de optimizar el margen máximo, optimizar la distribución del margen u optimizar el número de vectores de soporte.

El modelo más simple, y por lo tanto el primero en ser abordado en este trabajo, es el *clasificador de máximo margen*.

3.4.2. Clasificador de máximo margen.

Si se tiene un problema de clasificación de patrones de datos *linealmente separables* como el mostrado en la Figura 3. 1 el problema del clasificador de máximo margen es el de encontrar un hiperplano que separe ambas clases maximizando el margen m .

Se define el margen m como la distancia entre dos puntos dados, de ambas clases, más próximos a la superficie de decisión que, en nuestro caso, es un hiperplano.

Un problema de clasificación entre dos clases de datos puede ser de dos tipos: Cuando las clases son linealmente separables o cuando las clases no linealmente separables. El primero de ellos se corresponde con el caso más sencillo y, por este motivo, será detallado en primer lugar para luego hacer extensivo el método a los casos no linealmente separables.

3.4.2.1 Clasificador para datos linealmente separables.

Maximizar el margen m es un problema de optimización matemática con una función objetivo *cuadrática convexa* y puede ser resuelto mediante la proposición de un problema dual incorporando multiplicadores de Lagrange.

Sin ningún conocimiento del mapeo, un SVM encuentra el hiperplano óptimo utilizando el producto punto con funciones en el espacio de características que son llamadas *kernels*. La solución del hiperplano óptimo puede ser escrita como la combinación de un reducido número de puntos de entrada que son llamados *vectores de soporte*.

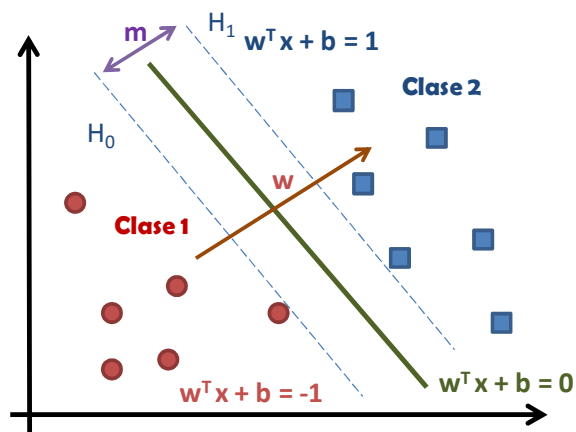


Figura 3. 1 - Fronteras de decisión.

Suponga el caso mostrado en la Figura 3. 2 en donde, para un mismo conjunto de datos, se encuentran dos planos que cumplen con la función de separar las dos clases presentes originalmente.

Como se puede ver, el margen m_1 encontrado en la figura Figura 3. 2 a) es más estrecho al margen m mostrado en la figura Figura 3. 2 b). Este ejemplo muestra que no hay un único plano que separa ambas clases de datos, sin embargo existe un único plano óptimo que maximiza el margen m .

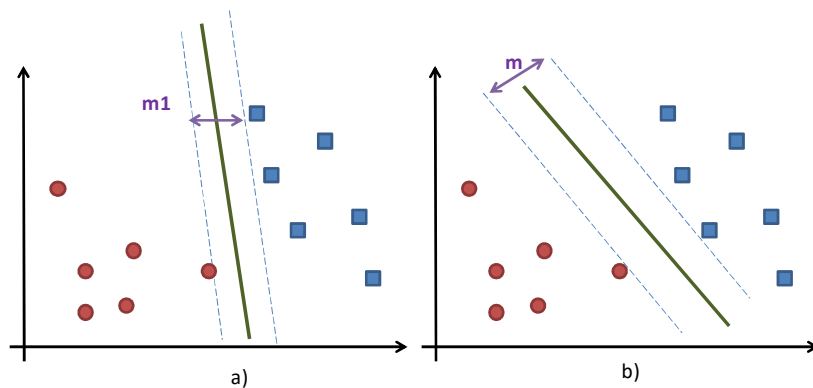


Figura 3. 2 - Caso linealmente separable.

La *teoría de aprendizaje* permite controlar la capacidad de generalización y la posibilidad de sobre-entrenamiento (*overfitting*) a través del control de los parámetros del hiperplano [89].

En este trabajo se considera la manera más común de encontrar el plano óptimo de separación el cual consiste en utilizar herramientas de optimización matemática para reducir la norma de un vector w también conocido como vector peso [85].

3.4.2.2 Límites de decisión.

Con el objetivo de brindar un soporte matemático a los conceptos anteriormente explicados se intenta definir las reglas de los límites de decisión como función del conjunto de muestras.

Considere un conjunto de puntos de entrada S , dado $x_i \in \mathbb{R}^n$ con $i = 1, 2, \dots, N$. Este conjunto de puntos, clasificados como muestras positivas (círculos) y muestras negativas (cuadrados), separados por una región de decisión, es mostrado en la Figura 3. 3.

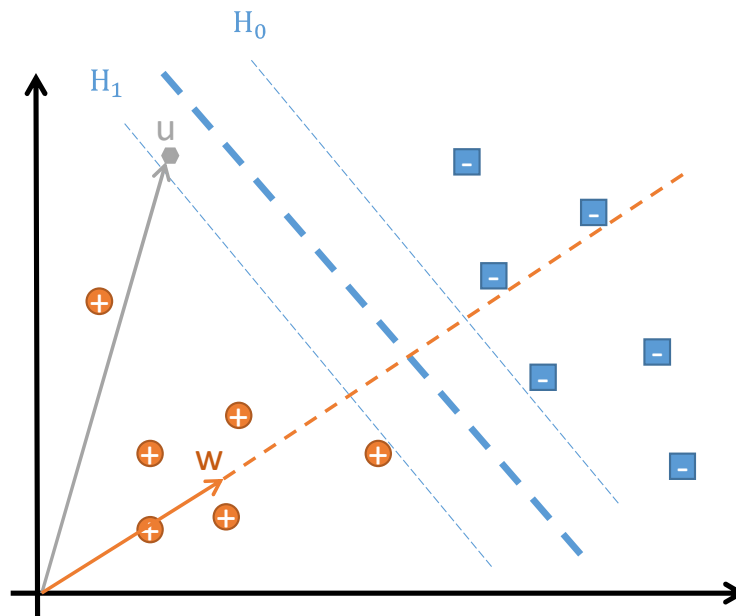


Figura 3. 3 - Límite de decisión.

Es importante destacar que H_0 y H_1 son paralelos (tienen la misma norma) y no hay ninguna muestra entre ellos.

Aunque aún no se conoce cuál de las infinitas posibles rectas que dividen el conjunto de muestras linealmente separables, en esta instancia, es posible definir un vector \vec{w} cuya particularidad principal es la de resultar perpendicular a la frontera de decisión.

Por otro lado, en la Figura 3. 3 también se observa un vector \vec{u} que se corresponde con un punto cualquiera del plano. Este punto puede encontrarse indistintamente en la zona de las muestras positivas o negativas.

Para resolver este interrogante se propone establecer una proyección del vector \vec{u} sobre el vector \vec{w} . El valor de la proyección, obtenido mediante el producto punto y representado mediante un escalar, brindará una medida de cuán cerca o lejos nos encontramos de la región de decisión. Para determinar la distancia de un punto cualquiera del plano a la frontera de decisión se define un umbral C cuya evaluación define en que zona la muestra se encuentra.

Matemáticamente esto se puede escribir como:

$$\vec{w} \cdot \vec{u} \geq C \quad (3.23)$$

Sin perder generalidad, podemos definir $C = -b$ y reescribir la ecuación (3.23) como:

$$\vec{w} \cdot \vec{u} + b \geq 0 \text{ entonces } u \in x_+ \quad (3.24)$$

Hasta aquí solo se sabe que \vec{w} es un vector perpendicular a la frontera de decisión y se presentó una ecuación que, en caso de cumplirse, permite identificar la región donde se encuentra un punto cualquiera del plano.

Para acotar el problema y continuar con el análisis se supone que se agrega una nueva restricción dada por:

$$\vec{w} \cdot \vec{x}_+ + b \geq 1 \text{ para } u \in x_+ \quad (3.25)$$

Esta restricción significa que, si bien el punto que denominamos con la letra u puede pertenecer a cualquier punto del plano, si u se encuentra en la región del plano perteneciente a la clases positiva, entonces la ecuación (3.24) será siempre mayor a 1 dejando la igualdad para el caso en que el punto u se encuentre sobre la frontera de decisión.

Un análisis análogo podría extenderse para las muestras negativas:

$$\vec{w} \cdot \vec{x}_- + b \leq -1 \text{ para } u \in x_- \quad (3.26)$$

Para simplificar las ecuaciones (3.25) y (3.26) se introduce una nueva variable y_i tal que:

$$y_i = \begin{cases} 1 & \text{para todo } u = x_+ \\ -1 & \text{para todo } u = x_- \end{cases} \quad (3.27)$$

Multiplicando ambos miembros de las ecuaciones (3.25) por y_i se tiene:

$$y_i(\vec{w} \cdot \vec{x}_+ + b) \geq y_i \text{ para } u \in x_+ \quad (3.28)$$

Como la ecuación (3.25) es válida solo para $u = x_+$, entonces $y_i = 1$ y la ecuación (3.28) puede re escribirse como:

$$y_i(\vec{w} \cdot \vec{x}_+ + b) \geq 1 \text{ para } u \in x_+ \quad (3.29)$$

De la misma manera, para la ecuación (3.26) se resuelve:

$$y_i(\vec{w} \cdot \vec{x}_- + b) \leq y_i(-1) \text{ para } u \in x_- \quad (3.30)$$

Pero como $y_i = -1$ para todos los valores $u = x_-$, y teniendo en cuenta que la desigualdad se invierte a la hora de multiplicar por un valor negativo se tiene:

$$y_i(\vec{w} \cdot \vec{x}_- + b) \geq (-1)(-1) \text{ para } u \in x_- \quad (3.31)$$

$$y_i(\vec{w} \cdot \vec{x}_- + b) \geq 1 \text{ para } u \in x_- \quad (3.32)$$

De esta manera, incorporando la variable y_i hace que (3.29) y (3.32) sean iguales y puedan fusionarse en una única ecuación dada por:

$$y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1 \quad \forall i = 1, 2, \dots, N \quad (3.33)$$

$$y_i(\vec{w} \cdot \vec{x}_i + b) - 1 \geq 0 \quad \forall i = 1, 2, \dots, N \quad (3.34)$$

La ecuación (3.34) es válida para ambas clases. Es importante destacar que aquellos puntos donde la desigualdad se transforma en igualdad representan los puntos que delimitan la los hiperplanos H_0 y H_1 . Para estos puntos la ecuación válida es:

$$y_i(\vec{w} \cdot \vec{x}_i + b) - 1 = 0 \quad \forall i = 1, 2, \dots, N \quad (3.35)$$

La ecuación (3.34) es una ecuación que permite definir en qué región se encuentra un punto cualquiera u , mientras que la (3.35) representan una restricción para aquellos puntos que definen la separación entre los hiperplanos H_0 y H_1 .

Sin olvidar que el objetivo perseguido es maximizar el margen se pretende encontrar una relación que represente la distancia que separa las muestras positivas y negativas como una función de dichas muestras. Para ello considere la Figura 3.4.

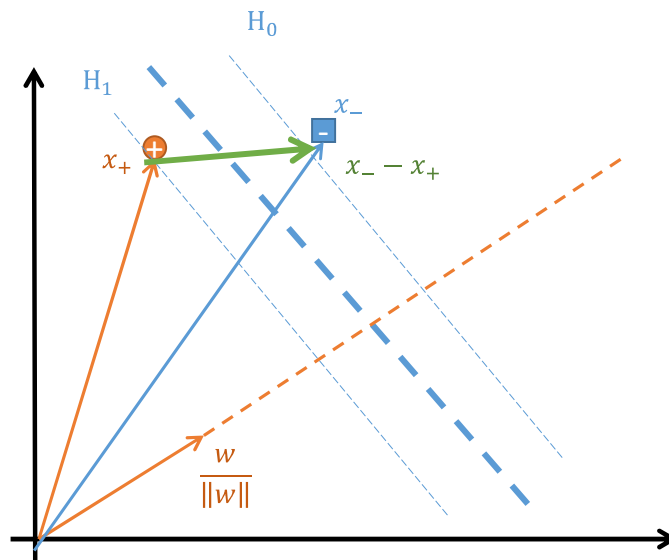


Figura 3.4 - Separación entre los hiperplanos H_0 y H_1 .

En la figura x_+ representa un vector hacia la muestra positiva más cercana al margen de decisión, idénticamente x_- indica un vector desde el origen hacia la muestra negativa más cercana al margen de decisión y $(x_- - x_+)$ indica el vector diferencia.

Para obtener la distancia entre la muestra positiva y la muestra negativa más cercana al margen de decisión se podría encontrar la proyección del vector diferencia con un vector unitario en dirección normal a la línea que representa la recta el margen de separación entre ambas clases.

$$Margen = (\vec{x}_- - \vec{x}_+) \cdot \frac{\vec{w}}{\|\vec{w}\|} = \left(\vec{x}_- \cdot \frac{\vec{w}}{\|\vec{w}\|} - \vec{x}_+ \cdot \frac{\vec{w}}{\|\vec{w}\|} \right) \quad (3.36)$$

El resultado de la ecuación (3.36) es un escalar que representa la separación entre ambas clases.

Como ya se mencionó, la ecuación (3.35) representa las muestras sobre los hiperplanos H_0 y H_1 . Reemplazando \vec{x}_i por \vec{x}_+ y $y_i = 1$ se obtiene:

$$(\vec{w} \cdot \vec{x}_+ + b) - 1 = 0 \quad (3.37)$$

$$\vec{w} \cdot \vec{x}_+ = 1 - b \quad (3.38)$$

De la misma manera, se puede fijar reemplazando en (3.35) \vec{x}_i por \vec{x}_- y $y_i = -1$ se obtiene:

$$-(\vec{w} \cdot \vec{x}_- + b) - 1 = 0 \quad (3.39)$$

$$\vec{w} \cdot \vec{x}_- = -1 - b \quad (3.40)$$

Re escribiendo la ecuación (3.36):

$$Margen = \left(\vec{x}_- \cdot \frac{\vec{w}}{\|\vec{w}\|} - \vec{x}_+ \cdot \frac{\vec{w}}{\|\vec{w}\|} \right) \quad (3.41)$$

$$Margen = \frac{-1 - b - (1 - b)}{\|\vec{w}\|} = \frac{-2}{\|\vec{w}\|} \quad (3.42)$$

La ecuación (3.32) representa una sorprendente conclusión ya que indica que la separación entre los hiperplanos, distancia que se desea maximizar, es una función que depende exclusivamente del vector w .

Como se trata de un problema de maximización, maximizar el margen descrito en la ecuación (3.42) es equivalente a:

$$Max(Margen) = cte. \cdot Max\left(\frac{1}{\|\vec{w}\|}\right) = Min(\|\vec{w}\|) = \min\left(\frac{1}{2}\|\vec{w}\|^2\right) \quad (3.43)$$

Por lo tanto, dado un conjunto de datos linealmente separables, representados de la siguiente manera $S = [(x_1, y_1), \dots, (x_l, y_l)]$, un hiperplano representado por $wx + b = 0$ puede ser encontrado solucionando el siguiente problema de optimización:

$$\text{función objetivo: } \text{Min } \left(\frac{1}{2} \|\vec{w}\|^2\right) \tag{3.44}$$

$$\text{Sujeto a } y_i(x_i w + b) \geq 1 \quad \forall i = 1, 2, \dots, l \tag{3.45}$$

Una solución para el caso bidimensional deberá tener la forma representada en la Figura 3. 5. Los puntos para los que aplica la igualdad dada en la ecuación (3. 45) son aquellos que están sobre los hiperplanos H_0 y H_1 y se ven resaltados de sus pares por medio de una representación sin color de relleno. Estos puntos son llamados *vectores de soporte* [90].

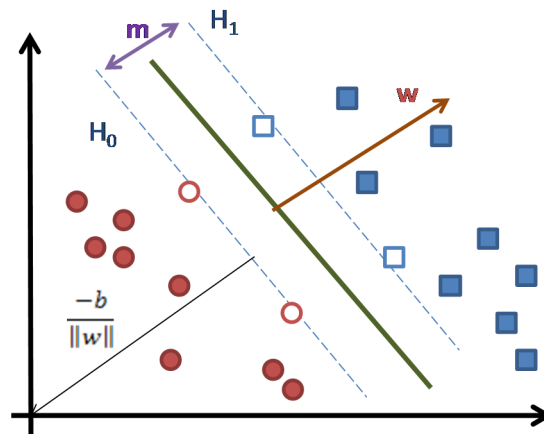


Figura 3. 5 - Vectores de soporte.

Para transformar el problema presentado en (3. 44) en un problema de optimización cuadrática convexa se debe realizar la siguiente transformación:

$$\text{minimizar}_{w,b} \frac{1}{2} (w \cdot w) \tag{3.46}$$

$$\text{sujeto a } y_i(x_i w + b) \geq 1 \quad \forall i = 1, 2, \dots, l$$

3.4.2.3 Optimización mediante multiplicadores de Lagrange.

Conforme lo presentado en la sección 3.3.4 se puede utilizar el método de los multiplicadores Lagrange para resolver el problema de optimización cuadrático presentado en la sección anterior.

- La restricción presentada en (3. 46) es sustituida por una nueva restricción que se define en función de los multiplicadores de Lagrange, lo que la hace más fácil de resolver computacionalmente.
- En esta formulación del problema, los datos de entrenamiento aparecen en forma de un producto interno entre vectores.

Introduciendo los multiplicadores de Lagrange $\alpha_i, i = 1, \dots, l$. Se realiza el producto entre la restricción de la ecuación (3. 25) y los multiplicadores de Lagrange positivos y ese producto es sustraído de la función objetivo para formar la función de Lagrange. Por lo tanto, para

solucionar el problema de optimización propuesto en la ecuación (3. 46) se debe encontrar el punto silla de la siguiente función de Lagrange:

$$\mathcal{L}_p = \frac{1}{2} (w \cdot w)^2 - \sum_{i=1}^l \alpha_i y_i (x_i w + b) + \sum_{i=1}^l \alpha_i \quad (3. 47)$$

De acuerdo a las ecuaciones (3. 13) y (3. 14) las derivadas de \mathcal{L}_p con relación a w y b debe ser nula y esto se corresponde al hecho que el punto óptimo tiene las siguientes ecuaciones de punto silla:

$$\frac{\partial \mathcal{L}_p(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^l y_i \alpha_i x_i = 0 \quad (3. 48)$$

$$\frac{\partial \mathcal{L}_p(w, b, \alpha)}{\partial b} = \sum_{i=1}^l y_i \alpha_i = 0 \quad (3. 49)$$

Sustituyendo las relaciones obtenidas se tiene:

$$w = \sum_{i=1}^l y_i \alpha_i x_i \quad (3. 50)$$

$$\sum_{i=1}^l y_i \alpha_i = 0 \quad (3. 51)$$

La ecuación (3. 50) tiene marcada relevancia ya que indica que el ancho de separación entre los hiperplanos es una función solo de las muestras (positivas y negativas). Para ciertos puntos en el plano α_i valdrá cero.

De acuerdo a lo expresado por el teorema de dualidad presentado en la sección 3.3.5 se puede transformar un problema primario en un problema dual fijando a cero la derivada del lagrangiano con respecto a las variables primarias y substituyendo las relaciones obtenidas en el lagrangiano. De esta manera se logra remover la dependencia de las variables primarias. Reescribiendo la ecuación (3. 47):

$$\mathcal{L}_D = \frac{1}{2} \sum_i \alpha_i \bar{x}_i y_i \sum_j \alpha_j \bar{x}_j y_j - \sum_i \alpha_i \bar{x}_i y_i \sum_i \alpha_i \bar{x}_i y_i - \sum_i \alpha_i y_i b + \sum_{i=1}^l \alpha_i \quad (3. 52)$$

En donde el tercer término es igual a cero de acuerdo a (3. 51). Reacomodando las sumatorias se tiene:

$$\mathcal{L}_D = \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \langle \vec{x}_i, \vec{x}_j \rangle - \sum_i \sum_j \alpha_i \alpha_j y_i y_j \langle \vec{x}_i, \vec{x}_j \rangle + \sum_{i=1}^l \alpha_i \quad (3.53)$$

Lo que finalmente deriva en:

$$\mathcal{L}_D = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \langle x_i x_j \rangle \quad (3.54)$$

Es importante destacar que \mathcal{L}_p es el problema primario y \mathcal{L}_D se denomina el problema dual. Una solución es encontrada por medio de la minimización de \mathcal{L}_p o la maximización de \mathcal{L}_D . Existe un multiplicador de Lagrange para cada punto del entrenamiento.

La resolución \mathcal{L}_p y \mathcal{L}_D se realizan mediante algoritmos de optimización matemática. A pesar que tanto el problema original como su problema dual se resuelven mediante técnicas similares, haber planteado el problema dual permite obtener la siguiente importante conclusión:

El problema de encontrar un plano de separación que maximice el margen depende solamente del producto punto entre las muestras.

Además de las restricciones arriba expuestas, las condiciones complementarias de *Karush-Kuhn-Tucker (KKT)* presentadas en la sección 3.3.4 proporcionan información útil sobre la estructura de la solución. Las condiciones dicen que, para una solución óptima α^* , w^* y b^* deben satisfacer:

$$\alpha_i^* [y_i(w^* x_i + b^*) - 1] = 0 \quad \text{para } i = 1, \dots, l \quad (3.55)$$

Para problemas convexos, las condiciones de *KKT* son necesarias y suficientes para que α^* , w^* y b^* sean solución, entonces, encontrar una solución para un *SVM* es equivalente a encontrar una solución para las condiciones de *KKT*.

Como los llamados vectores de soporte poseen α_i^* no nulos, ellos son los únicos involucrados en la expresión de los pesos w , por lo tanto, el vector que representa el hiperplano de margen máximo es calculado como una combinación lineal de la siguiente manera:

$$w^* = \sum_{i=1}^l y_i \alpha_i^* x_i \quad (3.56)$$

puede ser rescrita en función de los vectores de soporte:

$$w^* = \sum_{\text{vectores de soporte}} y_i \alpha_i^* x_i \quad (3.57)$$

Re escribiendo el problema nuevamente, ahora colocando la expresión para w^* en función del problema dual de Lagrange y teniendo en cuenta las condiciones de *KKT* el problema cuadrático convexo de *SVM* se transforma en el siguiente:

Dado un $\alpha_0 = (\alpha_1^0, \dots, \alpha_l^0)$ una solución para el problema arriba propuesto, entonces:

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j \langle x_i x_j \rangle \quad (3.58)$$

Sujeto a:

$$\sum_{i=1}^l y_i \alpha_i = 0 \quad (3.59)$$

$$\alpha_i \geq 0 \quad i = 1, \dots, l$$

una norma del vector w que corresponde al hiperplano óptimo es igual a:

$$\|w\|^2 = 2W(\alpha_0) = \sum_{\text{vectores de soporte}} \alpha_i^0 \alpha_j^0 \langle x_i x_j \rangle y_i y_j \quad (3.60)$$

La regla de separación, con base en el hiperplano óptimo es la función indicadora siguiente:

$$f(x) = \text{sgn} \left(\sum_{\text{vectores de soporte}} \alpha_i^0 \alpha_j^0 \langle x_i x_j \rangle y_i y_j \right) \quad (3.61)$$

en el cual x_i son los vectores de soporte y α_0 son los coeficientes de Lagrange correspondientes a los vectores de soporte y b_0 es un umbral constante dado por:

$$b_0 = \frac{1}{2} [\langle w_0 x^*(1) \rangle + \langle w_0 x^*(-1) \rangle] \quad (3.62)$$

donde $x^*(1)$ se corresponde a cualquier vector de soporte que pertenece a la primera clase y $x^*(-1)$ se corresponde a cualquier vector de soporte perteneciente a la segunda clase.

Un clasificador de mínimo margen queda restringido a aplicaciones cuando los datos son linealmente separables.

3.4.2.4 Clasificador para datos no linealmente separables.

Para tratar los casos linealmente no separables es necesario relajar las restricciones del problema. Si el conjunto S no es linealmente separable, violaciones a la clasificación deben ser permitidas en la formulación de la *SVM*.

De este modo, las restricciones presentadas en las ecuaciones (3.28) y (3.31) denominadas restricciones de margen duro (*hard margin*) son transformadas en restricciones de margen

blando (*soft margin*) mediante la incorporación de una variable de relajación ξ_i $i = 1, \dots, l$ de la siguiente manera:

$$x_i w + b \geq +1 - \xi_i \quad \text{para } y_i = +1 \quad (3.63)$$

$$x_i w + b \leq -1 + \xi_i \quad \text{para } y_i = -1 \quad (3.64)$$

Esta estrategia permite una cierta tolerancia a datos denominados *outliers*, Un *outlier* es un dato muy distante de los restantes puntos de la clase a la que pertenecen. Por lo tanto la expresión $\sum_{i=1}^N \xi_i$ representa el número de errores de tratamiento. Debido a esto, con el fin de poder representar los costos extras por los errores que resultan de la adición de las variables de holgura, existe la necesidad de cambiar la función objetivo presentada en la ecuación (3.46) para el caso *linealmente separable* de la siguiente manera:

$$\frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^N \xi_i \right)^k \quad (3.65)$$

donde el parámetro C es una constante que actúa como una función de penalidad previniendo que los *outliers* afecten el cálculo del hiperplano óptimo.

Un valor grande de C hace que se asume mayor penalidad por incorporar errores. Por tratarse de problemas de programación convexa k puede ser cualquier valor entero positivo en particular $k=1$ y $k=2$ se trata de un problema de programación cuadrática. Para el caso del valor más pequeño de $k>0$ tiene la ventaja de no ser necesario que ξ_i y sus multiplicadores de Lagrange aparezcan en el problema dual. El problema con esta modificación se vuelve:

$$\mathcal{L}_D \equiv \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \langle x_i x_j \rangle \quad (3.66)$$

sujeto a:

$$\sum_{i=1}^l y_i \alpha_i = 0 \quad (3.67)$$

$$0 \leq \alpha_i \leq C \quad i = 1, \dots, l$$

y la solución viene dada por

$$w = \sum_{\text{vectores de soporte}} \alpha_i y_i x_i \quad (3.68)$$

Así que la única diferencia para el caso del hiperplano óptimo es que α_i tiene un límite superior en C . Esta situación se ve representada en la Figura 3.6.

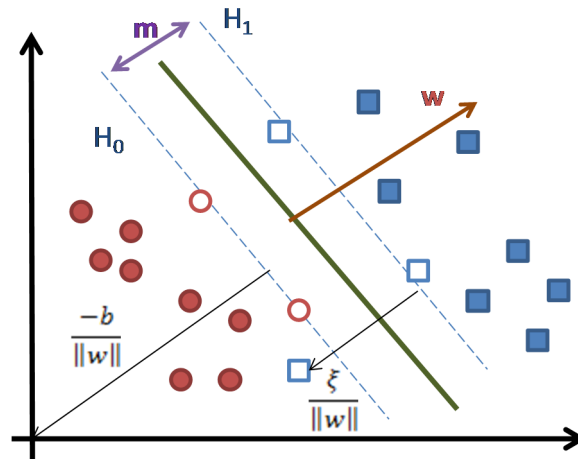


Figura 3. 6 - Hiperplano para el caso no linealmente separable.

3.4.3. Transformación de sub espacio usando funciones Kernel.

De acuerdo al teorema de Cover un problema complejo de clasificación de patrones tiene más probabilidades de ser separable linealmente mientras mayor sea la dimensión del espacio característico.

Considerando un espacio de entrada en donde los patrones no son linealmente separables, el teorema de Cover dice que ese espacio puede ser transformado en un nuevo espacio característico donde los patrones tienen mayor posibilidad de ser linealmente separables sobre las siguientes dos condiciones: La transformación nunca puede ser lineal y la dimensión del nuevo espacio característico debe ser de mayor que la dimensión del espacio característico de los datos de entrada.

Una vez transformado el espacio característico la idea nuevamente recae en construir un hiperplano de margen máximo de acuerdo a lo descrito en la sección 3.4.2 para clasificar las dos clases de datos pero con una diferencia fundamental:

El hiperplano de margen máximo ahora se construye como una función lineal de vectores del espacio característico transformado en lugar de hacerlo con los vectores del espacio característico original.

Vapnik [89] observó que, para la construcción del hiperplano de margen máximo en un espacio característico Z , no es necesario considerar todo el espacio de forma explícita. En su lugar, se puede simplemente calcular el producto interno entre los vectores de soporte y los vectores del espacio característico.

Suponiendo que se establecieran patrones $x \in \mathbb{R}^N$ donde la mayor parte de la información está contenida en los monomios de grado d donde:

$$x_{j_1}, \dots, x_{j_d}$$

donde $j_1, \dots, j_d \in \{1, \dots, N\}$

En tal situación, sería interesante que extraer primero los monomios (o productos de las posiciones específicas dentro de los vectores de características) y trabajar en un espacio de características Z en representación de todos los monomios de grado d . Tomando como ejemplo $x \in \mathbb{R}^2$ se puede seleccionar los monomios de grado 2 a través del siguiente mapeo no lineal:

$$\phi: \mathbb{R}^2 \rightarrow Z = \mathbb{R}^3 \quad (3.69)$$

$$(x_1, x_2) \rightarrow (x_1^2, x_2^2, x_1x_2) \quad (3.70)$$

En ciertas situaciones hay una manera de computar los productos internos sin realizar un mapeo a través de los *kernels no lineales*.

Para computar productos internos de la forma $\langle \phi(x) \cdot \phi(y) \rangle$ se propone la siguiente representación para un *kernel*.

$$K(x, y) = \langle \phi(x) \cdot \phi(y) \rangle \quad (3.71)$$

Supongamos que $N = d = 2$ por lo que los monomios deben ser de grado $d=2$:

$$C_2(x): \langle x_1x_2 \rangle \rightarrow (x_1^2, x_2^2, x_1x_2, x_2x_1) \quad (3.72)$$

Supongamos ahora la misma transformación para y :

$$C_2(y): \langle y_1y_2 \rangle \rightarrow (y_1^2, y_2^2, y_1y_2, y_2y_1) \quad (3.73)$$

el producto interno entre ambos monomios nos da:

$$\langle C_2(x)C_2(y) \rangle \rightarrow x_1^2y_1^2 + x_2^2y_2^2 + 2x_1x_2y_1y_2 = (x \cdot y)^2 \quad (3.74)$$

la distribución de los monomios brinda las condiciones para que el *kernel* K deseado se obtiene simplemente mediante el cuadrado del producto interno de los espacios de entrada.

Sea C_d una función que mapea $x \in \mathbb{R}^N$ para el vector $C_d(x)$, cuyas entradas son todos los posibles productos (monomios) ordenados de grado d de las entradas x . Entonces, una función *kernel* que es equivalente a realizar el producto interno de los vectores mapeados por C_d y:

$$k(x, y) = \langle C_d(x)C_d(y) \rangle = (x \cdot y)^d \quad (3.75)$$

El objetivo de este ejemplo es el de demostrar que un producto interno puede ser calculado sin el mapeo tradicional por medio de las funciones *kernel*.

Las representaciones *kernel* trabajan como proyecciones del espacio donde los datos residen a un espacio de características de más alta dimensión para permitir la clasificación de los datos que, en su espacio característico original, son no linealmente separables. En la Figura 3.7 se puede ver como un conjunto de datos *no linealmente separables* se transforma en un conjunto *linealmente separable* en un nuevo espacio característico, de mayor dimensión, creado por la función *kernel* ϕ .

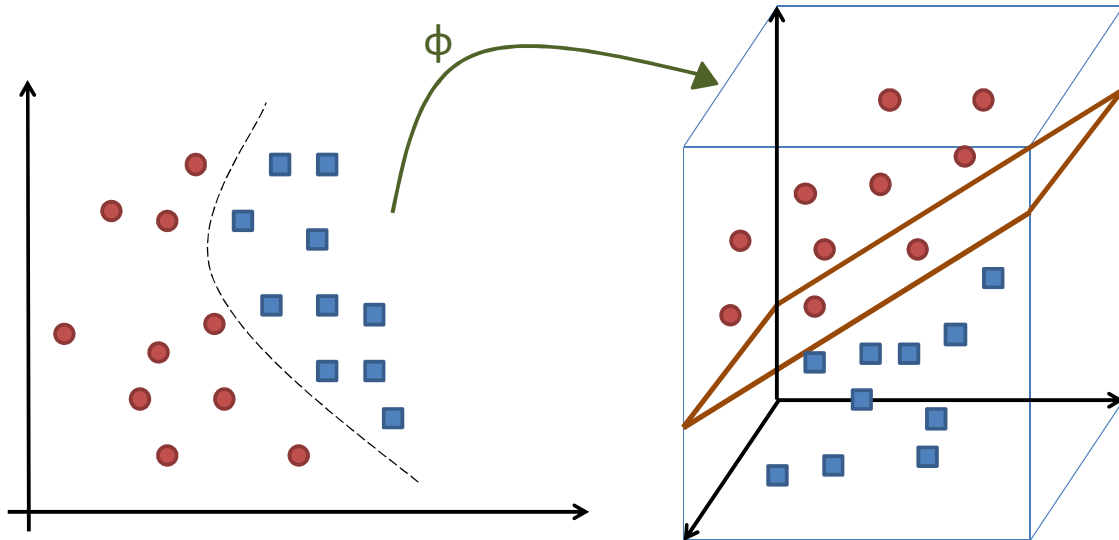


Figura 3. 7 - Transformación del espacio de características usando funciones kernel.

3.4.3.1 Funciones kernel.

Uno de los principales factores a tener en cuenta cuando se diseña un clasificador SVM es encontrar la función *kernel* que mejor se adapta al problema.

De manera análoga a lo que representa la configuración de una red neuronal artificial, para el caso de los SVM, la selección de la arquitectura es equivalente a seleccionar una apropiada función *kernel* [85]. Cada función *kernel* debe ser estudiada teniendo en cuenta el problema en particular.

Las condiciones que una función *kernel* debe cumplir son:

- Debe ser simétrica: $K(x, z) = \langle \phi(x), \phi(z) \rangle = K(z, x)$
- Debe cumplir con la desigualdad de *Cauchy-Schwarz*:

$$K(x, z)^2 = (\langle \phi(x), \phi(z) \rangle)^2 \leq \|\phi(x)\|^2 \|\phi(z)\|^2$$

$$= \langle \phi(x), \phi(x) \rangle \langle \phi(z), \phi(z) \rangle = K(x, x) K(z, z)$$

- Deben cumplir con el teorema de *Mercer*: La condición de *Mercer* garantiza que una función simétrica $K(x, z) = \sum_i \phi(x)_i \phi(z)_i$ es un *kernel* si la siguiente integral:

$$\iint K(x, z) g(x) g(z) \delta x \delta z \geq 0$$

para

$$\int g^2(x) \delta x \leq \infty$$

Las funciones kernel más comúnmente empleadas son:

- Máquinas de aprendizaje *polinomial*: Una máquina de aprendizaje *polinomial* de grado d está especificada por:

$$k(x, x_i) = [\langle x x_i \rangle + 1]^d \quad (3.76)$$

- Máquinas de funciones de base radial: Es producida por una red del tipo *Radial Basis Function (RBF, por sus siglas en inglés)* y está dada por:

$$k(x, x_i) = e^{\frac{-\|x-x_i\|^2}{2\sigma^2}} \quad (3.77)$$

- Redes neuronales de dos capas: se puede definir

$$k(x, x_i) = S[v(x, x_i) + c] \quad (3.78)$$

- donde v y c son parámetros escalares y $S(u)$ es la función sigmoidea.

Cada función *kernel* debe ser estudiada teniendo en cuenta el problema en particular.

3.4.3.2 Incorporación del kernel al SVM.

Con la introducción de la función kernel el problema de SVM no se ve alterado en gran medida. Para encontrar los α_i es preciso resolver el siguiente problema:

$$\text{maximizar } \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j K(x_i x_j) \quad (3.79)$$

Sujeto a:

$$\sum_{i=1}^l y_i \alpha_i = 0 \quad (3.80)$$

$$0 \leq \alpha_i \leq C \quad i = 1, \dots, l$$

$$\sum_{i=1}^l y_i \alpha_i = 0 \quad (3.81)$$

$$0 \leq \alpha_i \leq C \quad i = 1, \dots, l$$

Es importante observar que la forma de la ecuación es idéntica a la que se utilizó para encontrar el hiperplano de margen máximo excepto que el producto interno entre $x_i x_j$ ahora es realizado a través de la función *kernel* $K(x_i x_j)$.

3.4.4. Máquinas de Vectores de Soporte Multiclase.

En este capítulo primero se abordó el problema de la clasificación entre dos clases de datos linealmente separables. Debido a que este tipo de problemas rara vez se presenta en la vida real, se agregó la teoría de los márgenes relajados (*softmargin*) y las funciones *kernel* para la clasificación de patrones compuestos por dos clases que no cumplen con la condición de ser linealmente separables.

En esta sección se presentará una estrategia para hacer extensivo el uso de los *SVM* cuando el problema involucre más de dos clases de datos [91].

Los clasificadores *SVM* fueron originalmente diseñados para realizar tareas de clasificación binaria. La mejor manera de extender su aplicación a los casos donde más de una clase está presente es una tarea que aún se encuentra bajo la lupa de los investigadores.

Para lograr que los *SVM* sean capaces de clasificar problemas con múltiples clases dos aproximaciones hasta ahora fueron propuestas:

La primera se basa en utilizar un conjunto de clasificadores binarios estratégicamente relacionados. Esta aproximación tiene, como principales exponentes, los métodos conocidos como *Uno vs el Resto*, *Uno vs Uno* y *Árbol de decisión binaria*.

La segunda aproximación propone resolver los problemas multi clase mediante el desarrollo de un único problema de optimización. Este tipo de problemas presenta una compleja conformación matemática.

3.4.4.1 Uno contra el resto (*One versus Rest*).

En este método K diferentes clasificadores son construidos, uno para cada una de las posibles k clases que se encuentran presentes en el set de entrenamiento. El clasificador i -th se entrena sobre todo conjunto de datos para aprender a distinguir la clase i de las restantes clases. En este esquema, para el i -ésimo clasificador, los miembros de la i -ésima clase son rotulados con un $(+1)$ y los restantes datos son rotulados con (-1) [92].

Así, dado un set de entrenamiento compuesto por l muestras y K clases $(x_1, y_1), (x_2, y_2) \dots (x_l, y_l)$, donde $x_i \in \mathbb{R}^n$, $i = 1, 2, \dots, l$ e $y_i \in (1, \dots, k)$ es la clase de x_i , para el i -ésimo clasificador se resuelve de la siguiente manera:

$$\min_{w^i, b^i, \xi^i} = \frac{1}{2} (w^i)^T w^i + C \sum_{j=1}^l \xi_j^i \quad (3.82)$$

$$(w^i)^T \phi(x_j) + b^i \geq 1 - \xi_j^i \quad Si \ y_j = i \quad (3.83)$$

$$(w^i)^T \phi(x_j) + b^i \leq -1 + \xi_j^i \quad Si \ y_j \neq i \quad (3.84)$$

$$\xi_j^i \geq 0 \quad j = 1, 2, \dots, l \quad (3.85)$$

Donde cada una de las muestras presentes en el set de entrenamiento se mapean a un nuevo espacio de mayor dimensión mediante la función kernel ϕ y C se lo conoce como factor de penalización.

Después de resolver la ecuación (3. 82) se obtienen K funciones de decisión:

$$(w^i)^T \phi(x) + b^i \quad (3. 86)$$

⋮

$$(w^k)^T \phi(x) + b^k \quad (3. 87)$$

Es decir, que por cada muestra x que se desea probar se tomarán k decisiones. La clase de x se podrá decidir con el valor más grande de la función de decisión dada por:

$$\text{clase de } x = \operatorname{argmax}_{i=1,\dots,k} ((w^k)^T \phi(x) + b^i) \quad (3. 88)$$

3.4.4.2 Uno contra uno (One vs One).

Para cada una de las clases posibles un clasificador binario es entrenado. Cada clasificador es entrenado utilizando parte del conjunto de entrenamiento que contiene algunos ejemplos de dos de las K clases involucradas. En este caso una cantidad de $K(K-1)/2$ clasificadores son construidos y entrenados. Cada clasificador es entrenado por dos clases [93].

Para la clase i y j el problema de clasificación binario viene dado por:

$$\min_{w^{ij}, b^{ij}, \xi^{ij}} = \frac{1}{2} (w^{ij})^T w^{ij} + C \sum_t \xi_t^{ij} \quad (3. 89)$$

$$(w^{ij})^T \phi(x_t) + b^{ij} \leq -1 + \xi_t^{ij} \quad \text{Si } y_t \neq j \quad (3. 90)$$

$$(w^{ij})^T \phi(x_t) + b^{ij} \leq -1 + \xi_t^{ij} \quad \text{Si } y_t \neq j \quad (3. 91)$$

$$\xi_t^{ij} \geq 0 \quad (3. 92)$$

La principal desventaja de esta metodología es que la respuesta final viene dada por una estrategia de votación [94].

3.4.4.3 Árbol de jerarquía de clasificadores.

Este método se conoce también con las siglas *DAGSVM* (*Directed Acyclic Graph Support Vector Machine*). Mediante este método el problema de clasificación es descompuesto en una serie de sub-problemas de clasificación binaria organizados en un esquema jerárquico [95, 96].

La etapa de entrenamiento, al igual que en el método uno contra uno presentado en la sub sección 3.4.4.2, involucra la conformación de $K(K-1)/2$ clasificadores binarios. Sin embargo,

en la fase de prueba se emplea un esquema tipo árbol de decisión tal como se muestra en la Figura 3. 8.

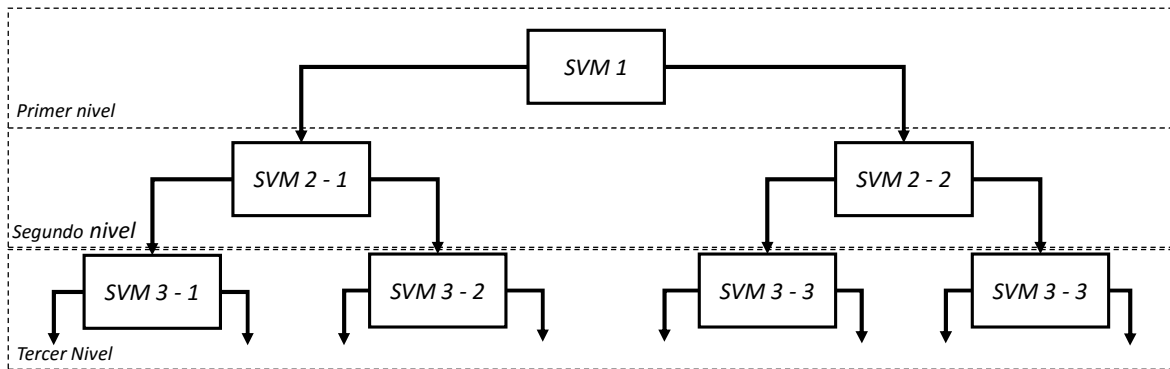


Figura 3. 8 - Arbol de decisión binaria.

Cada nodo del árbol es un clasificador *SVM* binario entrenado para distinguir la clase i de la j . Dado un determinado patrón de entrada, la clasificación comienza en el nodo raíz. En este nodo se toma la primera decisión binaria. Teniendo en cuenta la decisión del nodo raíz, el proceso continúa aguas abajo hasta completar la evaluación. La última etapa de decisión debería culminar en una de las K clases.

En el trabajo de Chih-Wei y Chih-Jen (2002) [97] se realiza una interesante comparación entre las diferentes opciones que se pueden adoptar para desarrollar un clasificador *SVM multiclase*. En esta publicación se concluye que el método *Uno contra Uno* y *Uno contra el Resto* presentan muy similares prestaciones en cuanto al porcentaje de aciertos y al tiempo de entrenamiento, teniendo el primero una pequeña ventaja. A pesar de estos resultados, para un problema donde necesita clasificar una combinación de clases, como es el caso de perturbaciones complejas, el método que mejor se adapta es *Uno contra el Resto*.

3.5. APLICACIÓN DE MAQUINAS DE VECTORES DE SOPORTE EN ESTUDIOS ELECTRICOS.

Las técnicas de *AI* pueden describirse como la automatización de actividades relacionadas con el pensamiento humano tales como su capacidad de aprendizaje a través de ejemplos, percepciones, razonamiento, la toma de decisiones y la resolución de problemas a través de la capacidad de generalizar. Las técnicas de *AI* empleadas para los problemas relacionados con la clasificación de perturbaciones eléctricas y la localización de fallas son variadas, dentro de las cuales las más destacadas son, las *ANN*, *FL*, *Sistemas Expertos (ES, por sus siglas en inglés)*, técnicas de *clustering* y la más novedosa, las *SVM*.

En las próximas sub secciones se presentan los aspectos y conclusiones más importantes de las publicaciones que adoptan *SVM* como técnica de clasificación.

3.5.1. Clasificación de perturbaciones.

Tiempo atrás, las *ANN*, en sus diferentes configuraciones, se convirtieron en la técnica más ampliamente adoptada por la comunidad científica para la clasificación de perturbaciones. Sin embargo, diferentes recientes trabajos indican que *SVM* es una herramienta superadora para la clasificación de patrones.

En este sentido Peisheng y Weilin (2006) presentan una comparación del rendimiento de un clasificador basado en una *Red Neuronal Artificial (ANN)* del tipo *backpropagation* y un clasificador basado en *Maquinas de Vectores de Soporte (SVM)* para la clasificación de las perturbaciones en un *SEP* [98].

En [99] Janik and T. Lobos (2006) presentan la comparación de un método basado en *Maquinas de Vectores de Soporte* y una red neuronal del tipo *RFB*.

Por su parte, Vega (2009) realiza una comparación del rendimiento y tiempo de entrenamiento de un sistema de clasificación de perturbaciones basado en una red neuronal perceptrón multi capa y un clasificador multi clases *One vs One SVM* [100].

Weiming, Xuelei, Jingbo y Zhiheng (2006) [32], publicaron un artículo donde se plantea la comparación de un clasificador *SVM* configurado en una arquitectura del tipo *árbol de decisión binaria basado* y una red neuronal para la clasificación de los eventos de calidad de energía.

En todos los trabajos arriba presentados se concluye que *SVM* es una herramienta superadora si se tiene en cuenta el tiempo de entrenamiento, tiempo de prueba, confiabilidad y exactitud como métricas de calidad.

En [97] se realiza una interesante comparación entre las diferentes opciones que se pueden adoptar para desarrollar un clasificador *SVM* multiclase. En esta publicación se concluye que el método *Uno contra Uno* y *Uno contra el Resto* presentan muy similares prestaciones en cuanto al porcentaje de aciertos y al tiempo de entrenamiento, teniendo el primero una pequeña ventaja. A pesar de estos resultados, para un problema donde necesita clasificar una

combinación de clases, como es el caso de *perturbaciones complejas*, el método que mejor se adapta es *Uno contra el Resto*.

Los sistemas de clasificación que basan su funcionamiento en métodos de *AI* deben ser entrenados antes de ser utilizados en un escenario real. Por este motivo, para aplicaciones comerciales, es esperable que el usuario no tenga que entrenar el sistema clasificador antes de utilizarlo. En su lugar dicho clasificador debería encontrarse configurado y entrenado por el fabricante. Teniendo en cuenta esta problemática los autores Axelberg, Gu y Bollen (2007) publicaron un trabajo donde se simula el comportamiento de un clasificador basado en *SVM* en situaciones reales cuando el algoritmo es entrenado con datos reales y datos simulados. La estrategia seleccionada en esta publicación para configurar un clasificador multiclase fue *uno contra todos* en una arquitectura tipo árbol de decisión binaria [101].

La arquitectura utilizada para los clasificadores multiclase basados en *SVM* usualmente utiliza un esquema *uno contra uno* o *uno contra el resto*. El tamaño de estas redes crece marcadamente a medida que se agregan nuevas clases a los datos a clasificar. Por este motivo, Whei-Min, Chien-Hsien, Chia-Hung y Fu-Sheng (2008) presentaron un sistema de procesamiento que utiliza la *WT* para la extracción de parámetros característicos de una señal y una estrategia denominada *disturbance vs normal (DVN por sus siglas en inglés)* para la clasificación de las perturbaciones. En este esquema un *SVM* por cada tipo de perturbación es entrenado para distinguir un evento de calidad de energía de una señal limpia de perturbaciones. El sistema de procesamiento propuesto por los autores permite, además de la detección y clasificación, la localización de un evento de calidad de energía en el SEP [102].

El trabajo de Milchevski y Taskovski (2011) presenta un método que utiliza la energía de los coeficientes de la *WT* para la extracción de parámetros característicos de una señal perturbada. Adicionalmente, para facilitar la tarea del sistema de clasificación, los autores proponen un método de umbralización de la amplitud de la señal cuya finalidad es la extracción de tres parámetros característicos adicionales. El sistema de clasificación propuesto por los autores se basa en un árbol de decisión binaria donde un modelo *SVM* es creado para cada nodo. [103]

Un reciente trabajo de Erisi y Demir (2012) presenta un completo sistema de procesamiento de perturbaciones basado en la *WT* y las *SVM* para la detección, localización y clasificación de las perturbaciones en un sistema eléctrico de potencia. El sistema de procesamiento fue desarrollado en MATLAB, utilizando una librería llamada *LibSVM* la cual utiliza el método *uno contra todos* para la clasificación de patrones multi clase. La función *kernel* elegida por los autores fue *RBF* y se utilizó un proceso de búsqueda basado grillas para encontrar los parámetros de configuración óptimos de *SVM*. [104].

En su trabajo, Biswall, Dash y Mishra presentan un esquema conformado por una transformada del dominio del tiempo al dominio del tiempo (Transformada T-T) para la clasificación de patrones de perturbación no estacionarios en señales provenientes de sistemas eléctricos de potencia. La Transformada T-T no es ni más ni menos que una variante de la *ST* y utiliza una nueva función para el ventaneo que es proporcional a la frecuencia de

la señal a analizar. Los parámetros característicos derivados de la transformada se utilizan para el reconocimiento automático empleando *SVM* basado en *clustering*. El trabajo también presenta un algoritmo de optimización inmune para mejorar el rendimiento del sistema Transformada T-T – SVM. [105].

3.5.2. Localización de fallas en sistemas de distribución.

No son muchas las investigaciones que adoptan métodos de *AI* para la localización de fallas en sistemas de distribución eléctrica.

Un método [63] basado en una *ANN* y *SVM*, para la localización de fallas en un sistema eléctrico radial es propuesto por Thukaram (2005). El método propuesto utiliza mediciones de tensión y corrientes obtenidas en una subestación así como el estado de seccionadores y relés como parámetros característicos para entrenar y probar el clasificador.

Un método que propone el análisis de patrones de corriente utilizando *SVM* para la localización de fallas en sistemas eléctricos de topología radial es propuesto por Mora-Flores (2008) [65]. Este método se vale en la información contenida en las mediciones de corriente producto del accionar de seccionadores y dispositivos de protección alocaados en distintos sitios de la red.

Apisit (2012) propone un método [64] basado en la combinación de *DWT* y *SVM* para la localización de fallas en cableado subterráneo. Este método propone la utilización del máximo valor de secuencia positiva calculado sobre la corriente como parámetro característico.

Lovisoló (2012) propone un método [62] basado en una técnica de procesamiento de señales conocida como sinusoides amortiguadas para la extracción de parámetros característicos. Los parámetros característicos son procesados por una etapa basada en *SVM* para identificar el tipo de falla y un segundo módulo, basado en *ANN* para su localización.

En [61], Agrawal (2013) propone un método que utiliza un clasificador *SVM* multi clase para localizar y diagnosticar fallas en sistemas eléctricos de distribución con alta penetración de energía distribuida. El método propuesto se basa en las mediciones de tensión y corriente en las tres fases obtenidas en las subestaciones donde las fuentes de generación distribuida se conectan con la red.

Ye Lei (2014), por su parte, propone un método que combina la *DWT* con un algoritmo de regresión basado en *SVM* para localizar fallas en sistemas de distribución radial subterráneos [69].

Recientemente, Deng (2015), propone un método basado en la *DWT* y *SVM* para la localización de fallas en una red de distribución del tipo bucle [66].

Finalmente, en el trabajo propuesto Ray y Mishra (2015) se presenta un método híbrido para detectar fallas en un sistema de distribución de 11 Kilo Volts [106]. La información colectada se procesa mediante el *Paquete de Transformada Wavelets (WPT)* con el objetivo

de extraer parámetros característicos. Un *Algoritmo Genético* se emplea para la selección de parámetros característicos. El sub set de parámetros óptimos es empleado para entrenar una *ANN* y una etapa *SVM* con el objetivo de localizar la falla.

3.6. CONCLUSIONES.

A pesar de resultar una herramienta de notoria complejidad matemática, la literatura indica que las *SVM* presentan resultados equivalentes o superiores frente a otras técnicas de clasificación.

En la primera parte del capítulo se presentaron las herramientas necesarias para el entendimiento de la matemática involucrada en el entrenamiento de una Máquina de Vectores de Soporte. Con tal fin se presentaron algunos conceptos relacionados con la *Teoría de Aprendizaje* y la *Optimización Matemática*.

En la segunda parte del capítulo se presentaron los conceptos teóricos y matemáticos involucrados en el uso y entrenamiento de un clasificador basado en *SVM*. En esta sección también fueron abordadas varias estrategias para transformar un clasificador *SVM binario* en un clasificador multiclase.

Finalmente fueron presentadas las publicaciones más relevantes donde se utilizan las *SVM* para conformar un sistema de clasificación de eventos de calidad de energía y la localización de fallas en un sistema eléctrico de potencia.

Segunda Parte

Desarrollo de la propuesta

Capítulo 4:
Marco Conceptual para el
Desarrollo de un
Clasificador basado en
Inteligencia Artificial

4.1 INTRODUCCION

En ciencias de la computación el aprendizaje automático o aprendizaje de máquinas es una rama de la *AI* cuyo objetivo es desarrollar técnicas que permitan a los algoritmos *aprender a resolver problemas* en lugar de *memorizar su solución*.

De forma más concreta, se trata de crear programas capaces de generalizar comportamientos a partir de una información no estructurada suministrada en forma de ejemplos. Por lo tanto se puede concluir que un algoritmo de aprendizaje debe ser entrenado para que aprenda a generalizar relaciones entre un conjunto de parámetros de entrada y las respuestas esperadas.

El algoritmo produce una función que establece una correspondencia entre las entradas y las salidas deseadas del sistema. Un ejemplo de esto es un problema donde se necesita clasificar en diferentes categorías o clases un determinado número de ejemplos. En este caso el sistema de aprendizaje trata de etiquetar (clasificar) cada ejemplo, representado por un vector de características de entrada, empleando una entre varias categorías predefinidas (clases). La base de conocimiento del sistema está formada por ejemplos etiquetados.

El desarrollo de un sistema de clasificación basado en *AI* se realiza en dos etapas: Una etapa de diseño y entrenamiento y la etapa de clasificación. Este concepto se resume en la Figura 4. 1.

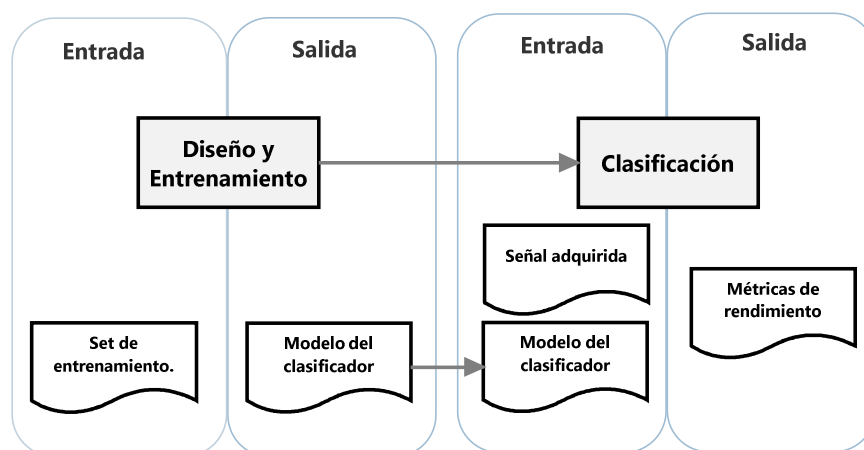


Figura 4. 1 - Sistema de clasificación.

La etapa de *Diseño y Entrenamiento* tiene como objetivo la obtención de un modelo del clasificador entrenado y listo para ser empleado en la etapa de clasificación.

El diseño del clasificador consiste en encontrar un conjunto de parámetros que gobiernan el clasificador que mejor se adaptan a cada problema en particular. Como entrada a este proceso se cuenta con un set de entrenamiento conformado por parámetros característicos que representan las clases que el clasificador deberá aprender a distinguir y sus respectivas etiquetas.

En una primera instancia se ajustan los parámetros del algoritmo para luego proceder con el algoritmo de entrenamiento supervisado. Como resultado se obtiene el modelo del clasificador entrenado para identificar cada clase presente en el set de entrenamiento.

En la etapa de *Clasificación*, el algoritmo emplea el modelo obtenido en la etapa de *Diseño y Entrenamiento* para determinar a qué clase pertenece el conjunto de parámetros presente en la entrada.

Es común emplear la tasa de aciertos (cantidad de aciertos con respecto a la cantidad de muestras total) como métrica de rendimiento del clasificador.

En las siguientes secciones se detallan los procesos desarrollados para diseñar, entrenar y probar un sistema de clasificación basado en *Máquinas de aprendizaje lineal*.

4.1.1 Objetivos del capítulo.

El objetivo de este capítulo es presentar el desarrollo de una metodología general para el diseño, entrenamiento y prueba de un sistema de clasificación basado en una sub categoría de la *AI* conocida como *Máquinas de Aprendizaje Lineal*, enfocado en aquellos métodos que utilizan técnicas de aprendizaje supervisado.

La metodología aquí presentada será posteriormente utilizada en esta tesis, en el *capítulo 6*, para el entrenamiento de un sistema de clasificación de perturbaciones eléctricas. Del mismo modo la metodología propuesta se emplea, en el *capítulo 7*, para el desarrollo de un sistema que permite la identificación del tipo de falla y su localización en un sistema eléctrico de distribución. En ambos casos, el sistema de clasificación se basa en una técnica de aprendizaje lineal conocida como *Máquinas de Vectores de Soporte*.

Este capítulo se organiza de la siguiente manera: La sección 4.2 presenta el desarrollo del algoritmo para el Diseño y Entrenamiento del clasificador, mientras que la sección 4.3 detalla el desarrollo del algoritmo destinado a probar y validar el método.

4.2 ALGORITMO DE DISEÑO Y ENTRENAMIENTO DEL CLASIFICADOR.

El entrenamiento de la mayoría de los algoritmos de *AI* se trata de una metodología que tiene por objeto deducir una función que vincule un conjunto de patrones de entrada con un determinado número de respuestas válidas.

Existen muchas clases de entrenamiento siendo el *entrenamiento supervisado* uno de los más intuitivos y ampliamente elegidos. El entrenamiento supervisado es un proceso que consiste en utilizar un par de objetos (un patrón de entrada y su correspondiente etiqueta) para crear una función capaz de decidir que etiqueta corresponde a cualquier objeto de entrada.

Un algoritmo de *AI* tiene la capacidad de *generalizar* si puede entregar respuestas adecuadas aún ante patrones que son similares, aunque no idénticos, a los empleados para el entrenamiento. La serie de ejemplos conocidos y su respectiva etiqueta se conoce como set de entrenamiento.

El algoritmo de diseño y entrenamiento mostrado en la Figura 4. 2 está conformado por una secuencia de procesos cuyo objetivo es:

- Encontrar el conjunto óptimo de parámetros característicos que describen a cada clase que se desea clasificar.
- Sintonizar los parámetros que gobiernan el comportamiento del método para optimizar su capacidad de interpretar la información de entrada y expresar una respuesta esperada.
- Generar, como salida, un modelo optimizado y entrenado del sistema que será posteriormente empleado en la etapa de prueba o puesta en producción para la clasificación de patrones propuestos como entrada.

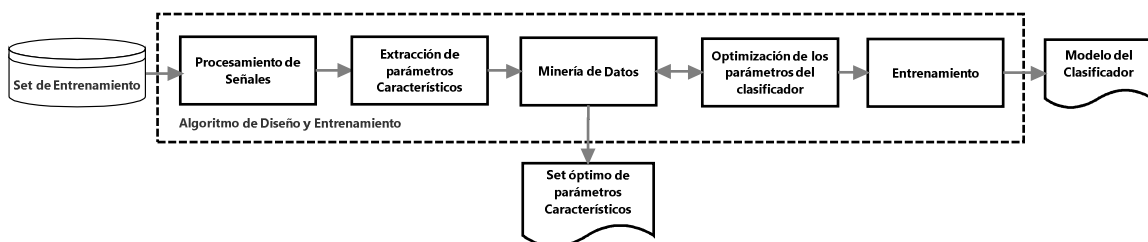


Figura 4. 2 - Algoritmo de diseño y entrenamiento.

Es importante aclarar que el algoritmo presentado en este capítulo se enfoca en el análisis de formas de onda de un sistema eléctrico de potencia por lo que su operación se basa en el procesamiento de señales para la obtención de parámetros que caracterizan a cada clase que se intenta clasificar.

En las siguientes subsecciones se explica detalladamente la función que cada proceso desempeña.

Para el desarrollo de las siguientes secciones se supone que se tiene, como entrada del algoritmo, un conjunto de m formas de onda las que se pueden clasificar en K diferentes tipos o clases.

4.2.1 Procesamiento de Señales.

Las técnicas que conforman las funcionalidades del módulo de procesamiento de señales se dividen, a su vez, en dos categorías, de acuerdo a las funciones que desempeñan. Estas categorías son: Técnicas de pre procesamiento y técnicas de procesamiento.

4.2.1.1 Técnicas de pre procesamiento.

Las técnicas de pre procesamiento son un conjunto de herramientas que tienen por objetivo el ajuste y normalización de las magnitudes de las formas de onda que conforman el set de entrenamiento de forma tal que, luego, el procesamiento de las mismas retorne valores normalizados y comparables entre sí. Los principales objetivos de las técnicas de pre procesamiento son:

- Disminuir el efecto del ruido en las mediciones: Debido a que las mediciones de campo inevitablemente se ven contaminadas con ruido, en esta etapa se implementan técnicas para mitigar su efecto. La comunidad científica llama a estos procedimientos como técnicas de *de-noising*, por su traducción al inglés [83].
- Fijar una frecuencia de muestreo apropiada: La frecuencia de muestro es un factor que determinará la cantidad de información a ser procesada en posteriores etapas del algoritmo. Por lo tanto, en esta etapa, se debe configurar un valor apropiado estableciendo valores no demasiado pequeños para evitar descartar información valiosa de eventos rápidos pero que, al mismo tiempo, no sea lo suficientemente grande como para generar gran cantidad de información redundante.
- Normalizar las magnitudes a valores por unidad. La amplitud en estado estacionario de una señal medida en un sistema es un factor que no debe influir en el entrenamiento del algoritmo clasificador. Por este motivo, en esta etapa se produce la normalización de las formas de onda para obtener una amplitud de salida normalizada a valores por unidad.
- Acotar la duración de las señales: En esta etapa se implementan técnicas de ventaneo de las señales para acotar la cantidad de información a ser procesada por el algoritmo.

La configuración de la frecuencia de muestreo y la longitud de la ventana seleccionada fijan la cantidad de muestras que representan a cada forma de onda. A este valor lo denominamos en esta instancia con la letra s .

Un set de entrenamiento, por lo tanto, estará conformado por una matriz de dimensión $[m \times s]$ donde m se corresponde con el número de formas de ondas empleadas para el entrenamiento y s la cantidad de muestras que representan cada perturbación.

4.2.1.2 Técnicas de procesamiento de señales.

La detección, clasificación y extracción de parámetros característicos rara vez se lleva a cabo directamente sobre el conjunto de muestras obtenidas de las mediciones de campo.

Dependiendo de la frecuencia de muestreo configurada y la longitud de la ventana de análisis seleccionada en la etapa de pre procesamiento, la cantidad de muestras a analizar resulta considerablemente grande y, sobre todo, redundante en cuanto a la información que acarrea.

Por este motivo, de acuerdo a lo expresado en el *capítulo 1* diferentes técnicas de procesamiento de señales fueron implementadas por la comunidad científica, con el objetivo de reducir la dimensión de los datos a procesar, potenciar determinadas características de las formas de onda para favorecer los algoritmos de detección y permitir la extracción de los parámetros característicos de cada clase en particular. Dentro de las técnicas más importantes se destacan: *ST* [18] [20], *FT* [107], *STFT* [108], *FFT* [13], *WT* [29], *WPT* [78], *Cálculo del perfil rms* [109], entre otras [11].

4.2.2 Extracción de parámetros característicos.

La información obtenida en el proceso de muestreo de una señal, estadísticamente contiene un alto grado de información redundante, ruidosa e irrelevante.

Los clasificadores basados en algoritmos de aprendizaje lineal se ven marcadamente afectados cuando son entrenados con abundante cantidad de datos que no siempre representan el evento que se desea identificar.

Por este motivo, este proceso representa la primera etapa del algoritmo donde se realiza la reducción de la dimensión de los datos que serán empleados para el entrenamiento del sistema de clasificación. Este proceso cuenta con dos etapas, el cálculo de los parámetros característicos y su posterior normalización. Ambos sub procesos son explicados a continuación.

4.2.2.1 Cálculo de parámetros característicos.

Los parámetros característicos se corresponden con valores numéricos obtenidos por medio de cálculos realizados sobre los datos resultantes del procesamiento de señales presentado en la sección 4.2.1.2 [51] o directamente de las muestras de la forma de onda normalizadas.

Si los parámetros característicos fueron seleccionados apropiadamente, los valores numéricos obtenidos mediante estos cálculos son representativos, con un cierto grado de aproximación, del conjunto total de las muestras.

El cálculo de los parámetros característicos no involucra el análisis de cada clase presente en el set de entrenamiento. Por este motivo un amplio número de parámetros característicos se calculan en esta instancia siendo tarea de posteriores procesos la de seleccionar aquellos que mejor describen a cada tipo o clase en particular.

Por lo tanto, el conjunto de parámetros característicos calculados deben permitir la representación de todas clases de eventos que pueden encontrarse en una forma de onda presente en el set de entrenamiento. A pesar que la elección de los parámetros característicos a extraer no se realiza de manera óptima, en esta etapa se consigue una drástica reducción de la dimensión de los datos a ser procesados por las siguientes etapas del algoritmo.

Como ya se mencionó para el entrenamiento del clasificador se ingresa al algoritmo con un conjunto de perturbaciones formas de onda organizadas como en una matriz de $[m \times s]$

Como resultado de esta etapa, se obtiene una matriz de dimensión $[m \times n]$, donde el valor de m continúa siendo referencia a las distintas formas de onda empleadas para entrenar el clasificador y n se corresponde con el número total de parámetros característicos calculados. Para una efectiva reducción de dimensión, n suele ser bastante menor a s .

4.2.2.2 Normalización de parámetros característicos.

Debido a que los parámetros característicos son calculados empleando diversas técnicas, el rango dinámico de cada columna de la matriz de entrenamiento puede resultar bastante amplio.

Cualquier método de aprendizaje lineal funciona correctamente cuando los parámetros a clasificar no presentan un amplio rango de valores. Por este motivo, en este módulo se aplican métodos para la normalización de cada columna de la matriz que contiene los vectores característicos de tal modo que el rango dinámico de los parámetros se ajusten dentro de cotas predeterminadas y similares entre sí.

4.2.3 Minería de datos.

En recientes estudios, diversos algoritmos para la selección del set de parámetros característico más apropiado se han implementado, aunque esta área se encuentra todavía notablemente poco explorada. Estos algoritmos suelen fundamentarse en herramientas como *Algoritmos Genéticos (GA)* [110, 111], *Particle Swarm Optimization (PSO)* [112], y *Ant Colony (AC)* [113].

Como ya se mencionó en la sección 4.2.2, el proceso de extracción de parámetros característicos permitió la transformación del set de entrenamiento, originalmente representado por una matriz de dimensión $[m \times s]$ en un arreglo de dimensión $[m \times n]$, donde $n < s$. También se enfatizó que el proceso de extracción de parámetros característicos se realizó de manera independiente a cada una de las N clases presentes en el set de entrenamiento.

También se destacó que la selección estratégica de los parámetros característicos empleados para describir a cada una de las clases representa uno de los principales objetivos de diseño de un clasificador basado en técnicas de inteligencia artificial.

El proceso etiquetado como *Minería de Datos* (*en inglés, Data Mining*) permite la selección de aquellos parámetros característicos que describen mejor cada clase, eliminando información redundante, irrelevante o ruidosa.

Esta etapa, una de las más importantes en el diseño del sistema de clasificación, es un proceso desarrollado para seleccionar un sub conjunto de j parámetros característicos del total de los n parámetros que ingresan a este proceso, con $j \ll n$.

Como se muestra en la Figura 4. 3 este proceso está compuesto por tres etapas: La generación de etiquetas, el filtrado heurístico y el algoritmo de búsqueda exhaustiva. A continuación se explican estos tres sub procesos:

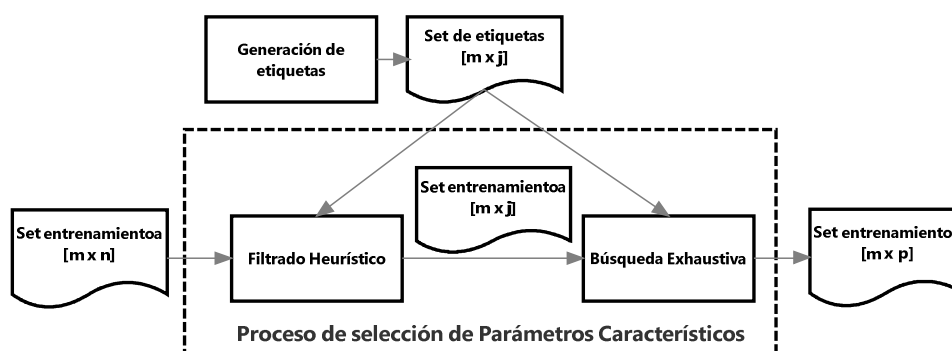


Figura 4. 3 - Minería de datos.

4.2.3.1 Generación de Etiquetas.

El proceso de generación de etiquetas consiste en asignar, en un arreglo vectorial independiente, un valor numérico que se corresponda con cada una de las m filas de la matriz que representa el set de entrenamiento. Estos valores numéricos se corresponden con etiquetas e indican a cuál de las N clases pertenece cada fila del set de entrenamiento.

Este proceso se puede llevar a cabo ya que a priori se conoce a que clase pertenece cada una de las m filas del set de entrenamiento.

El proceso de generación de etiquetas se resume en los siguientes pasos:

Paso 1: Se genera un arreglo de dimensión $m \times 1$ al que llamamos L_i donde $i=1,2,\dots, N$.

Paso 2: A cada componente del arreglo L_i se le asigna el valor i si la i -ésima fila de la matriz de parámetros característicos pertenece a la clase i .

4.2.3.2 Filtrado heurístico.

Existen dos tipos de algoritmos para la selección de parámetros según su forma de implementación, los métodos de filtrado, cuya selección se realiza mediante cálculos heurísticos, y los métodos conocidos como *wrapper*, que emplean la evaluación sistemática del algoritmo de clasificación para diferentes combinaciones de parámetros característicos.

Como se ve en la Figura 4. 3, el método de filtrado heurístico se ejecuta en una etapa previa al algoritmo de búsqueda exhaustiva con el objetivo de obtener una reducción dimensional necesaria para procesar, en una segunda instancia, el algoritmo de búsqueda exhaustiva.

Estos métodos evalúan el total de los parámetros característicos asignando un valor numérico a cada uno de ellos. Cuanto mayor sea este valor mayor será correlación del parámetro con un tipo o clase de datos. Se trata de un indicador que mide la importancia del parámetro a la hora de caracterizar una clase dada.

De esta manera es posible conformar un ranking de parámetros característicos que mejor describen a cada una de las clases.

De acuerdo a lo expresado en [114], no existe un marco teórico que justifique, a priori, cuál de los métodos de filtrado presentará mejores resultados.

Para prevenir esta desventaja, se propone realizar un ranking de parámetros característicos mediante la combinación de las siguientes tres técnicas:

- Algoritmo de selección de parámetros característicos *Chi-cuadrado*: Se trata de uno de los métodos más comúnmente empleados. La técnica evalúa el peso que un parámetro tiene para describir una clase mediante el cálculo del estadístico chi cuadrado [115]. El método propone la selección de una hipótesis inicial: supone que dos parámetros se encuentran no correlacionados y, de esta manera, realizar el test *chi-cuadrado* mediante la siguiente expresión:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (4. 1)$$

donde O_{ij} es el parámetro observado y E_{ij} es el parámetro esperado que afirma la hipótesis nula. Mientras mayor sea el valor del estadístico mayor será la evidencia que afirma que hipótesis es verdadera.

- Algoritmo de selección de parámetros característicos *Relief-F*: El algoritmo evalúa la proporción en la que un parámetro aporta para describir una clase por medio del repetido muestreo de una instancia teniendo en cuenta el valor de la instancia (vecina) más cercana de la misma clase y de diferentes clases. El algoritmo asigna un peso a cada parámetro característico basándose en la habilidad de un determinado parámetro para definir una clase. Luego selecciona aquellos parámetros característicos cuyos pesos superan un umbral pre definido por el usuario [116].
- Algoritmo de selección de parámetros basado en la *Asimetría de la Incertidumbre* (SU, por sus siglas en inglés): Se trata de un método que se basa en la entropía (cantidad de información o grado de incertidumbre) de las variables que describen una clase [117]. Definido un determinado criterio de impureza de un set de entrenamiento S , se puede definir una medida que refleje la cantidad de información

de Y , provista por X de tal manera que la entropía de Y disminuya. A esta información se la conoce como Ganancia de la Información (IG, por sus siglas en inglés):

$$IG = H(Y) - H\left(\frac{Y}{X}\right) = H(X) - H\left(\frac{X}{Y}\right) \quad (4. 2)$$

Donde H , en este caso, se conoce como entropía. El método conocido como *Incertidumbre Simétrica* compensa el sesgo inherente en el método de ganancia de la información dividiendo por la suma de ambas entropías.

$$SU = 2 \frac{IG}{H(Y) + H(X)} \quad (4. 3)$$

Como entradas del algoritmo de filtrado se tiene la matriz de parámetros característicos de dimensión $[m \times n]$ y el vector de etiquetas de dimensión $m \times 1$.

El resultado de cada uno de los métodos heurísticos se corresponde con una lista donde se ordenan, de mayor a menor, aquellos parámetros característicos que mejor describen a cada clase.

La estrategia para seleccionar la lista final de parámetros característicos que mejor describen a cada clase es la creación de un único ranking que combine los resultados de los tres métodos. En una instancia posterior se seleccionarán aquellos parámetros que serán analizados posteriormente por el algoritmo de búsqueda exhaustiva.

Finalmente, se seleccionan los j parámetros característicos que describen con más certeza a cada clase en particular donde j es, en general, menor a n y puede ser fijado mediante criterios elegidos en el diseño de la metodología.

Como resultado del sub proceso de filtrado heurístico, el set de entrenamiento de dimensión original $[m \times n]$ se reduce a una matriz de dimensión $[m \times j]$.

4.2.3.3 Algoritmo de búsqueda exhaustiva.

Una vez que se tiene un sub conjunto de parámetros característicos, cuya dimensión se redujo mediante el filtrado a un número considerablemente menor, se procede a la búsqueda de la combinación óptima que provoca en el clasificador el mejor rendimiento.

A diferencia de los métodos utilizados para el pre filtrado, el método propuesto utiliza el algoritmo de clasificación como parte del proceso de selección de parámetros característicos. A esta metodología se la conoce como *wrapper* [118].

El método propuesto consiste en probar el rendimiento del clasificador para las 2^j combinaciones posibles del conjunto de parámetros característicos de dimensión $[m \times j]$ obtenido en la etapa de filtrado heurístico.

Como entradas del algoritmo se tiene la matriz filtrada de dimensión $[m \times j]$ y el vector de etiquetas de dimensión $m \times 1$.

El sub proceso realiza las siguientes tareas:

Paso 1: Se genera la Tabla 4. 1 de j columnas y 2^j filas. Cada fila se corresponde a una de las 2^j posibles combinaciones de los parámetros característicos. Se inicializa una variable que llamamos *contador* en 1.

Paso 2: Se genera una matriz auxiliar cuyas columnas se obtienen de extraer de la matriz de entrenamiento aquellas columnas que se encuentran apuntadas por la variable *contador*. Así, por ejemplo, para la iteración número 2^{j-1} , la variable *contador* será igual a 2^{j-1} y la matriz auxiliar, tal como se muestra en la Tabla 4. 1, estará conformada por todas las columnas de la matriz de entrenamiento excepto la columna j .

Paso 3: La matriz auxiliar y el vector de etiquetas sirven de variables de entrada para invocar al proceso *Optimización de los parámetros del clasificador* que se explica en la próxima sección.

Paso 4: El proceso *Optimización de los parámetros del clasificador* devuelve los parámetros del clasificador y el porcentaje de precisión máxima obtenido. Los mismos son almacenados en la fila indicada por el valor de la variable *contador* tal como se muestra en Tabla 4. 1.

Paso 5: Se valida la condición de paro. La misma puede ser configurada hasta obtener una combinación que supere un porcentaje de acierto pre definido, hasta la expiración del número dado de iteraciones contabilizadas por la variable *contador* o hasta completar la tabla.

Tabla 4. 1 - Selección de parámetros característicos.

Parámetro/ Contador	1	2	...	j-1	j	Parámetro Kernel	C	Precisión
1	0	0	...	0	1	0.5	4000	98,8%
2	0	0	...	1	0	2	1000	99,6%
3	0	0	...	1	1	0.001	3000	100 %
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2^{j-1}	1	1	...	1	0	3	500	99,7%
2^j	1	1	...	1	1	50	500	99,7%

En la sección 4.4 se presenta el diagrama de flujo para el algoritmo de diseño y entrenamiento del clasificador que permite apreciar más directamente la relación entre el proceso de *Minería de datos* y el proceso *Optimización de parámetros del clasificador* que se detalla en la próxima sección.

4.2.4 Optimización de los parámetros del clasificador.

El rendimiento de un clasificador basado en técnicas de aprendizaje lineal depende, en gran medida, a la calidad descriptiva los datos empleados para su entrenamiento y de la óptima sintonización de los parámetros que lo gobiernan.

La tarea de la selección de parámetros característicos (descrita en la sección 4.2.3) y la óptima sintonización del algoritmo de aprendizaje lineal son procesos que presentan una fuerte dependencia entre sí. Por lo tanto, los resultados del algoritmo aquí presentado serán válidos para una combinación de parámetros característicos en particular. Este algoritmo será reiteradamente invocado por el proceso denominado *Minería de datos* como un modo de encontrar el mejor conjunto de los parámetros que gobiernan al clasificador para cada combinación de parámetros característicos.

Para describir el procedimiento debemos considerar un algoritmo de aprendizaje lineal en particular. Por lo tanto, este procedimiento se explicará utilizando como ejemplo un clasificador basado en *SVM*, destacando que este procedimiento es válido para cualquier técnica de aprendizaje lineal que se sintonice mediante el ajuste de *dos* parámetros.

De acuerdo a lo presentado en el *capítulo 3*, para un clasificador basado en *SVM*, independientemente de la función kernel seleccionada, existe un parámetro que gobierna el entrenamiento del clasificador. Este parámetro se conoce como factor de penalización *C* (*box constraint, en inglés*).

Luego, existe una serie de parámetros que están ligados a la función *kernel* empleada, como el parámetro (σ) sigma para la función *RBF* y la función *Gaussiana*, el parámetro grado (*d*) para un kernel *polinómico*, el parámetro alfa (α) para el kernel *hiperbólico*, entre otros.

No existe una solución heurística para el cálculo simultáneo de los valores de *C* y el *parámetro* particular de la función kernel, por lo que no es posible conocer de ante mano los valores óptimos para la resolución de un determinado problema.

Como consecuencia de esto, para encontrar los valores óptimos, dado un determinado problema, es necesario implementar algoritmos de búsqueda.

En este trabajo se propone el algoritmo de búsqueda tipo grilla, conocido como *grid search algorithm*.

El objetivo del algoritmo de búsqueda tipo grilla es encontrar los parámetros que hacen que el clasificador proporcione mayor precisión a la hora de clasificar un conjunto de datos sin que se pierda su capacidad de *generalización*.

Para evitar errores producidos por el sobre entrenamiento y la pérdida de capacidad de generalización es conveniente utilizar distintos set de señales para el entrenamiento y la validación del algoritmo. El procedimiento de dividir el conjunto de señales en un set de entrenamiento y un set de prueba es parte de lo que se conoce como “*validación cruzada*”.

Por lo tanto, para encontrar los parámetros óptimos del algoritmo presentado en este trabajo se implementó el *método de búsqueda basado en una grilla* [119] en combinación con técnicas de *validación cruzada*.

El algoritmo de búsqueda de grilla tiene por objetivo completar la Tabla 4. 2.

Para ello, como primera medida, se deben seleccionar los valores que podrán asumir los parámetros a explorar y se configura la fila etiquetada como *Parámetro del kernel* y las columnas tituladas *Parámetro C* de la Tabla 4. 2. Cada celda de la tabla representa una combinación particular de los mencionados parámetros.

Para la búsqueda de la mejor combinación el algoritmo necesita un set de entrenamiento que consiste en una matriz de parámetros característicos con su respectivo vector de etiquetas. En el caso propuesto, la matriz de entrenamiento y validación empleada se obtiene del proceso presentado en la sección 4.2.3.3 y el vector de etiquetas empleado es el generado en la sección 4.2.3.1.

Tabla 4. 2 - Algoritmo de búsqueda tipo grilla.

Parámetro del kernel	Parámetro C								
	0.1	1	5	10	25	50	...	1000	5000
$1e^{-5}$	94	94	95	98	93	97	...	96	96
$1e^{-4}$	90	89	92	98	92	97	...	96	96
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.1	92	92	93	97	100	100	...	96	96
1	93	93	94	95	99	100	...	95	95
10	93	93	94	95	99	100	...	96	90
100	93	93	94	95	97	99	...	95	90
$1e^4$	93	93	94	95	98	98	...	94	90

Luego, el algoritmo procede según los siguientes pasos:

Paso 1: Para garantizar la capacidad de generalización del algoritmo clasificador, partiendo de la matriz que representa el set de entrenamiento, se selecciona un porcentaje de las filas del set entrenamiento para generar una matriz de entrenamiento auxiliar y las restantes filas conforman un arreglo que se llamará matriz de validación.

Paso 2: Se selecciona el primer valor propuesto para el *Parámetro de la función kernel*. En el ejemplo propuesto se corresponde a $1e^{-5}$

Paso 3: Se selecciona el primer valor del *Parámetro C*. Para el ejemplo propuesto se corresponde con *0.1*.

Paso 4: Se configura el clasificador empleando los parámetros seleccionados ($1e^{-5}, 0.1$). El clasificador se entrena empleando la matriz de entrenamiento auxiliar y luego se prueba utilizando la matriz de validación. El porcentaje de aciertos obtenido de la validación cruzada se almacena en la correspondiente celda del arreglo. Este resultado, en el ejemplo propuesto se corresponde a un *94%*.

Paso 5: Se valida si todos los valores de *C* fueron evaluados. Si el resultado de la validación es negativo se incrementa el valor de *C* y se vuelve al *Paso 4*. Caso contrario el algoritmo continúa al *Paso 6*.

Paso 6: Se valida si se evaluaron todos los posibles valores del *Parámetro de la función kernel*. En caso de resultar negativa la validación el valor de dicho parámetro se actualiza y se regresa al *Paso 4*. Caso contrario, el algoritmo continúa al *Paso 7*.

Paso 7: En esta etapa la grilla se encuentra completa. El algoritmo busca la celda cuyo porcentaje de acierto es máximo y extrae la mejor combinación de parámetros y la máxima precisión alcanzada para el set de entrenamiento propuesto. En caso de encontrar más de una celda cuyo porcentaje de acierto sea máximo, pueden definirse condiciones de selección adicionales como aquella combinación cuyo valor de C sea mínimo, la combinación cuyo valor del parámetro del clasificador sea mínimo, entre otras.

Como resultado de la evaluación del algoritmo se obtienen los valores óptimos de los parámetros que gobiernan el clasificador y la máxima precisión alcanzada para el set de entrenamiento de entrada.

A modo de ejemplo, en la Tabla 4. 2 se presenta una grilla completa, en donde en cada celda se presenta el porcentaje de aciertos del clasificador para cada uno de los valores de C y los parámetro de la función kernel. En este ejemplo puede verse que para cuatro combinaciones de parámetros se obtiene un porcentaje de acierto del 100%. Si bien las cuatro combinaciones proporcionan el mismo resultado para ese set de entrenamiento en particular, el criterio que prima a la hora de elegir la mejor combinación de parámetros es seleccionar aquella que presenta el menor valor de C . Para el ejemplo $C = 25$ y el parámetro *del kernel* correspondiente es 0.1 .

4.2.5 Entrenamiento de las etapas del clasificador.

La Tabla 4. 1 se completa conjuntamente con los procesos presentados en las secciones 4.2.3 y 4.2.4.

En esta etapa del algoritmo se selecciona la combinación de parámetros característicos y los parámetros del clasificador que producen el mayor porcentaje de acierto. En caso que haya más de una combinación de parámetros característicos que provoquen el mismo resultado, se utilizarán otros criterios de selección como, por ejemplo, la mínima cantidad total de parámetros calculados, con el objetivo de reducir el costo computacional.

Finalmente, con los parámetros característicos óptimos para cada una de las etapas y sus correspondientes valores de C y los *parámetros de la función kernel* seleccionada se procede a entrenar el clasificador.

La salida de este proceso se corresponde el modelo de cada clasificador entrenado. En esta instancia se considera que el sistema de clasificación se encuentra completamente diseñado y se en condiciones de ser probado.

4.3 ALGORITMO DE CLASIFICACION.

Mediante el algoritmo presentado en la sección 4.2 se obtuvo como resultado el modelo optimizado del sistema de clasificación.

Para probar el rendimiento del modelo del clasificador en situaciones reales se implementa el algoritmo presentado en la Figura 4. 4.

Como se puede apreciar, muchos de los procesos que conforman el algoritmo de clasificación tienen funcionalidad similar a los presentados en la Figura 4. 2. Por este motivo, para evitar repetir información de manera innecesaria aquí solo se mencionarán aquellos aspectos que difieren en lo funcional a los procesos ya presentados.

Como entrada del proceso se tiene forma de onda a clasificar. Como salida de este proceso se obtiene una etiqueta que indica a que clase pertenece la mencionada señal.

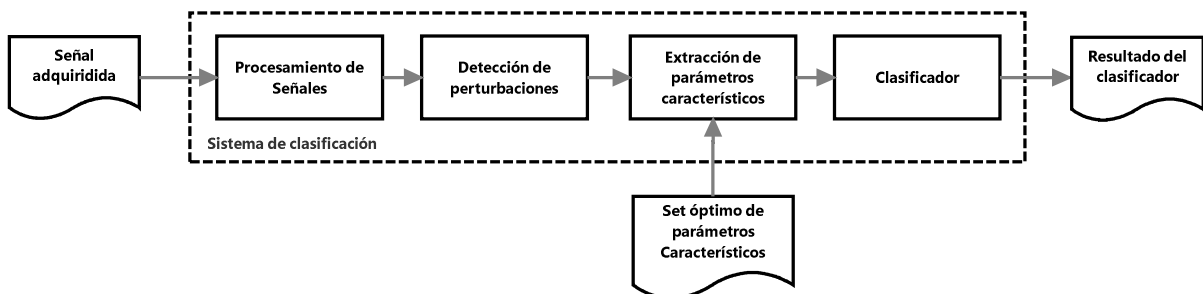


Figura 4. 4 - Algoritmo de clasificación.

4.3.1 Procesamiento de Señales.

Las técnicas de pre procesamiento aquí implementadas se deben corresponder con que se implementan en el algoritmo de entrenamiento. La configuración de la frecuencia de muestreo, longitud de la ventana, normalización de la señal y técnicas de de-noising se deben corresponder con las configuradas en la sección 4.2.1.1.

Como resultado de este proceso se obtiene un conjunto de muestras ordenadas en un arreglo de dimensión $[1 \times s]$ donde s representa la cantidad de muestras y está relacionado con la frecuencia de muestreo y la longitud de la ventana de análisis. Es importante destacar que este conjunto de muestras no necesariamente puede contener un evento que se desea clasificar.

De la misma manera, las técnicas de procesamiento de señales implementadas en este proceso guardan correspondencia con las implementadas en la sección 4.2.1.2 y tienen por objetivo permitir la detección de perturbaciones y extraer los parámetros característicos que serán empleados para la clasificación.

4.3.2 Detección de eventos.

En implementaciones prácticas el algoritmo desconoce a priori cuando una forma de onda contiene un evento que será, a posteriori, clasificado. Por este motivo el algoritmo continuamente monitorea las señales, adquiriendo un conjunto de muestras que luego son procesadas y, en caso de no encontrar anomalías, son descartadas. Solamente se tienen en cuenta aquellas señales que contienen los eventos o anomalías que se desean clasificar.

El objetivo de este módulo es detectar la presencia de una anomalía en las señales adquiridas.

4.3.3 Extracción de parámetros característicos.

Este proceso cumple la importante función de calcular aquellos parámetros característicos que representan a cada perturbación. Solamente se calculan, en esta etapa, aquellos parámetros característicos óptimos encontrados proceso de Data Mining presentado en la sección 4.2.3. De esta forma se evita, el desperdicio computacional que representa el cálculo de parámetros que resultan redundantes o irrelevantes para la clasificación.

Como salida de este proceso se obtiene un arreglo de dimensión $[1 \times j]$, donde j representa la cantidad de parámetros característicos representativos obtenidos del *Proceso de diseño y entrenamiento*, que serán utilizados por el sistema de clasificación para decidir a qué clase pertenece.

Por último, las ecuaciones y técnicas aplicadas en el proceso de normalización de parámetros característicos deben respetar concordancia con aquellas empleadas en la etapa de diseño y entrenamiento.

4.3.4 Clasificación.

El proceso de clasificación tiene por objetivo indicar a qué clase pertenece el conjunto de parámetros característicos calculados en por el proceso presentado en 4.3.3 empleando el modelo del clasificador que diseñado y entrenado por el algoritmo presentado la sección 4.2.

El resultado de este proceso se corresponde con una etiqueta que indica a que clase pertenece el conjunto de muestras obtenido de la medición y muestreo de una forma de onda.

4.4 DIAGRAMAS DE FLUJO.

Esta sección presenta los diagramas de flujo de los algoritmos presentados en la sección 4.2 y la sección 4.3.

La Figura 4. 5 muestra el diagrama de flujo del algoritmo de diseño y entrenamiento del clasificador.

La Figura 4. 6, por otro lado, muestra el diagrama de flujo del algoritmo destinado a la clasificación.

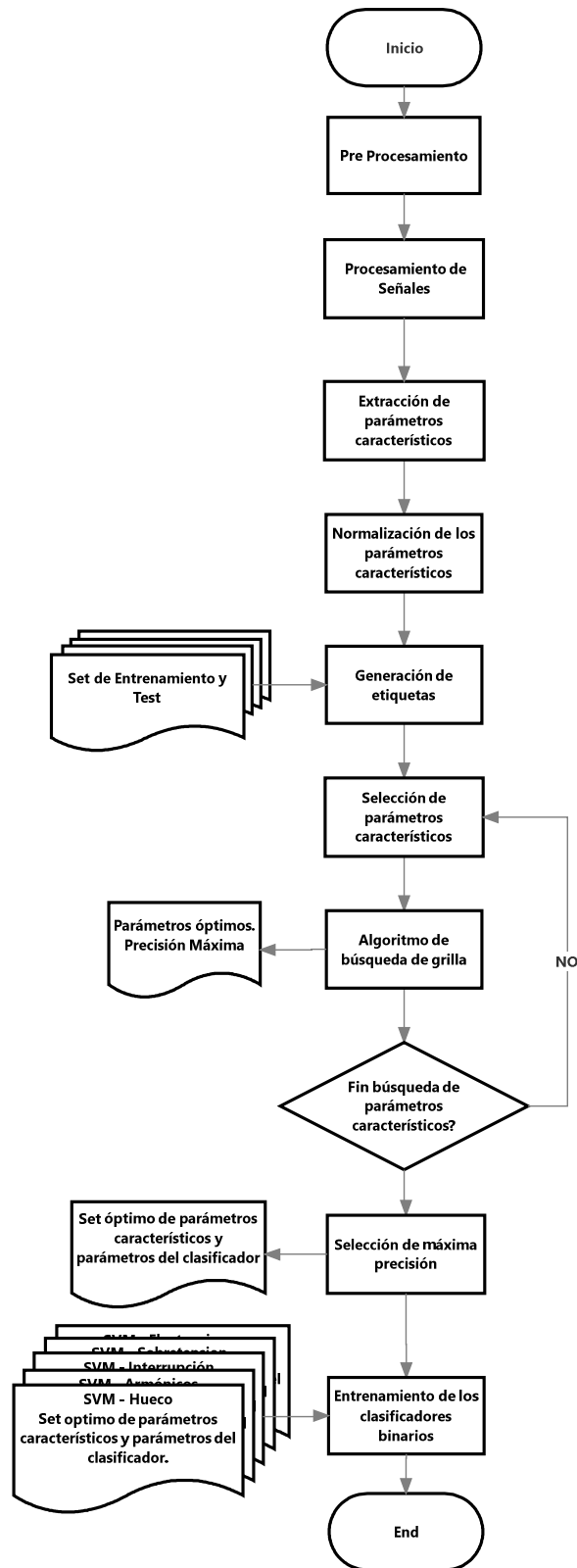


Figura 4. 5 - Diagrama de flujo del algoritmo de entrenamiento del clasificador.

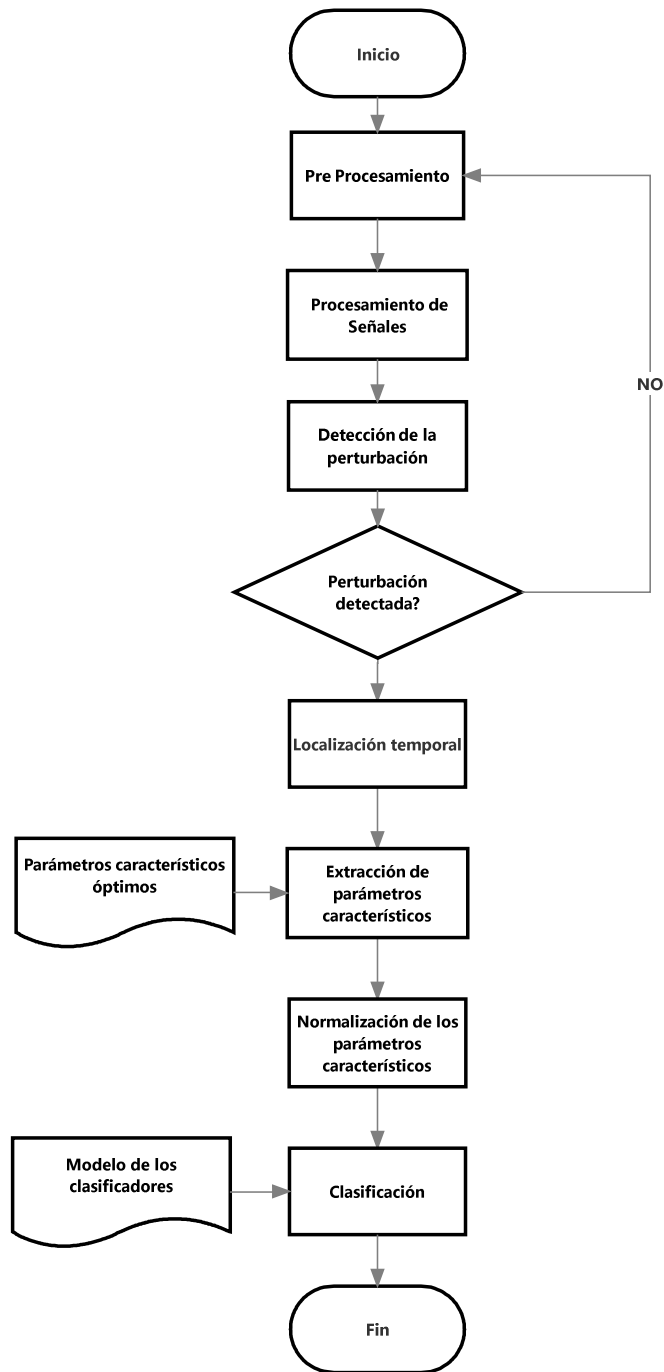


Figura 4. 6 - Algoritmo de clasificación.

4.5 CONCLUSIONES.

En la primera parte del capítulo se presentó un método para el diseño y entrenamiento de algoritmos de clasificación que se basan en el procesamiento de señales para la obtención de parámetros característicos y en técnicas de aprendizaje lineal.

En la segunda parte del capítulo, se presentó un algoritmo que permite la prueba del clasificador desarrollado empleando muestras de una forma de onda.

Las nociones explicadas en el presente capítulo son de carácter genérico y en su desarrollo se evitó ahondar en las técnicas particularmente empleadas en cada proceso ya que, como se verá en los *capítulos 6 y 7*, estas dependen de cada aplicación en particular.

Por este motivo en los futuros capítulos de la presente tesis se hará referencia a la metodología aquí presentada.

En cada caso se mencionará solamente las técnicas implementadas en cada uno de los procesos que conforman la metodología propuesta.

Capítulo 5: Software para el Análisis de Eventos de Calidad de Energía

5.1 INTRODUCCION

Los eventos y actores involucrados en la generación de perturbaciones de diversa naturaleza son cada día más comunes en un Sistema Eléctrico de Potencia (SEP). Paradójicamente, tanto para clientes industriales como en los consumidores minoristas, se ha incrementado drásticamente la utilización de dispositivos sensibles a las perturbaciones electromagnéticas.

Un aspecto crítico para la disminución del impacto de las perturbaciones en un sistema eléctrico radica en la habilidad de la entidad que realiza el monitoreo para detectar, localizar y clasificar una perturbación de manera automática.

Se entiende por sistema de monitoreo de calidad de energía a una combinación de hardware y software específicamente diseñados para la detectar y registrar eventos provenientes de mediciones de la red eléctrica. En el *capítulo 1* de este trabajo de tesis se detallan los aspectos fundamentales relacionados a un sistema de monitoreo.

En el presente capítulo se describen los algoritmos que han sido desarrollados con el objetivo de detectar y clasificar perturbaciones provenientes de las mediciones realizadas en un sistema eléctrico.

La herramienta, desarrollada en MATLAB, cuenta con una interfaz gráfica que permite el acceso de manera intuitiva a todas sus funcionalidades.

El software se ha desarrollado con tres objetivos:

- Permitir la generación masiva de formas de onda que permitan el diseño de nuevos algoritmos para la detección, clasificación y localización de perturbaciones en un sistema eléctrico.
- Permitir el desarrollo de una herramienta de monitoreo de eventos de calidad de energía en tiempo real.
- Contar con una herramienta académica para análisis distintas técnicas y algoritmos de análisis de eventos de calidad de energía en un sistema eléctrico de potencia.

Como más adelante se detallará en este capítulo el software permite el análisis de perturbaciones obtenidas mediante el muestreo de señales reales así como también de señales generadas de manera artificial.

A su vez, la herramienta cuenta con un módulo para la generación de perturbaciones simples y complejas permitiendo la configuración de las características más importantes que gobiernan a cada tipo de evento como, por ejemplo, el tiempo de duración, la magnitud o la frecuencia del evento.

Este capítulo también presenta el desarrollo de un algoritmo basado en la *WT* (Ver *capítulo 2*) el cual permite detectar la presencia de un evento calidad de energía y la extracción de parámetros característicos de dicha señal que son de utilidad para el proceso de clasificación.

Finalmente en este capítulo se exhibe el desarrollo de un módulo que realiza la clasificación de los eventos de calidad de energía de acuerdo a las características extraídas en el módulo de detección. Los algoritmos de clasificación están basados en *ANN* y *SVM* (Ver *capítulo 3*)

5.2 SOFTWARE PARA EL ANALISIS DE EVENTOS DE CALIDAD DE ENERGIA

El software fue desarrollado para permitir el análisis de distintos métodos de detección y clasificación de perturbaciones en un sistema eléctrico de potencia. A su vez el diseño de la herramienta involucró procedimientos para simular el comportamiento de los algoritmos para situaciones reales.

Por este motivo, la herramienta fue desarrollada para permitir:

- El entrenamiento del módulo clasificador.
- El análisis del comportamiento de los algoritmos ante situaciones reales.

El objetivo de esta sección es presentar el software desarrollado como parte de las tareas de investigación.

Este software, desarrollado con fines educativos [120] permite el entrenamiento y prueba de algoritmos destinados a la detección y clasificación de perturbaciones simples y complejas.

El procedimiento comprende las siguientes etapas

- Acceso a la herramienta.
- Generación de señales de entrenamiento y prueba.
- Extracción de parámetros característicos de las señales de entrenamiento y prueba.
- Entrenamiento y prueba del clasificador.
- Incorporación de señales reales.
- Prueba del clasificador.

5.2.1 Acceso a la herramienta.

Para acceder a la herramienta se debe ejecutar la aplicación *MATLAB* y configurar el directorio raíz donde se encuentran los scripts del software. En la ventana de comandos, como se muestra en la Figura 5. 1, se invoca al panel principal de la herramienta por medio de la siguiente instrucción:

```
>> Analizador_de_señales
```

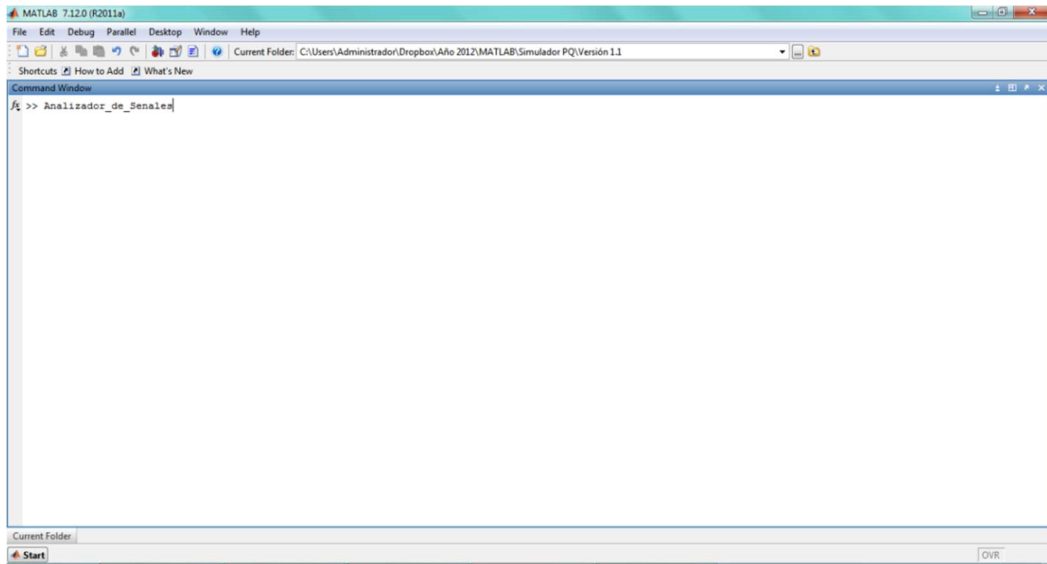


Figura 5. 1 - Acceso a la herramienta.

Al ejecutar este comando se despliega el panel principal de la aplicación mostrado en la Figura 5. 2. Esta interfaz permite al usuario el acceso a los siguientes módulos cuyas funcionalidades serán explicadas posteriormente en este capítulo:

- Módulo adquirente de Señales.
- Módulo generador de Señales.
- Módulo de detección de Eventos.
 - Transformada Wavelet.
- Módulo de clasificación de Eventos.
 - ANN BackPropagation.
 - Support Vector Machine.

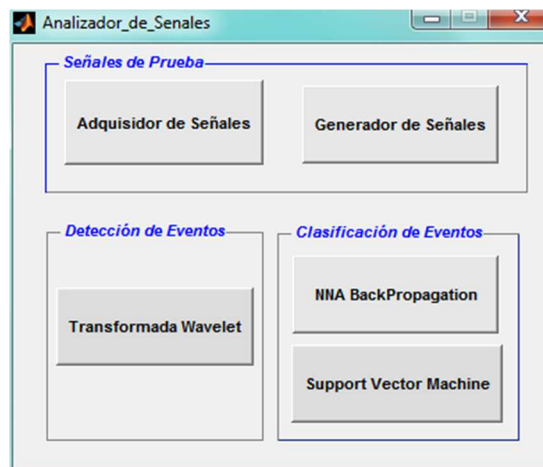


Figura 5. 2 - Panel principal

5.2.2 Generación de señales de Entrenamiento y Prueba.

Para generar las señales de entrenamiento y prueba se debe presionar la opción *Generador de Señales* del panel principal. Al realizar esta acción se despliega la interfaz gráfica que se muestra en la Figura 5. 3.

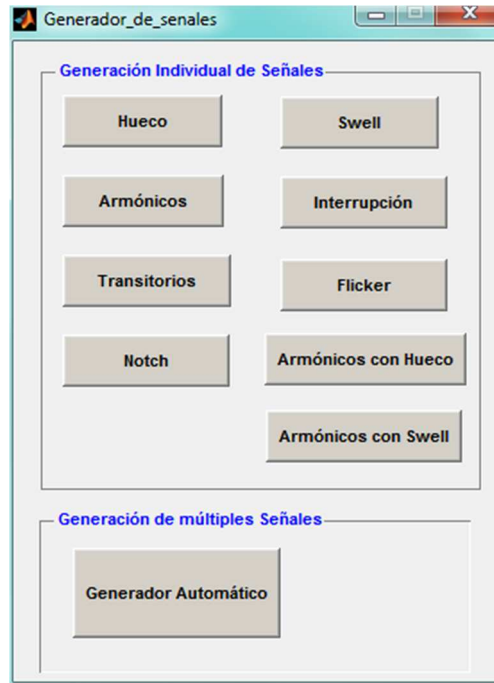


Figura 5. 3 - Generador de señales de prueba.

Presionando la opción *Generador Automático* se accede a la interfaz gráfica mostrada en la Figura 5. 4 cual permite la generación de señales de prueba y entrenamiento.

Para nuestro ejemplo se crea un set de entrenamiento conformado por 2500 señales de los distintos tipos de perturbaciones que el software permite simular. La frecuencia fundamental seleccionada es de 50 [Hz], la longitud de la ventana se configura en 20 ciclos y la frecuencia de muestreo en 10000 [muestras/segundo].

Para generar la señal se configuran todos los campos correspondientes a las perturbaciones que se van incluir en la señal resultante, se seleccionan los correspondientes check box y luego se presiona el botón *Generar Señal*. Este botón genera un vector que contiene una la cantidad de señales perturbadas de acuerdo a lo configurado en el campo *Cantidad de Señales* incluyendo solo aquellas perturbaciones que hayan sido debidamente seleccionadas. La herramienta fue desarrollada para que la generación de cualquier tipo de perturbación tenga la misma probabilidad. Como resultado se obtiene un vector que contiene las muestras de diferentes perturbaciones generadas de manera aleatoria, cuyas características principales se ven regidas de acuerdo a parámetros también generados aleatoriamente aunque respetando los valores máximos y mínimos configurados en la interfaz gráfica.

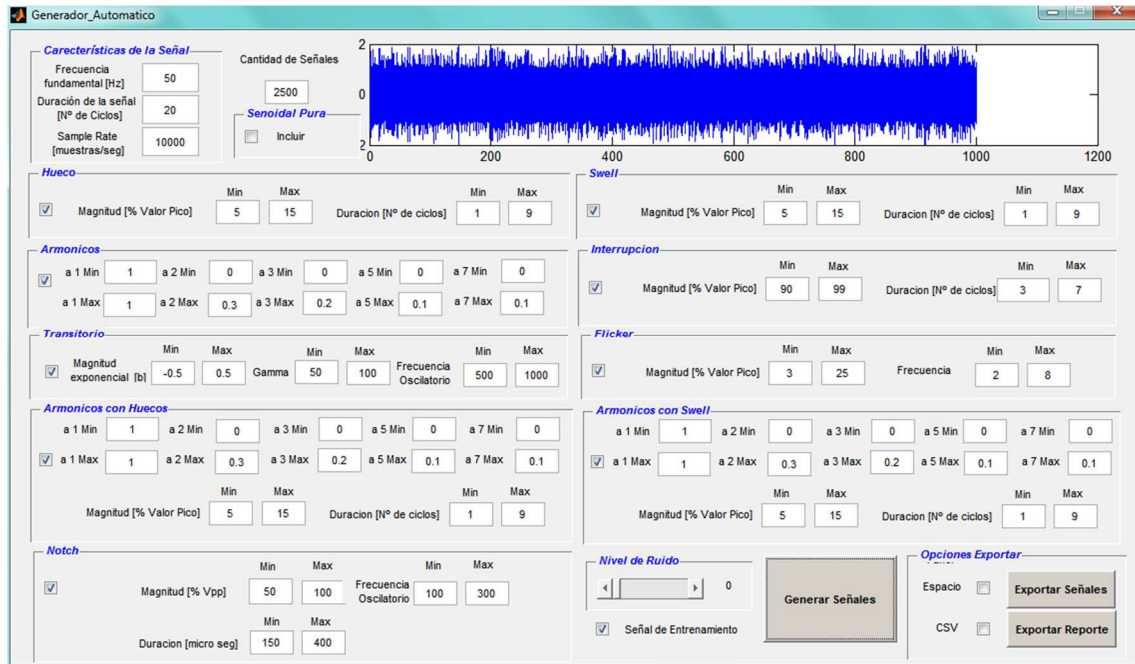


Figura 5. 4 - Generador automático de perturbaciones simples.

Para generar la señal de entrenamiento se debe seleccionar el check box denominado *Señal de Entrenamiento*. Del mismo modo, para generar una señal del tipo señal de prueba se debe cuidar que el mencionado check box no se encuentre marcado.

La Figura 5. 5 muestra una interfaz gráfica desarrollada para la generación de diferentes tipos de perturbaciones complejas. Esta interfaz gráfica presenta las mismas funcionalidades que la interfaz desarrollada para la generación de perturbaciones simples.

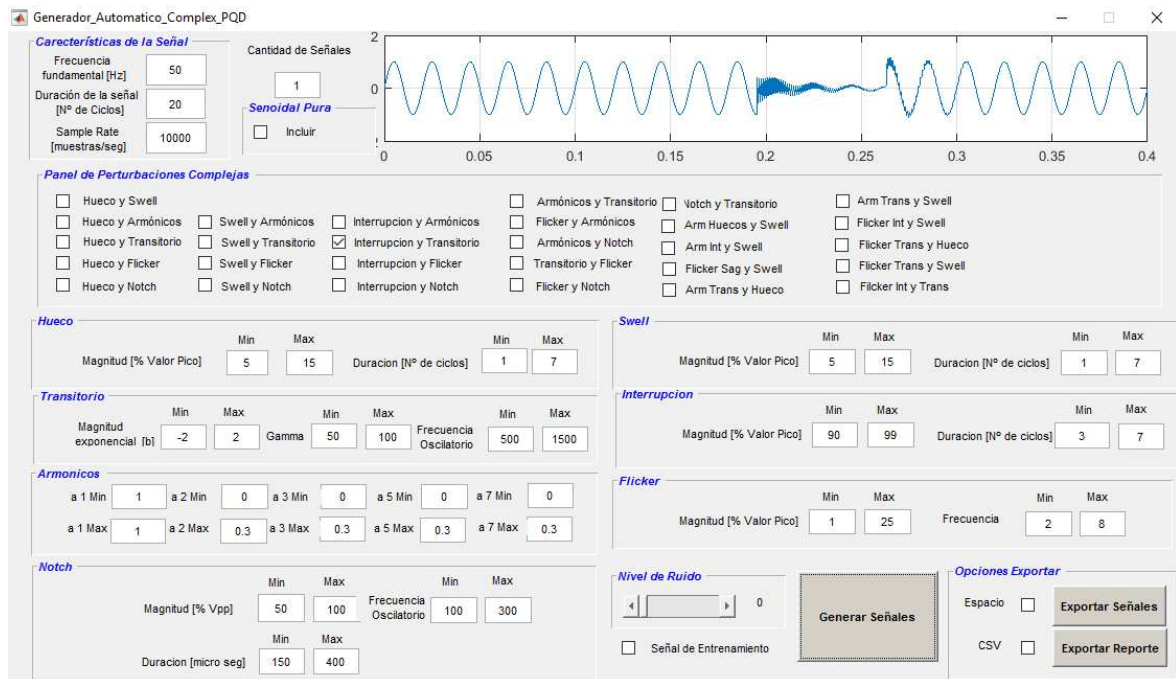


Figura 5. 5 - Generador automático de perturbaciones complejas.

5.2.3 Extracción de parámetros característicos de las señales de entrenamiento y prueba.

El acceso a la interfaz gráfica mostrada en la Figura 5. 6 se realiza seleccionando la opción *Transformada Wavelet* del Panel Principal presentado en la sección 5.2.1.

Para generar el vector característico de entrenamiento se deben introducir los parámetros relativos al set de entrenamiento generado en la sección anterior. Para nuestro ejemplo se debe introducir la frecuencia fundamental del set de entrenamiento la cual es $50 [Hz]$, la longitud de la ventana de $20 [ciclos]$, la frecuencia de muestreo se configura en $10000 [muestras/segundo]$ y, finalmente, la cantidad de señales 2500 .

Al presionar el botón *Importar Perturbación*, la herramienta automáticamente incorpora la señal de Entrenamiento generada en la etapa anterior. Para indicar que los vectores característicos generados se utilizarán para entrenar los algoritmos de clasificación se debe marcar el check box denominado *Señal de Entrenamiento*.

Para generar un set de prueba se debe acceder al módulo *Generador de Señales*, crear un conjunto de señales cuidando de no marcar el check box *Señal de Entrenamiento* y repetir el proceso explicado en esta sección.

La Figura 5. 6 muestra el panel desarrollado para la generación del vector característico.

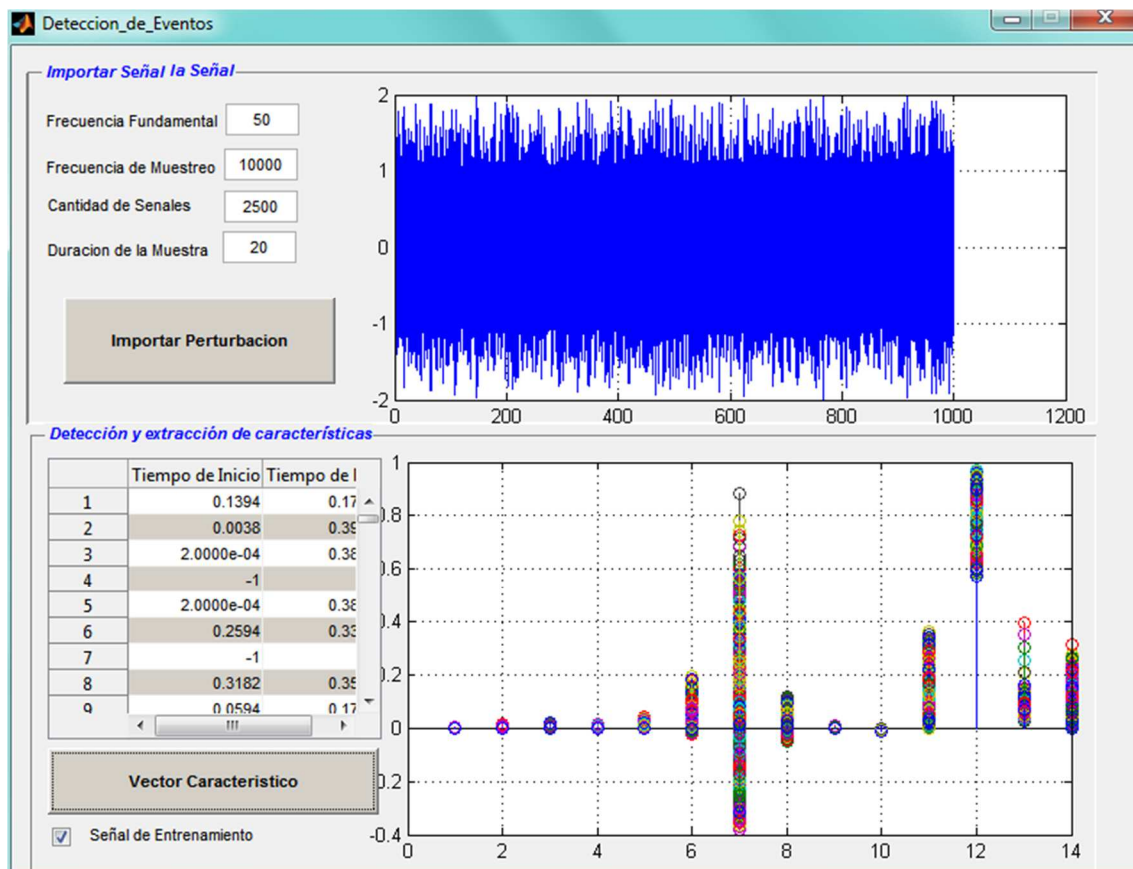


Figura 5. 6 - Panel de generación del vector característico.

Una vez que se generaron las señales y los vectores característicos de entrenamiento y prueba el Workspace de MATLAB se encuentra conformado por los siguientes arreglos:

- SeñalGenerada: Vector que contiene la señal de prueba.
- SeñalGeneradaEntrenamiento: Vector que contiene la señal de entrenamiento.
- reporteSeñal: Vector que contiene las etiquetas correspondientes al vector de las señales de prueba.
- reporteSeñalEntrenamiento: Vector que contiene las etiquetas correspondientes al vector de las señales de entrenamiento.
- vectorCaracterístico: Matriz cuyas columnas se corresponden a los parámetros característicos y las filas a cada perturbación generada. Esta matriz será empleada para la prueba de los clasificadores.
- vectorCaracterísticoEntrenamiento: Matriz conformada de manera idéntica a la denominada *vectorCaracterístico* que será empleada para el entrenamiento del clasificador.

5.2.4 Entrenamiento y prueba del clasificador.

Este módulo del software fue desarrollado con el objetivo permitir, de manera gráfica e intuitiva, la configuración de dos clasificadores, uno basado en una *Red Neuronal Backpropagation* y un segundo clasificador basado etapas *SVM binarias* organizadas en una arquitectura One vs Rest, como se explicó en la *sección 3.4.4.1*.

El acceso a cada uno de estos módulos se realiza a través del *Panel Principal* mencionado en la *sección 5.2.1*.

5.2.4.1 Red Neuronal Backpropagation.

En la Figura 5. 7 se muestra la interfaz gráfica desarrollada con el objetivo de permitir una configuración manual una red neuronal back propagation con dos capas ocultas. La interfaz gráfica permite el ingreso los parámetros básicos para configurar una red neuronal. Por defecto para la prueba se configuran dos capas ocultas conformadas correspondientemente por 12 y 10 neuronas. Además se configura como 1000 el número máximo de épocas y el error mínimo permitido es de 1×10^{-3} [121].

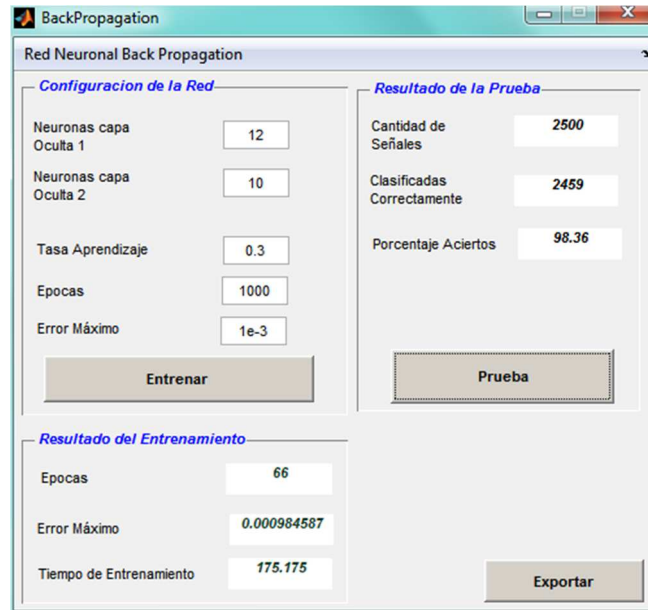


Figura 5. 7 - Clasificador basado en una Red Neuronal Backpropagation.

Al presionar el botón *Entrenar* el software conforma el set de entrenamiento utilizando la matriz *vectorCaracterísticoEntrenamiento* y el vector *reporteSeñalEntrenamiento*.

Una vez que el entrenamiento finaliza los campos destinados a los *Resultados del Entrenamiento* son poblados automáticamente. Para el ejemplo propuesto, el entrenamiento culminó alcanzando el error máximo configurado. El proceso del entrenamiento tomó 175.175 [segundos] y se emplearon 66 épocas para lograr un error de 0.000984587.

Para comprobar el rendimiento del clasificador se presiona el botón *Prueba*. El software utiliza la matriz *vectorCaracterístico* para obtener un vector con los resultados de la clasificación el cual es cotejado con el vector *reporteSenal* con el objetivo de detectar la cantidad de clasificaciones correctas y la cantidad de clasificaciones erróneas. Los resultados obtenidos para nuestro ejemplo indican que de 2500 señales de prueba 2458 fueron clasificadas correctamente. El porcentaje de efectividad del clasificador es de 98.36%.

Es importante destacar para corroborar la propiedad de *generalización* del clasificador no se utiliza el mismo conjunto de datos para el entrenamiento y prueba de los algoritmos de clasificación.

Si el porcentaje de efectividad del clasificador no se corresponde con un valor esperado es posible modificar los parámetros de configuración de la red neuronal y volver a ejecutar el entrenamiento.

Cuando la red neuronal alcanza el porcentaje de efectividad deseado se presiona el botón *Exportar*. Este botón genera un archivo denominado *parametrosRed.net* dentro del directorio donde se encuentra la aplicación el cual contiene los parámetros que caracterizan el comportamiento de la red (pesos, bias, arquitectura, etc.). Este archivo será utilizado por

un software de monitoreo en tiempo real cuyo clasificador se basa en una red neuronal del tipo backpropagation.

5.2.4.2 Máquinas de Vectores de Soporte.

Para entrenar algoritmos basados en SVM de soporte se presiona la opción *Support Vector Machine* del panel principal presentado en la sección 5.2.1.

El clasificador, por defecto, está conformado por nueve SVM las cuales serán entrenadas para diferenciar cada tipo de perturbación del resto. A pesar que la arquitectura es fija, la herramienta permite anular aquellas etapas que no se deseen implementar.

Como se explicará en el siguiente capítulo, tanto los parámetros característicos como los parámetros que gobiernan el rendimiento de cada etapa SVM (*sigma* y *C* para un clasificador cuyo kernel se corresponde a la función rbf) pueden ser optimizados para maximizar la precisión global del clasificador.

Los algoritmos desarrollados con el objetivo de obtener los parámetros óptimos serán explicados en detalle en el *capítulo 6* y *capítulo 7*.

Los parámetros de diseño del clasificador podrán ingresarse a través de la interfaz gráfica presentada en esta sección.

La Figura 5. 8 muestra que existen dos parámetros (rbf_sigma y box constraint) que gobiernan el comportamiento de cada una de las 9 etapas del clasificador.

Configuración de los SVM		Resultado de la Prueba	
	RBF Sigma	Box Constraint	
SVM Huecos	0.5	3	Cantidad de Señales: 2500
SVM Swell	0.3	3	Clasificadas Correctamente: 2426
SVM Armónicos	1	7	Porcentaje Aciertos: 97.04
SVM Interrupcion	1.3	7	
SVM Transitorios	4.5	8	
SVM Flicker	0.5	5	
SVM Notch	1	7	
SVM Armonicos y Huecos	5	3	
SVM Armonicos y Swell	5	10	
Tiempo de Entrenamiento	19.3616		

Figura 5. 8 - Clasificador basado en Máquinas de Vectores de Soporte.

Al presionar el botón *Entrenar* el software conforma el set de entrenamiento utilizando la matriz *vectorCaracterísticoEntrenamiento* y el vector *reporteSeñalEntrenamiento*.

Una vez que el entrenamiento finaliza el campo tiempo de Entrenamiento es poblado de manera automática. Para el ejemplo propuesto el proceso del entrenamiento tomó *19.3676 [segundos]*.

Para comprobar el rendimiento del clasificador se presiona el botón *Prueba*. Del mismo modo que el módulo basado en la *red neuronal backpropagation* el software utiliza la matriz *vectorCaracterístico* para obtener un vector con los resultados de la clasificación el cual se compara con el vector *reporteSenal* para detectar la cantidad de clasificaciones correctas. Los resultados obtenidos para nuestro ejemplo indican que de 2500 señales de prueba 2426 fueron clasificadas correctamente. El porcentaje de efectividad del clasificador es de 97.04%.

Al presionar el botón exportar se generan 9 archivos (el número nueve se corresponde a la cantidad de etapas binarias configuradas, por lo que, si se desean clasificar solo cinco clases, la cantidad de archivos generados será cinco y no nueve) denominados *SVM1.mat*, *SVM2.mat*, ..., *SVM9.mat* dentro del directorio donde se encuentra la aplicación. Estos archivos contienen los parámetros que caracterizan el comportamiento del clasificador. Estos archivos serán utilizados por el software de monitoreo en tiempo real cuyo clasificador se basa en una SVM.

5.2.5 Incorporación de señales reales.

La Figura 5. 9 ilustra una señal obtenida mediante el muestreo de una de las fases del sistema eléctrico de potencia de una planta de procesamiento de aceites.

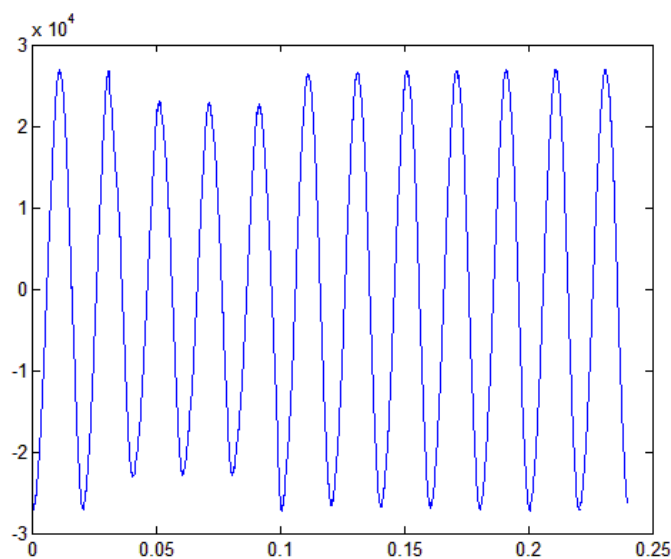


Figura 5. 9 - Hueco de tensión adquirido de la red.

5.2.6 Prueba del clasificador.

Para probar el comportamiento del módulo de detección y clasificación se desarrolló una interfaz de simulación, mostrada en la Figura 5. 10, que permite la incorporación de señales reales, la extracción de sus parámetros característicos y la clasificación de acuerdo a los clasificadores presentados en las secciones anteriores.

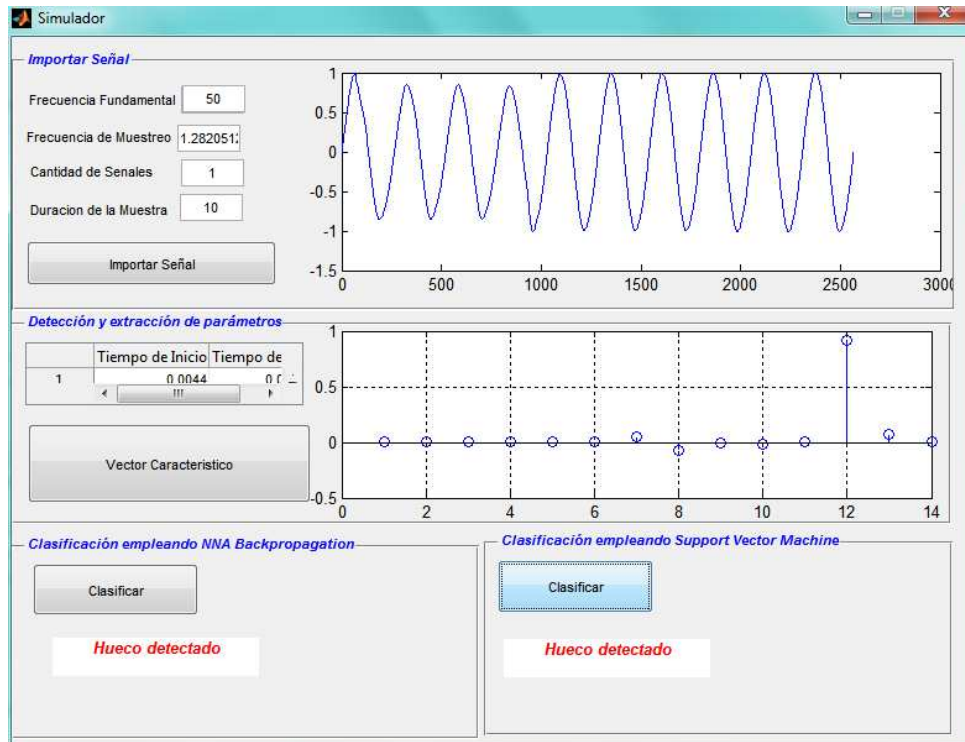


Figura 5. 10 - Localización y clasificación de una señal real.

Para importar una señal desde un archivo se presiona el botón *Importar Señal*. Esta acción hace que emerja un cuadro de diálogos que permite la selección de un archivo de texto que contiene la señal perturbada. Además se carga la frecuencia fundamental, la frecuencia de muestreo, la cantidad de señales que contiene el archivo de texto a analizar y la duración de la muestra (longitud de la ventana).

Luego se ejecuta el algoritmo de extracción de parámetros característicos presionando el botón *Vector Característico*.

El módulo de Simulación fue desarrollado para probar los algoritmos de clasificación diseñados y entrenados tal como se explicó en la sección 5.2.4. Cuando se presiona el botón *Clasificar* correspondiente a la *Red Neuronal Backpropagation* la herramienta busca en el directorio raíz el archivo denominado *parametrosRed.mat* que contiene los parámetros de la red neuronal y lo utiliza para clasificar el vector característico mostrando en pantalla el resultado de la clasificación.

De la misma manera, presionando el botón *Clasificar* correspondiente al clasificador basado en *SVM*, la herramienta automáticamente carga los archivos *SVM1.data*, *SVM2.data*, ..., *SVM9.data* los cuales contienen los parámetros del clasificador entrenado como se detalla en la sección 5.2.4. Estos archivos son empleados para realizar la clasificación cuyo resultado es mostrado en pantalla.

5.3 CONCLUSION

Debido a la dificultad de contar con una gran variedad de señales obtenidas por medio de la medición de una red eléctrica, la primera parte de este capítulo presenta el desarrollo de una herramienta que permite la generación de perturbaciones simples y complejas de manera artificial. Este módulo probó ser útil para generar una gran variedad de eventos de calidad de energía los cuales fueron destinados a la confección de conjuntos de señales de entrenamiento y prueba para ser usados en la sintonización del sistema de clasificación.

Esta herramienta es de gran utilidad, como se verá en el *capítulo 6*, para la optimización y diseño de un clasificador destinado a identificar perturbaciones complejas.

En este capítulo se detalla un *caso de uso* del software desarrollado donde se destacó la utilización de la herramienta para la generación de señales artificiales, la confección de los set de entrenamiento y prueba, el entrenamiento de los sistemas de clasificación, la detección de las perturbaciones, la localización temporal de los eventos de calidad de energía y la utilización del software para el análisis de una señal adquirida en un sistema eléctrico de potencia de una planta industrial.

La configuración de los sistemas de clasificación presentados en este capítulo se realiza de manera manual, a través de diferentes interfaces gráficas destinadas a tal efecto.

Los parámetros empleados para la configuración de estos clasificadores puede ser optimizada para maximizar su capacidad de acierto. Para ello, su configuración debe ser obtenida a través de algoritmos de optimización que se presentan en el *capítulo 6* y el *capítulo 7* de la presente tesis.

Es importante aclarar que todos los desarrollos mencionados en este capítulo fueron realizados en MATLAB.

Capítulo 6:

Clasificación de

Perturbaciones Complejas

6.1 INTRODUCCION.

La mayoría de los sistemas de clasificación proponen vincular un conjunto de parámetros característicos con perturbaciones simples de acuerdo a lo establecido en el estándar [122].

Sin embargo, debido a la variada naturaleza de las causas que generan un evento de calidad de energía, en las mediciones de campo frecuentemente se puede encontrar la coexistencia de dos o más tipos de perturbaciones en una única señal adquirida. Así pues, es posible hallar, por ejemplo, combinaciones de perturbaciones estacionarias (armónicos o fluctuaciones) con perturbaciones de corta duración (transitorios, huecos o sobre tensiones) o combinación de perturbaciones de corta duración, como ser huecos seguidos por sobre tensiones, transitorios oscilatorios combinados con huecos de tensión, etc. A este tipo de perturbaciones se las conoce como *Perturbaciones Complejas*.

En la Figura 6. 1, Figura 6. 2 y Figura 6. 3 se pueden ver ejemplos de perturbaciones complejas compuestas por un evento de corta duración y uno estacionario.

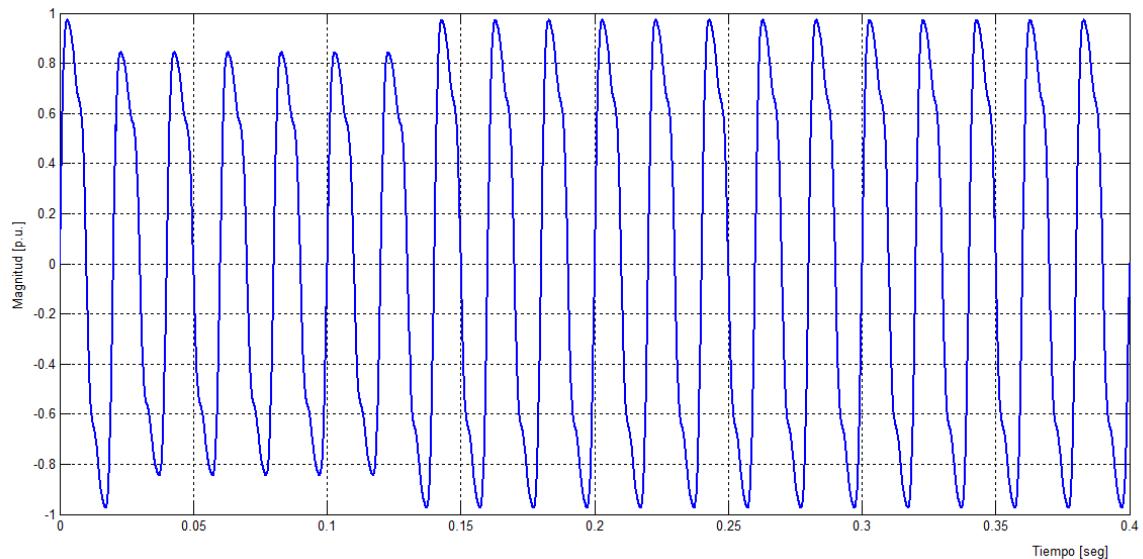


Figura 6. 1 - Armónicos y hueco.

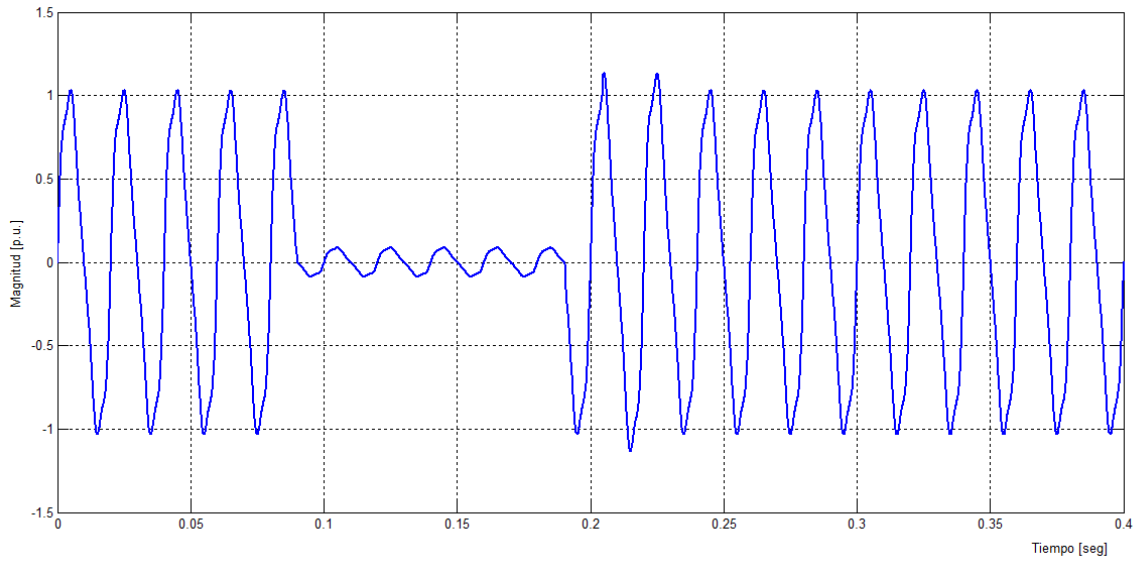


Figura 6. 2 - Armónicos e interrupción.

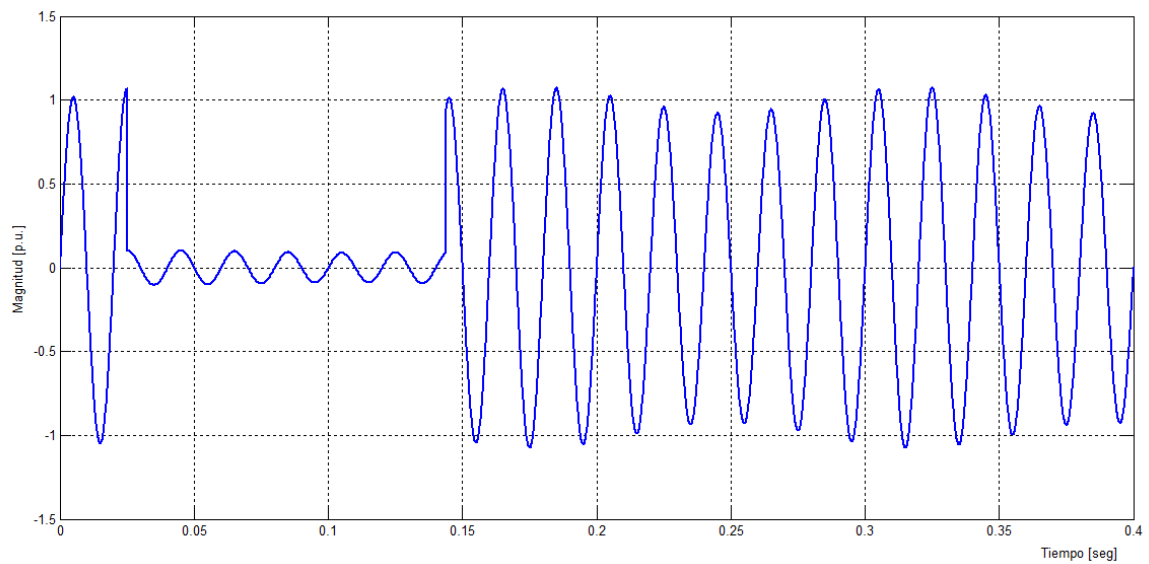


Figura 6. 3 - Fluctuaciones e interrupción.

Por otro lado, la Figura 6. 4 muestra una perturbación compleja doble compuesta por dos eventos de corta duración, mientras que la Figura 6. 5 presenta una perturbación compleja conformada por tres tipos de perturbación, dos de corta duración y una estacionaria.

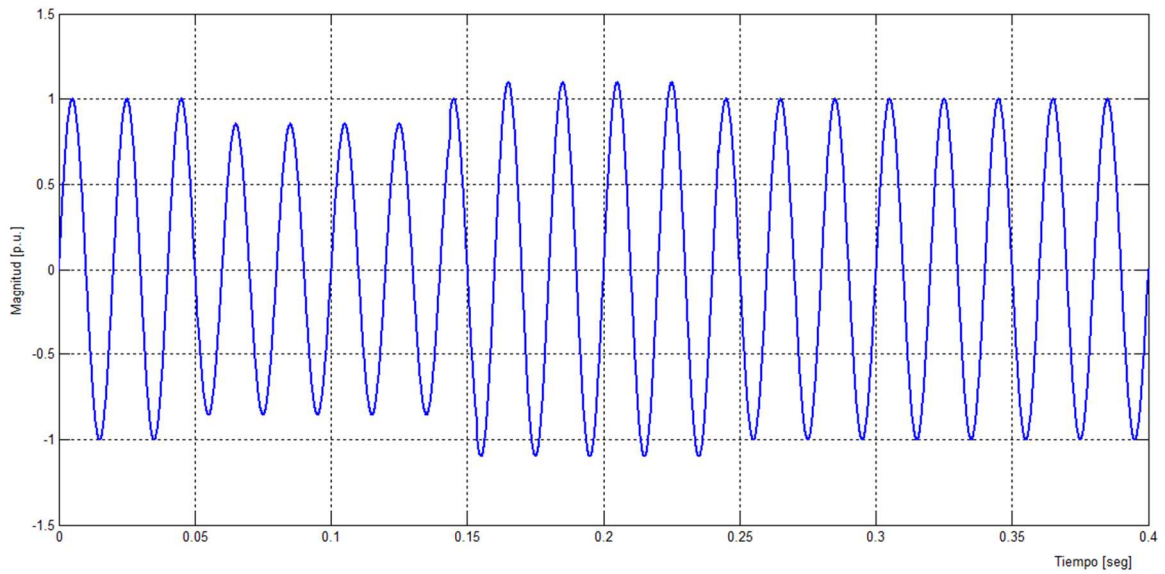


Figura 6. 4 - Hueco y sobre tensión.

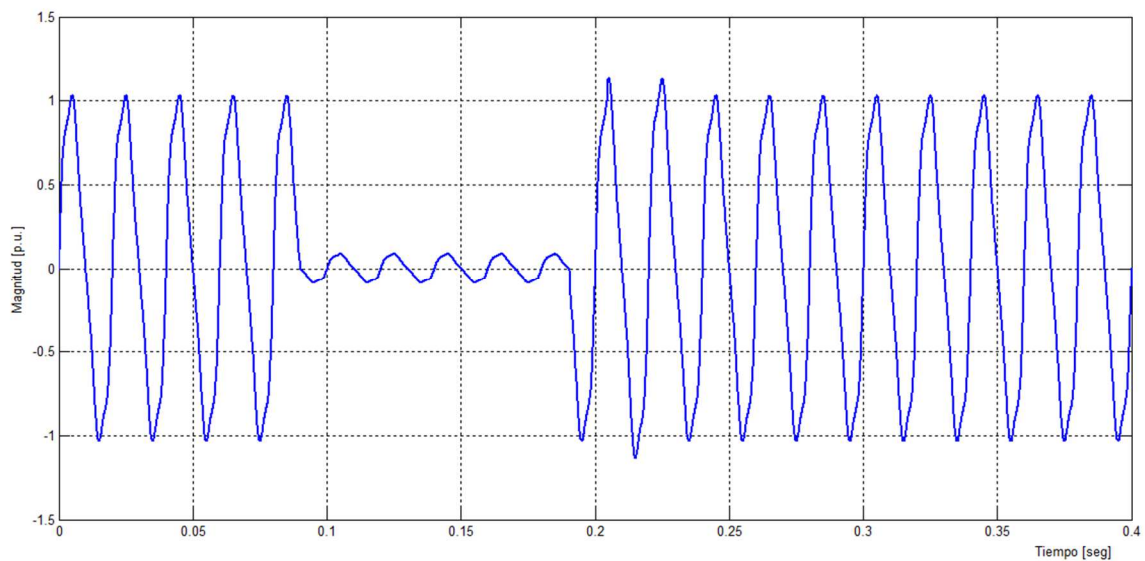


Figura 6. 5 - Armónicos, interrupción y sobre tensión.

La aparición de perturbaciones complejas generan grandes dificultades para la clasificación debido a la existencia simultánea de diferentes características que definen una perturbación. Estas dificultades incluyen mala caracterización del evento, la evaluación incorrecta de la misma y una notable disminución en el porcentaje de aciertos.

Por otro lado, la clasificación de perturbaciones complejas es una problemática escasamente estudiada. La mayoría de los trabajos relacionados con el estudio de calidad de energía se enfocan en perturbaciones simples y solo unos pocos abordan las perturbaciones complejas de manera lateral [11].

A continuación se presentan algunos trabajos que abordan, aunque indirectamente algunos de ellos, la problemática de perturbaciones complejas.

El trabajo presentado por Wei Min [123] se presenta la comparación de un clasificador basado en una *Red Neuronal Back Propagation* y un clasificador multi clase basado en *SVM* binarios organizados en una arquitectura *One vs One*. En el trabajo se concluye que el clasificador basado en *SVM* logra mejores resultados para un mismo escenario. Si bien el objetivo del trabajo no radica en el estudio de perturbaciones complejas, las mismas son consideradas como casos particulares. El método presenta resultados aceptables, sin embargo se debe destacar que el algoritmo precisa de mediciones en múltiples nodos de la red eléctrica lo cual no es viable en todos los casos debido a que no siempre es posible contar con un despliegue de dispositivos de medición e infraestructura de comunicaciones en un sistema eléctrico de potencia como el algoritmo precisa.

Por otro lado, el método presentado por Cheng-Long [124] propone un sistema que se basa diferentes técnicas de acuerdo a la naturaleza de la perturbación que se pretende identificar. Así, se propone un método basado en el análisis del *perfil rms* para analizar perturbaciones de corta duración (huecos, sobre-tensión e interrupciones), la *WT* de cuatro niveles para identificar perturbaciones con altos contenidos en frecuencia (transitorios) y, finalmente, se utiliza una *Red Neuronal Dinámica* para clasificar perturbaciones estacionarias (armónicos y fluctuaciones). El método propuesto logra un porcentaje de aciertos superior al 97%, incluso clasificando perturbaciones complejas compuestas por más de dos eventos. Sin embargo el algoritmo es difícil de implementar y puede resultar computacionalmente costoso debido a que se utilizan múltiples técnicas de procesamiento de señales y algoritmos basados en inteligencia artificial. Por otro lado, el coeficiente de detalle d_1 de la *WT* es altamente sensible al ruido. Debido a que el algoritmo propuesto utiliza este coeficiente para extraer parámetros característicos cuya finalidad es la de distinguir la presencia de un transitorio impulsivo, su capacidad de identificar este tipo de perturbación se verá netamente afectada si no se implementan técnicas de de-noising.

Recientemente diferentes aproximaciones fueron propuestas por Biswal, & Dash [22] donde se presenta un método que emplea la *ST* para la extracción de parámetros característicos de señales perturbadas y una metodología basada en un *árbol de decisión binaria* para la clasificación de los eventos de calidad de energía. A pesar que el método demuestra una notable eficiencia en la clasificación el resultado es obtenido después de la evaluación sucesiva de siete etapas de decisión, hecho que trae aparejado un elevado tiempo de procesamiento.

En la actualidad, la comunidad científica se ha inclinado por implementar clasificadores basados en *SVM* debido a sus marcadas ventajas frente a otros algoritmos de inteligencia artificial. En este sentido, en el trabajo presentado por L. Gang, L. Fanguang [125] se muestra el desarrollo de un método basado en la energía normalizada de diferentes coeficientes de la *WT* en combinación con la técnica *Análisis de Componente Principal (PCA, por su sigla en inglés)* y *Análisis de Componente Independiente (ICA, por sus siglas en inglés)* para la extracción de los principales parámetros característicos de una señal. La clasificación se

realiza utilizando un clasificador *SVM* multi clase. El método demuestra que la reducción dimensional de la matriz de parámetros empleando *PCA* mejora la capacidad del sistema de clasificación. Sin embargo este cálculo representa un costo computacional que podría evitarse seleccionando estratégicamente el conjunto de parámetros característicos, tal como se propone en el presente capítulo.

Sovan Dalai [126] propuso un método basado en la *Transformada Cross Hilbert- Huang (CHHT, por sus siglas en inglés)* para la extracción de parámetros característicos de señales perturbadas, *PCA* para la reducción dimensional de los parámetros característicos y un clasificador multi clase basado en *One vs All SVM*. El principal aporte del trabajo es el empleo de la *CHHT* como método de procesamiento de señales. El porcentaje de clasificaciones correctas supera el 98% aun utilizando señales contaminadas con ruido.

El método propuesto en [127] considera las perturbaciones complejas como un problema multi clase. En el trabajo, Z. Liu, Y. Cui and W. Li, proponen una técnica conocida como *Método Empírico de Descomposición Conjunta (EEMD, por sus siglas en inglés)* para la extracción de parámetros característicos de una perturbación y un sistema de clasificación multi clase conocido como *Rank Wavelet Support Vector Machine*. La principal ventaja de este trabajo es la preservación de la correlación entre las diferentes clases de perturbaciones. En el trabajo también se muestra que el método es notablemente inmune a la presencia de ruido en las muestras. Sin embargo, el máximo nivel de descomposición se configura en 11 para cubrir todos los parámetros característicos presentes en una perturbación compleja transformando al método computacionalmente costoso. Los resultados obtenidos, para el escenario planteado, no superan el 93% de precisión lo que se encuentra levemente por debajo de la mayoría de los trabajos aquí mencionados.

Es práctica común crear una clase individual por cada perturbación compleja [22, 123, 126] lo cual representa una importante desventaja ya que si se quiere incorporar una nueva perturbación se requiere modificar completamente la arquitectura del clasificador, sobre todo si la arquitectura empleada es del tipo árbol de decisión binaria.

Un importante aspecto que se encuentra poco explorado radica en las mejoras en los porcentajes de clasificación obtenidos en caso de encontrar los descriptores correctos para cada tipo de perturbación en particular. La mayoría de las publicaciones utilizan el mismo conjunto de parámetros característicos para describir los diferentes tipos de perturbación [22, 123-126] aunque sus características sean notablemente diferentes.

Finalmente se destaca que en muy pocos trabajos se aborda la sistematización para obtener los parámetros que gobiernan a la etapa de clasificación.

Teniendo en cuenta el análisis de las actuales necesidades y la evaluación de las diferentes metodologías estudiadas por la comunidad científica se puede inferir que el desarrollo de nuevos algoritmos capaces de clasificar perturbaciones simples y complejas, de fácil y rápida implementación y de bajo costo computacional representa un desafío y una marcada necesidad [11].

6.1.1 Objetivos.

El objetivo del presente capítulo es presentar el desarrollo de un clasificador multi clase basado en *SVM* organizados en una arquitectura *One vs Rest* para la clasificación de perturbaciones simples y complejas.

El sistema propuesto se configura como un arreglo de cinco clasificadores *SVM* binarios organizados en una arquitectura *One vs Rest*.

Como uno de los principales aportes de este capítulo se destaca el hecho que gracias a arquitectura propuesta es posible optimizar la extracción parámetros característicos de acuerdo al tipo de perturbación que se pretende clasificar y seleccionar la mejor combinación de parámetros que gobiernan el comportamiento de cada clasificador binario.

Así, por ejemplo, un clasificador binario encargado de clasificar los huecos de tensión, será entrenado con un set de parámetros característicos que, por su naturaleza, serán distintos al que se emplea para clasificar perturbaciones tipo armónicos.

Por otro lado, la arquitectura en paralelo del algoritmo propuesto permite también realizar cálculos de manera simultánea haciéndolo mucho más eficiente desde el punto de vista computacional si se compara con aquellos clasificadores multiclase del tipo árbol de decisión binaria en donde la clasificación tiene lugar luego de la evaluación secuencial de varias etapas binarias.

6.1.2 Metodología.

Los algoritmos desarrollados en el presente capítulo se entrenan y prueban utilizando señales simuladas mediante una herramienta desarrollada por el autor en el marco de la presentación la tesis de maestría [128]. Esta herramienta, presentada en el *capítulo 5*, permite la generación de una base de datos conformada por diversos tipos de perturbaciones, simples y complejas.

El diseño, entrenamiento y prueba del mismo se realiza mediante la metodología presentada en el *capítulo 4*. El presente capítulo enfoca en presentar las diferentes técnicas las técnicas adoptadas en cada sub-proceso que conforman los mencionados algoritmos.

En la primera parte del capítulo se detallan los métodos y técnicas empleadas para el diseño y entrenamiento del clasificador.

En la segunda parte del capítulo se presenta el resultado obtenido de la prueba del algoritmo desarrollado.

El rendimiento del clasificador se presenta en función del porcentaje de aciertos y los resultados obtenidos son ponderados con respecto a trabajos similares.

6.2 ENTRENAMIENTO Y DISEÑO DEL SISTEMA CLASIFICADOR.

En el *capítulo 4* se presenta una metodología genérica para el diseño y entrenamiento de un clasificador basado en técnicas de inteligencia artificial, haciendo hincapié en una en particular conocida como *Máquinas de Vectores de Soporte*.

En este capítulo se presenta una implementación práctica de las metodologías propuestas para la clasificación de perturbaciones simples y complejas.

El entrenamiento se realiza por medio de un completo conjunto de señales conformado por perturbaciones simples y complejas.

Para la prueba del método se proponen dos escenarios de prueba, uno para el caso de perturbaciones simples y un segundo para perturbaciones complejas.

6.2.1 Arquitectura del clasificador.

En el *capítulo 3* se mencionó que, si bien *SVM* es una técnica que originalmente fue diseñada para resolver problemas binarios, existen varias aproximaciones para lograr un clasificador *SVM multiclase*. En la sección 3.4.4 se presentaron diferentes técnicas que emplean un arreglo de clasificadores *SVM* binarios para lograr resolver problemas donde más de una clase está involucrada.

Este método, desarrollado para identificar cinco tipos de perturbaciones, se basa en una arquitectura *Uno versus el Resto (One vs Rest)*.

El método desarrollado permite la clasificación de las siguientes perturbaciones simples: Huecos, Sobre tensión, Interrupción, Armónicos y Fluctuaciones. Por otro lado, también es posible entrenar el sistema propuesto para clasificar perturbaciones complejas conformadas como una combinación de algunas estas cinco clases.

En el clasificador de la Figura 6. 6 se observan cinco clasificadores binarios que operan en paralelo de acuerdo a la arquitectura presentada en la sección 3.4.4.1.

En este gráfico se puede apreciar una etapa donde se seleccionan los parámetros característicos más apropiados de acuerdo a los resultados obtenidos en la etapa de *Minería de datos* que se presenta en la sección 6.2.3.3.

De acuerdo con lo expuesto, el clasificador llamado *SVM 1* es el encargado de identificar la presencia de un hueco de tensión en el vector característico de entrada. Siguiendo la misma lógica, el clasificador binario *SVM 2* distinguirá sobre tensiones, el *SVM 3* armónicos, el *SVM 4* interrupciones y el *SVM 5* fluctuaciones.

Gracias a la arquitectura propuesta, el método será capaz de clasificar perturbaciones simples y, con un entrenamiento adecuado, una determinada variedad de perturbaciones complejas sin la necesidad de incorporar etapas adicionales.

Los cinco clasificadores binarios son configurados con la función kernel conocida como *RBF* por su probada eficiencia en problemas de Calidad de Energía [97].

Los parámetros que gobiernan a cada clasificador son el parámetro C , conocido como factor de penalización y un parámetro relacionado con la función kernel denominado con la letra griega σ (*sigma*). La configuración simultánea de ambos parámetros determinará la precisión del clasificador.

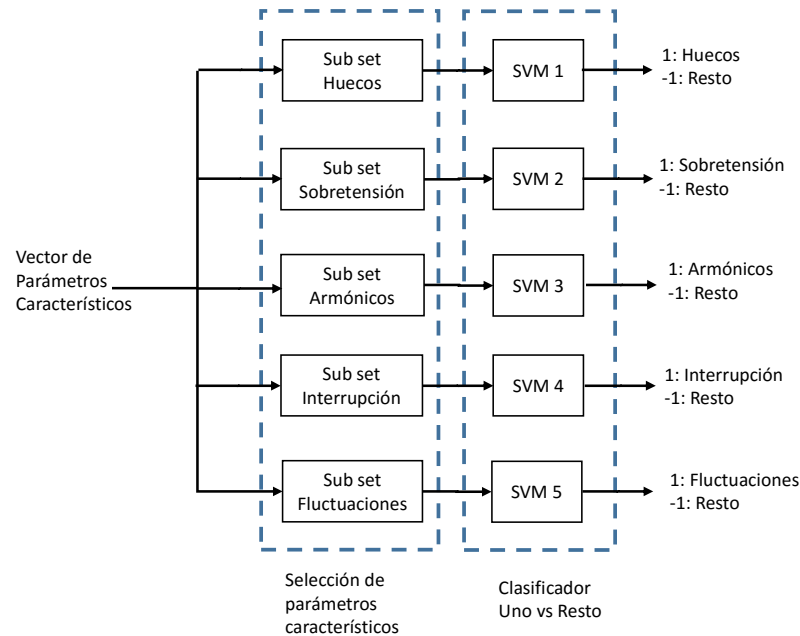


Figura 6. 6 - Arquitectura uno contra el resto.

6.2.2 Configuración del set de entrenamiento.

Para entrenar el clasificador 2600 señales contaminadas con perturbaciones de distinto tipo se simulan empleando una herramienta en MATLAB desarrollada por el autor [128, 129].

La Tabla 6. 1 y la Tabla 6. 2 presentan las perturbaciones simuladas tanto para el caso de perturbaciones simples como para perturbaciones complejas. A su vez se puede observar que las perturbaciones complejas no se representan como una clase en particular si no como una combinación de eventos simples. El set de entrenamiento que se utilizará para el desarrollo de las pruebas presentadas en este capítulo está conformado por la combinación de ambas tablas.

Como se explicó en el *capítulo 1* una perturbación puede ser simulada mediante una expresión matemática donde sus variables deben ajustarse dentro de ciertos rangos tabulados por estándares. En nuestro caso, el estándar adoptado es el *IEEE 1159 - Prácticas Recomendadas para el Monitoreo de la Calidad de Energía* [122]. Con el objetivo de no perder generalidad, las variables que gobiernan cada perturbación aquí simuladas fueron generadas de manera aleatoria acotándolas a los mencionados rangos.

Tabla 6. 1 - Set de perturbaciones simples.

Perturbaciones Simples	Clase	Total
Huecos	1	400
Sobre tensiones	2	400
Armónicos	3	400
Interrupciones	4	400
Fluctuaciones	5	400
Total		2000

Tabla 6. 2 - Set de perturbaciones complejas.

Perturbaciones complejas	Clase	Total
Armónicos y Huecos	3+1	100
Armónicos y Sobre tensiones	3+2	100
Armónicos e Interrupciones	3+4	100
Fluctuaciones y Huecos	5+1	100
Fluctuaciones y Sobre tensiones	5+2	100
Fluctuaciones e Interrupciones	5+4	100
Total		600

6.2.3 Resultados: Diseño y Entrenamiento del clasificador.

En el *capítulo 4* se presentó, de manera genérica, la metodología desarrollada para el diseño y entrenamiento de un clasificador basado en técnicas de inteligencia artificial.

Esta metodología será empleada para el entrenamiento de cada etapa *SVM binaria* que conforman el clasificador presentado en Figura 6. 6.

En la Figura 6. 7 se muestra un diagrama esquemático del algoritmo desarrollado para el diseño y entrenamiento del clasificador. En las siguientes sub secciones se presentarán los métodos implementados en cada proceso y los resultados obtenidos en el entrenamiento de los clasificadores propuestos como caso de estudio.

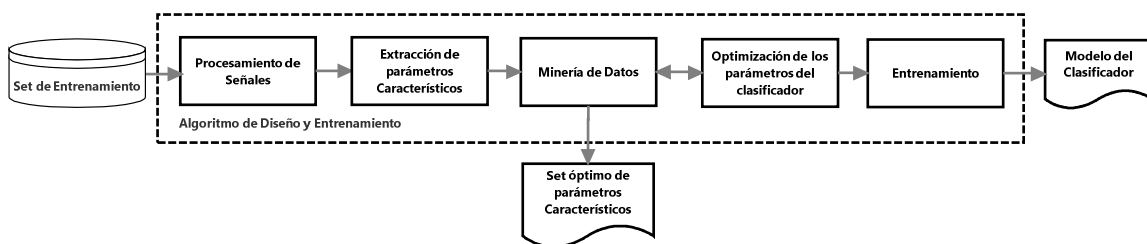


Figura 6. 7 - Algoritmo de diseño y entrenamiento.

6.2.3.1 Procesamiento de señales.

Como se mencionó en la sección 4.2.1, las metodologías implementadas se dividen en dos grupos funcionales: El pre procesamiento de las muestras y el procesamiento propiamente dicho.

En la etapa de pre procesamiento se configuran los siguientes parámetros:

- Longitud de la ventana: $400 [ms]$, lo que equivale a 20 ciclos de la señal sin perturbar.
- Frecuencia de muestreo: 10 Kilo muestras por segundo.
- Normalización de magnitud: La amplitud de cada forma de onda adquirida se normaliza en esta etapa a valores por unidad $[p.u.]$.
- Técnicas de de-noising: No se implementan técnicas de de-noising.

La cantidad de muestras de cada forma de onda es producto de la frecuencia de muestreo seleccionada y la longitud de la ventana la cantidad configurada. Cada señal, por lo tanto, estará representada por 4000 muestras.

Dado que el set de entrenamiento se configura como un arreglo matricial donde cada fila se corresponde a cada una de las señales presentadas en la sección 6.2.1 y cada columna representa cada muestra, la dimensión del arreglo que representa el set de entrenamiento pre procesado es de $[2600 \times 4000]$.

Una vez que se tiene el set de entrenamiento pre procesado, se procede a aplicar técnicas de procesamiento de señales para transformar el set de entrenamiento en un conjunto de datos que facilite la tarea de detectar la presencia de una perturbación y la extracción de parámetros característicos.

En este trabajo, dos técnicas de procesamiento de señales se aplican a tal efecto:

- *Perfil cuadrático medio*: El cálculo del perfil cuadrático medio (rms, por sus siglas en inglés) es una técnica de procesamiento de señales que permite detectar alteraciones en la magnitud de una forma de onda senoidal y extraer parámetros que resultan de extrema importancia a la hora de describir una perturbación electromagnética. El cálculo de perfil rms. se realiza de la siguiente se realiza mediante la siguiente ecuación:

$$x_{r.m.s.}[i] = \sqrt{\frac{1}{k} (x_i^2 + x_{i+1}^2 + \dots + x_{i+k-1}^2)} = \sqrt{\frac{1}{k} \sum_{j=i}^k x_j^2} \quad (6.1)$$

donde i se corresponde a la i -ésima muestra de la curva del perfil rms y k es la longitud de la ventana adoptada para el cálculo del perfil rms.

En la Figura 6. 8 se muestra el resultado del cálculo del perfil cuadrático medio para una perturbación compleja compuesta por un hueco de seguido por sobre tensión.

Para una señal senoidal sin perturbar, de magnitud unitaria, el valor cuadrático medio se corresponde a 0.707. En la figura se observa como el perfil cuadrático medio decae

por debajo del valor nominal para luego escalar a valores que superan el valor normal.

A simple vista se puede observar que el valor máximo y el valor mínimo del perfil cuadrático medio son parámetros que claramente sirven para caracterizar perturbaciones del tipo huecos de tensión, sobre tensiones o interrupción. Sin embargo, esta herramienta resulta inapropiada para extraer información de perturbaciones del tipo estacionaria, como fluctuaciones y armónicos.

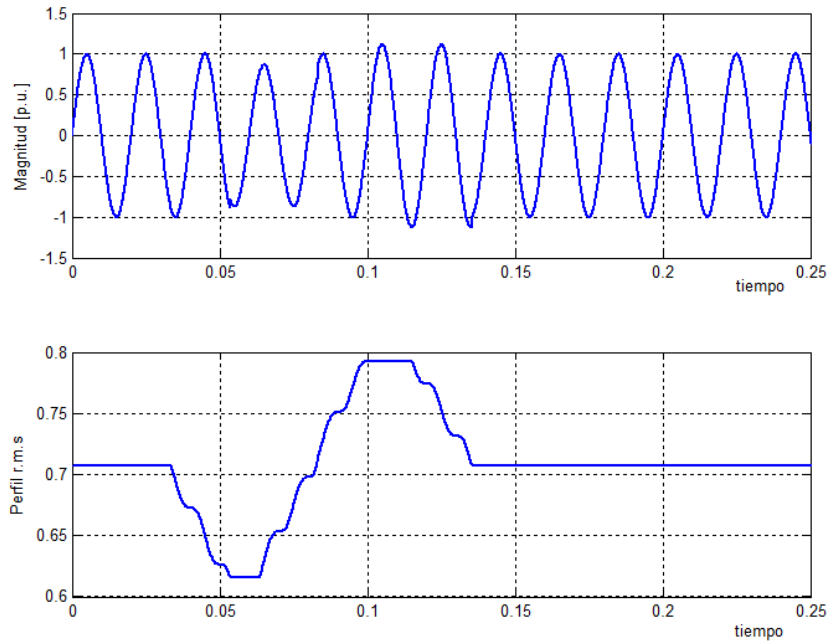


Figura 6. 8 - Perfil rms.

- *Transformada Wavelet (WT)*: Se propone el uso de la *WT* por su capacidad para segmentar información de la señal en diferentes bandas de frecuencia. En el presente trabajo se realiza el cálculo de la *DWT* de nueve 9 niveles, empleando la Wavelet madre Daubechie número 4.

A modo de ejemplo se propone una perturbación compleja compuesta por una perturbación estacionaria (*Armónicos*) y una perturbación de corta duración (Sobre tensión) como se ilustra en la Figura 6. 9.

En la Figura 6. 10, Figura 6. 11 y Figura 6. 12 se muestra el resultado del cálculo de la *WT* de 9 niveles.

Un análisis de los resultados muestra un transitorio impulsivo en los coeficientes D_1 - D_6 que tienen una correlación directa con los tiempos donde comienza y finaliza la perturbación del tipo sobre tensión.

A su vez, en los coeficientes D_7 - D_9 se pueden distinguir oscilaciones vinculadas con la perturbación armónica.

En su conjunto, como se demostrará en las siguientes secciones del presente capítulo, la *WT* de nueve niveles, en conjunto con el cálculo del perfil rms, son dos herramientas de procesamiento de señales que permiten extraer suficiente información para caracterizar perturbaciones de corta duración y perturbaciones estacionarias.

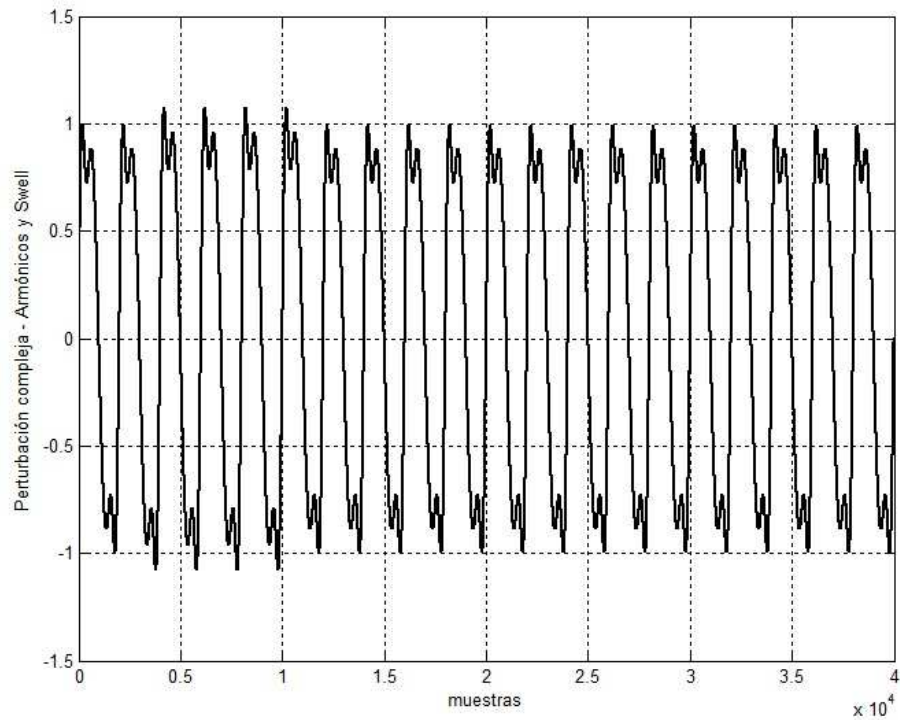


Figura 6. 9 - Perturbación compleja - Armónicos y sobre tensión.

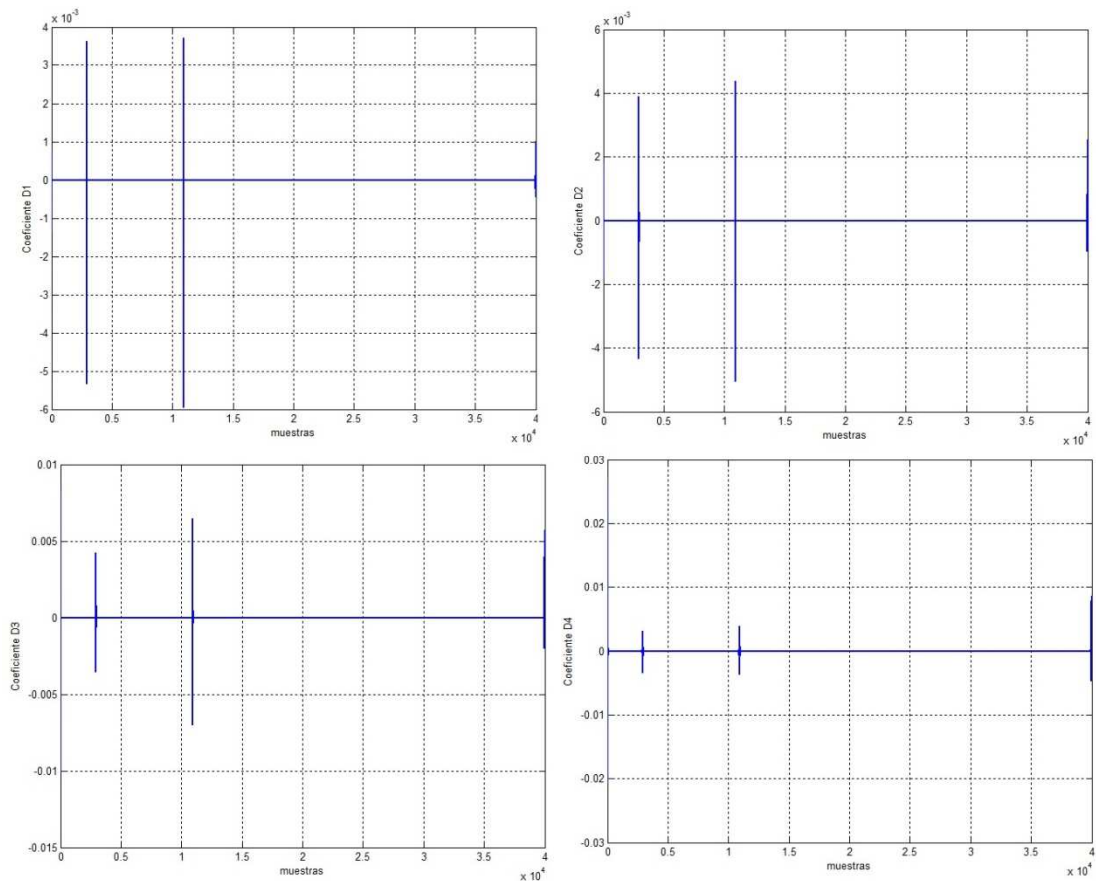


Figura 6. 10 - Coeficientes de detalle (D₁ - D₄).

En el gráfico correspondiente al coeficiente D₁ de la WT presentado en la Figura 6. 10 se puede observar variaciones rápidas, de marcada amplitud, en las muestras correspondientes con el inicio y fin de una perturbación de corta duración.

Por lo tanto se puede concluir que una perturbación de corta duración genera picos en las muestras de la WT y la identificación de estos abruptos cambios en el coeficiente D₁ permitirá identificar el inicio y el fin de dicha perturbación.

Comparar el coeficiente D₁ con un determinado umbral permite la detección del cien por cien de las perturbaciones de corta duración en condiciones ideales.

Sin embargo, en la práctica, las muestras de una forma de onda adquirida se encuentran corruptas por ruido aditivo.

Por su naturaleza, el ruido presenta contenido espectral en todas las frecuencias (ruido blanco) que afectan por igual todos los coeficientes de la WT. La presencia de ruido en el coeficiente D₁, bajo ciertas circunstancias, puede enmascarar las variaciones abruptas que delatan la presencia de una perturbación de corta duración. En la práctica, una perturbación no detectada no puede ser clasificada. Por este motivo la etapa de detección incide directamente en el porcentaje de aciertos de un clasificador.

Diferentes técnicas para eliminar el ruido, tanto en la etapa de hardware como en la de software son aplicadas para disminuir su impacto en la etapa de detección.

En esta tesis se consideran señales libres de ruido por lo que la etapa de detección no tendrá incidencia en los porcentajes de clasificación aunque se aclara que en la práctica se debe resolver la problemática del ruido blanco aditivo.

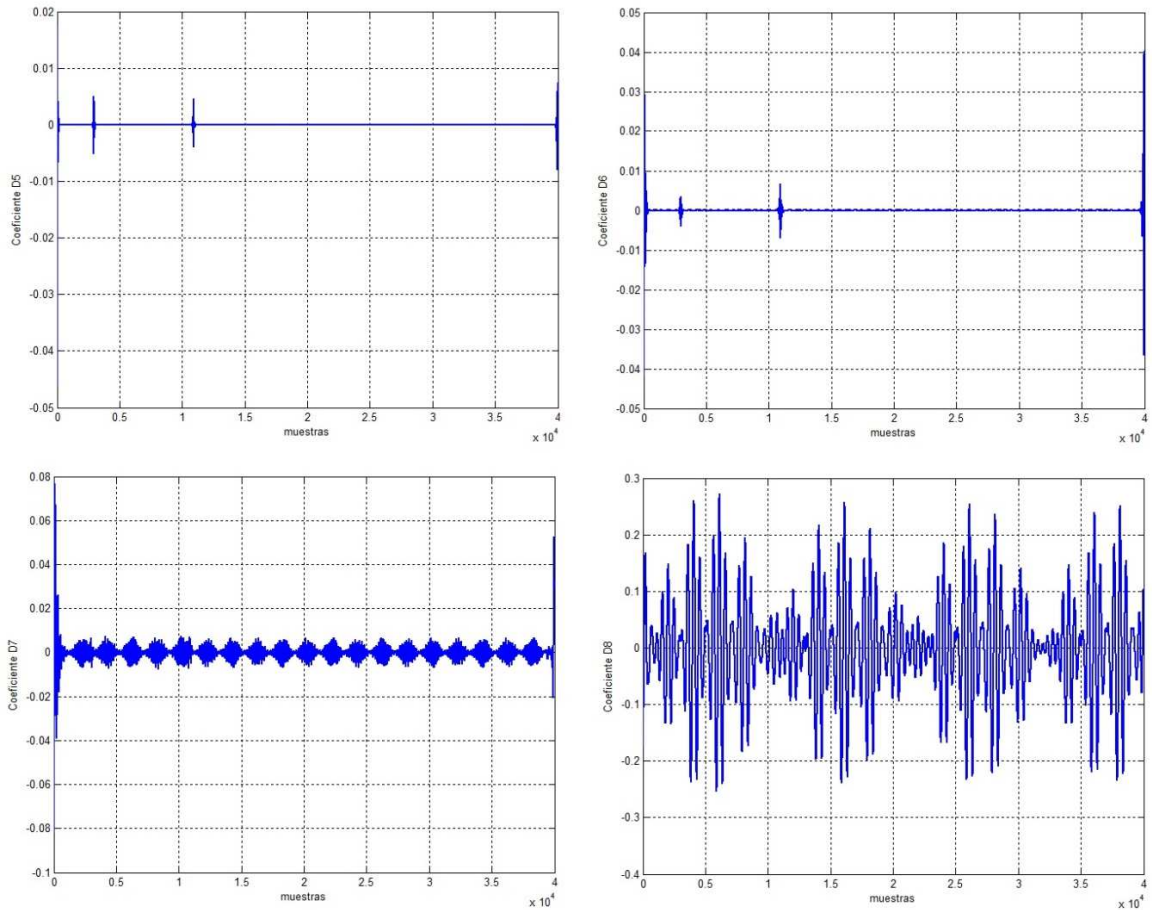


Figura 6. 11 - Coeficientes de detalle (D₅-D₈).

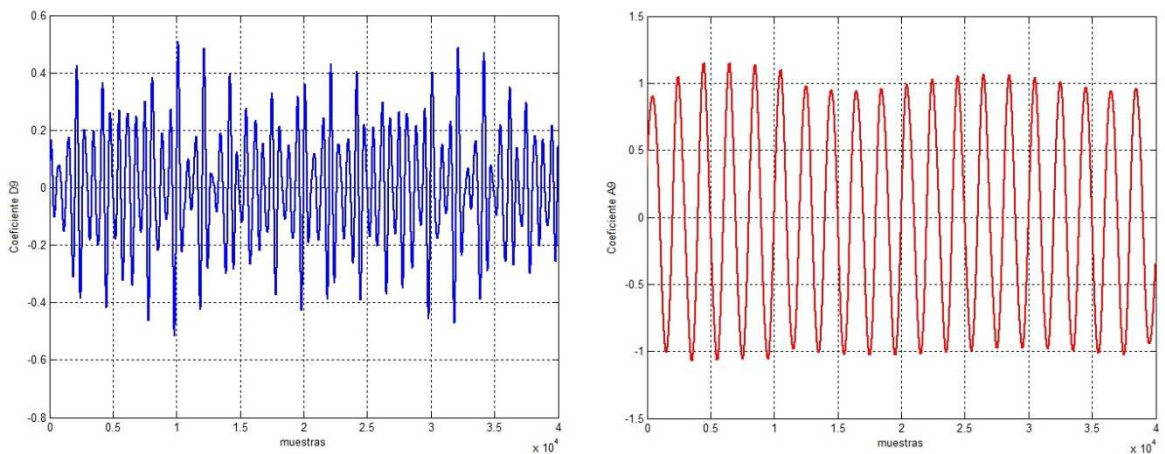


Figura 6. 12 - Coeficiente de detalle D₉ y de aproximación A₉.

6.2.3.2 Extracción de parámetros característicos.

Como se mencionó en el *capítulo 4* el proceso de extracción de parámetros característicos se divide en dos sub procesos, el cálculo de los parámetros característicos y su posterior normalización.

La extracción de los parámetros característicos es un proceso que consiste en la aplicación de cálculos con el objetivo de parametrizar el perfil rms y los coeficientes de la WT obtenidos en la etapa del procesamiento de señales. Los parámetros característicos calculados se muestran en la Tabla 6. 3.

Tabla 6. 3 - Parámetros característicos.

Técnica de Procesamiento de Señales	Parámetro característico n°	Parámetros característicos	Nomenclatura
Transformada Wavelet	1	Energía normalizada del coeficiente d1.	Energía d1
	2	Valor máximo del módulo del coeficiente d1.	Max mod d1
	3	Desviación estándar del coeficiente d1.	DStd d1
	4	Energía normalizada del coeficiente d2	Energía d2
	5	Valor máximo del módulo del coeficiente d2.	Max mod d2
	6	Desviación estándar del coeficiente d2.	DStd d2
	7	Energía normalizada del coeficiente d3.	Energía d3
	8	Valor máximo del módulo del coeficiente d3.	Max mod d3
	9	Desviación estándar del coeficiente d3.	DStd d3
	10	Energía normalizada del coeficiente d4.	Energía d4
	11	Valor máximo del módulo del coeficiente d4.	Max mod d4
	12	Desviación estándar del coeficiente d4	DStd d4
	13	Energía normalizada del coeficiente d5.	Energía d5
	14	Valor máximo del módulo del coeficiente d5.	Max mod d5
	15	Desviación estándar del coeficiente d5	DStd d5
	16	Energía normalizada del coeficiente d6.	Energía d6
	17	Valor máximo del módulo del coeficiente d6.	Max mod d6
	18	Desviación estándar del coeficiente d6.	DStd d6
	19	Energía normalizada del coeficiente d7	Energía d7
	20	Valor máximo del módulo del coeficiente d7.	Max mod d7
	21	Desviación estándar del coeficiente d7.	DStd d7
	22	Energía normalizada del coeficiente d8	Energía d8
	23	Valor máximo del módulo del coeficiente d8.	Max mod d8
	24	Desviación estándar del coeficiente d8.	DStd d8
	25	Energía normalizada del coeficiente d8	Energía d9
	26	Valor máximo del módulo del coeficiente d9.	Max mod d9
	27	Desviación estándar del coeficiente d9.	DStd d9
	28	Energía normalizada del coeficiente a9	Energía a9
	29	Valor máximo del módulo del coeficiente a9.	Max mod a9
	30	Desviación estándar del coeficiente a9.	DStd a9
Perfil rms	31	Valor máximo perfil rms.	Máx. rms
	32	Valor mínimo perfil rms.	Mín. rms.

Los 32 parámetros característicos son calculados independientemente del tipo de perturbación presente en el set de entrenamiento.

Este proceso provoca una reducción de la dimensión de los datos a ser procesados por las siguientes etapas del algoritmo ya que el set de entrenamiento, originalmente representado por una matriz de dimensión 2600×4000 es reducida a una matriz de dimensión 2600×32 . De esta manera, a pesar que el procesamiento se realiza sin evaluar a que clase pertenece cada parámetro, se logra reducir la dimensión del set de entrenamiento en un factor de 99,2%.

De acuerdo a lo expuesto, los parámetros característicos se obtienen por medio de cálculos realizados sobre los coeficientes de la *WT* y del *perfil rms*. Como ya se mencionó en el capítulo 4, con el objetivo de acotar el rango dinámico de cada columna de la matriz de parámetros característicos es necesario aplicar técnicas de normalización. En este trabajo se utiliza una metodología basada en el ajuste del valor medio y la desviación estándar para la normalización del set de entrenamiento.

El sub proceso de normalización recibe una matriz y ajusta su valor medio y su varianza a los valores deseados. La expresión matemática empleada por este algoritmo es:

$$y = (x - x_{av}) \frac{y_{std}}{x_{std}} + y_{av} \quad (6. 2)$$

Donde x_{av} y x_{std} representan el valor medio y la desviación estándar de cada columna de la matriz y y_{av} e y_{std} son los valores esperado y la desviación estándar de la matriz resultante.

En nuestro ejemplo, el valor medio y la desviación estándar esperados se proponen como cero y uno respectivamente.

Cuando se cuenta con un set de entrenamiento conformado por un determinado número de formas de onda, la normalización se realiza calculando el valor medio y la desviación estándar de cada columna del arreglo (cada columna representa un parámetro característico en particular) empleado para el entrenamiento y la validación.

Una vez que el clasificador se encuentra diseñado y entrenado, el modelo es utilizado en un sistema de monitoreo para identificar perturbaciones en las formas de onda adquiridas. En este caso, para guardar concordancia con los valores empleados en el entrenamiento, también se deben normalizar los parámetros característicos. Debido a que, en esta instancia no se cuenta con un arreglo al cual calcular el valor medio y la desviación estándar de cada parámetro característico lo que se hace es emplear los valores obtenidos en el entrenamiento del clasificador.

A la izquierda, en Figura 6. 13, se ilustra el resultado del cálculo de los treinta y dos parámetros característicos propuestos para la forma de onda mostrada en la Figura 6. 9. En la gráfica se puede observar que los parámetros característicos cuyo índice está comprendido entre el número 1 y el 20 presentan valores muy pequeños, cercanos a cero.

Por el contrario la gráfica situada a la derecha se corresponde al vector característico normalizado donde se puede observar que los valores son ajustados a un rango que va de -1 a 1.

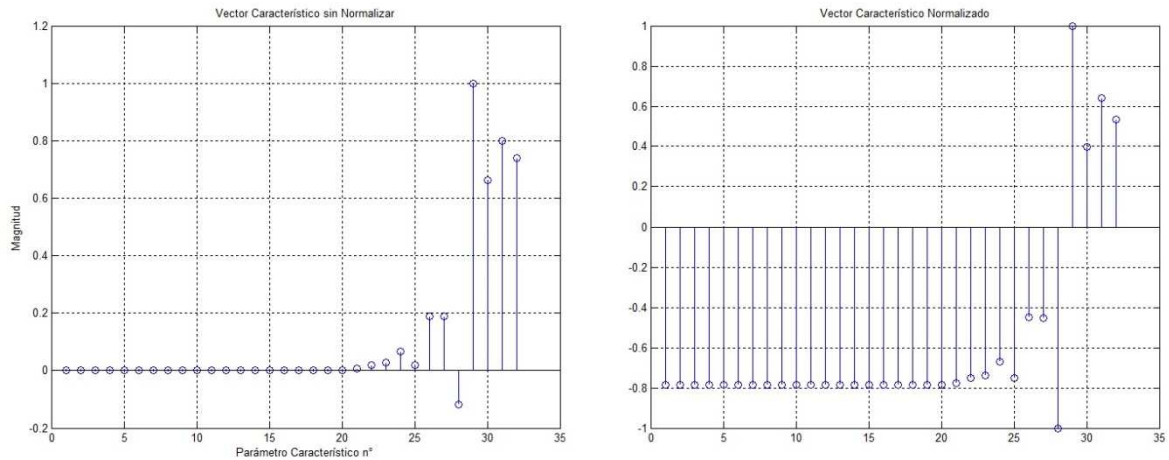


Figura 6. 13 - Resultados de la normalización.

6.2.3.3 Minería de Datos.

La etapa de minería de datos es un proceso desarrollado para seleccionar un pequeño sub conjunto de parámetros característicos del total de los parámetros que ingresan a este proceso. A diferencia del proceso de *Extracción de parámetros característicos*, el proceso de *Minería de datos* fundamentará sus cálculos en cada una de las 5 clases presentes en el set de entrenamiento original.

Como se explicó en el *capítulo 4* este proceso está compuesto por tres etapas: Generación de etiquetas, el filtrado heurístico y el algoritmo de búsqueda exhaustiva. A continuación se presentan los resultados de estos tres sub procesos:

Generación de etiquetas: La asignación de etiquetas se realiza de acuerdo a lo mostrado en la Tabla 6. 4. Como se puede observar a cada clasificador binario se le asigna un conjunto de etiquetas de acuerdo a cada perturbación simple.

De la misma manera, en la Tabla 6. 5 se muestra la estrategia seleccionada para asignar etiquetas a cada clasificador binario para las perturbaciones complejas estudiadas en este trabajo. Es importante aclarar que, en esta propuesta las perturbaciones complejas no son tratadas como una clase en si misma si no que, por el contrario, son tratadas como una superposición de dos perturbaciones simples. Por este motivo, no existe una etiqueta en particular para ninguna perturbación compleja.

Tabla 6. 4 - . Asignación de etiquetas para perturbaciones simples.

Perturbaciones Simples	Clase	Etiqueta SVM Huecos	Etiqueta SVM Sobretensión	Etiqueta SVM Armónicos	Etiqueta SVM Interrupción	Etiqueta SVM Fluctuaciones
Huecos	1	1	-1	-1	-1	-1
Sobre tensión	2	-1	1	-1	-1	-1
Armónicos	3	-1	-1	1	-1	-1
Interrupción	4	-1	-1	-1	-1	-1
Fluctuaciones	5	-1	-1	-1	1	1

Tabla 6. 5 - Asignación de etiquetas para perturbaciones complejas.

Perturbaciones Complejas	Clase	Etiqueta SVM Huecos	Etiqueta SVM Sobretensión	Etiqueta SVM Armónicos	Etiqueta SVM Interrupción	Etiqueta SVM Fluctuaciones
Armónicos y Huecos.	3+1	1	-1	1	-1	-1
Armónicos y Sobretensión.	3+2	-1	1	1	-1	-1
Armónicos e Interrupción	3+4	-1	-1	1	1	-1
Fluctuaciones y Huecos	5+1	1	-1	-1	-1	1
Fluctuaciones y Sobre tensión	5+2	-1	1	-1	-1	1
Fluctuaciones e Interrupción	5+4	-1	-1	1	-1	1

- Filtrado heurístico:* El filtrado heurístico tiene por objetivo el seleccionar un número acotado de parámetros característicos que mejor describen a cada clase. Para ello, el sub proceso realiza una serie de cálculos y asigna un factor numérico a cada parámetro característico según su grado de relevancia a la hora de describir una clase particular. Luego confecciona una lista que incluye, ordenados de mayor a menor relevancia, la totalidad de los parámetros característicos con su respectivo índice de relevancia. Este índice es un valor numérico que indica cuán importante resulta el parámetro para describir a una determina clase. La selección de parámetros se realiza sumando el porcentaje de relevancia de los j parámetros preponderantes hasta alcanzar un porcentaje del 90%. Este proceso se repite por cada una de las clases presentes en el set de entrenamiento. En la Tabla 6. 6 se presentan los resultados del proceso de filtrado heurístico.

Tabla 6. 6 - Resultados del filtrado heurístico de parámetros característicos.

Perturbación	Parámetro Característico Seleccionado (Umbral del 90%)	Dimensión
Huecos	Min RMS, Max RMS, Max Mod a ₉ , Energía d ₇ , Max mod d ₅ , Dst d ₄ .	2600 x 6
Sobre tensión	Min RMS, Max RMS, Std d ₉ , Energía d ₇ , Max mod d ₄ , Max mod d ₃ , Max mod d ₂ .	2600 x 7
Armónicos	Min RMS, Max RMS, Dst d ₉ , Max mod d ₉ , Dst d ₈ , Max mod d ₅	2600 x 6
Interrupción	Min RMS, Max RMS, Max mod a ₉ , Std d ₉ , Max mod d ₄ .	2600 x 5
Fluctuaciones	Min RMS, Max RMS, Max mod a ₉ , Std d ₉ , Max mod d ₉ , Std d ₈ , Max mod d ₈ .	2600 x 7

De la matriz de parámetros característicos de dimensión 2600×32 se obtienen cinco matrices cuya dimensión, en el peor de los casos es de 2600×7 . Por este motivo esta instancia representa una segunda etapa de reducción de dimensión.

- *Algoritmo de Búsqueda Exhaustiva:* Después que el número de parámetros característicos es significativamente reducido por la etapa de filtrado heurístico, es importante encontrar una combinación de los parámetros seleccionados que produce el mayor porcentaje de precisión en la etapa de entrenamiento y validación. Para entrenar el algoritmo se utilizó un 60% de las 2600 perturbaciones que conforman el set entrenamiento mientras que el restante 40% se utilizó para su validación. En la Figura 6. 14 se presenta el porcentaje de acierto alcanzado por el clasificador para cada una de las $2^6 = 64$ combinaciones de los parámetros característicos. La Tabla 6. 7 muestra los resultados del algoritmo de búsqueda exhaustiva para el clasificador SVM encargado de identificar los huecos de tensión. Se puede observar que se necesitan, al menos, cuatro parámetros característicos para obtener un porcentaje de acierto superior al 98%. Debido a que más de una combinación de parámetros característicos genera una precisión máxima, se selecciona aquella que emplea la menor cantidad de parámetros característicos. Si no se hubiera aplicado el algoritmo de búsqueda exhaustiva y se hubieran utilizado los 6 parámetros característicos obtenidos del filtrado heurístico, tal como se ve en la Figura 6. 14, la precisión máxima alcanzada se hubiera correspondido con valores por debajo del 93%. Es evidente que, desde la perspectiva del clasificador SVM, no todos los parámetros característicos contienen información descriptiva relevante para identificar una perturbación del tipo hueco de tensión.

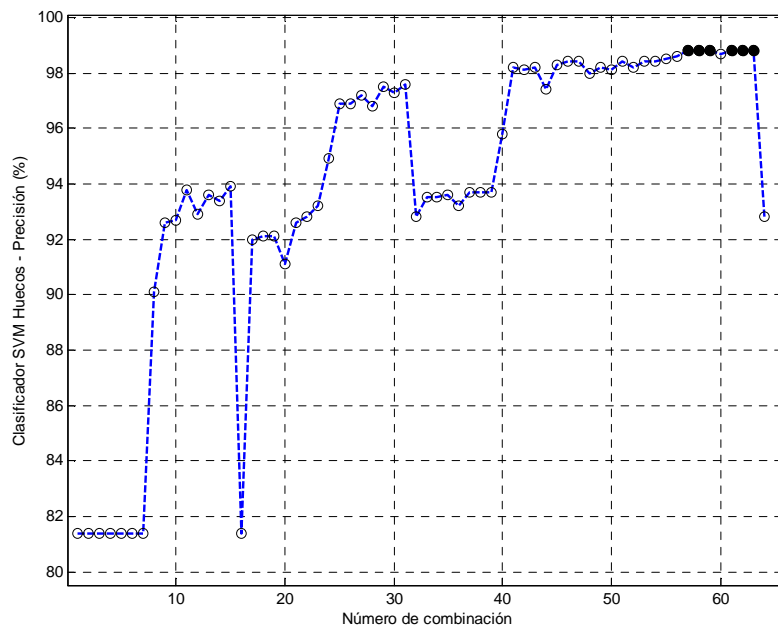


Figura 6. 14 - Resultado del algoritmo de búsqueda exhaustiva – SVM Huecos.

Tabla 6. 7 - Combinación de máxima precisión - Huecos.

Combinación Número	Min Rms	Max RMS	Max módulo a9	Energía D7	Max Modulo D5	Std D4	Métrica	Porcentaje
57	1	1	1	0	0	1	4	98.8
58	1	1	1	0	1	0	4	98.8
59	1	1	1	0	1	1	5	98.8
61	1	1	1	1	0	1	5	98.8
62	1	1	1	1	1	0	5	98.8
63	1	1	1	1	1	1	6	98.8

Las figuras comprendidas desde Figura 6. 15 a la Figura 6. 18 muestran los perfiles de precisión para los restantes clasificadores binarios.

De la misma manera, las tablas comprendidas desde la Tabla 6. 8 a la Tabla 6. 11 muestran las combinaciones de parámetros característicos que generan la máxima precisión para cada etapa binaria del clasificador multiclase.

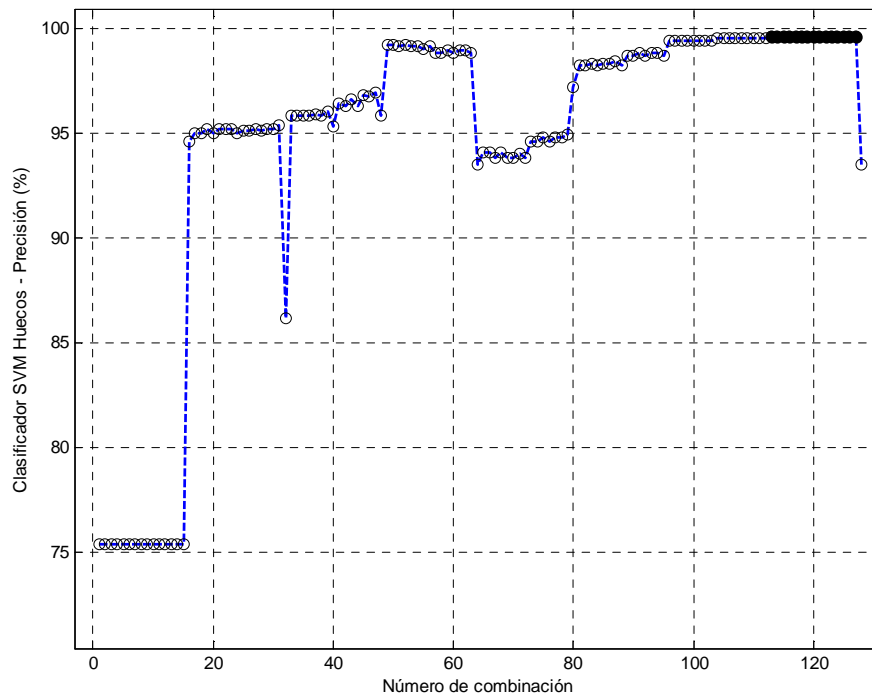


Figura 6. 15 - Resultado del algoritmo de búsqueda exhaustiva - SVM sobre tensión.

Tabla 6. 8 - Combinación de máxima precisión – Sobretensión.

Combinación Número	Min Rms	Max RMS	DSt D9	Energía D7	Max Modulo D4	Max Modulo D3	Max mod D2	Métrica	Porcentaje
113	1	1	1	0	0	0	1	4	99.6
114	1	1	1	0	0	1	0	4	99.6
115	1	1	1	0	0	1	1	5	99.6
116	1	1	1	0	1	0	0	4	99.6
117	1	1	1	0	1	0	1	5	99.6
118	1	1	1	0	1	1	0	5	99.6
119	1	1	1	0	1	1	1	6	99.6
120	1	1	1	1	0	0	0	4	99.6
121	1	1	1	1	0	0	1	5	99.6
122	1	1	1	1	0	1	0	5	99.6
123	1	1	1	1	0	1	1	6	99.6
124	1	1	1	1	1	0	0	5	99.6
125	1	1	1	1	1	0	1	6	99.6
126	1	1	1	1	1	1	0	6	99.6
127	1	1	1	1	1	1	1	7	99.6

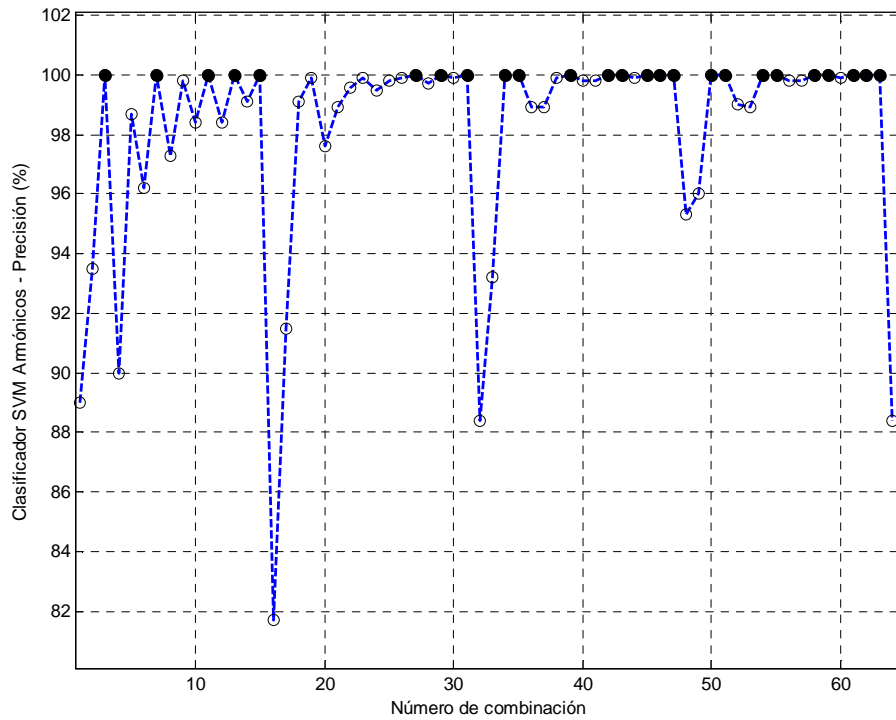


Figura 6. 16 - Resultado del algoritmo de búsqueda exhaustiva - Armónicos.

Tabla 6. 9 - Combinación de máxima precisión - Armónicos.

Combinación Número	Min Rms	Max RMS	DSt D9	Max Modulo D9	DSt D8	Max mod D5	Métrica	Porcentaje
3	0	0	0	0	1	1	2	100
7	0	0	0	1	1	1	3	100
11	0	0	1	0	1	1	3	100
13	0	0	1	1	0	1	3	100
15	0	0	1	1	1	1	4	100
27	0	1	1	0	1	1	4	100
29	0	1	1	1	0	1	4	100
31	0	1	1	1	1	1	5	100
34	1	0	0	0	1	0	2	100
35	1	0	0	0	1	1	3	100
39	1	0	0	1	1	1	4	100
42	1	0	1	0	1	0	3	100
43	1	0	1	0	1	1	4	100
45	1	0	1	1	0	1	4	100
46	1	0	1	1	1	0	4	100
47	1	0	1	1	1	1	5	100
50	1	1	0	0	1	0	3	100
51	1	1	0	0	1	1	4	100
54	1	1	0	1	1	0	4	100
55	1	1	0	1	1	1	5	100
58	1	1	1	0	1	0	4	100
62	1	1	1	1	1	0	1	100
63	1	1	1	1	1	1	0	100

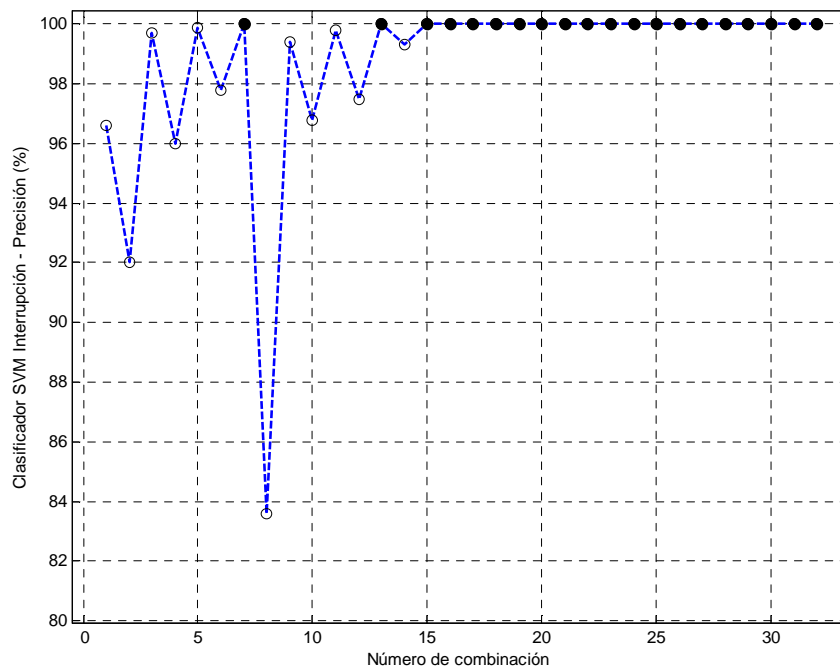


Figura 6. 17 - Resultado del algoritmo de búsqueda exhaustiva - Interrupción.

Tabla 6. 10 - Combinación de máxima precisión – Interrupción.

Combinación Número	Min Rms	Max RMS	Max módulo a9	DSt D9	Max Modulo D4	Métrica	Porcentaje
7	0	0	1	1	1	3	100
15	0	1	1	0	1	3	100
16	0	1	1	1	1	4	100
17	1	0	0	0	0	1	100
18	1	0	0	1	0	2	100
19	1	0	0	1	1	3	100
20	1	0	1	0	0	2	100
21	1	0	1	0	1	3	100
22	1	0	1	1	0	3	100
23	1	0	1	1	1	4	100
24	1	1	0	0	0	2	100
25	1	1	0	0	1	3	100
26	1	1	0	1	0	3	100
27	1	1	0	1	1	4	100
28	1	1	1	0	0	3	100
29	1	1	1	0	1	4	100
31	1	1	1	1	0	4	100
32	1	1	1	1	1	5	100

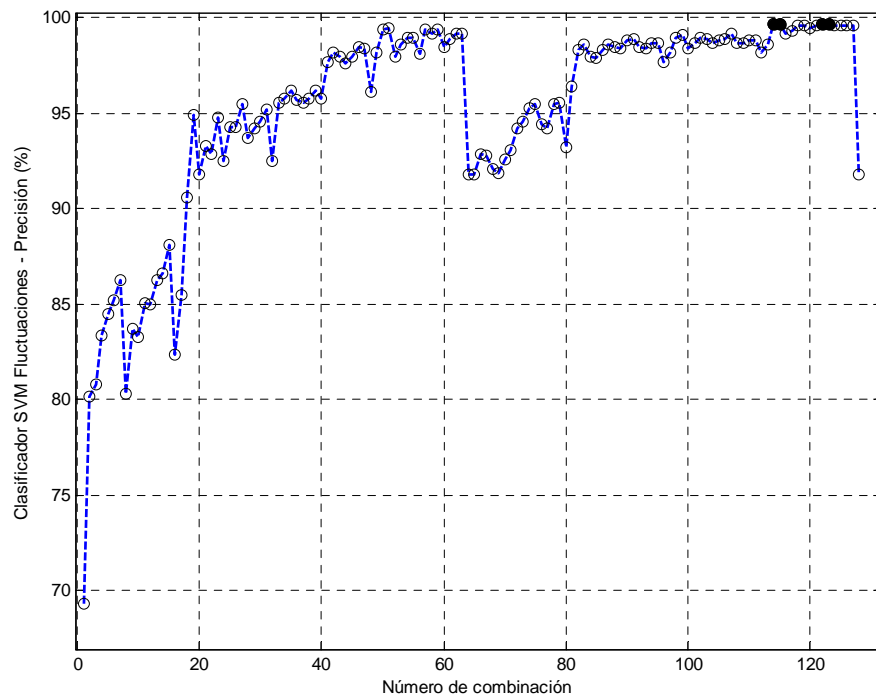


Figura 6. 18 - Resultado del algoritmo de búsqueda exhaustiva - Fluctuaciones.

Tabla 6. 11 - Combinación de máxima precisión – Fluctuaciones.

Combinación Número	Min Rms	Max RMS	Max módulo a9	DSt D9	Max Modulo D9	DSt D8	Max Modulo D8	Métrica	Porcenta je
114	1	1	1	0	0	1	0	4	99.7
115	1	1	1	0	0	1	1	5	99.7
122	1	1	1	1	0	1	0	5	99.7
123	1	1	1	1	0	1	1	6	99.7

En la Tabla 6. 12 se resumen los resultados del algoritmo de búsqueda exhaustiva. Como se puede apreciar, el número de parámetros característicos que se utilizarán para el entrenamiento y prueba del clasificador se ven drásticamente reducidos gracias al proceso *Minería de Datos*.

Tabla 6. 12 - Resultado del algoritmo de búsqueda exhaustiva.

Perturbación	Parámetro característicos seleccionados	Precisión de Validación	Dimensión
Huecos	Min RMS, Max RMS, Max mod a ₉ , DSt d ₄ .	98,8%	2600 x 4
Sobre tensión	Min RMS, Max RMS, Dst d ₉ , Max mod d ₂ .	99,6%	2600 x 4
Armónicos	DSt d ₈ , Max mod d ₅ .	100 %	2600 x 2
Interrupción	Min RMS.	100%	2600 x 1
Fluctuaciones	Min RMS, Max RMS, Max mod a ₉ , DSt d ₈ .	99,7%	2600 x 4

6.2.3.4 Optimización de parámetros del clasificador.

El entrenamiento de un clasificador basado en *SVM* posee un parámetro de ajuste conocido como factor de penalización y es representado por la letra *C* y uno (o más) parámetros vinculados con la función kernel seleccionada.

Los clasificadores empleados en el presente trabajo de tesis utilizan la función kernel *RBF* por su probada eficiencia en previos trabajos de investigación.

Una función *RBF* se describe matemáticamente como sigue:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (6. 3)$$

Como se puede apreciar, la función *RBF* depende solo de las muestras y de la variable sigma.

Por lo tanto, el entrenamiento de un clasificador *SVM* que utiliza una función *RBF* como función *kernel*, queda definido por la combinación de valores que asumen los parámetros *C* y *sigma*. No existe una ecuación matemática para el cálculo de los valores óptimos de *C* y *sigma*, por lo tanto no es posible conocer a priori los valores óptimos para la resolución de un determinado problema.

Uno de los métodos para encontrar la mejor combinación de estos parámetros es el método de búsqueda tipo grilla presentado en el *capítulo 4*. Este método, en combinación con la técnica de validación cruzada, son empleados para encontrar los resultados presentados en la Tabla 6. 13.

En la tabla se pueden encontrar la combinación de parámetros óptima que gobiernan cada clasificador binario SVM con su respectiva precisión máxima alcanzada. Es importante aclarar que estos parámetros se logran para un set de entrenamiento y validación conformado solo para aquellas combinaciones óptimas de los parámetros característicos presentadas en la Tabla 6. 12.

Tabla 6. 13 - Resultados del algoritmo de búsqueda de grilla.

Etapa binaria	Sigma	C	Precisión de validación
1 – Huecos.	2	4000	98,8%
2 – Sobre tensiones.	2	1000	99,6%
3 – Armónicos.	3.5	3000	100 %
4 – Interrupción.	1e-5	50	100%
5 – Fluctuaciones.	4	500	99,7%

6.2.4 Entrenamiento.

Una vez que se obtienen los set de parámetros característicos que mejor representan a cada perturbación, los cuales se presentan en la Tabla 6. 12 y el par óptimo de parámetros que gobiernan a cada etapa del clasificador, los cuales se resumen en la Tabla 6. 13, cada clasificador binario que conforman la arquitectura One vs Rest es entrenada. Para ello, en el presente desarrollo se emplea la librería LibSVM [130].

6.2.4.1 Resultados: Pruebas de rendimiento.

En la Figura 6. 19 se muestra un diagrama del algoritmo de clasificación presentado en el capítulo 4. Este algoritmo será el empleado para probar el rendimiento del arreglo mostrado Figura 6. 6 en la cuyos clasificadores binarios fueron diseñados y entrenados de acuerdo a lo presentado en la sección 6.2.3.

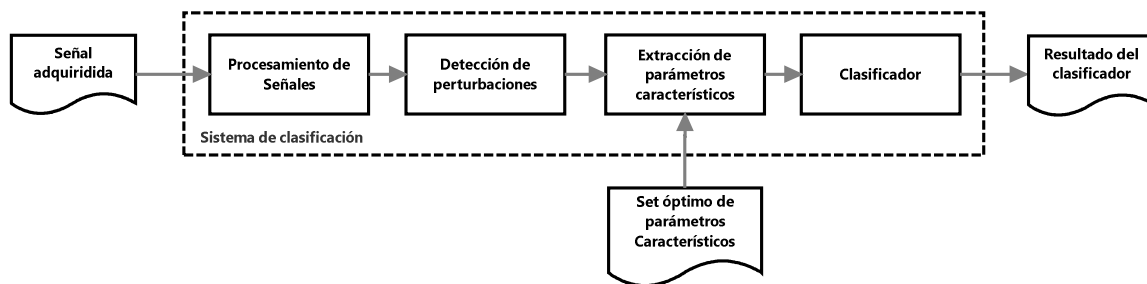


Figura 6. 19 - Algoritmo de clasificación.

En las siguientes sub secciones se presentarán los resultados obtenidos para dos posibles escenarios, uno planteado para probar el rendimiento ante perturbaciones simples y el segundo, y más importante, para probar el porcentaje de aciertos del método propuesto cuando se emplean perturbaciones complejas.

6.2.4.2 Prueba 1 – Perturbaciones Simples.

Si bien el trabajo se enfoca en estudiar el rendimiento del método propuesto para perturbaciones complejas, es necesario probar que el algoritmo clasifique correctamente perturbaciones simples.

Para la prueba se generaron 1000 señales, 200 por cada tipo de perturbación. Los parámetros que gobiernan a cada perturbación fueron generados de manera aleatoria dentro de los rangos establecidos por la norma [122].

La matriz de confusión presentada en la Tabla 6. 14 representa los resultados del clasificador para un escenario de prueba conformado solo por perturbaciones simples.

Tabla 6. 14 - Matriz de confusión - Perturbaciones simples.

	Huecos	Sobre tensión	Armónicos	Interrupción	Fluctuación	Perdidas	Dobles
1 – Huecos	200	0	0	0	0	0	0
2 – Sobretension	0	200	0	0	0	0	0
3 –Armónicos	0	0	198	0	0	1	1 (2+3)
4- Interrupción	0	0	0	199	0	0	1 (4+3)
5 –Fluctuaciones	0	0	0	0	200	0	0

La columna de la tabla titulada con la tabulada con la etiqueta *Dobles* se corresponde a aquellas perturbaciones simples que fueron erróneamente clasificadas como perturbaciones complejas. El número que aparece fuera de los paréntesis indica cantidad y los números indicados entre paréntesis se corresponden con las clases que conforman la perturbación compleja.

Así, por ejemplo, para la fila número 3 de la matriz de confusión correspondiente a la perturbación armónicos, la columna etiquetada como doble presenta un resultado indicado por $1(2 + 3)$ que significa una única perturbación compleja doble conformada por la clase *Sobre tensión* (2) y *Armónicos* (3).

Lo mismo ocurre para el caso de la clasificación doble presente la fila que representa las Interrupciones, en cuyo caso la respuesta del clasificador se corresponde con una única perturbación compleja conformada por la suma de Armónicos e Interrupción.

Es importante destacar que en ambos casos, una doble clasificación podría interpretarse como una clasificación parcialmente correcta.

Como principal conclusión del escenario desarrollado se puede destacar que el método propuesto presenta resultados destacados ya que fue capaz de clasificar correctamente el 99,7% de las perturbaciones de prueba.

6.2.4.3 Prueba 2 – Perturbaciones complejas

De manera similar que en el caso presentado en la sección anterior, se generaron 1200 señales, 200 por cada tipo de perturbación compleja. Los parámetros que gobiernan a cada

perturbación fueron generados de manera aleatoria dentro de los rangos establecidos por la norma [9].

La matriz de confusión presentada en la Tabla 6. 15 representa los resultados del clasificador.

Tabla 6. 15 - Matriz de confusión - Perturbaciones complejas.

	Arm.y Huecos.	Arm y Sobre tensión.	Arm e Interr.	Fluct. y Huecos.	Fluct. y Sobre tensión.	Fuct. e Interr.	Simple.	Triple.
Arm y Huecos.	197	0	0	1	0	0	1 (3)	1 (1+3+5)
Arm y Sobre tensión.	0	198	0	0	0	0	1 (3)	0
Arm e Int.	0	0	194	0	0	1	4 (3)	1 (3+4+5)
Fluct Huecos	0	0	0	190	0	0	10 (5)	0
Fluct. y Sobre tensión.	0	0	0	0	189	0	11 (5)	0
Fluct e Int.	0	0	0	0	0	199	1 (5)	0

Los valores entre paréntesis indican el tipo de perturbación, o una combinación de las mismas que fueron clasificadas.

De las 1200 señales complejas generadas solo 33 fueron clasificadas erróneamente lo que da un porcentaje de acierto de 97, 25%.

Analizando las clasificaciones erróneas se puede encontrar que, en la mayoría de los casos, el clasificador fue capaz de clasificar una de las dos perturbaciones que conforman el evento complejo. En este caso, el porcentaje de acierto alcanzaría el 98,58%.

6.2.4.4 Comparación de los resultados.

En la Tabla 6. 16 se presentan los resultados obtenidos en los trabajos más relevantes en lo relativo a la clasificación de *perturbaciones complejas*.

Si bien la *precisión* del clasificador es un parámetro que se puede establecer para realizar comparaciones con trabajos similares, existen muchos otros aspectos a tener en cuenta a la hora de cotejar la técnica propuesta con otras publicadas. Algunos de estos aspectos pueden ser la complejidad del algoritmo, el tiempo de procesamiento, el costo computacional, entre otros posibles parámetros.

Por otro lado, para que las comparaciones sean resulten más indicativas deberían todas emplear el mismo set de entrenamiento y prueba.

Teniendo esto último en cuenta se procede a cotejar los resultados obtenidos con otros algoritmos desarrollados con el mismo fin que el aquí presentado.

El porcentaje obtenido por el algoritmo propuesto, que alcanza el 98,54%, supera a casi todos los trabajos del área para la clasificación de perturbaciones complejas. Este porcentaje es solo superado por el algoritmo propuesto en [125].

Sin embargo, es importante mencionar, que el clasificador SVM propuesto por Gang Liu, es entrenado con el resultado de un proceso de reducción de dimensión conocido como es Análisis de Componente Principal (PCA). Esto provoca que, cada vez que un set de parámetros característicos necesite ser clasificado se deberá, previamente, realizar esta transformación. Esto genera una carga computacional adicional con respecto al método propuesto en la presente tesis.

Tabla 6. 16 - Comparación de resultados – Perturbaciones complejas.

Autores	Año	Procesamiento de señales	Inteligencia Artificial	Selección de parámetros. característicos	Precisión (%) Perturbación Compleja
Liu, Zhigang; Cui, [127]	2015	EEMD	RankSVM	N/A	95,36
Biswal, B; Biswal, M. K; Dash, [22]	2013	FDST	DT	ES	98,19
Cheng-Long, Chuang; Yen-Ling [124]	2005	WT	DS-ANN	N/A	98,4875
Gang, Liu; Fanguang, Li; [125]	2013	WT	SVM	ICA	99,75
Dalai, S; Dey, D; Chatterjee, B [126].	2013	CHT	ESVM	PCA	97,6

6.3 CONCLUSIONES

En este capítulo se presenta un método simple, rápido y fácil de entrenar que permite la clasificación de perturbaciones simples y complejas.

La metodología se basa en la *WT* y el cálculo del *perfil rms* para la extracción de parámetros característicos, técnicas de minería de datos para la selección del set óptimo de parámetros característicos y una arquitectura *SVM One vs Rest* binarios para la clasificación de perturbaciones simples y complejas.

La arquitectura propuesta emplea clasificadores que deben tomar decisiones binarias, lo que los hace mucho más simples que sus pares entrenados para tomar decisiones multiclase.

Además, el método diseñado utiliza solo una etapa por cada perturbación simple que se desea clasificar. Esto se debe a que cada perturbación compleja no es tratada como una clase nueva si no que cada clasificador binario es entrenado y optimizado para distinguir un tipo particular de evento de calidad de energía aunque el mismo se encuentre involucrado como parte de una perturbación compleja. Esto representa una clara ventaja frente a aquellas arquitecturas que configuran una clase individual por cada tipo de perturbación compleja.

La arquitectura propuesta permite, a su vez, la paralización de los cálculos, lo que permitiría reducir considerablemente el tiempo de procesamiento y costos computacionales.

El algoritmo emplea una etapa de selección de parámetros característicos que reduce considerablemente la cantidad de parámetros característicos necesarios para lograr destacados resultados. En este sentido se muestra que, con un número pequeño de parámetros característicos (4 para los Huecos, Sobretensión y Fluctuaciones, 2 para Armónicos e incluso un único parámetro para el caso de las Interrupciones) se logran resultados destacados.

Por otro lado, la exploración de los parámetros que gobiernan cada clasificador empleando el algoritmo de búsqueda por grilla consigue optimizar el comportamiento de cada clasificador binario.

Una primera experiencia con perturbaciones simples demuestra que el algoritmo propuesto presenta una notable capacidad ya que logra un porcentaje de acierto de 99,7%.

Una segunda experiencia fue diseñada para probar la capacidad de clasificación en el caso de perturbaciones complejas. Los resultados obtenidos son destacados ya que se logra un porcentaje que supera el 97.5%. Además, el restante 2,5 % está compuesto, en un alto número, por clasificaciones parcialmente correctas. Esto se logra gracias a la estructura de clasificadores binarios en paralelo.

Debido a que es complejo contar con perturbaciones complejas para el entrenamiento del clasificador, en el futuro se debe estudiar cual es el número óptimo dichas perturbaciones necesarias en el entrenamiento del clasificador para lograr resultados aceptables.

Por otro lado, la metodología propuesta introduce dos fuentes de error no presentes en clasificadores multiclase, como lo son las clasificaciones perdidas o clasificaciones

múltiples. Al igual que la mayoría de los métodos basados en máquinas de aprendizaje lineal, el clasificador propuesto deberá ser completamente re entrenado si se presenta un nuevo tipo de perturbación compleja, aunque no necesariamente este hecho involucra un re diseño del mismo si la nueva clase introducida está conformada por una combinación dos o más de las cinco perturbaciones simples empleadas en este trabajo.

Capítulo 7:

Localización de Fallas

7.1 INTRODUCCION

Las exigencias de los usuarios en lo referente a la calidad y continuidad de servicio de energía eléctrica se han incrementado drásticamente en los últimos años. Esto obliga a las prestadoras a contar con herramientas que les permitan restablecer el suministro eléctrico lo más rápidamente posible. Mientras menores sean los tiempos de restauración, menores serán las pérdidas económicas que los clientes y las prestadoras deberán afrontar.

Por este motivo, el monitoreo del estado de la red eléctrica se volvió crucial para mejorar índices de servicio que permitan a las empresas proveedoras disminuir las pérdidas económicas generadas por el creciente rechazo de los clientes ante un servicio deficiente.

Los algoritmos presentados en el *capítulo 6* se especializan en detectar una desviación en los parámetros que caracterizan una forma de onda sinusoidal (Amplitud, frecuencia y fase), sin importar si se trata de una señal de tensión o corriente, y clasificar un evento de acuerdo al tiempo, los parámetros involucrados y los rangos en los que estos parámetros se desvían de los valores considerados como *normales*.

Sin embargo, una perturbación electromagnética medida en una forma de onda de tensión y/o corriente está siempre vinculada con algún fenómeno físico en algún sitio de la red eléctrica. Así, un *transitorio oscilatorio* puede estar vinculado con un arco de tensión o con la conexión de un banco de capacitores, un fenómeno tipo *muescas* se puede relacionar a la presencia de tiristores o las *fluctuaciones* muchas veces se la conexión de bombas hidráulicas en algún sitio de la red.

Una falla de baja impedancia se corresponde con la pérdida temporal de suministro eléctrico en alguna (o todas) de las fases de la red eléctrica. Esta pérdida de servicio puede deberse a la ruptura de una línea y su desconexión del sistema eléctrico producto de algún fenómeno climático, natural o alguna operación errónea por parte de los operarios de la red. La duración de la falla dependerá si la misma puede despejarse automáticamente o si se necesita la intervención humana para el restablecimiento del servicio.

Cuando una falla ocurre en algún nodo de la red eléctrica, la misma es percibida por los nodos vecinos, dependiendo de su severidad, como una perturbación del tipo *interrupción* o *hueco* presentes en las formas de onda de tensión. Dependiendo de las características de la perturbación detectada, un algoritmo de localización de fallas puede localizar fenómeno físico que provoca dicha perturbación.

De acuerdo a lo presentado en la sección 1.3.5 del *capítulo 1*, existen diversos tipos de sistemas de *localización de fallas*, estableciendo una clasificación primaria que los separa en dos grandes grupos de acuerdo al tipo de datos que procesan para luego profundizar en la clasificación según la cantidad de nodos de medición que emplean.

También se mencionó que por los costos de despliegue, en las redes de distribución, no es común contar con múltiples nodos de medición, por lo que, aunque no es una característica

excluyente, un método *single-end* sería la solución más apropiada para este tipo de red eléctrica.

Aunque no existe una precisa definición de los alcances o funcionalidades de un localizador de fallas, en la mayoría de los artículos se presentan las siguientes como las principales:

- i) Identificación del sitio de la falla: Esta funcionalidad se corresponde con la identificación precisa, con un cierto margen de error, del sitio donde la falla se produce. Este tipo de funcionalidad es muy importante para líneas de gran longitud o líneas subterráneas donde identificar la avería resulta una tarea costosa. El grado de precisión del algoritmo generalmente está vinculado a una mayor cantidad de nodos de medición.
- ii) La identificación del tipo de falla: Existe una amplia clasificación de la falla de acuerdo a sus características como, por ejemplo, fallas de alta impedancia o fallas de baja impedancia. Dentro de las fallas de baja impedancia se puede encontrar una sub clasificación de acuerdo a la cantidad de fases involucradas como, por ejemplo, *una fase a tierra, dos fases a tierra, tres fases a tierra, fase a fase* [67, 68].
- iii) La identificación de la línea o zona donde la falla se produce: Cuando la longitud de las líneas no es de gran envergadura y no se encuentran emplazadas bajo tierra, una aproximación aceptable es identificar la línea o la zona donde la falla se produce. A diferencia de la característica presentada en el inciso i) con un nodo de medición se puede tener un área de cobertura bastante grande.

En el *capítulo 1* también se mencionó que recientemente se ha adoptado la *AI* en los algoritmos destinados a identificar y localizar fallas en sistemas eléctricos de potencia [60].

En [60] se presenta un método *double-end* para la localización de fallas en un sistema de distribución donde la *WT* se emplea para la extracción de parámetros característicos, una etapa de seis SVM se utiliza para la clasificar el tipo de falla y una etapa de seis *ANN* adaptiva se desarrolla para localizar la falla.

En [61] se propone un algoritmo que utiliza un clasificador basado en SVM para la localización de fallas en una red radial con generación distribuida. Este método obtiene buenos resultados aunque, al tratarse de un método *double-end*, necesita de múltiples nodos de medición, al menos uno por cada generador distribuido.

El método propuesto en [131] utiliza un arreglo de cuatro *ANN* multicapa para determinar la localización de la falla en un sistema de distribución radial. El método necesita mediciones de las corrientes en las tres fases de cada nodo donde se conecta un generador distribuido así como en la subestación. Esto implica un sistema de comunicación y múltiples nodos de medición lo cual encarece su implementación. En el método no se muestra ninguna alternativa para optimizar la configuración de la red neuronal.

La mayoría de los métodos presentados que emplean métodos *single-end* para sistemas de distribución son implementados utilizando técnicas de *ondas viajeras* [57, 132-136].

El método conocido ondas viajeras (*traveling waves*) localiza la falla mediante la medición de ondas de alta frecuencia que se reflejan en el sitio donde se produce el evento que la genera. Estos métodos presentan dificultades cuando se aplican a sistemas eléctricos de distribución ya que estas redes suelen presentar múltiples ramas laterales que también provocan componentes reflejadas de altas frecuencias. Las ondas reflejadas producto de las ramas laterales se montan sobre los componentes que se reflejan producto de la falla lo que suele provocar alteraciones en las predicciones de los algoritmos.

A pesar que los métodos de *AI* se pueden adaptar bien a este tipo de problemas, han sido poco aplicadas en métodos *single-end* para la localización de fallas en sistemas de distribución.

En [63] un conjunto de *ANN* son entrenadas para definir diferentes niveles de fuentes de corto circuito, para luego, mediante clasificadores basados en *SVM*, estimar la localización de la falla en un sistema eléctrico de distribución. Este algoritmo emplea mediciones de tensión y corriente disponibles en una única subestación. A pesar de ello, el algoritmo necesita conocer los estados de diferentes seccionadores e interruptores de la red lo cual lo hace que no se trate de un método *single-end* puro.

En [62] se presenta un algoritmo para la detección, clasificación y localización de fallas que provocan huecos de tensión de corta duración en las mediciones de voltaje de un sistema de distribución. El método propuesto utiliza sinusoides amortiguadas para la extracción de parámetros característicos, una etapa *SVM* para la clasificación del tipo de falla y una etapa de *ANN* para la localización de la misma.

En la estadística, un parámetro estadístico, estadístico muestral o simplemente estadístico es una medida cuantitativa, derivada de un conjunto de datos de una muestra, con el objetivo de estimar o inferir características de una población o modelo estadístico.

La mayoría de las publicaciones proponen el cálculo de diferentes parámetros estadísticos para la obtención de descriptores (parámetros característicos) aunque muy pocos emplean técnicas para la selección de aquellos parámetros que mejor representan un evento en particular.

Finalmente se destaca que en pocas publicaciones se presenta una metodología para el ajuste de los parámetros que gobiernan el rendimiento de un clasificador.

7.1.1 Objetivos.

En el presente capítulo se plantea, como objetivo principal, proponer un método *single-end* que permita la detección, clasificación y localización de fallas en un sistema eléctrico de distribución, de topología radial de tamaño reducido, empleando mediciones en el *dominio del tiempo*.

Los algoritmos desarrollados en este capítulo se basan en el hecho que cuando una falla ocurre, las mediciones capturadas en el nodo alimentador presentarán diferentes patrones en las señales medidas (huecos de tensión, picos de corriente) cuyos patrones serán diferentes en función de cuán lejos se encuentra el nodo medidor del sitio donde la falla se encuentra localizada.

Para ello, técnicas de procesamiento de señales se implementan para la detección de una falla y la extracción de parámetros característicos, procedimientos de minería de datos para la selección óptima de los parámetros que caracterizan los eventos estudiados, un algoritmo de búsqueda de grilla para la selección de los parámetros que gobiernan al clasificador y algoritmos basados en *SVM* para la clasificación y localización de la fallas en un sistema eléctrico radial.

7.1.2 Metodología.

En primer lugar se presenta el modelo de un sistema eléctrico de distribución, de topología radial, de tamaño reducido. El modelo desarrollado permite la simulación de diversos tipos de fallas en diferentes secciones de las líneas o en las barras.

Utilizando este modelo se genera una base de datos conformada por diversos tipos y situaciones de falla.

El diseño, entrenamiento y prueba del algoritmo de localización de fallas se realiza empleando el marco conceptual propuesto en el *capítulo 4* y exitosamente implementado para la clasificación de perturbaciones simples y complejas tal como se presenta en el *capítulo 6* de la presente tesis.

En la primera parte del capítulo se detallan los métodos y técnicas empleadas para el diseño y entrenamiento del clasificador.

En la segunda parte del capítulo se presentan los resultados obtenidos de la prueba del algoritmo desarrollado.

Los parámetros característicos se obtienen mediante el cálculo del *perfil rms* y la obtención de la *WT* de dos niveles utilizando la *Wavelet Madre Daubechies número 4*. Para el entrenamiento y diseño del clasificador se implementa un proceso de minería de datos que permite seleccionar los parámetros característicos que mejor describen las perturbaciones en las formas de onda de tensión y corriente adquiridas post falla.

El algoritmo de análisis de falla consiste en un clasificador basado en *SVM* multiclase para la clasificación del tipo de falla y una segunda etapa, también basada en *SVM* para localizar la línea, dentro del sistema eléctrico de potencia, donde la falla se produce.

7.2 ALGORITMO PARA LOCALIZACIÓN E IDENTIFICACION DEL TIPO DE FALLA

El método para la localización de fallas propuesto en esta tesis se fundamenta en el hecho que cuando una falla ocurre, las mediciones de voltaje y corriente obtenidas en nodos estratégicos de la red presentarán características particulares según la ubicación geográfica donde la falla se produce. En este sentido, si una falla se produce en cercanías del punto de medición los datos adquiridos presentarán determinadas características que serán propias y distintivas del sitio donde el evento ocurre.

La implementación del módulo de clasificación y el módulo de localización se fundamentan en dos clasificadores SVM multiclase.

Estos algoritmos, al pertenecer al grupo de técnicas de aprendizaje lineal, deben ser diseñados y entrenados para que aprendan a relacionar patrones con distintas clases predefinidas.

En la Figura 7. 1 se muestra un diagrama esquemático del algoritmo desarrollado para el diseño, entrenamiento y prueba de las etapas que conforman el sistema de localización de fallas.

Como se puede observar, dicho algoritmo se fundamenta en las metodologías presentadas, de manera genérica, en el *capítulo 4*.

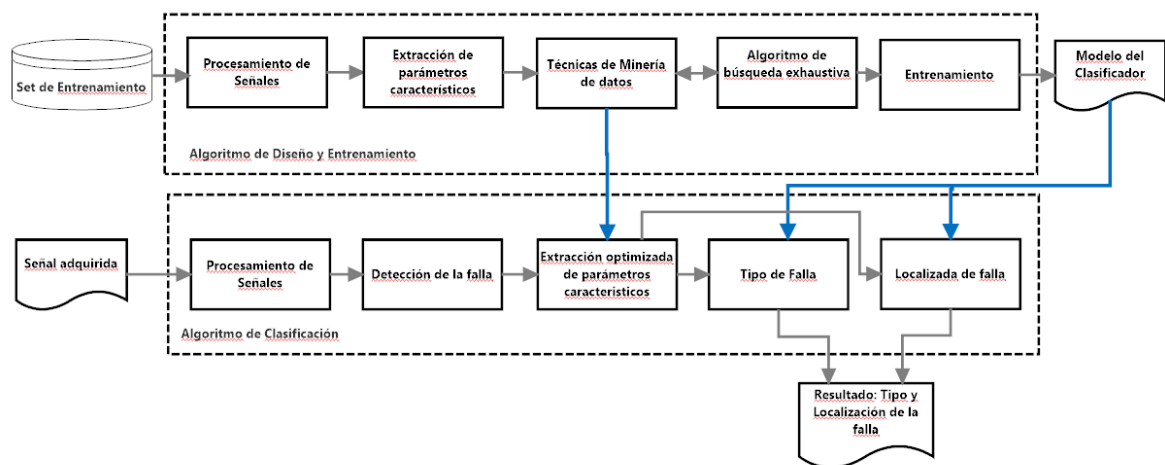


Figura 7. 1 - Diseño, Entrenamiento y Prueba del sistema de localización de fallas.

7.2.1 Modelo del Sistema Eléctrico de Potencia.

Cuando una red de gran dimensión y topología compleja necesita ser monitoreada, en general, un conjunto grande de dispositivos inteligentes de adquisición y procesamiento de datos, se distribuyen en diferentes nodos para monitorear los equipos, subestaciones, líneas o clientes.

Para redes de distribución de gran tamaño, el costo de implementar un número grande dispositivos de medición puede resultar el principal limitante para el despliegue de un sistema de localización de fallas. En este sentido se hace necesario encontrar estrategias que puedan emplear la tecnología de medición pre existente y que minimicen la cantidad de nueva tecnología a ser implementada.

En lugar de diseñar algoritmos que realicen el análisis de la totalidad de una red compleja, una estrategia alternativa consiste en dividir la red de distribución en pequeñas sub-redes, colocar estratégicamente dispositivos de medición encargados de monitorear una zona delimitada y efectuar la localización de la falla dentro de ese sub-sistema, identificando la zona o línea donde la falla se produce. En este esquema, una etapa de mayor jerarquía será la encargada decidir la línea donde la falla se produce, en caso que dos o más medidores alerten sobre la presencia de una falla dentro de su zona de cobertura.

Como primer paso de desarrollo, el presente capítulo se enfoca en el diseño de un método que sea capaz de detectar, clasificar y localizar fallas en las líneas de un sistema eléctrico de distribución de topología radial, de pequeñas dimensiones, empleando mediciones en el dominio del tiempo adquiridas en un único nodo.

El sistema eléctrico utilizado para el desarrollo y prueba del algoritmo se muestra la Figura 7. 2.

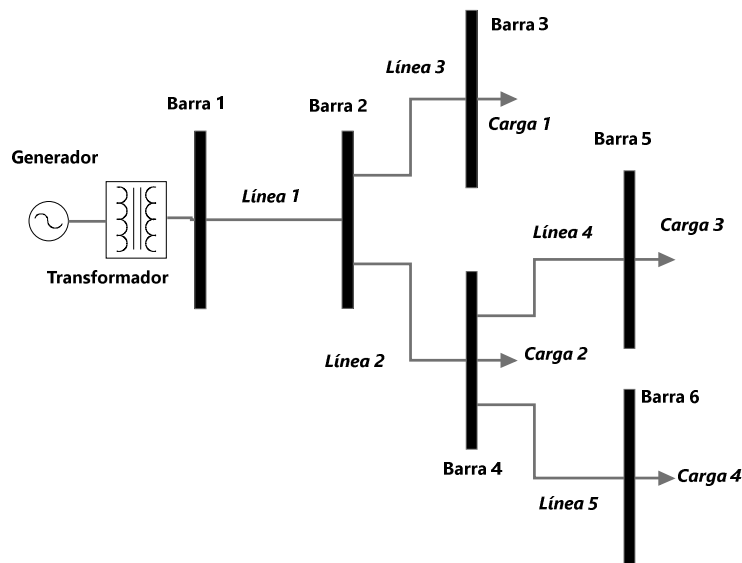


Figura 7. 2 - Sistema eléctrico de potencia de entrenamiento y prueba.

Los parámetros más importantes del sistema del sistema son:

- Cargas: Se consideró solo el caso de cargas con potencia activa.
- Frecuencia fundamental: El sistema se configura con una frecuencia fundamental de $50 [Hz]$.

- Líneas: Los parámetros distribuidos de las líneas se resumen en la Tabla 7. 1 mientras que en la Tabla 7. 2 se presentan la longitud de las líneas y los valores de las cargas para cada escenario de prueba propuesto.

Tabla 7. 1 - Parámetros de las líneas.

Línea	Longitud [Km]	R ₀ [ohm/Km]	R ₁ [ohm/Km]	L ₀ [Hy/Km]	L ₁ [Hy/Km]
Línea 1	2	0.2	0.36	0.0014	0.006
Línea 2	1	0.2	0.36	0.0014	0.006
Línea 3	5	0.2	0.36	0.0014	0.006
Línea 4	1	0.2	0.36	0.0014	0.006
Línea 5	3	0.2	0.36	0.0014	0.006

Tabla 7. 2 - Longitud de las líneas y potencia de las cargas.

	Longitud Línea [Km]					Potencia Activa [MW]			
	Línea 1	Línea 2	Línea 3	Línea 4	Línea 5	Carga 1	Carga 2	Carga 3	Carga 4
Escenario 1	2	1	5	1	3	0.1	1	2	2.5
Escenario 2	2	1	5	1	3	1	1	1	1

7.2.2 Generación del Set de Entrenamiento y Prueba.

La generación del set de entrenamiento se realiza mediante la implementación del modelo presentado en la sección 7.2.1 utilizando la toolbox *Simulink de MATLAB*.

La metodología empleada consiste en la simulación de diversas fallas en distintas líneas del sistema, para luego adquirir las formas de onda de tensión y corriente perturbadas medidas en cada fase de la *barra 1*. Como resultado, por cada falla, se estarán adquiriendo seis formas de onda.

Para generar un set de entrenamiento y prueba, que represente la mayor cantidad de casos posible, la simulación se propone respetando las siguientes condiciones:

- Tipos de falla: Para entrenar y probar las etapas de clasificación, de acuerdo a lo propuesto en [5] cuatro tipos de falla fueron simulados: *Falla Tipo A* (una fase a tierra), *falla Tipo B* (dos fases a tierra), *falla tipo C* (fase a fase), *falla tipo D* (tres fases a tierra). La distribución se realizó respetando lo propuesto en la Tabla 7. 3.
- Localización de fallas. La Tabla 7. 4 presenta la distribución de las fallas en el sistema eléctrico mostrado en la Figura 7. 2. Cada celda de indica el número total de fallas simuladas para determinada localización de una línea dada, que incluye fallas Tipo A, B, C y D.

Por ejemplo, para la Línea 1, 50 fallas se simularon a un 20% de la longitud de la línea (cuyos parámetros se definen en la Tabla 7. 1). Estas 50 fallas pueden ser del tipo A, B, C y D.

- Tiempo de inicio de falla: El tiempo de inicio de la falla, también conocido como ángulo de inepción, se simula de manera aleatoria.

- Resistencia de falla: Para limitar la cantidad de señales se simula una resistencia de falla muy pequeña, tendiente a cero.

Tabla 7. 3 - Tipos de falla.

Tipo de Falla	Detalle	Cantidad	Total
Tipo A	Fase A a tierra	100	300
	Fase B a tierra	100	
	Fase C a tierra	100	
Tipo B	Fases A y B a tierra	100	300
	Fases A y C a tierra	100	
	Fases B y C a tierra	100	
Tipo C	Fase A con Fase B	100	300
	Fase A con Fase C	100	
	Fase B con Fase C	100	
Tipo D	Fases A,B y C a tierra	100	100
Total			1000

Tabla 7. 4 - Localización de la falla.

Línea	1	2	3	4	5	Total
Localización						
20%	50	50	50	50	50	250
40%	50	50	50	50	50	250
60%	50	50	50	50	50	250
80%	50	50	50	50	50	250
Total	200	200	200	200	200	1000

El set de entrenamiento, diseñado para el entrenamiento y prueba del algoritmo, propuesto se conforma simulando 1000 eventos diferentes.

Debido a que seis formas de onda se emplean para representar cada evento (Una señal de tensión y una de corriente por cada fase), 6000 señales, en total, conforman el set de entrenamiento.

7.2.3 Diseño y Entrenamiento del Clasificador.

En esta sección se presentarán las técnicas utilizadas, métodos implementados y los resultados obtenidos en cada etapa del proceso mostrado en la Figura 7. 1.

7.2.3.1 Procesamiento de Señales.

Las tareas llevadas a cabo por este proceso se dividen en el pre procesamiento y procesamiento de las formas de onda adquiridas.

- Pre procesamiento de señales: La longitud de las señales de tensión y corriente empleadas como entrada a las restantes etapas del algoritmo será reducida en esta etapa. Para las experiencias propuestas las formas de onda son muestreadas con una frecuencia de 10 [Kilo muestras/seg]. La longitud de la ventana de tiempo es de $0,04$ [seg], medio ciclo de una señal de frecuencia de 50 [Hz], por lo tanto, el algoritmo deberá ser capaz clasificar y localizar la falla aun cuando la misma no haya sido despejada. Teniendo en cuenta el tiempo de muestreo y la longitud de la ventana, cada forma de onda queda representada por 400 muestras.

Las mediciones obtenidas, como ya se mencionó, se corresponden con los valores de tensión y corriente de cada fase medidas en un único nodo, en este caso la barra 1. En la Figura 7. 3 se muestran las mediciones de tensión y corriente realizadas en las tres fases de la barra 1. Las perturbaciones observadas en las mediciones son efecto de una falla del tipo *Tipo B* ocurrida en la línea 3. Por lo tanto, cada falla estará definida por seis formas de onda de 400 muestras cada una (tensión en las fases A, B y C y corriente en las fases A, B y C).

Como resultado del pre procesamiento el set de entrenamiento, en esta etapa del algoritmo se corresponde con una matriz dimensión $[1000,2400]$ donde cada fila se corresponden con las formas de onda de tensión y corriente medidas en el único nodo de medición y donde las columnas se corresponden con las muestras obtenidas para cada una de las señales.

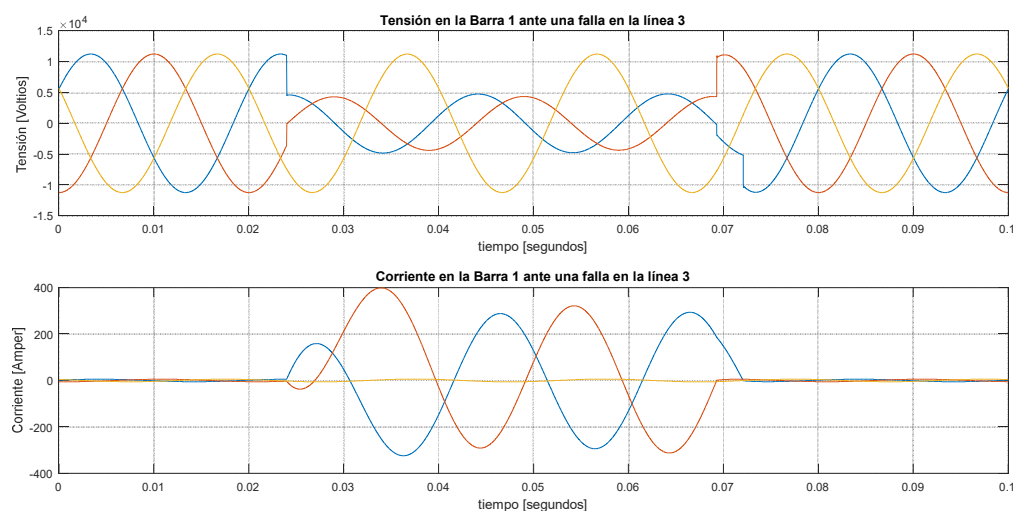


Figura 7. 3 - Falla en línea 3 - Medición en barra 1.

- **Procesamiento de Señales:** Tal como se explicó en la sección 4.2.1 cuando se aplican técnicas de procesamiento de señales se produce una transformación del set de entrenamiento en un conjunto de datos que, por sus características, hace más fácil la tarea de detectar la presencia de una perturbación y la extracción de parámetros característicos. En este trabajo, las técnicas de procesamiento de señales que se aplican son:
 - **Perfil cuadrático medio (rms, por sus siglas en inglés):** El cálculo del perfil rms es una herramienta sumamente útil que facilita la tarea de detección de una perturbación (analizando su derivada) y la extracción de parámetros que resultan de extrema importancia a la hora de describir una perturbación.
 - **Transformada Wavelet:** En esta etapa se realiza el cálculo de la *WT* de 2 niveles, empleando la wavelet madre db-4.

Los procedimientos detallados en la sección 6.2.3.1 para el cálculo del *perfil rms* y la *WT* son idénticos a los aquí adoptados.

7.2.3.2 Extracción de parámetros característicos.

Los parámetros característicos calculados se muestran en la Tabla 7. 5. Este proceso representa una primera etapa de reducción de la dimensión de los datos a ser procesados por las siguientes etapas del algoritmo.

El set de entrenamiento, originalmente representado por una matriz de dimensión $[1000,2400]$ es reducida a una matriz de dimensión $[1000,54]$. A pesar que el procesamiento se realiza sin evaluar a que clase pertenece cada parámetro, se logra reducir la dimensión del set de entrenamiento en un factor de 99,97%.

Tabla 7. 5 - Parámetros característicos.

Técnica de Procesamiento de Señales	Parámetros característicos	Nomenclatura	Fase	Tensión o Corriente
Caracterización de las formas de onda	Valor medio de la Tensión	mediaTension	A,B,C	Tensión
	Valor medio Corriente	mediaCorriente	A,B,C	Corriente
	Valor máximo Corriente	maxCorriente	A,B,C	Corriente
Perfil rms	Valor mínimo del perfil RMS.	minRMS	A,B,C	Ambos
	Valor máximo del perfil RMS.	maxRMS	A,B,C	Ambos
	Profundidad Hueco Tensión	profundidadHueco	A,B,C	Tensión
Transformada Wavelet	Valor máximo del módulo del coeficiente d_1 .	Max mod d1	A,B,C	Ambos
	Desviación estándar del coeficiente d_1 .	DSt d1	A,B,C	Ambos
	Energía normalizada del coeficiente d_1 .	Energía d1	A,B,C	Ambos
	Valor máximo del módulo del coeficiente A.	Max mod A	A,B,C	Ambos
	Desviación estándar del coeficiente A.	DSt 9	A,B,C	Ambos
	Energía normalizada del coeficiente A.	Energía A	A,B,C	Ambos

7.2.3.3 Minería de Datos.

Como se muestra en la Figura 7. 1, el sistema de análisis de fallas está conformado por dos clasificadores, uno encargado de clasificar el tipo de falla y el segundo que indicará la línea donde la falla se presenta.

En el modelo conceptual presentado en el *capítulo 4* se explicó que esta etapa está conformada por tres sub procesos: Generación de etiquetas, filtrado heurístico y algoritmo de búsqueda exhaustiva.

En esta sub sección se presenta los resultados obtenidos para cada etapa de los dos clasificadores.

- Generación de etiquetas: Para obtener resultados, los procesos de minería de datos necesitan incluir en el análisis de los datos la variable *clase*. Para ello, en una primera instancia, se genera un vector de etiquetas para indicar al algoritmo a que clase pertenece cada entrada del set de entrenamiento. En el caso del clasificador encargado de identificar el tipo de falla, el vector etiqueta asumirá el valor *1* para las fallas de *tipo A*, *2* para fallas *tipo B*, *3* para las fallas *tipo C* y *4* para las fallas *tipo D*.

De la misma manera, el vector de etiquetas que representa la localización de la fallas asume el valor *1* cuando las fallas se presenten en algún lugar de la *línea 1*, el valor

2 para una falla en la *línea 2*, 3 para fallas en la *línea 3*, 4 para fallas en la *línea 4* y, por último la etiqueta 5 para una falla en la *línea 5*.

- Filtrado Heurístico: El filtrado heurístico se realiza de acuerdo a lo explicado en 4.2.3.2. Los resultados de este proceso se presentan en la Tabla 7. 6, tanto para el clasificador del tipo de falla como para el algoritmo encargado de su localización.

Tabla 7. 6 - Resultado de filtrado heurístico.

Clasificador	Parámetros característicos seleccionado	Dimensión
Tipo	(1) maxRMSTensión, (2) minRMSTensión, (3) DStATensión, (4) Energía A Tensión, (5) maxRMSCorriente, (6) maxCorriente, (7)Dst A Corriente, (8) Energía A Corriente.	1000 x 8
Localización	(1) maxRMSTensión, (2) minRMSTensión, (3) Energía A Tensión, (4) minRMSCorriente, (5) mediaCorriente, (6) maxCorriente, (7) Dst A Corriente..	1000 x 7

- Algoritmo de Búsqueda exhaustiva: Como se explicó detalladamente en la sección 4.2.3.3 el método de búsqueda exhaustiva consiste en generar un set de entrenamiento y validación para cada combinación de parámetros característicos. Luego, estos sets de entrenamiento, se emplean para entrenar y validar el rendimiento del clasificador y seleccionar aquella combinación que mejores resultados presentan. Para que los resultados sean consistentes con el rendimiento óptimo del sistema de clasificación, por cada combinación de parámetros característicos se invoca el proceso *Optimización de los parámetros del clasificador* presentado en la siguiente sección (Ver diagrama de flujos presentado en la Figura 4. 5).

Los porcentajes de aciertos obtenidos, para todas las posibles combinaciones de parámetros característicos se muestran en las Figura 7. 4 y Figura 7. 5. También se presentan, en las Tabla 7. 7 y la Tabla 7. 8.

Como se puede ver en la Figura 7. 4, para el clasificador destinado a identificar el tipo de fallas, de las $2^8=256$ posibles combinaciones, solamente la combinación número 140, correspondiente a los parámetros característicos: 1) maxRMSTensión, 5) maxRMSCorriente y 6) maxCorriente alcanza el un máximo correspondiente al 100% de efectividad.

Tabla 7. 7 - Resultado del algoritmo de búsqueda exhaustiva - Tipo de falla.

#	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	Métrica	Precisión
140	1	0	0	0	1	1	0	0	3	100

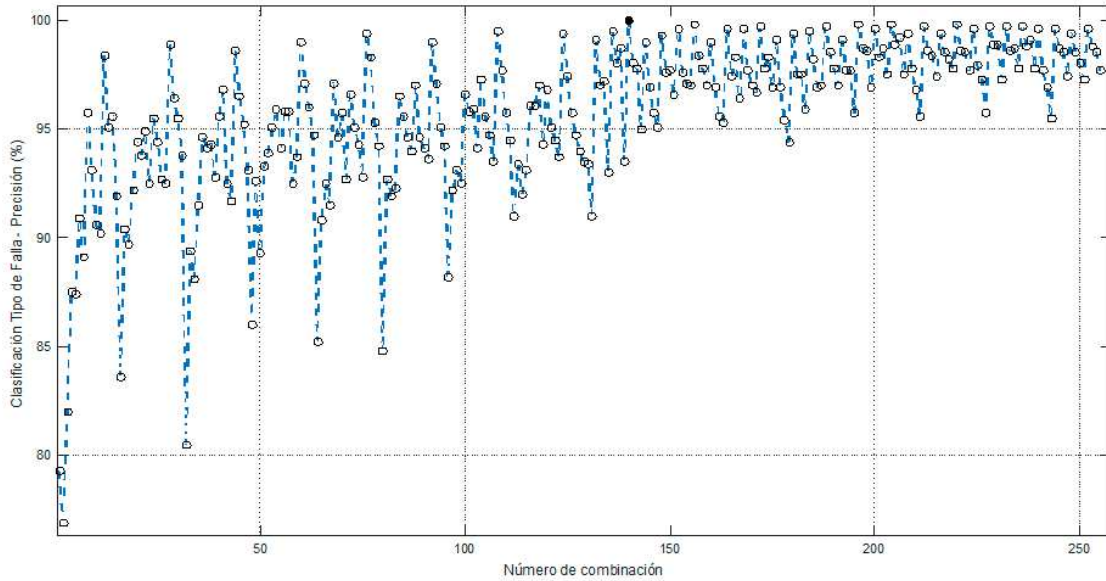


Figura 7. 4 - Resultado del algoritmo de búsqueda exhaustiva – Tipo de falla.

En la Figura 7. 5 se observa que de las $2^7=128$ posibles combinaciones, cinco de ellas alcanzan el mismo valor máximo de 98,8% de precisión. En este caso para entrenar el clasificador SVM que se encarga de la localización de fallas se selecciona aquella combinación que emplea la menor cantidad de parámetros característicos. En la Tabla 7. 8 se muestra remarcada la combinación seleccionada.

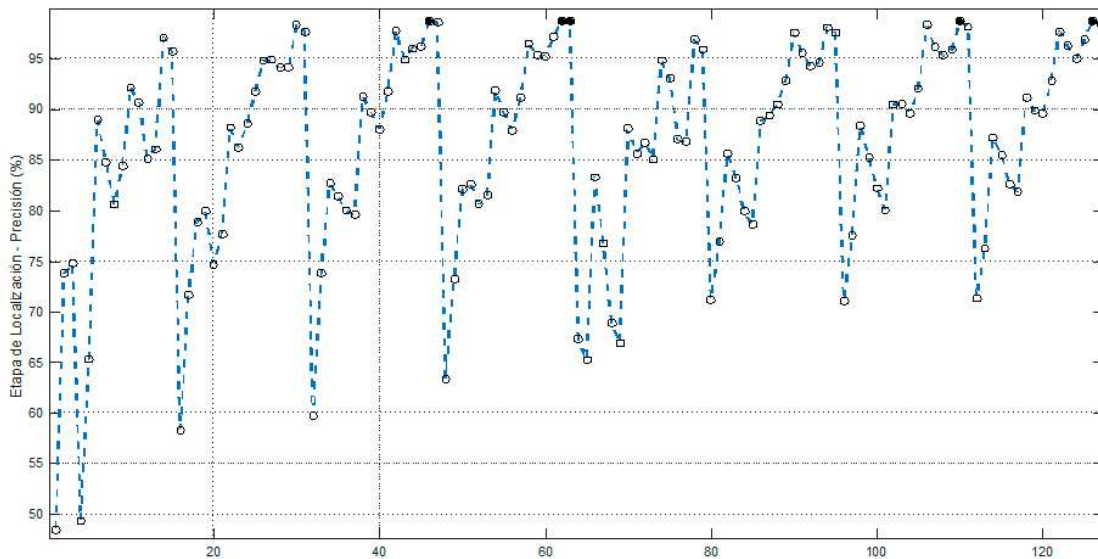


Figura 7. 5 - Resultado del algoritmo de búsqueda exhaustiva – Localización de falla.

Tabla 7. 8 - Resultado del algoritmo de búsqueda exhaustivo - Localización de falla.

#	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Métrica	Precisión
46	0	1	0	1	1	1	0	4	98.8
62	0	1	1	1	1	1	0	5	98.8
63	0	1	1	1	1	1	1	6	98.8
110	1	1	0	1	1	1	0	5	98.8
126	1	1	1	1	1	1	0	6	98.8

Un análisis de los resultados indica que el set de entrenamiento de dimensión $[1000,54]$ se reduce a una matriz de $[1000,9]$ (donde la dimensión 9 se corresponde a $maxRMSTensión$, $maxRMSCorriente$ y $maxCorriente$ para las fases A, B y C) para el clasificador destinado a identificar el tipo de falla, y una matriz de dimensión $[1000,12]$ (donde la dimensión 12 se obtiene de $minRMSTensión$, $minRMSCorriente$, $mediaCorriente$ y $maxCorriente$ para las fases A,B y C) para el clasificador encargado de la localización de falla.

7.2.3.4 Optimización de los parámetros del clasificador.

Los clasificadores empleados esta tesis son configurados con una función kernel *RBF*.

La función RBF se describe matemáticamente en la ecuación 6.1 presentada en el capítulo anterior.

Como se puede apreciar, la función RBF depende únicamente de la variable σ .

No existe una ecuación matemática para el cálculo de los valores óptimos de C y σ , por lo tanto no es posible conocer a priori los valores óptimos para la resolución de un determinado problema.

Para ello se implementa un algoritmo de búsqueda de grilla, explicado en la sección 4.2.4 cuyos resultados se muestran en la Tabla 7. 9.

Los resultados presentados en esta sección se corresponden a los valores óptimos de C y σ obtenidos para las combinaciones de los parámetros característicos presentadas en la sección 7.2.3.3 cuyo porcentaje de validación es máximo.

Tabla 7. 9 - Resultados del algoritmo de búsqueda tipo grilla.

Clasificador SVM	# Combinación parámetros característicos.	Sigma	C	Precisión de validación (%)
Tipo de Falla	140	0.5	30	100
	46	0.5	25	98.8
Localización	62	0.1	500	98.8
	63	0.1	55	98.8
	110	0.5	75	98.8
	126	0.1	60	98.8

Los clasificadores son entrenados con los parámetros obtenidos en las secciones 7.2.3.3 y 7.2.3.4.

7.2.4 Clasificador.

El proceso de clasificación, de acuerdo a lo presentado en la Figura 7. 1, está compuesto por una etapa de procesamiento de señales, seguido por un algoritmo de detección de falla, un proceso de extracción de parámetros característicos y, finalmente, la etapa clasificación. A continuación se presentan las técnicas implementadas y resultados de obtenidos en cada sub proceso.

7.2.4.1 Procesamiento de Señales.

Las técnicas de pre procesamiento y procesamiento de señales implementadas en esta etapa guardan correspondencia con las empleadas para el entrenamiento y diseño del clasificador las cuales fueron presentadas en la sección 7.2.3.1.

7.2.4.2 Detección de Fallas.

Las mediciones de tensión de las fases involucradas en una falla presentan una notoria caída en el valor rms. En la Figura 7. 6 se puede ver un hueco de tensión en una forma de onda medida en la línea 1 producto de una falla en la línea 5 y se distingue claramente una caída en el perfil rms de tensión en cada fase involucrada en la falla. Por este motivo, una metodología para detectar una falla es la de monitorear el *perfil rms* de tensión de cada fase con el objetivo de detectar una caída por debajo de un umbral determinado.

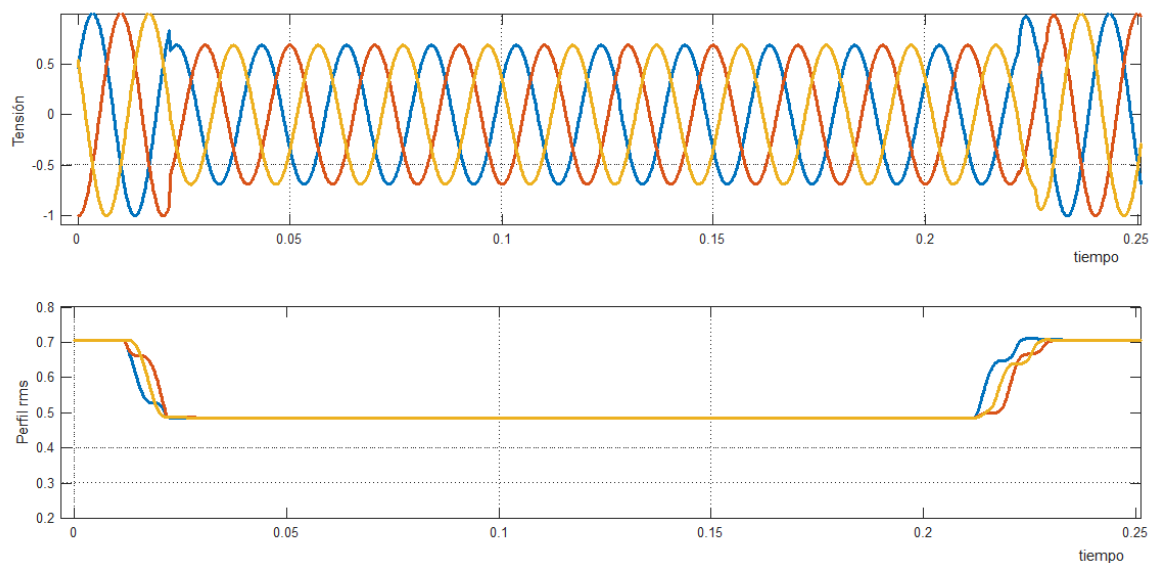


Figura 7. 6 - Perfil rms de tensión ante una falla.

7.2.4.3 Clasificación.

Para respetar concordancia con las formas de onda empleadas para el entrenamiento de los clasificadores, los parámetros que modelan la falla (localización, tipo, resistencia y duración de falla) se simularon de acuerdo a la metodología presentada la sección 7.2.2.

Los parámetros del sistema se simularon teniendo en cuenta los siguientes escenarios:

- Escenario 1: Cargas distintas, líneas distintas.
- Escenario 2: Cargas distintas, líneas iguales.

En la Tabla 7. 2 se muestra la longitud de las líneas y la potencia activa de las cargas para cada uno de los escenarios propuestos.

Para cada escenario se generan 1000 fallas, por lo que, el set de prueba estará conformado con una distribución idéntica a la presentada en la sección 7.2.2 para el set de entrenamiento.

Cada set de prueba se emplea como entrada del proceso de clasificación presentado en la Figura 7. 1.

Para ambos escenarios el sistema de localización de fallas fue capaz de identificar el tipo de falla e indicar correctamente la línea donde la falla se produce en el 100% de los casos.

La importancia de estos resultados radica en el hecho que los sistemas eléctricos en donde las cargas y las longitudes de las líneas difieren entre sí son los que más comúnmente pueden presentarse en un sistema eléctrico real. Para una configuración de este tipo se concluye que el método propuesto presenta resultados destacados.

7.2.4.4 Limitaciones observadas.

Un problema muy difícil de resolver, por cualquier metodología, se presenta cuando se tienen dos ramas con exactamente la misma longitud y las mismas condiciones de carga. Simulaciones bajo estas condiciones permitieron corroborar que el método probó ser capaz de identificar el tipo de falla para el 100% de los casos propuestos pero solamente en un 78,2 % de los casos el sistema acertó a la hora de identificar la línea donde la falla se produce.

Realizando un análisis de las clasificaciones erróneas, para este escenario, se encontró que las confusiones se presentan en posiciones cercanas a una barra.

Sin embargo, es conveniente aclarar, que es prácticamente imposible que un sistema real esté conformado por líneas cuya longitud sean iguales y las condiciones de carga, en todas las barras, sean idénticas.

En la práctica, en caso de presentarse esta situación, el problema podría resolverse agregando un medidor en el nodo donde estas líneas gemelas se bifurcan proporcionando, de este modo, mayor información al algoritmo de localización de fallas.

7.3 CONCLUSIONES

El presente capítulo propone un método *single-end* que utiliza mediciones en el dominio del tiempo para la detección, clasificación y localización de fallas en un sistema eléctrico de distribución de topología radial.

El método propuesto emplea el cálculo del *perfil rms* y la *WT* como técnicas de procesamiento de señales para la extracción de parámetros característicos. Luego, un módulo compuesto por dos *SVM* multiclase se encargan de la clasificación del tipo de falla y de su localización.

El método planteado se vale de la información inherente que acarrean las perturbaciones sufridas por las formas de onda de tensión y corriente en un único nodo de medición dentro del sistema eléctrico de potencia. Los parámetros característicos extraídos de estas formas de onda serán empleados por algoritmo para identificar el tipo de falla y localizar la línea donde la misma se produce.

Debido a que la habilidad del clasificador depende en gran medida de la correcta selección de los parámetros característicos que describen a las formas de onda de tensión y corriente adquiridas se propuso la incorporación de técnicas de minería de datos a la etapa de diseño de los clasificadores para la selección de los parámetros característicos que describen mejor al tipo y localización de las fallas.

Esta incorporación representa una clara ventaja dado que la mayoría de los trabajos que utilizan técnicas de *AI* para la localización de fallas no profundizan en la selección de parámetros característicos.

Una técnica de búsqueda tipo grilla se implementó para encontrar la combinación de los parámetros *C* y *sigma* que optimizan el rendimiento de ambos clasificadores.

El algoritmo de localización se entrenó y probó para dos escenarios diferentes: El primero supone que la longitud de las líneas y la potencia activa de las cargas son diferentes entre sí. El segundo, un poco más exigente supone que las líneas son todas de igual longitud aunque las cargas son diferentes.

Para ambos escenarios, el algoritmo desarrollado fue capaz de identificar el tipo de falla y localizar la línea donde la misma se produce para el 100% de los casos de prueba.

Un escenario de poco probable ocurrencia en la vida real se utilizó con el objetivo de medir las limitaciones del método propuesto. En este escenario la longitud de las líneas y las condiciones de carga se configuraron con idénticos parámetros. Para este caso el método desarrollado permitió clasificar el 100% del tipo de falla pero su rendimiento cae a un 78,2% a la hora de identificar la línea donde la falla se produce. Estos resultados son esperables ya que resulta muy difícil, para cualquier metodología, identificar dos situaciones que son prácticamente idénticas. Sin embargo, el análisis de las clasificaciones erróneas, para este escenario, permitió concluir que en todos los casos el clasificador confunde la línea por otra geográficamente cercana. En ningún caso se obtiene una clasificación con errores gruesos.

Los resultados obtenidos son aún más valiosos si se tiene en cuenta que las señales utilizadas para el entrenamiento y prueba del algoritmo se obtuvieron midiendo solamente medio ciclo de las señales de tensión y corriente en un solo nodo del sistema propuesto. Esta es una característica fundamental del algoritmo ya que, en la práctica, solo se necesitaría instalar un único sistema de medición para obtener, con marcada rapidez, resultados en cualquier línea del sistema presentado en este estudio.

Finalmente, una importante conclusión se obtiene del proceso de minería de datos incluido en el entrenamiento de los clasificadores. Estos resultados indican que solamente calculando tres parámetros característicos para el clasificador destinado a identificar el tipo de falla (*Valor máximo del perfil rms de tensión, valor máximo del perfil rms de corriente y el valor máximo de la corriente medida*) se obtiene un porcentaje de acierto del 100%. Un análisis análogo permite concluir que, para localizar la falla, se necesita calcular solamente cuatro parámetros característicos (*Valor mínimo del perfil rms de tensión, valor mínimo del perfil rms de corriente, valor medio de la señal de corriente y el valor máximo de la señal de corriente*) el clasificador destinado a identificar la línea donde se produce la falla obtiene los porcentajes de acierto ya mencionados para cada uno de los escenarios propuestos.

Una desventaja del algoritmo propuesto radica en el hecho que, si la topología de la red se modifica, o sus los valores de las carga se ven alterados considerablemente, el módulo de localización debe ser re entrenado para tener en cuenta estos cambios en el estado de la red.

Capítulo 8:

Conclusiones, Aportes y Trabajos Futuros

8.1. RESUMEN DEL CONTENIDO DE LOS CAPITULOS DE LA TESIS.

A lo largo de la presente tesis se ha destacado la relevancia que tiene un sistema monitoreo de eventos de calidad de energía para la operación y mantenimiento de un sistema eléctrico de potencia.

En el *capítulo 1* se presenta al software de análisis de señales perturbadas como una herramienta fundamental para aliviar las pérdidas económicas producto de los eventos de calidad de energía tanto para las empresas proveedoras de energía eléctrica como así también para sus clientes.

Con el fin de tipificar las perturbaciones que se presentan en un sistema eléctrico de potencia, de acuerdo a las variaciones relativas de los parámetros que las gobiernan, en esta tesis se adopta el estándar *IEEE Std 1159-2009*. Este estándar fue también empleado como un marco de trabajo para el desarrollo de los distintos componentes que conforman el sistema de monitoreo propuesto en este trabajo de tesis.

En el *capítulo 1* también se presenta el modelo esquemático de un sistema de análisis de perturbaciones que será el utilizado como referencia para el desarrollo de las herramientas de clasificación de perturbaciones y localización de fallas propuestas en de manera genérica en el *capítulo 4*, y aplicada a para resolver problemáticas particulares tal como propone en el *capítulo 6* y *capítulo 7*.

Finalmente, se presentó una detallada revisión bibliográfica sobre las técnicas propuestas por la comunidad científica para el *procesamiento de señales*, la *detección de perturbaciones*, la *extracción de parámetros característicos* y las técnicas de *AI* propuestas para la *clasificación de perturbaciones* y la *localización de fallas* en un sistema eléctrico de potencia.

En *capítulo 2* se abordaron los conceptos básicos de la *Transformada de Fourier (FT)* y la *Transformada de Fourier dependiente del tiempo (STFT)* para el análisis de señales estacionarias y no estacionarias demostrando las limitaciones y desventajas que ambos métodos presentan frente a la *WT*. En este capítulo se presentaron también algunas nociones teóricas y analíticas relativas a la *Transformada Wavelet continua (CWT)*, la *Transformada Wavelet Discreta (DWT)* y una útil interpretación de la *CWT* por un banco de filtros. En la última sección del *capítulo 2* se citaron los trabajos más destacados que adoptan la *Transformada Wavelet* y sus derivadas para la detección de perturbaciones, la extracción de parámetros característicos, la compresión de datos y la mitigación del efecto del ruido en las mediciones.

Por las sólidas bases matemáticas en las que esta técnica está fundada, la habilidad de sus algoritmos de entrenamiento para evitar el atascamiento en un mínimo local, su menor propensión a el sobre-entrenamiento y su menor costo computacional las *SVM* han representado uno de los métodos más adoptados por la comunidad científica presentando resultados equivalentes o superiores frente a otras técnicas de clasificación. En el *capítulo 3* se presentaron los conceptos teóricos y matemáticos involucrados en el uso y entrenamiento

de un clasificador basado en *SVM*. En este capítulo también se abordaron varias estrategias para transformar un clasificador *SVM* binario en un clasificador multiclase.

Por último, en el *capítulo 3* se realizó un recorrido de las publicaciones más relevantes que utilizan *SVM* como herramienta para la clasificación de perturbaciones y localización de fallas.

Una vez presentados los aspectos teóricos y las publicaciones más importantes de los métodos propuestos, en la segunda parte de esta tesis, se presentó el desarrollo de una metodología para el entrenamiento de un *ES* que permite la clasificación de perturbaciones simples y complejas así como la localización de fallas en un sistema eléctrico de potencia.

En el *capítulo 4* de la presente tesis se mostró una metodología genérica para el diseño, entrenamiento y prueba de un sistema de clasificación basado *Inteligencia Artificial*, enfocado en aquellos métodos que utilizan técnicas de aprendizaje supervisado. La metodología presentada en este capítulo sirvió de marco teórico para el diseño, entrenamiento y prueba de un algoritmo de clasificación de perturbaciones eléctricas presentado en el *capítulo 6*. Esta metodología fue también empleada, en el *capítulo 7*, para el desarrollo de un sistema de clasificación y localización de fallas en un sistema eléctrico de potencia.

En el *capítulo 5* se exhibió una interfaz gráfica desarrollada con el objetivo de permitir, de manera visual e intuitiva, la configuración, entrenamiento y prueba de dos clasificadores, uno basado en una red neuronal backpropagation y un segundo basado en 9 etapas *SVM* binarias organizadas en una arquitectura *One vs Rest*.

La optimización de los clasificadores es una tarea que está estrechamente vinculada a cada problemática en particular. La configuración óptima de un algoritmo destinado a clasificar fallas puede no resultar ser la mejor alternativa para un clasificador destinado a identificar el tipo de falla en un sistema eléctrico.

Los *capítulos 6* y *7* muestran una metodología para optimizar el rendimiento del clasificador de acuerdo a las problemáticas que estos están destinados a resolver.

Por lo tanto, valiéndose del modelo propuesto en el *capítulo 4*, en el *capítulo 6* se presentó la metodología para el diseño y prueba de un clasificador destinado a detectar y clasificar perturbaciones simples y perturbaciones complejas.

En el bloque de procesamiento de señales, se implementa la *WT* de nueve niveles empleando la *Db-4* como wavelet madre.

Diferentes estadísticos se aplicaron a los coeficientes de la *WT* para la extracción de parámetros característicos.

Técnicas de filtrado heurístico y un algoritmo de búsqueda exhaustiva se implementaron en el módulo responsable de la selección de los parámetros característicos que describen, de manera óptima, cada clase presente en el set de entrenamiento y prueba.

Se desarrolló un clasificador multiclase con una arquitectura *One vs Rest* que permitió la clasificación de perturbaciones simples y complejas.

Un algoritmo de búsqueda exhaustiva, en combinación con técnicas de validación cruzada, fueron desarrolladas para sintonizar, de manera óptima, cada clasificado garantizando un equilibrio entre el porcentaje máximo de acierto y su capacidad de generalización.

Utilizando una herramienta desarrollada por el autor [129] se generó una base de datos que permitió el entrenamiento y prueba de los métodos desarrollados.

El método probó ser capaz de clasificar perturbaciones simples y complejas con una altísima eficiencia. Los porcentajes de acierto obtenidos superan a la mayoría de los trabajos publicados en el área.

Utilizando el mismo marco conceptual presentado en el *capítulo 4*, en el *capítulo 7* se propuso un método que permitió la detección, clasificación y localización de fallas en un sistema eléctrico de distribución de topología radial.

Para el entrenamiento y prueba del módulo de localización de fallas se desarrolló el modelo de una red eléctrica de distribución, de topología radial. Este modelo permitió incorporar fallas en diferentes líneas y tomar mediciones de su impacto en diferentes nodos de la red.

En el módulo de procesamiento de señales se implementó el cálculo del *perfil rms* y la *WT* de dos niveles empleando la wavelet madre Db-4.

En el bloque de extracción de parámetros característicos se procedió al cálculo de varios estadísticos sobre los coeficientes de la *WT* y el *perfil rms*.

Con varios tipos de falla, localizadas en distintas líneas se realizó el entrenamiento de un módulo compuesto por dos etapas *SVM* multiclase que se encargaron tanto de la clasificación del tipo de falla como su localización.

Se desarrollaron técnicas de minería de datos en la etapa de entrenamiento de los clasificadores para la selección de los parámetros característicos que mejor permiten identificar el tipo de falla y facilitan su localización geográfica.

Una técnica de búsqueda tipo grilla se presentó con el objetivo de encontrar la combinación óptima de los parámetros *C* y *sigma* que maximizaron el rendimiento de ambos clasificadores.

El algoritmo de localización se entrenó y probó para diferentes escenarios probando ser un 100% efectivo tanto para clasificar, como para localizar la línea donde la falla se encuentra.

Los resultados obtenidos son aún más valiosos si se tiene en cuenta que las señales utilizadas para el entrenamiento y prueba del algoritmo se obtuvieron midiendo solamente medio de las señales de tensión y corriente en un solo nodo del sistema propuesto. Esta es una característica fundamental del algoritmo ya que, en la práctica, solo se necesitaría instalar

un único sistema de medición para obtener, con marcada rapidez, resultados en cualquier línea del sistema radial, de pequeñas dimensiones, tal como el presentado en este estudio.

8.2. RESUMEN DE LAS MOTIVACIONES.

El rendimiento de las herramientas para la clasificación basadas en técnicas de *AI* depende tanto de su correcta configuración como de la selección de los parámetros característicos empleados para la su entrenamiento [11].

A pesar de su importancia, la mayoría de los autores proponen el cálculo de una cantidad determinada de parámetros característicos sin profundizar en el estudio de técnicas para la selección de los descriptores que mejor se adaptan a cada evento que se pretende clasificar.

En la mayoría de las publicaciones se presenta un clasificador multiclase diseñado en una única unidad de procesamiento. Esta arquitectura presenta una serie de desventajas que se listan a continuación:

- La complejidad de problema de optimización empleado para el entrenamiento del clasificador se incrementa mientras aumenta la cantidad clases diferentes involucradas en el entrenamiento.
- El clasificador deberá ser entrenado con descriptores que sirvan para caracterizar todas las perturbaciones presentes en el set de entrenamiento. Esto puede afectar el rendimiento del clasificador para identificar ciertas clases ya que un parámetro característico puede resultar adecuado para describir un clase pero completamente irrelevante para describir una segunda clase.
- Un clasificador multiclase, implementado en una única unidad de procesamiento, permitirá el ajuste solamente de dos parámetros cuya combinación óptima definirá el rendimiento del clasificador.
- Cada perturbación compleja, aunque se trate de una combinación de dos clases individuales, debe ser incorporada como una clase adicional, complejizando aún más el entrenamiento del clasificador.

Finalmente, desde el punto de vista práctico, son muy pocos los autores que se enfocaron en el estudio de perturbaciones complejas, a pesar que las mismas se presenten con regularidad en la vida real.

En lo relativo a la localización de fallas, es importante destacar que la mayoría de los autores se han enfocado en el estudio de la problemática aplicada a sistemas eléctricos de transmisión. Las redes de distribución son completamente diferentes en cuanto a la topología, dimensiones, eventos que generan la falla, presupuesto y tecnología para la adquisición de mediciones. Por lo tanto, es de esperar que los desarrollos enfocados en la localización de fallas en una red de transmisión sean difícilmente aplicables a una red de distribución.

Enfocándonos en las publicaciones que aborda la localización de fallas en sistemas de distribución es importante apreciar que, sin importar las técnicas empleadas, muchas de ellas emplean métodos *double-end*. La necesidad de contar con múltiples nodos de medición

implica mayor cantidad de dispositivos de adquisición de datos, el despliegue de un sistema de sincronización de las muestras y una tecnología de comunicaciones que permita la centralización de la información colectada. Todo esto deviene en mayores necesidades de inversión para el despliegue de un sistema de localización de fallas.

En este sentido es deseable que un algoritmo de localización de fallas pueda alcanzar resultados apropiados con la menor cantidad de nodos de medición posible.

Por último, es importante destacar que son muy escasos los artículos que, empleando técnicas de inteligencia artificial, adoptan procedimientos de minería de datos para seleccionar el mejor conjunto de parámetros característicos que describen los eventos a clasificar. Son muy escasas también, las publicaciones donde se presente una metodología para el ajuste de los parámetros que sintonizan el clasificador.

8.3. APORTES REALIZADOS.

8.3.1. Desarrollo de un software de monitoreo

En el *capítulo 5* se presentó el desarrollo de una herramienta gráfica que permite la detección, clasificación y localización de perturbaciones presentes en una señal. Esta herramienta fue desarrollada para permitir:

- La generación perturbaciones sintética: Este módulo probó ser útil para generar una gran variedad de perturbaciones simples y complejas, las cuales fueron destinadas a la confección de conjuntos de señales de entrenamiento y prueba para ser usados en la sintonización del sistema de clasificación.
- Extracción de parámetros característicos de las perturbaciones y su localización temporal: El algoritmo desarrollado está basado en la *WT* que utiliza los coeficientes de alta frecuencia para la detección de las perturbaciones y el cálculo del tiempo de inicio y duración de cada evento de calidad de energía.
- Generación de los vectores característicos: La interfaz gráfica desarrollada permite, mediante la *WT* y el cálculo del perfil rms, la extracción de una serie de parámetros característicos. Los parámetros característicos que se desean calcular pueden seleccionarse a través de una interfaz gráfica desarrollada a tal efecto.
- Clasificación de las perturbaciones: Se desarrollaron dos algoritmos para la clasificación de perturbaciones. El primero de ellos se fue basado en las *ANN* y el segundo en *SVM*. El software desarrollado permite el entrenamiento y la prueba de estos algoritmos.

El principal aporte de este capítulo residió en contar con una herramienta de uso intuitivo que puede ser usada como software educativo y para la prueba de los algoritmos desarrollados en los siguientes capítulos.

8.3.2. Desarrollo de una metodología para el diseño y entrenamiento de un clasificador basado en técnicas de inteligencia artificial.

En el *capítulo 4* esta tesis se presenta un marco conceptual para el desarrollo, entrenamiento y prueba de algoritmos basados en técnicas de procesamiento de señales e inteligencia artificial, destinados a la detección, clasificación y localización eventos a partir de un conjunto de muestras.

La metodología propuesta presenta las siguientes características:

- Puede ser extendida a cualquier problema que involucre la adquisición de señales, la extracción de descriptores y la clasificación de patrones.
- La metodología propuesta está definida modularmente de acuerdo a las funciones que cada etapa deben llevar a cabo. Gracias a esta característica, se puede fácilmente incorporar nuevas técnicas de procesamiento de señales, extracción de parámetros

característicos, minería de datos y clasificación sin la necesidad de grandes alteraciones en el flujo de procesos.

- La metodología propuesta permite la selección de parámetros característicos a través de dos métodos que trabajan de manera secuencial: Filtrado heurístico y Algoritmo de búsqueda exhaustiva. Esta metodología permite optimizar la tarea de búsqueda de los descriptores que mejor se adaptan a cada problema en particular.
- El encontrar un set óptimo de parámetros característico permitirá ahorrar cálculos innecesarios en un sistema de clasificación en tiempo real ya que solamente los descriptores encontrados en la etapa de entrenamiento serán extraídos. La etapa de minería de datos brinda, por lo tanto dos ventajas, maximiza el rendimiento del clasificador y minimiza los cálculos realizados en el procesamiento de las muestras.
- Un proceso importante de la metodología propuesta es el algoritmo de búsqueda de los parámetros gobiernan el comportamiento del clasificador. Durante la etapa de diseño del clasificador, este módulo es continuamente invocado por el algoritmo de Búsqueda Exhaustiva con el objetivo de encontrar, simultáneamente, la mejor combinación de parámetros característicos y la mejor combinación de los parámetros que gobiernan al clasificador. Esta característica permite obtener excelentes resultados como se concluye en los capítulos 6 y 7.
- La metodología propuesta fue probada con éxito en dos problemas marcadamente diferentes relacionados con la calidad de energía: La clasificación de perturbaciones y la localización de fallas.

8.3.3. Implementación de un clasificador de perturbaciones simples y complejas.

Se desarrolló un algoritmo avanzado que permite la detección de perturbaciones simples y complejas mediante la combinación de técnicas de procesamiento y algoritmos de inteligencia artificial.

El clasificador multiclase desarrollado, el cual se presenta en detalle en el *capítulo 6*, se conformó por cinco clasificadores binarios organizados en una arquitectura *One vs Rest*. Cada clasificador, destinado a identificar una única perturbación simple, se diseñó individualmente para operar en paralelo y de manera totalmente independiente de sus pares.

El clasificador propuesto resultó ser una opción superadora en los siguientes aspectos:

- La arquitectura propuesta permitió, mediante el algoritmo de minería de datos desarrollado a tal efecto, seleccionar de manera óptima los parámetros característicos que describen a cada perturbación. Esto elimina información redundante e irrelevante en el entrenamiento del clasificador mejorando notablemente su rendimiento.

- Se comprobó que solamente un conjunto reducido de parámetros característicos es suficiente para describir cada tipo de perturbación. Esto resulta en un importante ahorro a la hora de calcular los descriptores.
- La arquitectura propuesta también permite configurar los parámetros que gobiernan a cada etapa (parámetros C y σ) de acuerdo a cada clase presente en el set de entrenamiento. La sintonía del clasificador se puede realizar de manera más precisa.
- Permite procesar los datos en paralelo lo cual resulta en una notable reducción de los tiempos de cálculo. Esta es una característica ideal para aplicaciones en tiempo real.
- Debido a que cada clasificador binario trabaja de manera independiente al resto de los clasificadores que conforman la arquitectura *SVM multiclase*, cuando se necesitan clasificar perturbaciones complejas, no es necesario crear una nueva clase e incorporar un clasificador binario para identificarla. La arquitectura propuesta permitirá, mediante la incorporación de este tipo de perturbaciones en el set de entrenamiento, que cada clasificador binario identifique el evento para el cual fue entrenado aunque el mismo presente en una perturbación compleja.

Como se puede apreciar en el *capítulo 6*, los resultados obtenidos son destacados ya que la precisión lograda a la hora de clasificar perturbaciones complejas supera la mayoría de la alcanzada por la escasa publicación avocada a la clasificación de perturbaciones complejas. Esto se logra sin que el algoritmo pierda capacidad de clasificar perturbaciones simples.

8.3.4. Implementación de un localizador de fallas en sistema de distribución radial.

Se desarrolló un algoritmo un algoritmo preciso, de fácil implementación, de bajo costo computacional, de bajo requerimiento en cuanto a la cantidad de mediciones para la detección, clasificación y localización de una falla en un sistema eléctrico radial de pequeñas dimensiones.

El software desarrollado mostró ser capaz de:

- Detectar una falla a través de una anomalía en las formas de onda de tensión y corriente.
- Estimar la localización geográfica del siniestro que provoca la perturbación dentro de la red, identificando la línea y la fase donde la falla se genera.

Los aportes realizados en lo relativo a la localización de fallas se destacan en las experiencias presentadas en el *capítulo 7* y se resumen a continuación:

- El algoritmo propuesto obtiene excelentes resultados a partir de mediciones realizadas en un único nodo de la red lo que resulta una característica importante a la hora de calcular los costos y tiempos de despliegue del sistema de localización de fallas.

- La habilidad del clasificador depende en gran medida de la correcta selección de los parámetros característicos que describen a las formas de onda de tensión y corriente adquiridas. Por lo tanto, la incorporación de técnicas de minería de datos, en la etapa de entrenamiento de los clasificadores, para la selección de los parámetros característicos que describen mejor al tipo y localización de las fallas representa una clara ventaja frente a la mayoría de los trabajos que utilizan técnicas de *AI* para la localización de fallas.
- Los resultados obtenidos son valiosos si se tiene en cuenta que las señales utilizadas para el entrenamiento y prueba del algoritmo se obtuvieron midiendo solamente medio de las señales de tensión y corriente en un solo nodo del sistema propuesto. Esta es una característica fundamental del algoritmo ya que, en la práctica, solo se necesitaría instalar un único sistema de medición para obtener, con marcada rapidez, resultados en cualquier línea del sistema presentado en este estudio.
- Los resultados del proceso de minería de datos indican que solamente calculando 3 parámetros característicos para el clasificador destinado a identificar el tipo de falla se obtiene un porcentaje de acierto del 100%. Un análisis análogo permite concluir que calculando solamente cuatro parámetros característicos el clasificador destinado a identificar la línea donde se produce la falla obtiene excelentes resultados.

8.4. HERRAMIENTAS UTILIZADAS.

La mayoría de los desarrollos mencionados en este trabajo fueron realizados en MATLAB. Las Toolbox empleadas fueron:

- Wavelet Transform Toolbox
- Simulink.

La librería LibSVM [137] fue compilada en MATLAB y empleada para el entrenamiento y prueba de los clasificadores.

El entorno de Minería de Datos Weka [138] se utilizó para implementar las técnicas filtrado heurístico.

8.5. TRABAJOS FUTUROS.

En esta sección se presentan algunos aspectos de la problemática abordada en este trabajo en donde se podría trabajar en un futuro.

Como la herramienta fue desarrollada de manera modular, se podrían implementar, sin demasiadas modificaciones en su estructura, técnicas de procesamiento como la *ST* o la *HHT*. La implementación de nuevas técnicas de procesamiento de señales permitirá explorar otros parámetros característicos.

Se propone, como trabajo futuro, la implementación de técnicas de minería de datos alternativas a las propuestas en este trabajo. *Algoritmos genéticos (GA)*, *Particle Swarm Optimization (PSO)* y *Ant Colony Optimization (ACO)* son algunas de las técnicas de minería de datos propuestas por la comunidad científica.

De acuerdo a lo que se ha explicado en el *capítulo 6* de este trabajo, las formas de onda utilizadas para el entrenamiento y prueba de los sistemas de clasificación fueron señales obtenidas mediante la simulación matemática de los eventos de calidad de energía. Estudios futuros podrían contemplar la simulación de eventos en el modelo de una red para el entrenamiento y prueba de la herramienta desarrollada para este trabajo de tesis. Esta propuesta permitirá incorporar el comportamiento de las señales de corriente ante un evento de calidad de energía, aspecto que se ha escasamente explorado.

Para el estudio y comparación del rendimiento de los clasificadores se podría incorporar un caso en donde las señales ingresadas al sistema de monitoreo se encuentren contaminadas con ruido.

Debido a que, muchas veces, es complejo contar con un número grande de formas de onda para el entrenamiento de los clasificadores, en el futuro se debe estudiar cual es el tamaño óptimo del set de entrenamiento para lograr resultados aceptables.

Un estudio futuro debería enfocarse en corroborar el rendimiento de los algoritmos desarrollados en lo referente al consumo del tiempo. Esto permitiría contar con información importante para depurar el software y volverlo óptimo en cuanto al consumo de tiempo. El código optimizado podría formar parte de un sistema de monitoreo de señales en tiempo real.

En trabajos futuros se podría contemplar la incorporación de otras técnicas de *AI* como, por ejemplo, *Extreme Learning Machine (ELM)*, cuyo rendimiento, en lo relativo a los porcentajes de acierto y velocidad de entrenamiento se han destacado en otras áreas pero su implementación aún no se ha adoptado en la clasificación de perturbaciones y localización de fallas.

Resulta interesante también comprobar el rendimiento de *ANN* especializadas en el reconocimiento de patrones mediante el aprendizaje no supervisado como las redes neuronales *Adaptive Resonance Theory (ANN ART)* y las *ANN* conocidas como *Self Organized Map (ANN SOM)*.

Nuevamente, debido a que las herramientas aquí presentadas fueron desarrolladas en forma modular, es posible la incorporación de nuevas técnicas de *AI* sin modificar demasiado la estructura de los códigos aquí propuestos.

En el *capítulo 7* se presentó un algoritmo que permitió localizar la línea donde la falla se produce. Como continuación a las investigaciones presentadas se plantea incorporar algoritmos de regresión para identificar, lo más exactamente posible, la distancia comprendida entre el inicio de la línea y el sitio donde la falla se produce.

Adicionalmente se podría comparar el rendimiento del clasificador propuesto empleando diferentes funciones kernel.

Para probar el algoritmo de localización de fallas se propone incrementar el tamaño del modelo de la red eléctrica a sistemas con mayor cantidad de líneas.

En modelos eléctricos de gran tamaño seguramente serán necesarias mediciones en más de un nodo de la red. En un escenario de este tipo, será necesario desarrollar algoritmos de optimización para minimizar la cantidad de nodos de medición necesarios e identificar las zonas geográficas donde estos nodos de medición deberán ser instalados para obtener un determinado porcentaje de acierto a la hora de localizar fallas en sistemas de distribución de gran dimensión.

8.6. PUBLICACIONES REALIZADAS.

8.6.1. Revistas indexadas

1) Educational Software for Power Quality Analysis [129]

Autores: D. De Yong, C. Reineri, and F. Magnago

Revista: IEEE Latin America Transactions, vol. Vol. 11, No. 1, pp. 479-485.

Año de publicación: 2013.

Idioma: Español.

Cantidad de citas: 4 (Al 20 de febrero de 2016).

2) An effective Power Quality classifier using Wavelet Transform and Support Vector Machines [139]

Autores: D. De Yong, S. Bhowmik, and F. Magnago

Revista: Expert Systems with Applications, vol. 42, pp. 6075-6081, 9,

Año de publicación: 2015.

Idioma: Inglés.

Cantidad de citas: 5 (Al 20 de febrero de 2016)

3) Optimized Complex Power Disturbance Classifier Using One vs. Rest Support Vector Machines [140]

Autores: D. De Yong, S. Bhowmik, and F. Magnago

Revista: International Journal of Electrical Power and Energy Systems

Idioma: Inglés.

Estado: Aceptado con correcciones mejores. Al 20 de febrero del 2016 se enviaron las correcciones.

8.6.2. Conferencias

1) Software Educativo para el Análisis de Calidad de Energía [120]

Autores: D. De Yong, C. Reineri, and F. Magnago,

Conferencia: Biennial Congress of Argentina (ARGENCON).

<http://www.ewh.ieee.org/reg/9/etrans/ieee/issues/vol11/vol11issue1Feb.2013/Vol11issue1Feb.2013TLA.htm>

Año: 2012.

Lugar: Córdoba, Argentina.

2) Sistema de detección de eventos de calidad de energía basado en Maquinas de Vectores de Soporte [141].

Autores: D. De Yong, S. Bhowmik, and F. Magnago,

Conferencia: Biennial Congress of Argentina (ARGENCON), IEEE, pp. 108-113,

Año: 2014.

Lugar: Bariloche, Argentina.

<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6868480>

3) Utilización de redes auto expansibles en una Smart Grid [142]

Autores: F. Aguirre, D. D. Yong, E. Tardivo, and F. Magnago,

Conferencia: Seminario Nacional CIGREAC 2014, Buenos Aires, Argentina, 2014.

8.6.3. Trabajos de Tesis

Análisis de perturbaciones en Sistemas Eléctricos de Potencia utilizando Transformada Wavelets, Redes Neuronales y Máquinas de Vectores de Soporte [128].

Tesis de Maestría,

Facultad de Ingeniería, Universidad Nacional de Río Cuarto, Río Cuarto

Año: 2013.

REFERENCIAS

- [1] ITBA, CADER, FARN, AGGERA, UBA, ITBA, *et al.*, *Plataforma Escenarios Energéticos Argentina 2035*, 2014.
- [2] I.E.A., "Energy efficiency market report 2013 – Market trends and medium-term prospects," 2013.
- [3] N. Research. *Installed Base of Smart Meters in China to Reach Nearly 380 Million by 2020*. Available: <http://www.navigantresearch.com/research/smart-grid-in-china>
- [4] R. C. Dugan, S. Santoso, M. F. McGranaghan, and H. W. Beaty, *Electrical Power Systems Quality*: Mcgraw-hill, 2002.
- [5] M. H. Bollen and I. Gu, *Signal Processing of Power Quality Disturbances*: Wiley, 2006.
- [6] S. R. Mohanty, P. K. Ray, N. Kishor, and B. K. Panigrahi, "Classification of disturbances in hybrid DG system using modular PNN and SVM," *International Journal of Electrical Power & Energy Systems*, vol. 44, pp. 764-777, 1// 2013.
- [7] *Scopus*. Available: <http://www.scopus.com/>
- [8] B. W. Jaekel, "Description and classification of electromagnetic environments - revision of IEC 61000-2-5," in *Electromagnetic Compatibility, 2008. EMC 2008. IEEE International Symposium on*, 2008, pp. 1-4.
- [9] "IEEE Recommended Practice for Monitoring Electric Power Quality," *IEEE Std 1159-2009 (Revision of IEEE Std 1159-1995)*, pp. c1-81, 2009.
- [10] B. K. Panigrahi and V. R. Pandi, "Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm," *Generation, Transmission & Distribution, IET*, vol. 3, pp. 296-306, 2009.
- [11] S. Khokhar, A. A. B. Mohd Zin, A. S. B. Mokhtar, and M. Pesaran, "A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances," *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 1650-1663, 2015.
- [12] A. J. Roscoe, G. M. Burt, and J. R. McDonald, "Frequency and fundamental signal measurement algorithms for distributed control and protection applications," *Generation, Transmission & Distribution, IET*, vol. 3, pp. 485-495, 2009.
- [13] Z. Fusheng, G. Zhongxing, and Y. Wei, "The algorithm of interpolating windowed FFT for harmonic analysis of electric power system," *Power Delivery, IEEE Transactions on*, vol. 16, pp. 160-164, 2001.
- [14] G. T. Heydt, P. S. Fjeld, C. C. Liu, D. Pierce, L. Tu, and G. Hensley, "Applications of the windowed FFT to electric power quality assessment," *Power Delivery, IEEE Transactions on*, vol. 14, pp. 1411-1416, 1999.
- [15] Y. H. Gu and M. H. J. Bollen, "Time-frequency and time-scale domain analysis of voltage disturbances," *Power Delivery, IEEE Transactions on*, vol. vol.15, no.4, pp.1279-1284, Oct 2000.
- [16] A. A. Abdelsalam, A. A. Eldesouky, and A. A. Sallam, "Classification of power system disturbances using linear Kalman filter and fuzzy-expert system," *International Journal of Electrical Power & Energy Systems*, vol. 43, pp. 688-695, 2012.
- [17] M. Caujolle, M. Petit, G. Fleury, and L. Berthet, "Reliable power disturbance detection using wavelet decomposition or harmonic model based Kalman filtering," in *Harmonics and Quality of Power (ICHQP), 2010 14th International Conference on*, 2010, pp. 1-6.
- [18] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: the S transform," *Signal Processing, IEEE Transactions on*, vol. 44, pp. 998-1001, 1996.
- [19] P. K. Dash, B. K. Panigrahi, and G. Panda, "Power quality analysis using S-transform," *Power Delivery, IEEE Transactions on*, vol. 18, pp. 406-411, 2003.
- [20] Z. Fengzhan and Y. Rengang, "Power-Quality Disturbance Recognition Using S-Transform," *Power Delivery, IEEE Transactions on*, vol. 22, pp. 944-950, 2007.
- [21] P. K. Dash and M. V. Chilukuri, "Hybrid S-transform and Kalman filtering approach for detection and measurement of short duration disturbances in power networks," *Instrumentation and Measurement, IEEE Transactions on*, vol. 53, pp. 588-596, 2004.

- [22] M. Biswal and P. K. Dash, "Detection and characterization of multiple power quality disturbances with a fast S-transform and decision tree based classifier," *Digital Signal Processing*, vol. 23, pp. 1071-1083, 2013.
- [23] H. Erişti, Ö. Yıldırım, B. Erişti, and Y. Demir, "Automatic recognition system of underlying causes of power quality disturbances based on S-Transform and Extreme Learning Machine," *International Journal of Electrical Power & Energy Systems*, vol. 61, pp. 553-562, 10// 2014.
- [24] R. Kumar, B. Singh, D. Shahani, A. Chandra, and K. Al-Haddad, "Recognition of power-quality disturbances using S-transform-based ANN classifier and rule-based decision tree," *Industry Applications, IEEE Transactions on*, vol. 51, pp. 1249-1258, 2015.
- [25] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 1998, pp. 903-995.
- [26] B. Biswal, M. Biswal, S. Mishra, and R. Jalaja, "Automatic classification of power quality events using balanced neural tree," *Industrial Electronics, IEEE Transactions on*, vol. 61, pp. 521-530, 2014.
- [27] C. Soo-Hwan, J. Gilsoo, and K. Sae-Hyuk, "Time-Frequency Analysis of Power-Quality Disturbances via the Gabor–Wigner Transform," *Power Delivery, IEEE Transactions on*, vol. 25, pp. 494-499, 2010.
- [28] T. A. Kawady, N. I. Elkalashy, A. E. Ibrahim, and A.-M. I. Taalab, "Arcing fault identification using combined Gabor Transform-neural network for transmission lines," *International Journal of Electrical Power & Energy Systems*, vol. 61, pp. 248-258, 2014.
- [29] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," *Power Delivery, IEEE Transactions on*, vol. 11, pp. 924-930, 1996.
- [30] D. C. Robertson, O. I. Camps, J. S. Mayer, and W. B. Gish, "Wavelets and electromagnetic power system transients," *Power Delivery, IEEE Transactions on*, vol. 11, pp. 1050-1058, 1996.
- [31] P. Pillay and A. Bhattacharjee, "Application of wavelets to model short-term power system disturbances," *Power Systems, IEEE Transactions on*, vol. 11, pp. 2031-2037, 1996.
- [32] T. Weiming, S. Xuelei, L. Jingbo, and Z. Zhiheng, "Detection and Classification of Power Quality Disturbances Based on Wavelet Packet Decomposition and Support Vector Machines," in *Signal Processing, 2006 8th International Conference on*, 2006.
- [33] S.-J. Huang, C.-T. Hsieh, and C.-L. Huang, "Application of wavelets to classify power system disturbances," *Electric Power Systems Research*, vol. 47, pp. 87-93, 1998.
- [34] A. K. Ghosh and D. L. Lubkeman, "The classification of power system disturbance waveforms using a neural network approach," *Power Delivery, IEEE Transactions on*, vol. 10, pp. 109-115, 1995.
- [35] S. Santoso, E. J. Powers, W. M. Grady, and A. C. Parsons, "Power quality disturbance waveform recognition using wavelet-based neural classifier. I. Theoretical foundation," *Power Delivery, IEEE Transactions on*, vol. 15, pp. 222-228, 2000.
- [36] S. Santoso, E. J. Powers, W. M. Grady, and A. C. Parsons, "Power quality disturbance waveform recognition using wavelet-based neural classifier. II. Application," *Power Delivery, IEEE Transactions on*, vol. 15, pp. 229-235, 2000.
- [37] J. V. Wijayakulasooriya, G. A. Putrus, and P. D. Minns, "Electric power quality disturbance classification using self-adapting artificial neural networks," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 149, pp. 98-101, 2002.
- [38] I. Monedero, C. Leon, J. Roperro, A. Garcia, J. M. Elena, and J. C. Montano, "Classification of Electrical Disturbances in Real Time Using Neural Networks," *Power Delivery, IEEE Transactions on*, vol. 22, pp. 1288-1296, 2007.
- [39] M. B. I. Reaz, F. Choong, M. S. Sulaiman, F. Mohd-Yasin, and M. Kamada, "Expert System for Power Quality Disturbance Classifier," *Power Delivery, IEEE Transactions on*, vol. 22, pp. 1979-1988, 2007.
- [40] S. Mishra, C. N. Bhende, and K. B. Panigrahi, "Detection and Classification of Power Quality Disturbances Using S-Transform and Probabilistic Neural Network," *Power Delivery, IEEE Transactions on*, vol. 23, pp. 280-287, 2008.

- [41] I. C. Cheng, "Virtual Multifunction Power Quality Analyzer Based on Adaptive Linear Neural Network," *Industrial Electronics, IEEE Transactions on*, vol. 59, pp. 3321-3329, 2012.
- [42] C. Jaehak, E. J. Powers, W. M. Grady, and S. C. Bhatt, "Power disturbance classifier using a rule-based method and wavelet packet-based hidden Markov model," *Power Delivery, IEEE Transactions on*, vol. 17, pp. 233-241, 2002.
- [43] H. Jiansheng, M. Negnevitsky, and D. T. Nguyen, "A neural-fuzzy classifier for recognition of power quality disturbances," *Power Delivery, IEEE Transactions on*, vol. 17, pp. 609-616, 2002.
- [44] P. Dash, S. Mishra, M. Salama, and A. Liew, "Classification of power system disturbances using a fuzzy expert system and a Fourier linear combiner," *Power Delivery, IEEE Transactions on*, vol. 15, pp. 472-477, 2000.
- [45] Y. Liao and J.-B. Lee, "A fuzzy-expert system for classifying power quality disturbances," *International journal of electrical power & energy systems*, vol. 26, pp. 199-205, 2004.
- [46] P. Dash, M. Padhee, and T. Panigrahi, "A hybrid time-frequency approach based fuzzy logic system for power island detection in grid connected distributed generation," *International Journal of Electrical Power & Energy Systems*, vol. 42, pp. 453-464, 2012.
- [47] "IEEE Guide for Determining Fault Location on AC Transmission and Distribution Lines," *IEEE Std C37.114-2004*, pp. 1-44, 2005.
- [48] O. o. E. T. a. D. D. O. U.S. Department of Energy. (2011, GRID 2030' A National Vision for Electricity's Second 100 Years. Available: <http://energy.gov/oe/downloads/national-vision-electricity-s-second-100-years>
- [49] M. Kezunovic, "Smart Fault Location for Smart Grids," *Smart Grid, IEEE Transactions on*, vol. 2, pp. 11-22, 2011.
- [50] A. Yadav and Y. Dash, "An Overview of Transmission Line Protection by Artificial Neural Network: Fault Detection, Fault Classification, Fault Location, and Fault Direction Discrimination," *Advances in Artificial Neural Systems*, vol. 2014, p. 20, 2014.
- [51] Y.-G. Zhang, Z.-P. Wang, J.-F. Zhang, and J. Ma, "Fault localization in electrical power systems: A pattern recognition approach," *International Journal of Electrical Power & Energy Systems*, vol. 33, pp. 791-798, 3// 2011.
- [52] K. Srinivasan and A. St.-Jacques, "A new fault location algorithm for radial transmission lines with loads," *Power Delivery, IEEE Transactions on*, vol. 4, pp. 1676-1682, 1989.
- [53] J. Mora-Flòrez, J. Meléndez, and G. Carrillo-Caicedo, "Comparison of impedance based fault location methods for power distribution systems," *Electric Power Systems Research*, vol. 78, pp. 657-666, 4// 2008.
- [54] C. Myeon-Song, L. Seung-Jae, L. Duck-Su, and J. Bo-Gun, "A new fault location algorithm using direct circuit analysis for distribution systems," *Power Delivery, IEEE Transactions on*, vol. 19, pp. 35-41, 2004.
- [55] C. Myeon-Song, L. Seung-Jae, L. Seong-II, L. Duck-Su, and Y. Xia, "A Direct Three-Phase Circuit Analysis-Based Fault Location for Line-to-Line Fault," *Power Delivery, IEEE Transactions on*, vol. 22, pp. 2541-2547, 2007.
- [56] R. K. Aggarwal, Y. Aslan, and A. T. Johns, "New concept in fault location for overhead distribution systems using superimposed components," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 144, pp. 309-316, 1997.
- [57] F. H. Magnago and A. Abur, "A new fault location technique for radial distribution systems based on high frequency signals," in *Power Engineering Society Summer Meeting, 1999. IEEE*, 1999, pp. 426-431 vol.1.
- [58] M. Gilany, D. K. Ibrahim, and E. S. T. Eldin, "Traveling-Wave-Based Fault-Location Scheme for Multiend-Aged Underground Cable System," *Power Delivery, IEEE Transactions on*, vol. 22, pp. 82-89, 2007.
- [59] H. Livani and C. Y. Evrenosoglu, "A Machine Learning and Wavelet-Based Fault Location Method for Hybrid Transmission Lines," *Smart Grid, IEEE Transactions on*, vol. 5, pp. 51-59, 2014.
- [60] J. A. Jiang, C. L. Chuang, W. Yung-Chung, H. Chih-Hung, W. Jiing-Yi, L. Chien-Hsing, *et al.*, "A hybrid framework for fault detection, classification, and location Part I: Concept, structure and methodology," in *Power and Energy Society General Meeting (PES), 2013 IEEE*, 2013, pp. 1-1.

- [61] R. Agrawal and D. Thukaram, "Identification of fault location in power distribution system with distributed generation using support vector machines," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, 2013, pp. 1-6.
- [62] L. Lovisolò, J. A. Moor Neto, K. Figueiredo, L. de Menezes Laporte, and J. C. dos Santos Rocha, "Location of faults generating short-duration voltage variations in distribution systems regions from records captured at one point and decomposed into damped sinusoids," *Generation, Transmission & Distribution, IET*, vol. 6, pp. 1225-1234, 2012.
- [63] D. Thukaram, H. P. Khincha, and H. P. Vijaynarasimha, "Artificial neural network and support vector Machine approach for locating faults in radial distribution systems," *Power Delivery, IEEE Transactions on*, vol. 20, pp. 710-721, 2005.
- [64] C. Apisit, C. Pothisarn, and A. Ngaopitakkul, "An Application of Discrete Wavelet Transform and Support Vector Machines Algorithm for Fault Locations in Underground Cable," in *Innovations in Bio-Inspired Computing and Applications (IBICA), 2012 Third International Conference on*, 2012, pp. 89-92.
- [65] J. Mora-Florez, J. Bedoya-Ceballos, and L. Perez-Hernandez, "Selection of currents patterns using SVMs for locating faults in radial power systems," in *Transmission and Distribution Conference and Exposition: Latin America, 2008 IEEE/PES*, 2008, pp. 1-7.
- [66] X. Deng, R. Yuan, Z. Xiao, T. Li, and K. L. L. Wang, "Fault location in loop distribution network using SVM technology," *International Journal of Electrical Power & Energy Systems*, vol. 65, pp. 254-261, 2// 2015.
- [67] H. Livani and C. Y. Evrenosoglu, "A fault classification method in power systems using DWT and SVM classifier," in *Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES*, 2012, pp. 1-5.
- [68] I. Baqui, I. Zamora, J. Mazón, and G. Buigues, "High impedance fault detection methodology using wavelet transform and artificial neural networks," *Electric Power Systems Research*, vol. 81, pp. 1325-1333, 7// 2011.
- [69] L. Ye, D. You, X. Yin, K. Wang, and J. Wu, "An improved fault-location method for distribution system using wavelets and support vector regression," *International Journal of Electrical Power & Energy Systems*, vol. 55, pp. 467-472, 2// 2014.
- [70] A. Srivastava, A. Siddharth, and A. K. Singh, "Some observations on selection of most suitable mother wavelet for measurement of power system harmonics using CWT," in *TENCON 2009 - 2009 IEEE Region 10 Conference*, 2009, pp. 1-6.
- [71] V. Vega, C. Duarte, G. Ordóñez, and N. Kagan, "Selecting the best wavelet function for power quality disturbances identification patterns," in *Harmonics and Quality of Power, 2008. ICHQP 2008. 13th International Conference on*, 2008, pp. 1-6.
- [72] Y. M. M. Miiti, G. Openheim y J. M. Pogy, "Wavelet Toolbox, User's Guide Version 2," *The Math Works, Inc 2000*, 2000.
- [73] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, pp. 674-693, 1989.
- [74] I. Daubechies, "Ten Lectures on Wavelets," *Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania*, 1992.
- [75] O. Poisson, P. Rioual, and M. Meunier, "New signal processing tools applied to power quality analysis," *Power Delivery, IEEE Transactions on*, vol. 14, pp. 561-566, 1999.
- [76] A. M. Gaouda, M. M. A. Salama, M. R. Sultan, and A. Y. Chikhani, "Power quality detection and classification using wavelet-multiresolution signal decomposition," *Power Delivery, IEEE Transactions on*, vol. 14, pp. 1469-1476, 1999.
- [77] A. M. Gaouda, M. M. A. Salama, M. R. Sultan, and A. Y. Chikhani, "Application of multiresolution signal decomposition for monitoring short-duration variations in distribution systems," *Power Delivery, IEEE Transactions on*, vol. 15, pp. 478-485, 2000.
- [78] L. Liyan and Z. Zhezhaò, "The detection and location of power quality disturbances based on orthogonal wavelet packet transform," in *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, 2008, pp. 1831-1835.
- [79] Z. Liu, Y. Cui, and W. Li, "Combined Power Quality Disturbances Recognition Using Wavelet Packet Entropies and S-Transform," *Entropy*, vol. 17, p. 5811, 2015.

- [80] S. Santoso, E. J. Powers, and W. M. Grady, "Power quality disturbance data compression using wavelet transform methods," *Power Delivery, IEEE Transactions on*, vol. 12, pp. 1250-1257, 1997.
- [81] H. Zhengyou, G. Shibin, C. Xiaoqin, Z. Jun, B. Zhiqian, and Q. Qingquan, "Study of a new method for power system transients classification based on wavelet entropy and neural network," *International Journal of Electrical Power & Energy Systems*, vol. 33, pp. 402-410, 2011.
- [82] M. A. Masoum, S. Jamali, and N. Ghaffarzadeh, "Detection and classification of power quality disturbances using discrete wavelet transform and wavelet networks," *Science, Measurement & Technology, IET*, vol. 4, pp. 193-205, 2010.
- [83] Y. Hong-Tzer and L. Chiung-Chou, "A de-noising scheme for enhancing wavelet-based power quality monitoring system," *Power Delivery, IEEE Transactions on*, vol. 16, pp. 353-360, 2001.
- [84] L. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*: Pearson Education, 2006.
- [85] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*: Cambridge University Press, 2000.
- [86] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (3rd Edition)*: Prentice Hall, 2009.
- [87] G. B. Dantzig, *Linear programming and extensions*: PRINCETON University Press, 1998.
- [88] L. Lasdon, *Optimization theory for large systems*: Dover Publications, Incorporated, 1970.
- [89] V. Vapnik, *The Nature of Statistical Learning Theory*: Springer, 2000.
- [90] P. Winston. (2010, Accessed 30 Oct, 2015). 6.034 Artificial Intelligence. *Fall 16*. Available: <http://ocw.mit.edu>
- [91] J. Watson, "Support Vector Machine for Multi-Class Pattern Recognition," *Department of Computer Science, Royal Holloway, University of London*, 1999.
- [92] L. Bottou, C. Cortes, J. S. Denker, H. Drucker, I. Guyon, L. D. Jackel, *et al.*, "Comparison of classifier methods: a case study in handwritten digit recognition," in *Pattern Recognition, 1994. Vol. 2 - Conference B: Computer Vision & Image Processing., Proceedings of the 12th IAPR International. Conference on*, 1994, pp. 77-82 vol.2.
- [93] U. H.-G. Kre, #223, and *el*, "Pairwise classification and support vector machines," in *Advances in kernel methods*, S. Bernhard, I. Kopr, J. C. B. Christopher, and J. S. Alexander, Eds., ed: MIT Press, 1999, pp. 255-268.
- [94] J. H. Friedman, "Another approach to polychotomous classification," 1996.
- [95] J. C. Platt, N. Cristianini, and J. Shawe-Taylor, "Large Margin DAGs for Multiclass Classification," in *Proceedings of Neural Information Processing Systems, NIPS'99*, ed: MIT Press, 2000, pp. 547-553.
- [96] H. Osman, "Novel Multiclass SVM-Based Binary Decision Tree Classifier," in *Signal Processing and Information Technology, 2007 IEEE International Symposium on*, 2007, pp. 880-883.
- [97] H. Chih-Wei and L. Chih-Jen, "A comparison of methods for multiclass support vector machines," *Neural Networks, IEEE Transactions on*, vol. 13, pp. 415-425, 2002.
- [98] G. Peisheng and W. Weilin, "Power Quality Disturbances Classification using Wavelet and Support Vector Machines," in *Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on*, 2006, pp. 201-206.
- [99] P. Janik and T. Lobos, "Automated classification of power-quality disturbances using SVM and RBF networks," *Power Delivery, IEEE Transactions on*, vol. 21, pp. 1663-1669, 2006.
- [100] V. Vega, N. Kagan, G. Ordóñez, and C. Duarte, "Automatic power quality disturbance classification using wavelet, Support Vector Machine and Artificial Neural Network," in *Electricity Distribution - Part 1, 2009. CIRED 2009. 20th International Conference and Exhibition on*, 2009, pp. 1-4.
- [101] P. G. V. Axelberg, I. Y. H. Gu, and M. H. J. Bollen, "Support Vector Machine for Classification of Voltage Disturbances," *Power Delivery, IEEE Transactions on*, vol. 22, pp. 1297-1303, 2007.

- [102] L. Whei-Min, W. Chien-Hsien, L. Chia-Hung, and C. Fu-Sheng, "Detection and Classification of Multiple Power-Quality Disturbances With Wavelet Multiclass SVM," *Power Delivery, IEEE Transactions on*, vol. 23, pp. 2575-2582, 2008.
- [103] A. Milchevski and D. Taskovski, "Classification of power quality disturbances using wavelet transform and SVM decision tree," in *Electrical Power Quality and Utilisation (EPQU), 2011 11th International Conference on*, 2011, pp. 1-5.
- [104] Eris, x, H. ti, and Y. Demir, "Automatic classification of power quality events and disturbances using wavelet transform and support vector machines," *Generation, Transmission & Distribution, IET*, vol. 6, pp. 968-976, 2012.
- [105] B. Biswal, M. K. Biswal, P. K. Dash, and S. Mishra, "Power quality event characterization using support vector machine and optimization using advanced immune algorithm," *Neurocomputing*, vol. 103, pp. 75-86, 2013.
- [106] P. Ray and D. P. Mishra, "Signal processing technique based fault location of a distribution line," in *Recent Trends in Information Systems (ReTIS), 2015 IEEE 2nd International Conference on*, 2015, pp. 440-445.
- [107] C.-Y. Lee and S. Yi-Xing, "Optimal Feature Selection for Power-Quality Disturbances Classification," *Power Delivery, IEEE Transactions on*, vol. 26, pp. 2342-2351, 2011.
- [108] J. Han, W.-K. Kim, J.-W. Lee, and C.-H. Kim, "Fault type classification in transmission line using STFT," in *Developments in Power Systems Protection, 2012. DPSP 2012. 11th International Conference on*, 2012, pp. 1-5.
- [109] D. Kostadinov and D. Taskovski, "Automatic voltage disturbance detection and classification using wavelets and multiclass logistic regression," in *Instrumentation and Measurement Technology Conference (I2MTC), 2012 IEEE International*, 2012, pp. 103-106.
- [110] B. Panigrahi and V. R. Pandi, "Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm," *Generation, Transmission & Distribution, IET*, vol. 3, pp. 296-306, 2009.
- [111] K. Manimala, K. Selvi, and R. Ahila, "Optimization techniques for improving power quality data mining using wavelet packet based support vector machine," *Neurocomputing*, vol. 77, pp. 36-47, 2012.
- [112] M. Hajian and A. A. Foroud, "A new hybrid pattern recognition scheme for automatic discrimination of power quality disturbances," *Measurement*, vol. 51, pp. 265-280, 2014.
- [113] B. Biswal, P. K. Dash, and S. Mishra, "A hybrid ant colony optimization technique for power signal pattern classification," *Expert Systems with Applications*, vol. 38, pp. 6368-6375, 2011.
- [114] N. Jasmina, S. Perica, and B. Dusan, "Toward optimal feature selection using ranking methods and classification algorithms," *Yugoslav Journal of Operations Research*, 2011.
- [115] L. Huan and R. Setiono, "Chi2: feature selection and discretization of numeric attributes," in *Tools with Artificial Intelligence, 1995. Proceedings., Seventh International Conference on*, 1995, pp. 388-391.
- [116] M. Robnik-Šikonja and I. Kononenko, "Theoretical and Empirical Analysis of ReliefF and RReliefF," *Machine Learning*, vol. 53, pp. 23-69, 2003/10/01 2003.
- [117] M. A. H. a. L. A. Smith, "Practical Feature Subset Selection for Machine Learning," 1998.
- [118] M. A. H. a. L. A. Smith, "Feature Selection for Machine Learning: Comparing a Correlation-based Filter Approach to the Wrapper," *Department of Computer Science University of Waikato Hamilton New Zealand*, 1999.
- [119] C.-C. C. Chih-Wei Hsu, and Chih-Jen Lin. (2010, A Practical Guide to Support Vector Classification. *Department of Computer Science National Taiwan University*, .
- [120] D. De Yong, C. Reiner, and M. Fernando, "Software Educativo para el Análisis Eventos de Calidad de Energía Eléctrica," presented at the Biennial Congress of Argentina (ARGENCON), 2012., Córdoba, Argentina, 2012.
- [121] B. a. H. Demuth, "Neural Network Toolbox " *MATLAB*, 2009.
- [122] "IEEE Recommended Practice for Monitoring Electric Power Quality," in *IEEE Std 1159-2009 (Revision of IEEE Std 1159-1995)*, ed, 2009, pp. c1-81.
- [123] L. Whei-Min, W. Chien-Hsien, L. Chia-Hung, and C. Fu-Sheng, "Classification of Multiple Power Quality Disturbances Using Support Vector Machine and One-versus-One Approach," in *Power System Technology, 2006. PowerCon 2006. International Conference on*, 2006, pp. 1-8.

- [124] C. Cheng-Long, L. Yen-Ling, H. Tsong-Liang, H. Ying-Tung, and J. Joe-Air, "Recognition of Multiple PQ Disturbances Using Wavelet-based Neural Networks - Part 2: Implementation and Applications," in *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, 2005, pp. 1-6.
- [125] L. Gang, L. Fanguang, W. Guanglei, N. Shangkun, and Z. Siguo, "Classification of power quality disturbances based on independent component analysis and support vector machine," in *Wavelet Analysis and Pattern Recognition (ICWAPR), 2013 International Conference on*, 2013, pp. 115-123.
- [126] S. Dalai, D. Dey, B. Chatterjee, S. Chakravorti, and K. Bhattacharya, "Cross Hilbert-Huang transform based feature extraction method for multiple PQ disturbance classification," in *Condition Assessment Techniques in Electrical Systems (CATCON), 2013 IEEE 1st International Conference on*, 2013, pp. 314-317.
- [127] Z. Liu, Y. Cui, and W. Li, "A Classification Method for Complex Power Quality Disturbances Using EEMD and Rank Wavelet SVM," *Smart Grid, IEEE Transactions on*, vol. PP, pp. 1-1, 2015.
- [128] D. De Yong, "Análisis de perturbaciones en Sistemas Eléctricos de Potencia utilizando Transformada Wavelets, Redes Neuronales y Máquinas de Soporte de Vectores," Magister Magister, Facultad de Ingeniería, Universidad Nacional de Río Cuarto, Río Cuarto, 2013.
- [129] D. De Yong, C. Reineri, and F. Magnago, "Educational Software for Power Quality Analysis," *IEEE Latin America Transactions*, vol. Vol. 11, No. 1, pp. 479-485, 2013.
- [130] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 1-27, 2011.
- [131] S. A. M. Javadian, A. M. Nasrabadi, M. R. Haghifam, and J. Rezvantalab, "Determining fault's type and accurate location in distribution systems with DG using MLP Neural networks," in *Clean Electrical Power, 2009 International Conference on*, 2009, pp. 284-289.
- [132] F. H. Magnago and A. Abur, "Fault location using wavelets," *Power Delivery, IEEE Transactions on*, vol. 13, pp. 1475-1480, 1998.
- [133] F. Yan, Z. Chen, Z. Liang, Y. Kong, and P. Li, "Fault location using wavelet packets," in *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, 2002, pp. 2575-2579.
- [134] H. Hizman, P. Crossley, P. Gale, and G. Bryson, "Fault section identification and location on a distribution feeder using travelling waves," in *Power Engineering Society Summer Meeting, 2002 IEEE*, 2002, pp. 1107-1112.
- [135] A. Borghetti, S. Corsi, C. A. Nucci, M. Paolone, L. Peretto, and R. Tinarelli, "On the use of continuous-wavelet transform for fault location in distribution power systems," *International Journal of Electrical Power & Energy Systems*, vol. 28, pp. 608-617, 2006.
- [136] A. Borghetti, M. Bosetti, C. Nucci, M. Paolone, and A. Abur, "Integrated use of time-frequency wavelet decompositions for fault location in distribution networks: theory and experimental validation," *Power Delivery, IEEE Transactions on*, vol. 25, pp. 3139-3146, 2010.
- [137] C. Chang and C. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, May// 2011.
- [138] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," *SIGKDD Explorations*, vol. 11, 2009.
- [139] D. De Yong, S. Bhowmik, and F. Magnago, "An effective Power Quality classifier using Wavelet Transform and Support Vector Machines," *Expert Systems with Applications*, vol. 42, pp. 6075-6081, 9// 2015.
- [140] D. D. Yong, S. Bhowmik, and F. Magnago, "Optimized Complex Power Disturbance Classifier Using One vs. Rest Support Vector Machines.," *International Journal of Electrical Power & Energy Systems*, 2015.
- [141] D. De Yong, S. Bhowmik, and F. Magnago, "Sistema de deteccion de eventos de calidad de energía basado en Maquinas de Vectores de Soporte," in *Biennial Congress of Argentina (ARGENCON), 2014 IEEE*, 2014, pp. 108-113.
- [142] F. Aguirre, D. D. Yong, E. Tardivo, and F. Magnago, "Utilizacion de redes auto expansibles en una smart grid," in *Seminario Nacional CIGREAC 2014*, Buenos Aires, Argentina, 2014.