



CIATEC

EVALUACIÓN DE LA CALIDAD DEL AISLAMIENTO DE ALTO  
VOLTAJE UTILIZANDO MÁQUINAS DE SOPORTE VECTORIAL

# Tesis

QUE PARA OBTENER EL GRADO ACADÉMICO DE

Maestro en Ciencia y Tecnología  
en la Especialidad de Ingeniería  
Industrial y de Manufactura

PRESENTA

**Sergio Humberto Almanza Ruiz**

Director de Tesis

Dr. Javier Yáñez Mendiola



EN  
CIENCIA Y TECNOLOGIA

León, Guanajuato, Diciembre de 2012

# Agradecimientos

Al CONACYT por la beca otorgada para la realización de mis estudios.

Al PICYT por el apoyo académico y administrativo.

A mi asesor de tesis Dr. Javier Yañez Mendiola , por su asesoría y claridad.

A Anabel,mi esposa por su apoyo.

Asímismo agradezco a mis maestros, compañeros y a toda aquella persona que de alguna manera contribuyó al término de mis estudios de maestría.

# Resumen

En este trabajo se construyó un clasificador de descargas parciales para evaluar la calidad del aislamiento eléctrico del cableado subterráneo de alto voltaje, utilizando la técnica de maquina de soporte vectorial. La clasificación del tipo de descarga parcial es un estándar para evaluar la calidad del aislamiento de alto voltaje.

Actualmente la clasificación de descargas parciales se realiza por inspección visual de patrones bidimensionales de descarga parcial, aunque ha habido esfuerzos para automatizar la clasificación mediante sistemas expertos, redes neuronales entre otras técnicas de aprendizaje de máquina.

El objetivo de esta tesis es construir una herramienta de evaluación del aislamiento de alto voltaje para diagnosticar su funcionalidad. Más específicamente construir una herramienta que determine el tipo de descarga parcial teniendo como información los reportes de mediciones de descarga parcial en fase resuelta.

Un clasificador es una herramienta de aprendizaje de máquina que tiene por entrada conjuntos de datos y los coloca en alguna de  $k$  clases o categorías. En este trabajo se construyó el clasificador siguiendo la metodología de aprendizaje supervisado, teniendo como datos de entrada patrones numéricos extraídos de reportes de mediciones en sitio de descarga parcial y como clases el tipo de descarga parcial. De las técnicas existentes de aprendizaje supervisado se eligieron la maquina de soporte vectorial en base a resultados que otros investigadores han obtenido con datos generados mediante la simulación física de descargas parciales y una clasificación de las mismas con la técnica mencionada.

Se implementaron algoritmos para transformar los reportes de medición de descargas parciales en vectores de etiquetas y atributos, posteriormente se implemento un algoritmo para transformar los vectores de etiquetas y atributos en vectores de etiquetas y atributos tipo I y tipo II

Teniendo como entrada los vectores de etiquetas y atributos tipo I y tipo II se procedió a obtener modelos de clasificación mediante la técnica de maquina de soporte vectorial con la implementación SVM light. En la fase de prueba se midió la eficiencia de clasificación de los modelos generados en la fase de entrenamiento, variando el parámetro  $C$  y el tipo kernel polinomial o de base radial los resultados de la fase de prueba se concluye que el mejor modelo de clasificación es el que se genero utilizando vectores de etiquetas y atributos de tipo II, kernel polinomial de grado 1, para todos los valores de  $C$  entre 0.001 y 100 que en comparación de la literatura se realizo para mediciones de descarga parcial en sitio.

# Introducción

En ingeniería industrial mediante la aplicación del reconocimiento de patrones y el aprendizaje de maquina se han elaborado herramientas que construyen modelos partiendo directamente de los datos operativos. Dichos modelos pueden actualizarse tan pronto como se generen nuevos datos y así reflejar las condiciones cambiantes, este enfoque tiene amplia aplicabilidad en áreas que van desde la manufactura de procesos y desempeño del producto (Pfungsten, 2007) e investigación científica hasta finanzas y negocios.

Actualmente CFE proyecta la construcción de redes eléctricas subterráneas de alta tensión (115, 230 y 400 kV) el aislamiento polimérico del cable de dichas redes presenta poca tolerancia a la degradación debido al fenómeno de descargas parciales, teniendo como consecuencia la suspensión del suministro de energía eléctrica. Por ello, y aunque actualmente es muy complicado predecir la vida útil de un cable o su tiempo estimado antes del fallo, se cuida al máximo que los cables y las instalaciones estén libres de descargas parciales.

La interpretación de los patrones de descarga parcial puede revelar la fuente y causa de su ocurrencia, por tanto ha sido usada como una condición de monitoreo y control de calidad en la industria manufacturera (IEC60270, 2000). La clasificación de patrones de descarga parcial es realizada por un experto mediante identificación visual del patrón, no obstante ha habido esfuerzos para realizar la clasificación con sistemas inteligentes como sistemas expertos, redes neuronales, mapas autoorganizados y máquinas de soporte vectorial.

Este trabajo tiene por objetivo construir una herramienta de clasificación de descargas parciales empleando la metodología de aprendizaje supervisado de maquina concretamente con la técnica maquina de soporte vectorial.

# Índice general

Agradecimientos	I
Resumen	II
Introducción	III
<b>1. PLANTEAMIENTO Y DELIMITACIÓN DEL PROBLEMA</b>	<b>1</b>
<b>2. ESTADO DEL ARTE</b>	<b>3</b>
2.1. Introducción . . . . .	3
2.1.1. Aprendizaje de máquina en la ingeniería industrial . . . . .	3
2.2. Aprendizaje de máquina aplicadas a la evaluación de la calidad del aislamiento	4
<b>3. MARCO TEÓRICO</b>	<b>6</b>
3.1. Descargas parciales. . . . .	6
3.1.1. Definiciones básicas . . . . .	6
3.1.2. Patrones de descarga parcial . . . . .	7
3.2. aprendizaje de máquina y Reconocimiento de patrones . . . . .	8
3.2.1. Aprendizaje de máquina y clasificadores . . . . .	9
3.3. Máquina de soporte vectorial . . . . .	10
3.3.1. Kerneles y datos no linealmente separables . . . . .	12
<b>4. JUSTIFICACIÓN Y OBJETIVOS</b>	<b>14</b>
4.1. Objetivos . . . . .	14
4.1.1. Propósitos y objetivos de la investigación . . . . .	14
4.1.2. Hipótesis . . . . .	15
<b>5. METODOLOGÍA</b>	<b>16</b>
5.1. Diseño de la investigación . . . . .	16
5.2. Descripción de la implementación de la metodología. . . . .	17
5.3. Acondicionamiento de los datos para la fase de prueba y de entrenamiento de la herramienta de evaluación del aislamiento. . . . .	18
5.3.1. Transformación de los patrones numéricos de descarga parcial en vectores de etiquetas y atributos . . . . .	20
5.3.2. Fase de entrenamiento . . . . .	22
5.4. Fase de prueba . . . . .	23
	IV

<b>6. RESULTADOS</b>	<b>24</b>
6.1. Resultados obtenidos durante la fase de prueba . . . . .	24
6.2. CONCLUSIONES . . . . .	25
<b>Referencias</b>	<b>27</b>

# Índice de figuras

1.1. Patrón de descarga parcial . . . . .	2
3.1. Mediciones de descarga parcial en fase resuelta. . . . .	7
3.2. Patrones de descarga parcial en dos dimensiones . . . . .	7
3.3. Esquema general del reconocimiento de patrones. . . . .	9
3.4. Aprendizaje supervisado . . . . .	10
3.5. Clasificadores lineales . . . . .	10
3.6. Máximo margen separador . . . . .	11
3.7. Separación de clases. . . . .	12
4.1. Visión general del proyecto . . . . .	14
5.1. Visión general de la metodología . . . . .	16
5.2. Visión general de la parte experimental . . . . .	18
5.3. Obtención de mediciones de descarga parcial. . . . .	19
5.4. Mediciones transformadas a vectores de etiquetas y atributos . . . . .	21
5.5. Fase de entrenamiento de la máquina de soporte vectorial . . . . .	22
5.6. Parámetros para la generación de modelos de clasificación de descarga parcial	23
5.7. Prueba de la eficiencia de clasificación los modelos obtenidos . . . . .	23

# Índice de cuadros

5.1. Patrones de descarga parcial en forma de datos numéricos . . . . .	19
6.1. Porcentaje de error utilizando método de entrada I y kernel polinomial . . .	24
6.2. Porcentaje de error utilizando método de entrada II y kernel polinomial . .	24
6.3. Porcentaje de error utilizando método de entrada I y kernel función de base radial . . . . .	25
6.4. Porcentaje de error utilizando método de entrada I y kernel función de base radial . . . . .	25
6.5. Mínimo porcentaje de error alcanzado . . . . .	25



# CAPÍTULO 1

## PLANTEAMIENTO Y DELIMITACIÓN DEL PROBLEMA

Actualmente CFE proyecta la construcción de redes eléctricas subterráneas. Como todo equipo eléctrico de alta tensión (115, 230 y 400 kV) dichas redes presentan una degradación en el aislamiento debido a descargas parciales, tal que provoca la suspensión del suministro de energía eléctrica.

Las descargas parciales son descargas eléctricas que ocurren dentro del aislamiento de equipos de alto voltaje debido generalmente a la presencia de cavidades o contaminación que son el resultado de las fallas en los procesos de manufactura, esfuerzo mecánico o envejecimiento (Salama y Bartnikas, 2002), además el análisis de descarga parcial se ha establecido como una herramienta confiable de diagnóstico para evaluar la integridad y las deficiencias en el diseño de los sistemas de aislamiento.

En la industria eléctrica las pruebas de descarga parcial son generalmente muy directas ya que se llevan a cabo en laboratorios con los arreglos necesarios para reducir interferencias (J. Kuffel y Zaengl, 2000).

Por otra parte en instalaciones de alto voltaje las mediciones de descargas parciales son mucho más difíciles de realizar debido a las interferencias usualmente encontradas como señales de radio, mecanismos de interrupción, electrónica de potencia, maquinaria rotatoria y otras fuentes de descarga parcial (Satish y Nazneen, 2003)

Debido a estas disparidades es que las mediciones de descarga parcial requieren del uso de técnicas de procesamiento de señal más elaboradas.

La interpretación de los patrones de descarga parcial puede revelar la fuente y causa de su ocurrencia, por tanto ha sido usada como una condición de monitoreo y control de calidad en la industria manufacturera (IEC60270, 2000)

Para el presente trabajo se cuenta con mediciones de descarga parcial digitalizadas como patrones de descarga parcial de análisis de fase resuelta, tanto visuales como en forma de reportes numéricos.

En la Figura 1.1, se muestra un patrón bidimensional de descargas parciales en fase resuelta, en el eje de las abscisas se grafica la fase o ángulo, en la ordenada la carga aparente, mientras el color representa la frecuencia con que ocurre determinada medición. Existen patrones base asociados a los respectivos tipos de descarga parcial, obtenidos por investigadores

en el área de descargas parciales.

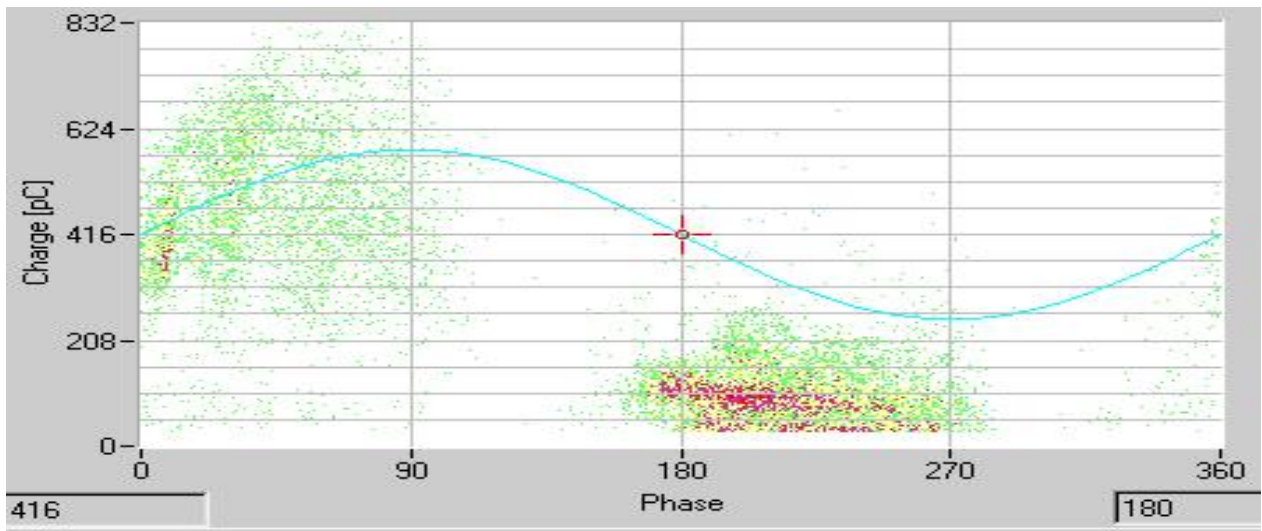


Figura 1.1: Patrón de descarga parcial .

Actualmente la clasificación de descarga la realiza un experto parcial mediante inspección visual de patrones visuales como los descritos en el párrafo anterior ,no obstante ha habido esfuerzos para realizar la clasificación con sistemas inteligentes como sistemas expertos, redes neuronales, mapas autoorganizados y maquinas de soporte vectorial.

En la literatura existen algunos artículos de clasificación de descargas parciales mediante maquinas de soporte vectorial dando una eficiencia de predicción de 98 % con sus respectivas condiciones, en el presente trabajo se emplearan. maquinas de soporte vectorial para la clasificación del tipo de descarga parcial que difieren principalmente en las condiciones del presente trabajo en que

- Los patrones de descarga parcial provienen de mediciones en sitio.
- No se emplea técnicas de eliminación de ruido posteriores a la digitalización de los patrones.
- Se cuenta con patrones base de descarga parcial.
- Se clasificará exclusivamente los tipos de descarga parcial, no así los tipos de falla asociados.

# CAPÍTULO 2

## ESTADO DEL ARTE

### 2.1 INTRODUCCIÓN

Actualmente los avances tanto en la técnica como en los negocios generan un creciente número de nuevos procesos a ser comprendidos y controlados. En estos procesos el incremento en el nivel de detalle y complejidad a ser dominada para satisfacer las fuerzas del mercado requiere la captación y asimilación de conocimiento a través de novedosas técnicas de modelado. Los enfoques convencionales de formulación de modelos son muchas veces inadecuados dadas las presiones actuales de tiempo y costo. Para lograr un perfil competitivo la ingeniería industrial del siglo XXI ha aumentado su acervo tradicional de modelado de procesos desarrollando mediante la aplicación del reconocimiento de patrones y el aprendizaje de máquina nuevas herramientas que construyen modelos partiendo directamente de los datos operativos. Dichos modelos pueden actualizarse fácilmente tan pronto como se generen nuevos datos para poder reflejar las condiciones cambiantes, este enfoque tiene amplia aplicabilidad en áreas que van desde la manufactura de procesos, desempeño del producto e investigación científica hasta finanzas y negocios.

#### 2.1.1. Aprendizaje de máquina en la ingeniería industrial

Una de las primeras herramientas de aprendizaje de máquina construidas para resolver un problema en la industria consistió en un clasificador en base a redes neuronales aplicado a la calidad de acero laminado en caliente (Browne, 1999), en dicho clasificador las diferentes calidades de acero laminado fueron consideradas como las clases y como parámetros la fuerza de los rodillos, temperatura de lámina, esfuerzos mecánicos entre otros.

Otro trabajo donde se emplean más ampliamente los métodos de aprendizaje de máquina para resolver problemas encontrados típicamente en el análisis de datos desde el diseño de un producto hasta su producción en masa es en (Pfungsten, 2007). En dicho trabajo se discute en detalle el caso de estudio de la producción de circuitos integrados y sistemas micromecánicos. Además de los fallos aleatorios, los errores sistemáticos pueden ocasionar pérdidas adicionales. Es difícil encontrar la causa raíz para las pérdidas no vistas anteriormente, ya que las interrelaciones físicas difícilmente pueden ser reconstruidas en instalaciones manufactureras complejas y existen generalmente un gran número de causas potenciales de error. El trabajo

de Pfingsten muestra que usando selección de características, los datos de chequeos de calidad pueden ser combinados con los datos de manufacturación para construir un mecanismo automático de localización de fallos. Pfingsten desarrolló un marco de referencia para realizar análisis estadístico de sensibilidad y optimización robusta que toma en cuenta tolerancias de proceso. Usando regresión gaussiana no paramétrica, el análisis de sensibilidad se pudo desarrollar eficientemente, siendo modelos probabilísticos, los procesos gaussianos permiten un diseño experimental óptimo, y por tanto reducen el número de corridas de simulación requeridas. Además aplico una regresión gaussiana de proceso de aprendizaje activo y se validó experimentalmente.

## 2.2 APRENDIZAJE DE MÁQUINA APLICADAS A LA EVALUACIÓN DE LA CALIDAD DEL AISLAMIENTO

JIAO Shang-bin (2006) Realizaron evaluación del grado de contaminación en aislamiento mediante máquina de soporte vectorial con mínimos cuadrados. Utilizaron el valor pico de la corriente de fuga, la amplitud y el tiempo de dicha corriente, temperatura y humedad asociada como las cinco variables de entrada y como salidas el grado de contaminación. Realizaron el ajuste de los parámetros de la máquina de soporte vectorial mediante validación cruzada, con kernel polinomial y base de función radial. Encontraron resultados satisfactorios en la predicción del grado de contaminación hallando mejor desempeño con el kernel de base radial.

Chang-Hsing Lee y Haeng (2008) Desarrollaron una serie de pruebas para asociar las propiedades de la descarga parcial que se presenta en las terminaciones del cable, la cual es causada por los distintos defectos como mala posición del cono, vacíos entre otros.

Hao y Lewin (2010) utilizaron un sistema de medición de descargas parciales en línea de banda ancha, para facilitar la identificación automática de descarga parcial. Emplearon tres modelos artificiales para simular las fuentes de descarga parcial típicas que se presentan en sistemas de potencia, aplicaron análisis de ondículas para preprocesar los datos de las mediciones obtenidas. Posteriormente estos datos fueron procesados utilizando análisis de correlación para clasificar las descargas en diferentes grupos. Mediante máquinas de soporte vectorial identificaron las diferentes fuentes de descarga parcial. La máquina de soporte vectorial fue entrenada para distinguir entre diferentes características inherentes de cada señal de fuente de descarga parcial.

Erişti y Demir (2010) Realizaron una clasificación automática de eventos de potencia, basados en transformaciones de onduletas y maquinas de soporte vectorial. Mediante dicho enfoque se utilizaron vectores de características de señales de eventos trifásicas se extrajeron después de que la señal fue segmentada y normalizada. Los parámetros de kernel y parametrización de la maquina de soporte vectorial se determinaron por validación cruzada. Conservaron el conjunto de parámetros que dio el menor porcentaje de fallos en la classifica-

ción. Utilizaron el modelo ATP/EMT para los seis tipos de eventos en sistemas de potencia, fallo de fase a tierra fallo de fase a fase, fallo trifásico, fallo en el encendido de la carga, fallo en el encendido del capacitor, y energización del transformador. Utilizaron datos con ruido y sin ruido en el algoritmo propuesto. Los resultados que obtuvieron indican que la clasificación automática propuesta es un algoritmo robusto y tiene la habilidad de distinguir fácilmente diferentes tipos de eventos de la calidad de la potencia.

Chou et al. (2010) Realizaron un sistema de detección predictivo de fallas en maquinas para la industria semiconductora. Teniendo en cuenta que no es posible monitorear en línea el sistema de producción y prescindieron de políticas de revisión rutinaria las cuales causan interrupción. Además de que empato las mediciones de calidad hechas en tiempo real , también detecto la degradación del desempeño de las maquinas correspondiente mediante la información del proceso de manufactura. Mediante análisis de componentes principales lograron incrementar la precisión del kernel, a mayor descripción de los datos incremento la precisión de la predicción, utilizaron algoritmos genéticos para ajustar los parámetros de la maquina de soporte vectorial mejorar el aprendizaje y la fase de prueba. Los resultados arrojaron que maquina de soporte vectorial supera a las redes neuronales de base radial para la misma tarea. Concluyen que utilizar sistema de fallas es mejor que las metodológicas estadísticas para control de fallas.

Mousavi y Butler-Purry (2009) Desarrollaron un sistema automático de evaluación de la condición de la instalación distribución subterránea basado en mediciones, utilizando clasificadores máquina de soporte vectorial y de generación reglas alimentando el sistema con mediciones de tiempo y onduletas. Obtuvieron resultados satisfactorios con la predicción del deterioro progresivo del sistema subterráneo.

Kulkarni (2009) Desarrollaron un sistema de prueba de descargas parciales para cable de alto voltaje de 1000kv y 765 kv

Lai y Blackburn (2010) implementaron una red neuronal con retro propagación, un mapa auto-organizado y una máquina de soporte vectorial para predecir tipos de falla localizada en aislamiento. Utilizaron análisis de fase para la descarga parcial así como dos métodos para los datos de entrada. A saber, el uso de cuatro momentos estadísticos y análisis de componentes principales (PCA por sus siglas en inglés). Compararon las tres técnicas y obtuvieron como resultado que la máquina de soporte vectorial es mejor en términos de exactitud 98 % y tiempo de proceso 30 s.

# CAPÍTULO 3

## MARCO TEÓRICO

### 3.1 DESCARGAS PARCIALES.

#### 3.1.1. Definiciones básicas

Una descarga parcial es un fenómeno de ruptura eléctrica que está confinado y localizado en la región de un medio aislante, entre dos conductores que se encuentran a diferente potencial (Kind y Feser, 2001)

Estas descargas comúnmente ocurren en pequeños vacíos dentro de un aislante o en las capas entre aislantes, Al pasar el tiempo las descargas parciales pueden degradar y dañar aislamiento en equipo de eléctrico de alto voltaje (Otte, 2006)

De acuerdo con la literatura en descargas parciales en general se tiene los siguientes tipos de descarga parcial:

- Internas
- Superficiales
- Efecto corona

Descargas internas ocurren en cavidades, en general gaseosas, al interior de un dieléctrico sólido. Estas pueden presentarse de varias formas.

- En el exterior, es decir la cavidad se produce entre los revestimientos de dos conductores paralelos.
- En el interior del dieléctrico, se presentan como una pequeña cavidad de aire.
- Entre dieléctrico y aislante.

Descargas superficiales ocurren en gas, ocasionalmente en líquido, en la superficie de un dieléctrico o interfase entre dos dieléctricos.

Las descargas corona ocurren en las superficies metálicas o que tienen un pequeño radio de curvatura, en las cuales el campo eléctrico supera la rigidez dieléctrica del gas. La descarga limitada a la zona de alto campo se extingue en el medio circundante lo que no origina una falla franca.

### 3.1.2. Patrones de descarga parcial

Actualmente existen equipos de detección de descargas parciales que registran patrones de descarga parcial en fase resuelta. Un patrón de descargas parciales de fase resuelta es un conjunto pulsos relacionados en magnitud ( $Q$ ), fase ( $\Phi$ ) y tasa de repetición ( $N$ ) producidos por descargas parciales. Estos patrones se registran en un equipo de computo, como se describe en la Figura 3.1 en la que para cada ciclo desde el ciclo 1 hasta  $N$  todas las cargas aparentes  $Q$  se colocan en su correspondiente posición en fase  $\Phi$ .

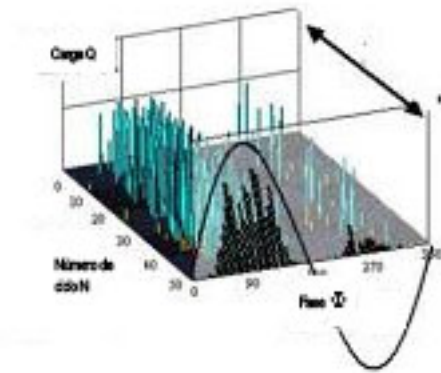


Figura 3.1: Mediciones de descarga parcial en fase resuelta.

Posteriormente mediante el software empotrado en los equipos de detección de descargas parciales dichos patrones se muestran como en la Figura 3.2 en la cual se observa lo siguiente:

- El número de repetición ( $N$ ) que se representa por un color
- La magnitud de la carga ( $Q$ ) que proporciona la abscisa al origen
- El ángulo de fase ( $\Phi$ ) donde ocurren las descargas, con respecto al instante en que el voltaje aplicado cambia de polaridad hacia positivo.

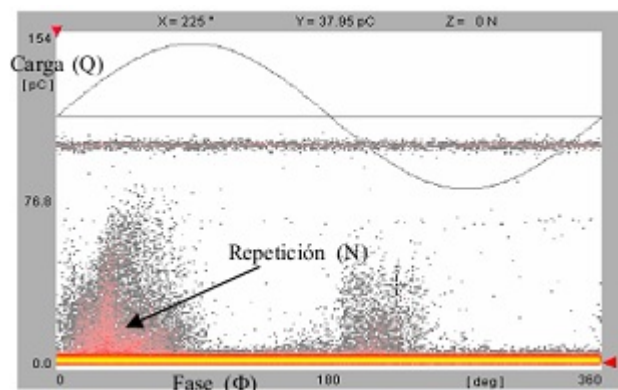


Figura 3.2: Patrones de descarga parcial en dos dimensiones

Investigadores en el área de descargas parciales han construido estándares donde se asocia alguno de los tres tipos de descarga parcial a patrones bidimensionales como el de la Figura 3.2, para se realice la identificación del tipo de descarga parcial visualmente.

## 3.2 APRENDIZAJE DE MÁQUINA Y RECONOCIMIENTO DE PATRONES

El problema de búsqueda de patrones en datos es fundamental, tiene una larga y exitosa historia. Ejemplo de ello fueron las observaciones astronómicas de Tycho Brahe en el siglo XVI que permitieron a Johannes Kepler descubrir leyes empíricas del movimiento planetario catapulta para el desarrollo de la mecánica clásica. Análogamente el descubrimiento de regularidades en el espectro atómico jugó un papel clave en el desarrollo y verificación de la física cuántica de principios del siglo XX .

Mediante la tecnología actual de sensores, capacidad de intercomunicación , procesamiento, y almacenamiento de los equipos de cómputo se obtiene una gran cantidad de mediciones o datos relevantes para muchas aplicaciones en ciencia e ingeniería como identificación y clasificación de patologías, reconocimiento de óptico de caracteres , asignación de créditos, identificación de correo electrónico tipo ‘spam’, entre otras.

El aprendizaje de máquina se enfoca en la tarea de lograr que las computadoras se autoprogramen partiendo de una experiencia y una estructura inicial (Mitchell, 2007) además mientras la estadística se aboca al problema de que conclusiones es posible inferir en base a los datos , el aprendizaje de máquina aborda problemas adicionales como: cuales arquitecturas de computadora y algoritmos son los mas efectivos para capturar, almacenar, indizar, recuperar, combinar datos, como organizar varias subtarefas de aprendizaje en un sistema mas grande, asi como problemas de tratabilidad computacional.

El campo de reconocimiento de patrones se aboca al descubrimiento automático de regularidades en los datos y con el uso de dichas regularidades realizar acciones como clasificar los datos en diferentes categorías.

El reconocimiento de patrones tiene sus orígenes en la ingeniería, mientras que el aprendizaje de máquina surge de la ciencia computacional. Ambas actividades pueden ser vistas como dos facetas del mismo campo.

La Figura 3.3 muestra el esquema general de un sistema de reconocimiento de patrones en la etapa inicial Señal: se tienen señales de entrada pertenecientes a cualquier tipo de proceso físico por ejemplo: señales eléctricas , voz , imágenes entre otros. En la siguiente etapa Sensor se capta la señal de entrada convirtiéndola en una serie de datos numéricos, que son aptos para ser procesados. Luego en la etapa de Extracción de características esta serie de datos numéricos se transforma a un vector multidimensional, en esta misma etapa se suele realizar una reducción de la dimensionalidad, para eliminar información redundante; cabe mencionar



que debe existir un compromiso entre la reducción de dimensionalidad y la información relevante para la etapa clasificación. La última etapa el Clasificador asigna a los vectores multidimensionales de la etapa anterior una categoría o clase.

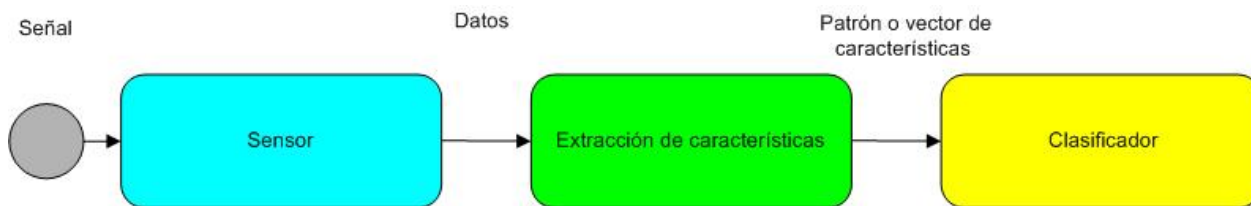


Figura 3.3: Esquema general del reconocimiento de patrones.

### 3.2.1. Aprendizaje de máquina y clasificadores

Un clasificador es una herramienta de aprendizaje de máquina que tiene por entrada conjuntos de datos y los coloca en alguna de  $k$  clases o categorías. Para construir un clasificador existen técnicas de aprendizaje de máquina que se dividen principalmente en tres grandes grupos: aprendizaje supervisado, semisupervisado y no supervisado.

Un algoritmo de aprendizaje supervisado busca aprender la relación entrada-salida  $f(x)$  (dependencia o función) usando un conjunto de datos de entrenamiento  $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)$  donde las entradas  $\mathbf{x}$  son vectores  $m$ -dimensionales y las etiquetas (respuestas del sistema) son discretas en el caso de clasificación y continuas en el caso de la regresión. Las redes neuronales artificiales y la máquina de soporte vectorial son dos de las técnicas más populares en esta área.

Hay dos tipos de algoritmos supervisados, a saber, clasificación (reconocimiento de patrones) y regresión (aproximación de funciones). En el problema de clasificación, los datos de entrenamiento consisten en ejemplos de clases distintas. El problema de clasificación más simple es uno binario que consiste en ejemplos de entrenamiento de dos clases distintas (clase  $+1$  o  $-1$ ), donde las salidas  $y_i \in \{1, -1\}$  representan la clase a la que pertenecen las etiquetas correspondientes a los respectivos vectores de entrada  $\mathbf{x}_i$ .

Los vectores  $\mathbf{x}_i$  consisten en mediciones o características que son usadas para distinguir ejemplos de diferentes clases. La tarea de aprendizaje en un problema de clasificación es construir clasificadores que puedan clasificar ejemplos no vistos  $\mathbf{x}_j$ . En otras palabras, la máquina tiene que primero aprender de ejemplos y luego deberá tomar decisiones complejas basadas en dicho aprendizaje.

A continuación se muestra el esquema general de implementación general de las técnicas de aprendizaje supervisado mediante la Figura 3.4, en la parte superior de dicha figura se tiene la Fase de entrenamiento, los datos de entrenamiento que generalmente son vectores de características etiquetados se transfieren a un algoritmo de aprendizaje supervisado como

redes neuronales o maquina de soporte vectorial,este algoritmo genera como salida un modelo o clasificador.Luego en la Fase de prueba Figura 3.4,se ingresa un conjunto de datos no vistos o de prueba al modelo obtenido en la etapa anterior y dicho modelo arroja como salida los datos clasificados.

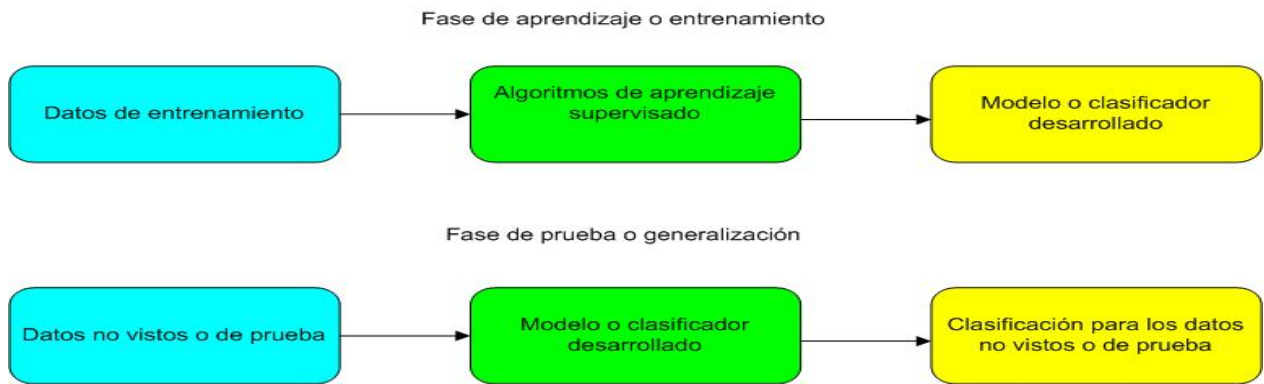


Figura 3.4: Aprendizaje supervisado

### 3.3 MÁQUINA DE SOPORTE VECTORIAL

El problema de clasificación puede restringirse a considerar exclusivamente dos clases sin pérdida de generalidad. En este problema el objetivo es separar dos clases mediante una función que se induce de los ejemplos disponibles. El objetivo es producir un clasificador que trabaje bien con ejemplos no vistos. Considere el ejemplo de la Figura 3.5 donde para los conjuntos de puntos azules y rojos existen muchos clasificadores lineales que separaran los datos (líneas grises y línea verde), pero existe solamente uno que maximiza el margen, (es decir maximizar la distancia entre el y los puntos más cercanos de cada clase). Este separador lineal es llamado máximo hiperplano separador buscando con ello obtener una mejor separación que con las otras fronteras .

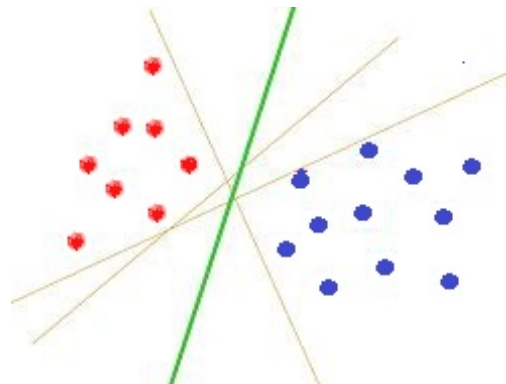


Figura 3.5: Clasificadores lineales

Analizando lo expuesto en el párrafo anterior en la Figura 3.6 de lado izquierdo se tiene

que el conjunto de puntos verdes y el conjunto de puntos rojos están separados tanto por línea verde como por la gris, y en la misma Figura 3.6 lado derecho se muestra como el margen de la línea verde formado por las líneas negras paralelas a esta, muestra un margen de separación considerablemente mayor ya que de hecho la línea verde es la línea con el máximo margen separador.

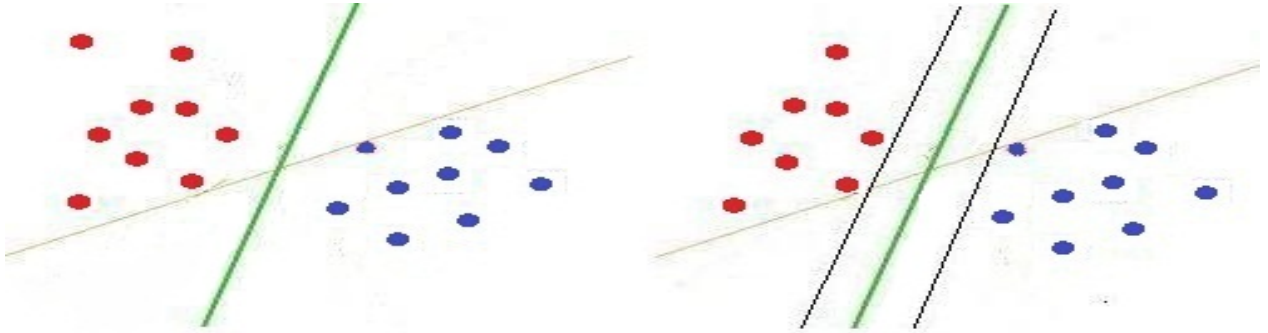


Figura 3.6: Máximo margen separador

El problema de encontrar un clasificador que maximice el margen de separación entre clases se puede ver geoméricamente como encontrar el vector unitario  $w$  y el parámetro  $b$ , con los cuales, se maximiza la distancia de los puntos de ambas clases al hiperplano. Lo cual se puede plantear de manera analítica de la siguiente forma:

Suponiendo que se tienen  $n$  ejemplos  $(x_1, y_1), \dots, (x_n, y_n)$  con etiquetas de clase  $y_i \in \{1, -1\}$ . Encontrar el hiperplano  $\langle w, x \rangle + b = 0$  (de parámetros  $(w, b)$ ) que cumpla las siguientes condiciones:

1. La escala de  $(w, b)$  es fija, es decir  $\{x_1, \dots, x_n\}$ . es decir,

$$\min_{i \leq n} |\langle w, x_i \rangle + b| = 1$$

2. El plano con parámetros  $(w, b)$  separa los  $+1$ 's de los  $-1$ 's. es decir

$$y_i(\langle w, x_i \rangle + b) \geq 0 \text{ para todo } i \leq n$$

3. El plano tiene el máximo margen  $\rho = 1/|w|$ . minimizando  $|w|^2$ .

Mediante las dos primeras condiciones se tiene que

$$y_i(\langle w, x_i \rangle + b) \geq 1 \text{ para todo } i \leq n.$$

Por tanto se busca resolver el siguiente problema de optimización

$$\text{minimizar } \frac{1}{2}|w|^2$$

para toda  $w \in R^d$  y  $b \in R$  sujetos a,

$$y_i(\langle w, x_i \rangle + b) - 1 \geq 0 \text{ para toda } i \leq n.$$

### 3.3.1. Kerneles y datos no linealmente separables

Sin embargo existen problemas de clasificación donde los datos no son linealmente separables en el espacio de entradas, como en el lado izquierdo de la Figura 3.7 donde se muestran dos conjuntos de puntos no linealmente separables a los cuales se le aplica la siguiente transformación  $\phi(x_1, x_2) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$ .

Esta transformación de  $R^2$  en  $R^3$  tiene la propiedad de realizar el calculo del producto interno de los vectores característicos  $\phi(x)$  y  $\phi(w)$  en  $R^3$  mediante la elevación del producto interno de al cuadrado de  $x$  y  $w$  en  $R^2$

$$\begin{aligned} K(x, w) &= \langle \phi(x), \phi(w) \rangle \\ &= x_1^2w_1^2 + 2x_1x_2w_1w_2 + x_2^2w_2^2 \\ &= (x_1w_1 + x_2w_2)^2 \\ &= (\langle x, w \rangle)^2 \end{aligned}$$

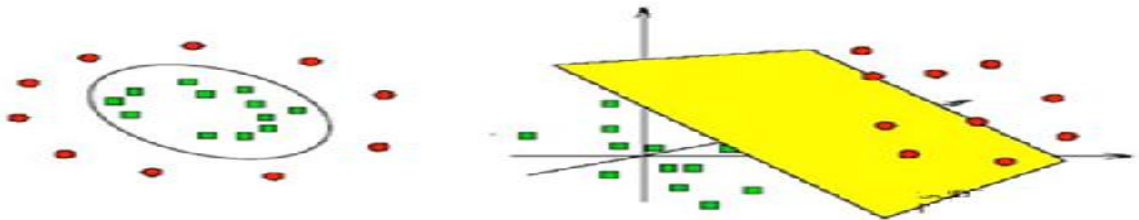


Figura 3.7: Separación de clases.

La maquina de soporte vectorial combina de manera exitosa tres importantes ideas.

- Del aprendizaje de maquina la idea de encontrar un clasificador que maximice el margen de separación entre clases, esto maximización esta en términos de un problema de optimización cuadrática.
- De programación matemática explota los algoritmos de programación cuadrática para conjuntos convexos
- Del análisis matemático la idea de representación de kernel para que los algoritmos de programación cuadrática sean computacionalmente factibles de ejecutar así como para obtener clasificadores no lineales.

# CAPÍTULO 4

## JUSTIFICACIÓN Y OBJETIVOS

### 4.1 OBJETIVOS

#### 4.1.1. Propósitos y objetivos de la investigación

La Figura 4.1 muestra una visión general del proyecto en la parte de la izquierda se describe como se planteo el objetivo y en derecha se describe la implementación del mismo.

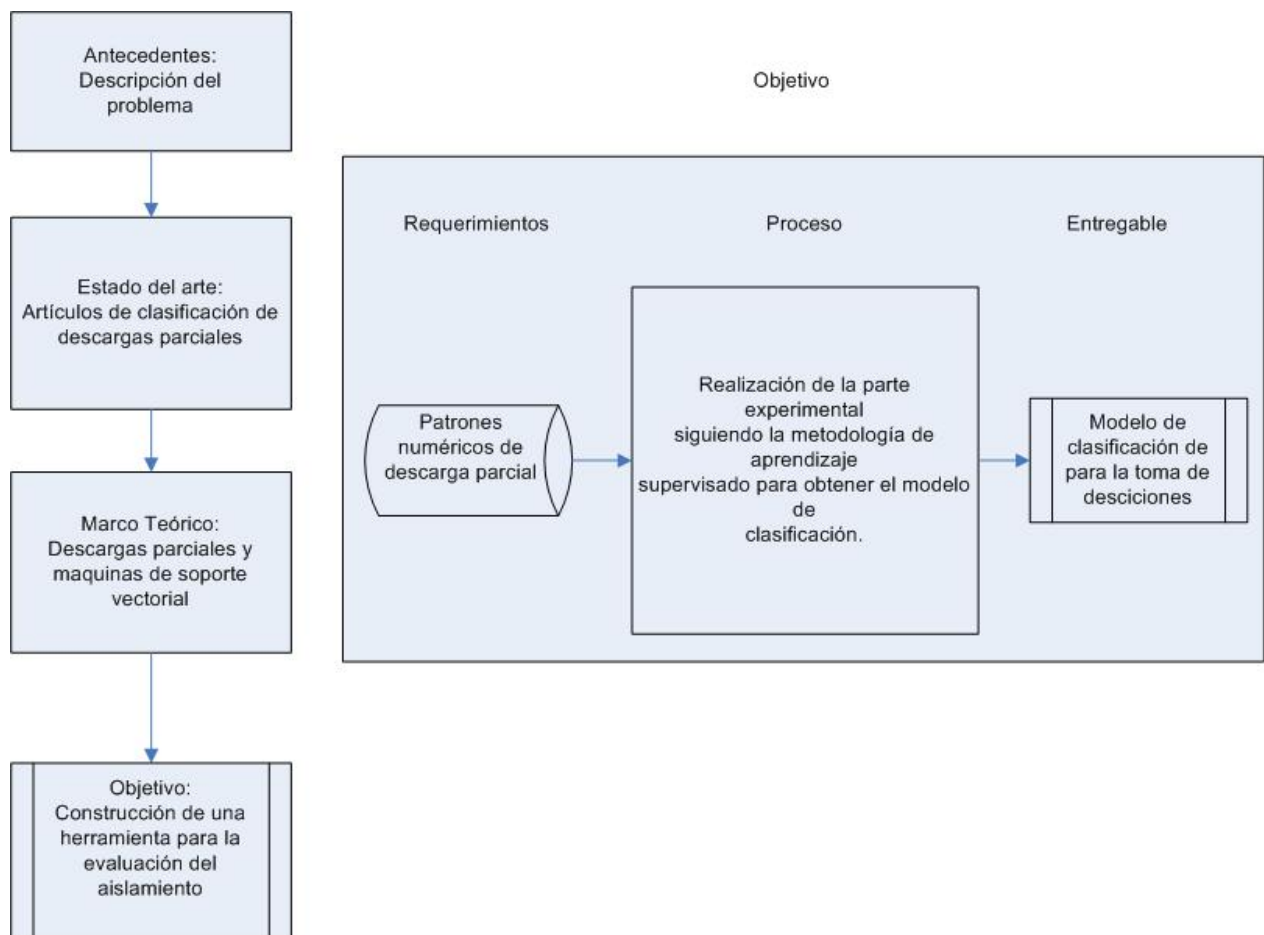


Figura 4.1: Visión general del proyecto

**Objetivo general** Construir una herramienta de evaluación del aislamiento de alto voltaje, para diagnosticar su funcionalidad.

### **Objetivos Específicos**

- Diseñar una máquina de soporte vectorial adecuada a los tipos de descarga parcial y la cantidad de datos disponibles para poder asociar el tipo de descarga parcial a los conjuntos de mediciones.
- Implementar la máquina de soporte vectorial para correr la fase de entrenamiento y la de prueba.
- Correr la fase de prueba la máquina de soporte vectorial para medir su eficiencia y tiempo de procesamiento con nuevos datos
- Adaptar la salida del máquina de soporte vectorial en términos de evaluación de la calidad para dar una herramienta de evaluación de la calidad del aislamiento.

#### **4.1.2. Hipótesis**

La máquina de soporte vectorial multiclase clasifica los patrones de descarga parcial provenientes de mediciones en sitio, maximizando la eficiencia de predicción.

# CAPÍTULO 5

## METODOLOGÍA

### 5.1 DISEÑO DE LA INVESTIGACIÓN

La Figura 5.1 muestra esquemáticamente la metodología del proyecto a la izquierda se tiene que la entrada al bloque central son los patrones numéricos de descarga parcial, el diagrama de flujo del bloque central describe brevemente el procesamiento de la información de entrada, dicho procesamiento siguió la metodología de aprendizaje supervisado para obtener un clasificador de descargas parciales que se muestra como resultado en el bloque de la derecha.

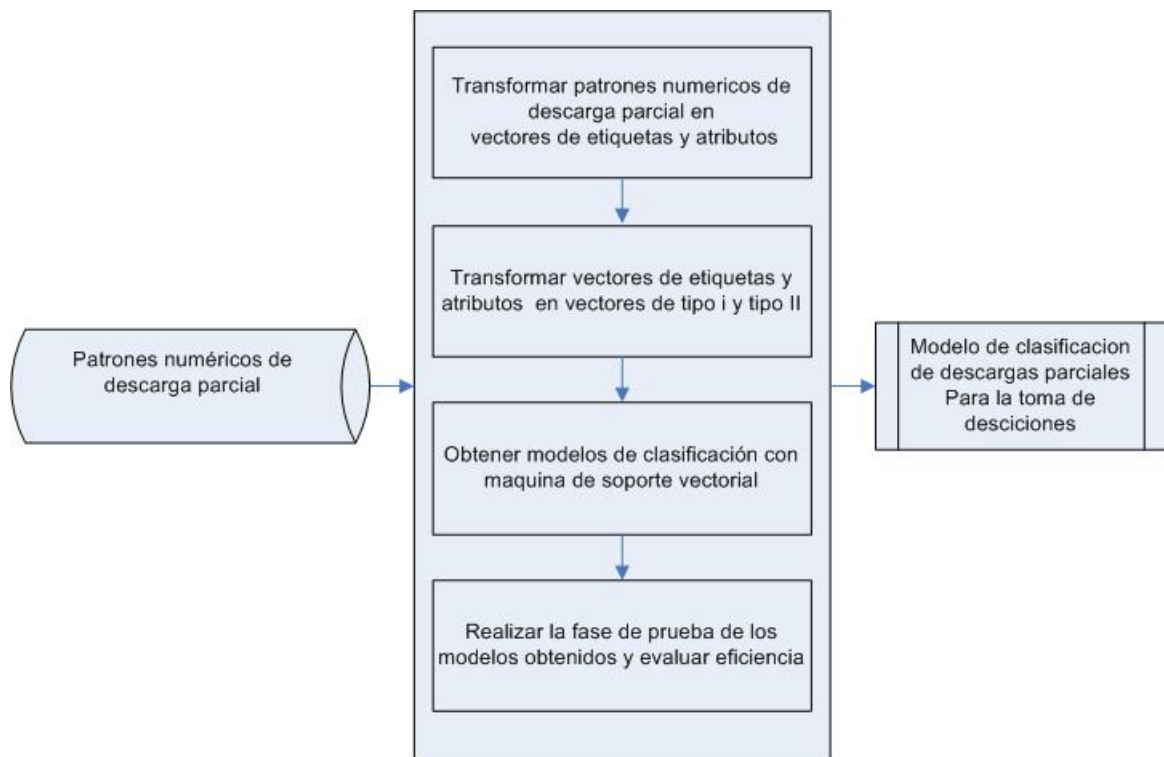


Figura 5.1: Visión general de la metodología

A continuación se detalla la metodología para la construcción de la herramienta.

- Se diseñó una máquina de soporte vectorial adecuada al número de clases a clasificar, número de atributos y cantidad de datos.



- Se implementó una maquina de soporte vectorial de acuerdo a las características del diseño.
- Se seleccionó adecuadamente los datos para la etapa de entrenamiento de la máquina de soporte vectorial.
- Se realizaron las corridas de la etapa de entrenamiento con los datos previamente seleccionados.
- Se evaluó la efectividad y el tiempo de procesamiento de la máquina de soporte vectorial
- Se realizo el ajuste de datos de la máquina de soporte vectorial.
- Se seleccionaron adecuadamente los datos para la etapa de prueba de la máquina de soporte vectorial.
- Se realizaron las corridas de la etapa de prueba de la máquina de soporte vectorial.
- Se evaluó la capacidad de predicción de la máquina de soporte vectorial.
- Se reportaron resultados en términos de evaluación de la calidad del aislamiento eléctrico.

## **5.2 DESCRIPCIÓN DE LA IMPLEMENTACIÓN DE LA METODOLOGÍA.**

La Figura 5.2 muestra esquemáticamente las actividades que se realizaron siguiendo la metodología de aprendizaje supervisado para aplicar maquina de soporte vectorial y construir el clasificador de descargas parciales, en tres etapas consecutivas, acondicionamiento de la información, la etapa de construcción de los modelos y la etapa de prueba.

A continuación se dará un listado de las condiciones experimentales y otro de las actividades realizadas para la construcción de la herramienta de evaluación ,posteriormente se da una descripción detallada de los puntos.

Se emplearon maquinas de soporte vectorial para la clasificación del tipo de descarga parcial en las siguientes condiciones:

- Los patrones de descarga parcial provienen de mediciones en sitio
- No se emplea técnicas de eliminación de ruido posteriores a la digitalización. de los patrones.
- Se clasificaron exclusivamente tipos de descarga parcial, no así los tipos de falla asociados.

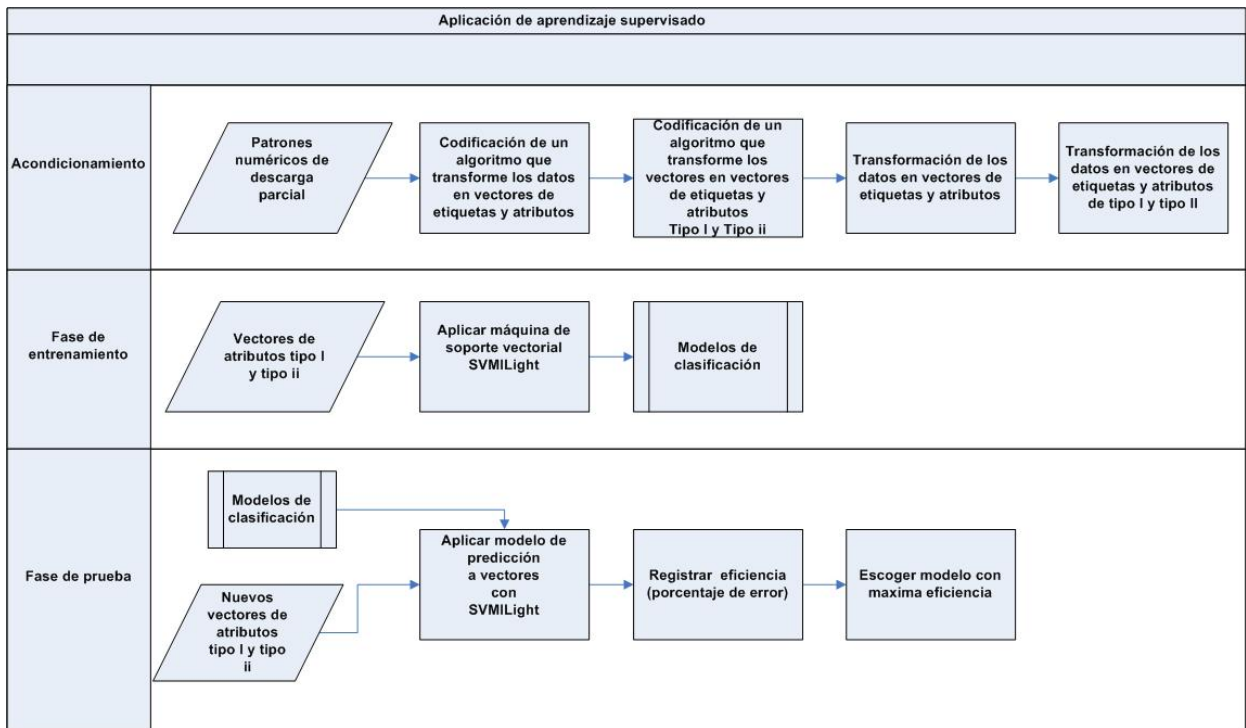


Figura 5.2: Visión general de la parte experimental

El listado de las actividades que se realizaron en la parte experimental es el siguiente:

- Se seleccionó un subconjunto de datos de mediciones de descarga parcial correspondiente a 3 tipos de patrones de descarga parcial.
- Se convirtieron los datos de mediciones a vectores de etiquetas y atributos.
- Se procesaron los vectores de etiquetas y atributos para formar vectores de etiquetas y atributos tipo I y tipo II.
- Se seleccionaron 300 vectores de entrada para entrenamiento y el resto para prueba.
- Se obtuvieron 96 modelos en la fase de entrenamiento.
- Se probaron los 96 modelos con los vectores asignados para dicho propósito.
- Se registraron los resultados de las 96 pruebas.

### 5.3 ACONDICIONAMIENTO DE LOS DATOS PARA LA FASE DE PRUEBA Y DE ENTRENAMIENTO DE LA HERRAMIENTA DE EVALUACIÓN DEL AISLAMIENTO.

Los patrones numéricos de descarga parcial provienen de mediciones en sitio las cuales se obtuvieron como se describe en la Figura 5.3 un sistema de pruebas resonante mide descarga

parcial (1) de una red subterránea (2) , posteriormente una interfaz con el equipo de computo (3) muestra algún patrón bidimensional de descargas parciales como el mostrado en la parte (4).

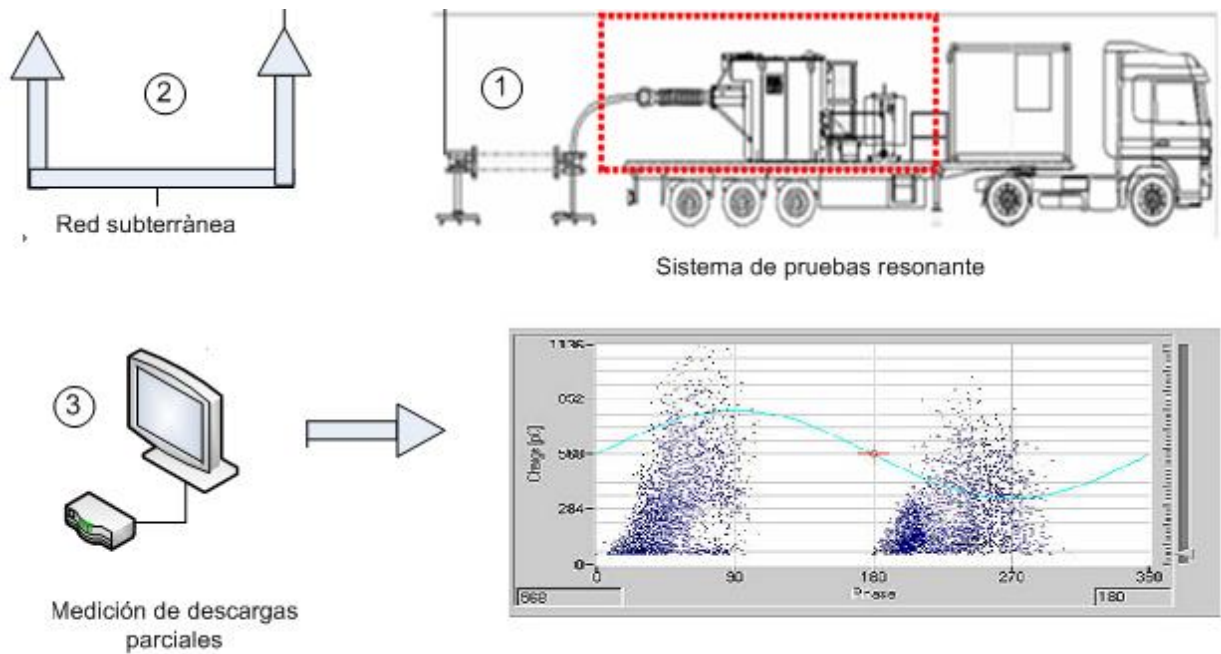


Figura 5.3: Obtención de mediciones de descarga parcial.

LAPEM proporcionó reportes de mediciones de descarga parcial en fase resuelta correspondientes a descarga parcial superficial, interna y efecto corona ,con un formato como el mostrado en la Tabla 5.1 ,donde la primera columna corresponde al ciclo  $n$  ,la segunda a la fase  $\phi$  , la tercera a la carga aparente  $q$  . y la cuarta al voltaje.

period number	phase [deg]	charge q [pC]	voltage u [kV]
679	272.0	48.89	-4.19
680	268.4	105.30	-4.18
681	74.5	42.62	3.95
681	261.2	104.67	-4.12
682	75.4	38.23	3.98
682	271.1	65.81	-4.22
⋮	⋮	⋮	⋮
2768	54.9	33.22	3.53
2768	263.7	72.71	-4.38

Tabla 5.1: Patrones de descarga parcial en forma de datos numéricos

### 5.3.1. Transformación de los patrones numéricos de descarga parcial en vectores de etiquetas y atributos

Para entrenar y probar la máquina de soporte vectorial se transformaron los datos del formato de Tabla 5.1, en vectores de etiquetas y atributos, donde las etiquetas corresponden a cada tipo de descarga parcial y los atributos al valor de la descarga aparente en cada ángulo  $\phi$  considerando  $\phi$ . Es decir para cada ciclo  $n$  se genera un vector como el que se muestra enseguida

$$(número\ de\ clase, valor\ de\ q\ en\ \phi = 1^\circ \dots valor\ de\ q\ en\ \phi = 360^\circ)$$

Donde:

- El número de clase toma valores 1, 2 o 3 según las mediciones sean de efecto corona, descarga superficial y descarga interna respectivamente.
- Atributo número  $i := entero\ mas\ cercano\ a\ \phi$
- Valor de  $i$ -ésimo atributo := *valor de carga aparente en  $\phi$*

Para utilizar la implementación de máquina vectorial del software SVMLight se generaron vectores de características con el formato siguiente:

$$Número\_de\_clase, Num\_atributo\_1 : valor\_atributo\_1, \dots, Num\_atributo\_n : \\ valor\_atributo\_n$$

Además en dicho formato cuando se omite algún atributo SVMLight lo considera de valor cero.

Para generar los vectores de características en el formato de SVMLight a partir del formato mostrado en la Tabla 5.1 se implementó un algoritmo en Excel. A continuación se da un ejemplo de un vector generado a partir de las mediciones del ciclo 681 que se encuentra en la Tabla 5.1. Como los datos de dicha tabla provienen de mediciones de descarga parcial superficial el número de clase es 2, y como las mediciones registradas corresponden a 74.5 y 261.2, y el algoritmo redondea estos valores de fase a los valores más próximos entonces obtenemos el siguiente vector.

$$2, 74 : 42,62, \dots, 261 : 104,67$$

La Figura 5.4 muestra el resultado de generar los vectores de características a partir de la muestra de datos que se encuentra en la Tabla 5.1

Adicionalmente se transformaron los vectores de etiquetas y atributos en vectores de etiquetas y atributos tipo I y tipo II, para generar los 96 modelos de predicción en la fase de entrenamiento y encontrar el que maximice la predicción en la fase de prueba, a continuación se proporciona una descripción de la mencionada transformación.

period number	phase [deg]	charge q [pC]	voltage u [kV]				
				2	272:48.89		
679	272.0	48.89	-4.19	2	268:105.30		
680	268.4	105.30	-4.18	2	74:42.62	261:104.67	
681	74.5	42.62	3.95	2	75:38.23	271:65.81	
681	261.2	104.67	-4.12				.
682	75.4	38.23	3.98				.
682	271.1	65.81	-4.22				.
⋮	⋮	⋮	⋮	2	55:33.22	254:72.71	
2768	54.9	33.22	3.53				
2768	263.7	72.71	-4.38				

Figura 5.4: Mediciones transformadas a vectores de etiquetas y atributos

### Vectores de etiquetas y atributos tipo I

A partir de 300 vectores de 360 atributos a los que se denotará por  $X^i$  se generaron 300 vectores de método de entrada I con 360 atributos a los que se denotará por  $E_1^i$  de la siguiente manera:

Sea  $k$  el número de atributo de  $X^i$  y en  $E_1^i$  y por  $X^i(k)$  y  $E_1^i$  el valor del  $k$ -ésimo atributo en  $X^i$  y  $E_1^i$  respectivamente luego

---

#### Pseudocódigo 1

---

Para cada vector de características  $X^i$

Calcular el promedio y desviación estándar  $\mu_1, \sigma_1$  de las entradas  $X^i(k)$  tales que  $1 \leq k \leq 180$

Calcular el promedio y desviación estándar  $\mu_2, \sigma_2$  de las entradas  $X^i(k)$  tales que  $181 \leq k \leq 360$

Para  $k = 1$  hasta 180

Hacer las entrada  $E_1^i(k) = Z(X^i(k), \mu_1, \sigma_1)$

Para  $k = 181$  hasta 360

Hacer las entrada  $E_1^i(k) = Z(X^i(k), \mu_2, \sigma_2)$

---

Es decir las primeras 180 entradas de  $E_1^i$  las primeras 180 normalizadas de  $X^i(k)$  análogamente las entradas de la 181 a la 360 de  $E_1^i$  son las entradas normalizadas de  $X^i(k)$  de 181 a la 360.

### Vectores de etiquetas y atributos tipo II

Por cada  $n$  vectores de 360 atributos (para el presente trabajo  $n = 300$ ) a los que se denotará por  $X^i$  se generó un vector de método de entrada II de 720 atributos al que se denotará por  $E_2$  de la siguiente manera:

---

## Pseudocódigo 2

---

Sea  $k$  el número de atributo de  $X^i$  y  $E$  y  $X^i(k)$  y  $E_1^i(k)$  el valor del  $k$ -ésimo atributo en  $X^i$  y  $E_2$  respectivamente y sea  $N_k$  el número de veces que  $X^i(k)$  es diferente de cero.

Para  $k = 1$  hasta 360

Hacer las entrada  $E_2(k) = (\sum_{i=1}^n X^i(k))/n$  el promedio del  $k$ -ésimo atributo.

Para  $k = 1$  hasta 360

Hacer las entrada  $E_2(360 + k) = N_k$

---

Las primeras 360 entradas de  $E_2$  son el promedio de  $k$ -ésimo atributo de los  $n = 300$  vectores  $X^i$  y los restantes 360 atributos el número de veces que el número de veces que  $X^i(k)$  es diferente de cero.

### 5.3.2. Fase de entrenamiento

Con los datos de mediciones de descarga parcial transformados a vectores de etiquetas y atributos se procedió a realizar la fase de entrenamiento de la máquina de soporte vectorial donde se obtuvieron los modelos de clasificación Figura 5.5



Figura 5.5: Fase de entrenamiento de la máquina de soporte vectorial

Los modelos de clasificación se obtuvieron variando el tipo de vectores de etiquetas y atributos, tipo de kernel con sus respectivos parámetros y el parámetro  $C$ , estas 96 posibilidades se muestran mediante un diagrama de árbol Figura 5.6

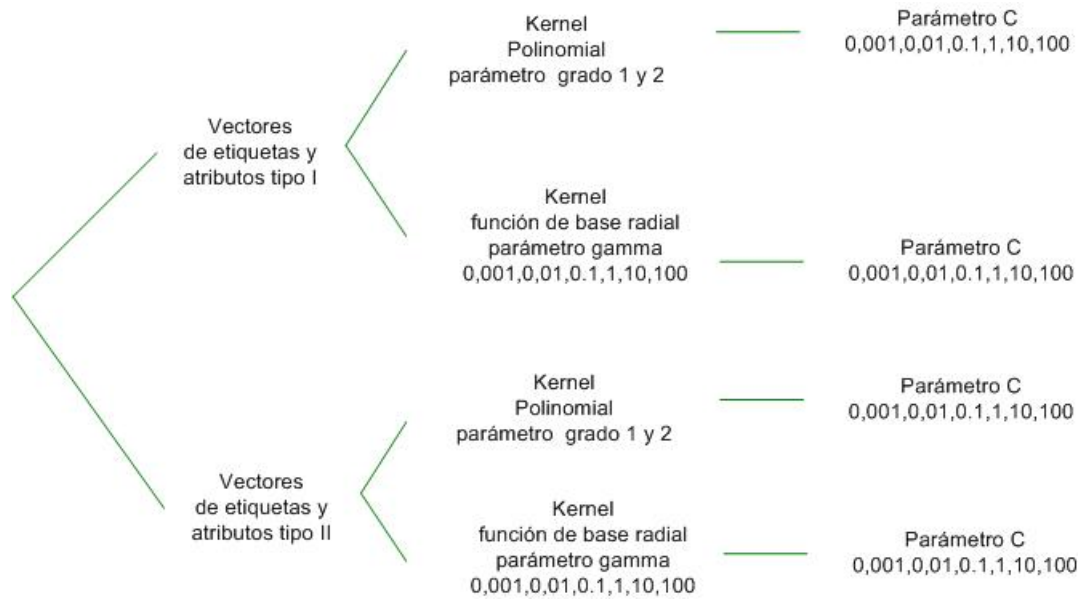


Figura 5.6: Parámetros para la generación de modelos de clasificación de descarga parcial

## 5.4 FASE DE PRUEBA

En esta etapa se probó la eficiencia de los modelos obtenidos durante la fase de entrenamiento aplicando cada modelo obtenido a un conjunto de vectores de etiquetas y atributos de tipo I y II, Figura 5.7

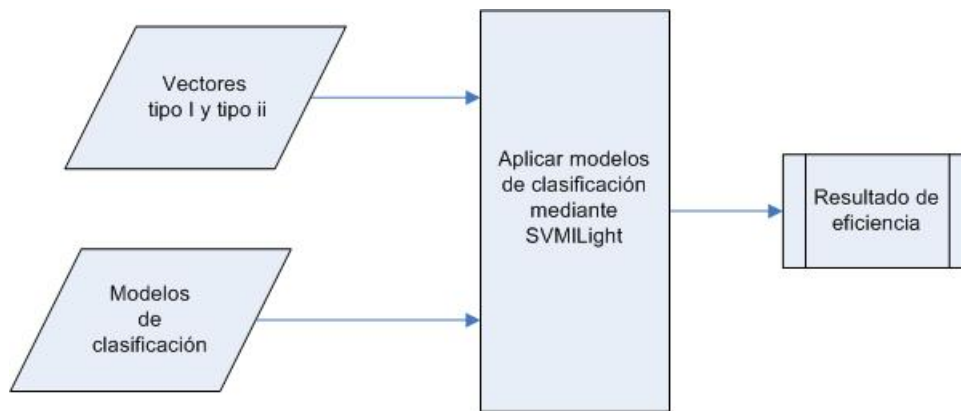


Figura 5.7: Prueba de la eficiencia de clasificación los modelos obtenidos

Este procedimiento se repitió para cada uno de los modelos generados. Considerando los datos de generación del modelo y el tipo de vectores de etiquetas y atributos entrada se registraron los resultados de esta etapa.

# CAPÍTULO 6

## RESULTADOS

### 6.1 RESULTADOS OBTENIDOS DURANTE LA FASE DE PRUEBA

En la fase de prueba se midió la eficiencia de predicción de los modelos generados en la fase de entrenamiento, la eficiencia de predicción se expresa mediante el porcentaje de error, es decir a menor porcentaje de error mayor eficiencia de predicción.

La Tabla 6.1 muestra los porcentajes de error obtenidos en la fase de prueba para los 12 modelos de predicción generados usando vectores de tipo I ,variando el parámetro C y el parámetro del kernel polinomial dando como resultado que el minino porcentaje de error es de 50.63 % y se obtiene para el modelo con grado 2, y parámetro C igual a 100.

Tabla 6.1: Porcentaje de error utilizando método de entrada I y kernel polinomial

Parámetros	C=0.001	C=0.01	C=0.1	C=1	C=10	C=100
grado=1	66.64 %	66.64 %	66.64 %	66.18 %	66.64 %	66.71 %
grado=2	66.64 %	52.37 %	65.24 %	65.24 %	61.49 %	<b>50.63 %</b>

La Tabla 6.2 muestra los porcentajes de error obtenidos en la fase de prueba para los 12 modelos de predicción generados usando vectores de tipo II ,variando el parámetro C y el parámetro del kernel polinomial dando como resultado que el minino porcentaje de error es de 0.0 % y se obtiene para todos los modelos con grado I.

Tabla 6.2: Porcentaje de error utilizando método de entrada II y kernel polinomial

Parámetros	C=0.001	C=0.01	C=0.1	C=1	C=10	C=100
grado=1	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %
grado=2	6.67 %	6.67 %	6.67 %	6.67 %	6.67 %	6.67 %

La Tabla 6.3 muestra los porcentajes de error obtenidos en la fase de prueba para los 30 modelos de predicción generados usando vectores de tipo I ,variando el parámetro C y el parámetro de kernel de base radial dando como resultado que el porcentaje de error es de 53.33 % para todos los obtenidos de esta forma.

La Tabla 6.4 muestra los porcentajes de error obtenidos en la fase de prueba para los 36 modelos de predicción generados usando vectores de tipo I ,variando el parámetro C y el



Tabla 6.3: Porcentaje de error utilizando método de entrada I y kernel función de base radial

Parámetros	$\gamma = 0,001$	$\gamma = 0,01$	$\gamma = 0,1$	$\gamma = 1$	$\gamma = 10$
c=0.001	53.33 %	53.33 %	53.33 %	53.33 %	53.33 %
C=0.01	53.33 %	53.33 %	53.33 %	53.33 %	53.33 %
C=0.1	53.33 %	53.33 %	53.33 %	53.33 %	53.33 %
C=1	53.33 %	53.33 %	53.33 %	53.33 %	53.33 %
C=10	53.33 %	53.33 %	53.33 %	53.33 %	53.33 %
C=100	53.33 %	53.33 %	53.33 %	53.33 %	53.33 %

parámetro de kernel de base radial dando como resultado que el minino porcentaje de error es de 40.09 % para gamma igual a 100 y para los valores de C de 10 y 100.

Tabla 6.4: Porcentaje de error utilizando método de entrada I y kernel función de base radial

Parámetros	$\gamma = 0,001$	$\gamma = 0,01$	$\gamma = 0,1$	$\gamma = 1$	$\gamma = 10$	$\gamma = 100$
C=0.001	66.77 %	66.71 %	66.64 %	66.18 %	56.10 %	40.36 %
C=0.01	66.71 %	66.71 %	66.64 %	66.18 %	56.10 %	40.29 %
C=0.1	66.77 %	66.77 %	66.64 %	66.18 %	56.10 %	40.23 %
C=1	66.77 %	66.64 %	66.64 %	66.18 %	56.10 %	40.23 %
C=10	66.64 %	57.18 %	66.36 %	66.18 %	56.37 %	40.09 %
C=100	54.50 %	53.71 %	66.44 %	66.18 %	56.44 %	40.09 %

La Tabla 6.5 muestra los mininos porcentajes de error obtenidos variando los tipos de entrada así como el kernel y sus parámetros ,encontrando el minino porcentaje de error se obtuvo para el modelo generado utilizando vectores de etiquetas y atributos de tipo II, kernel polinomial de grado 1,y para todos los valores de C.

Tabla 6.5: Mínimo porcentaje de error alcanzado

	kernel función de base radial	kernel polinomial
Método de entrada I	50.63 % , Para toda $\gamma, c$	40.04 % grado=2,c=100
Método de entrada II	53.33 % Para toda $\gamma, c$	<b>0 % ,grado=1,c</b>

## 6.2 CONCLUSIONES

En las redes eléctricas subterráneas de alta tensión (115, 230 y 400 kV) que proyecta CFE ,el aislamiento polimérico del cable presenta poca tolerancia a la degradación debido al fenómeno de descargas parciales, teniendo como consecuencia la suspensión del suministro de energía eléctrica.Por ello es indispensable determinar el estado de la calidad del aislamiento eléctrico de dichas redes mediante la clasificación del tipo de descarga parcial ya que este es un estándar para evaluar la calidad del aislamiento de alto voltaje.

Actualmente la clasificación de descargas parciales se realiza por inspección visual de patrones bidimensionales de descarga parcial,aunque se han realizado esfuerzos para auto-

matizar la clasificación mediante sistemas expertos, redes neuronales entre otras técnicas de aprendizaje de máquina.

En el estado del arte existen diversas aplicaciones de las técnicas de aprendizaje de máquina en ingeniería industrial como el clasificador en base a redes neuronales aplicado a la calidad de acero laminado en caliente (Browne, 1999), en particular en las aplicaciones de aprendizaje de máquina a la evaluación del aislamiento se encontró que se han utilizado redes neuronales y máquinas de soporte vectorial como en Lai y Blackburn (2010), estos últimos utilizados con datos de descarga parcial producidos en laboratorio.

Debido a lo anterior se determinó como objetivo del presente trabajo construir una herramienta de evaluación del aislamiento de alto voltaje para diagnosticar su funcionalidad, es decir construir un clasificador de descargas parciales y por lo encontrado en la literatura se construyó con la técnica de aprendizaje supervisado conocida como máquina de soporte vectorial.

Durante la implementación de la metodología se programaron algoritmos para transformar los reportes de medición de descargas parciales en vectores de etiquetas y atributos tipo I y tipo II, posteriormente teniendo como entrada dichos vectores se procedió a obtener modelos de clasificación mediante la técnica de máquina de soporte vectorial con la implementación SVM light.

De los resultados de la fase de prueba se concluye que el mejor modelo de clasificación es el que se generó utilizando vectores de etiquetas y atributos de tipo II, kernel polinomial de grado 1, para todos los valores de  $C$  entre 0.001 y 100, como en Lai y Blackburn (2010) se obtuvieron mejores resultados con kernels polinomiales que usando base radial y en el presente trabajo se utilizaron mediciones en sitio de descarga parcial.

# Referencias

- Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., 2006.
- Will Browne. *The Development of an Industrial Learning Classifier System for Application to a Steel Hot Strip Mill*. Tesis doctoral, University of Wales, Cardif, 1999.
- Min-Yen Chiu Huang Chih-Hsien Shih-Shong Yen Chang-Hsing Lee, Lin Yu-Chih y Chiang Haeng. Recognition of partial discharge defects in cable terminations. En *International Conference on Condition Monitoring and Diagnosis*. 2008.
- Pao-Hua Chou, Menq-Jiun Wu, y Kuang-Ku Chen. Integrating support vector machine and genetic algorithm to implement dynamic wafer quality prediction system. *Expert Syst. Appl.*, 37(6):4413–4424, 2010.
- Hüseyin Erişti y Yakup Demir. A new algorithm for automatic classification of power quality events based on wavelet transform and svm. *Expert Syst. Appl.*, 37(6):4094–4102, 2010. ISSN 0957-4174.
- Hao y Lewin. Partial discharge source discrimination using a support vector machine. *IEEE Transactions*, 17:189 – 197, 2010.
- International Electrotechnical Commission Standard IEC60270. High voltage testing - partial discharge measurements switzerland,third edition. 2000.
- E. Kuffel J. Kuffel y W. S. Zaengl. *High Voltage Engineering Fundamentals*. Newnes, 2000.
- LIU Ding JIAO Shang-bin. Forecasting the esdd of insulator based on least squares support vector machine. En *Proceeding of the CSEE 2006*, págs. 149–153. 2006.
- Dieter Kind y Kurt Feser. *High Voltage Test Techniques*. Newnes, 2001.
- Parekh B.R. Kulkarni, S.V. Partial discharge in high voltage equipments-hv cable. En *IEEE 9th International Conference on the Properties and Applications of Dielectric Materials*. 2009.
- Phung BT Lai, KX y Blackburn. Application of data mining on partial discharge part i: Predictive modelling classification. *IEEE Transactions on Dielectrics and Electrical Insulation*, 17:852 – 860, 2010.

- Tom Mitchell. The discipline of machine learning. Inf. téc., School of Computer Science, Carnegie Mellon University, 2007.
- Mirrasoul J. Mousavi y Karen L. Butler-Purry. A Novel Condition Assessment System for Underground Distribution Applications. *IEEE Transactions on Power Systems*, 24:1115–1125, 2009.
- Ryan Richard Otte. Properties of non-uniform electric fields that contribute to partial discharges. 2006.
- Jens Tobias Pfingsten. *Machine Learning for Mass Production and Industrial Engineering*. Tesis doctoral, Fakultät für Mathematik und Physik der Eberhard-Karls-Universität zu Tübingen, 2007.
- Magdy M. A. Salama y R. Bartnikas. Determination of neural-network topology for partial discharge pulse pattern recognition. *IEEE Transactions on Neural Networks*, 13:446–456, 2002.
- L. Satish y B. Nazneen. Wavelet-based denoising of partial discharge signals buried in excessive noise and interference. *IEEE Transactions on Dielectrics and Electrical Insulation*, 10:354–367, 2003.