



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA

INGENIERÍA ELÉCTRICA – SISTEMAS ELÉCTRICOS DE POTENCIA

DESARROLLO Y APLICACIÓN DE UN MODELO PREDICTIVO, PARA GESTIÓN DE  
MANTENIMIENTO EN SUBESTACIONES DE DISTRIBUCIÓN DE ENERGÍA ELÉCTRICA EN REDES  
DE MEDIA TENSIÓN

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN INGENIERÍA

PRESENTA:

ADRIEL RIEGO MARTINEZ

TUTOR PRINCIPAL

DR. CESAR ÁNGELES CAMACHO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA

CIUDAD DE MÉXICO, SEPTIEMBRE, 2018



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**JURADO ASIGNADO:**

Presidente: Dr. Juan Ramón Rodríguez Rodríguez  
Secretario: Dr. Gonzalo Sandoval Rodríguez  
1er. Vocal: Dr. César Angeles Camacho  
2do Vocal: Dr. Claudio Rubén Fuerte Esquivel  
3er. Vocal: Dr. Daniel Guillen Aparicio

Lugar o lugares donde se realizó la tesis: Instituto de Ingeniería, UNAM

**TUTOR DE TESIS:**

Dr. Cesar Ángeles Camacho

-----  
**FIRMA**

PENSAMIENTO

*“Nunca consideres el estudio como un deber, sino  
como una oportunidad para penetrar en el  
maravilloso mundo del saber”*

*Albert Einstein*

## DEDICATORIA

- *A mis padres: Con todo el cariño y amor para las personas que hicieron todo en la vida para que yo pudiera lograr mis sueños, entregándome en cada momento amor, ternura, seguridad y confianza, a ustedes por siempre mi corazón, los amo.*
- *A mi esposa Dianelis: Por hacerme el hombre más feliz del mundo, gracias por entenderme, apoyarme, ser mi compañera, amiga, confidente y permitirme ser parte de tu vida.*
- *A mi abuela Mima: Que, ya no es parte de este mundo e inculcó valores, amor, cariño y ternura en mi persona, por lo que muy a pesar de que no está presente en este momento de mi vida, muchas gracias, pues esté donde esté ha seguido cuidándome y guiándome a través de este largo camino que es la vida.*
- *A mis amados abuelos Papi y Haydee, a mi tío Oscar, a mi hermano, a mis tías Dunia, Alina, Ileana y familia en general: Por estar siempre que los necesito brindándome su apoyo incondicional y compartir conmigo buenos y malos momentos.*
- *A todos los amigos que han influido en mi formación, en especial a mi hermano Luis Miguel.*

## AGRADECIMIENTOS

- *A mi tutor César Ángeles Camacho por su dedicación, interés y apoyo incondicional para la realización de este trabajo.*
- *Agradezco encarecidamente a mis familiares y amigos por su constante preocupación.*
- *A mis compañeros de estudio, por compartir los momentos buenos y difíciles durante nuestra trayectoria estudiantil.*
- *Al personal de CFE por la colaboración brindada y en especial al M.I. Nabucodonosor Solís Ramos por el apoyo incondicional hacia mí.*
- *A la Universidad Nacional Autónoma de México y al Instituto de Ingeniería, por haberme dado la oportunidad de continuar con mi formación académica profesional.*
- *Al Consejo Nacional de Ciencia y Tecnología (CONACyT), por proporcionarme los recursos económicos durante este posgrado, sin los cuales no hubiese sido posible la realización de esta tesis de maestría.*
- *A todas las personas que de una forma u otra han colaborado para el desarrollo de este trabajo.*

*A todos muchas gracias*

## RESUMEN

El costo de mantenimiento en equipos de subestaciones eléctricas es una parte importante del costo total del ciclo de vida de éstos. Estimar de manera precisa la condición de estos equipos va a permitir la planificación de mantenimiento basada en la condición de estos activos. En esta tesis se realiza una investigación sobre la gestión de activos y su relación con análisis de riesgo y gestión de mantenimiento.

Se analizan los métodos predictivos utilizados en la literatura para detección de fallas en transformadores. Se propone la utilización de redes neuronales artificiales para la implementación de un algoritmo para el monitoreo de la condición en transformadores

El algoritmo considera la indicación del deterioro de varios sistemas de monitoreo de la condición para formular una estrategia de mantenimiento óptima para el componente dañado. Se propone un enfoque de monitoreo de condición basado en una red neuronal artificial (RNA) utilizando los datos almacenados en el sistema de control, supervisión y adquisición de datos (SCADA). El algoritmo propuesto utiliza una medida de distancia estadística llamada distancia de Mahalanobis para identificar cualquier operación anormal del componente monitoreado.

El enfoque de monitoreo de condición basado en RNA propuesto se aplica para el monitoreo de la temperatura en el aceite de dos transformadores similares, con datos SCADA reales, de dos transformadores del mismo fabricante, con una potencia nominal de 36 MVA. Los resultados muestran que el algoritmo propuesto es capaz de detectar daños en el transformador antes de que ocurra una falla. La aplicación del algoritmo de monitoreo de la condición propuesto tiene el potencial de reducir la probabilidad de fallas inesperadas y aumenta la confiabilidad del equipo.

## ABSTRACT

The cost of maintenance in the equipment of electrical substations is an important part of the total life-cycle cost of them. Estimating accurately the equipment conditions will allow maintenance planning based on the conditions of these assets. In this thesis, an investigation of assets management and its relationship with risk analysis and maintenance management is carried out.

The predictive methods used in the literature for the detection of faults in transformers are analyzed. The use of artificial neural networks is proposed for the implementation of an algorithm for the monitoring of the condition of electrical transformers.

The algorithm considers the indication of the deterioration of several conditions monitoring systems to formulate an optimal maintenance strategy for the damaged component. A condition monitoring approach based on an artificial neural network (ANN) using the data stored in the control, supervision, and data acquisition system (SCADA) is proposed. The proposed algorithm uses a measure of statistical distance called Mahalanobis distance to identify any abnormal operation of the monitored component.

The proposed RNA-based approach for condition monitoring is applied for monitoring the oil temperature of two similar transformers, with real SCADA data, of two transformers from the same manufacturer, both with a nominal power of 36 MVA. The results show that the proposed algorithm is able to detect damages in the transformer before a fault occurs. The proposed condition monitoring algorithm has the potential to reduce the probability of unexpected failures and increase the reliability of the equipment.



## ÍNDICE

PENSAMIENTO .....	III
DEDICATORIA .....	IV
AGRADECIMIENTOS .....	V
RESUMEN.....	VI
ABSTRACT .....	VII
ÍNDICE.....	VIII
ÍNDICE DE FIGURAS.....	XI
ÍNDICE DE TABLAS .....	XIII
ACRÓNIMOS .....	XIV
VARIABLES.....	XV
CAPÍTULO I: INTRODUCCIÓN.....	1
1.1.    Introducción .....	1
1.2.    Antecedentes y planteamiento del problema.....	1
1.3.    Hipótesis.....	3
1.4.    Objetivo general y objetivos específicos.....	4
1.5.    Justificación.....	4
1.6.    Metodología .....	5
1.7.    Organización de la tesis.....	6
1.8.    Metas .....	6
CAPÍTULO II: ASPECTOS GENERALES DE TEORÍA SOBRE GESTIÓN DE ACTIVOS, RIESGO Y MANTENIMIENTO .....	6
2.1.    Introducción .....	6
2.2.    Gestión de activos .....	6
2.2.1. ¿Qué es un activo?.....	6
2.2.2. Beneficios de la gestión de activos .....	7
2.3.    Gestión de activos en redes eléctricas. Clasificación. ....	8
2.4.    Gestión de riesgo.....	11
2.4.1. Análisis de riesgo.....	12
2.4.2. Estándares de gestión de riesgos .....	14
2.5.    Gestión de mantenimiento.....	15
2.5.1. Conceptos y definiciones .....	15
2.5.2. Gestión de mantenimiento en redes eléctricas .....	19
2.6.    Gestión de mantenimiento en transformadores de potencia.....	20
2.6.1. Diagnóstico y mantenimiento en los transformadores de potencia .....	21
2.6.2. Estadísticas de fallas en transformadores de potencia.....	21
2.6.3. Identificación de los sistemas y componentes críticos.....	22

2.6.4. Selección de las variables para fundamentar el mantenimiento predictivo basado en la condición.....	23
2.6.5. Definición de los modos de fallas de los componentes críticos.....	23
CAPÍTULO III: INTRODUCCIÓN AL ESTUDIO DE MÉTODOS PREDICTIVOS Y SU APLICACIÓN EN REDES ELÉCTRICAS DE POTENCIA.....	27
3.1. Introducción.....	27
3.2. Cadenas de Márkov.....	27
3.2.1. Procesos Estocásticos.....	27
3.2.2. Cadenas de Márkov.....	28
3.2.3. Aplicación de las Cadenas de Márkov en transformadores.....	29
3.3. Método de Montecarlo.....	30
3.3.1. Descripción del algoritmo.....	30
3.3.2. Aplicación del método de Montecarlo en REP.....	31
3.4. Aprendizaje Automático.....	32
3.4.1. Datos de muestra y modelos de construcción.....	33
3.4.2. Aprendizaje no supervisado.....	34
3.4.3. Aprendizaje supervisado.....	34
3.5. Problemas inherentes al Aprendizaje Automático.....	37
3.5.1. Sobre- y bajo-ajuste.....	37
3.5.2. Validación cruzada.....	38
3.5.3. Análisis de series temporales.....	38
3.5.4. Normalización.....	39
3.5.5. Maldición de dimensionalidad y selección de características.....	39
3.6. Modelos de Aprendizaje Automático.....	40
3.6.1. Redes neuronales.....	40
3.6.1.1. El Perceptrón.....	40
3.6.1.2. Red neuronal multicapa con conexiones hacia adelante.....	41
3.6.1.3. Red recurrente multicapa.....	42
3.6.2. Métodos de aprendizaje en RNA.....	43
3.6.2.1. Aprendizaje por lotes.....	43
3.6.2.2. Aprendizaje en línea.....	43
3.6.2.3. Algoritmo de aprendizaje de Levenberg-Marquardt.....	44
3.6.2.4. Funciones de activación.....	45
3.6.3. Máquinas de soporte vectorial.....	46
3.6.4. Máquina de puntos Bayes.....	47
3.6.5. Regresión logística.....	48
3.6.6. Conjunto de aprendizaje.....	48
3.6.7. Árboles de decisión.....	49
3.6.8. Bosques aleatorios.....	49
3.6.9. Jungla de decisión.....	49
3.6.10. Árbol rápido.....	49

CAPÍTULO IV: APLICACIÓN DE REDES NEURONALES ARTIFICIALES (RNA) UTILIZANDO EL SISTEMA DE MONITOREO DE TRANSFORMADORES DE DISTRIBUCIÓN .....	50
4.1. Introducción .....	50
4.2. Filosofía del desarrollo .....	50
4.3. Método propuesto para el monitoreo de la condición en transformadores.....	51
4.3.1. Matriz de correlación y correlaciones parciales .....	51
4.3.2. Modelado de RNA para el monitoreo de la condición .....	53
4.3.3. Estructura del modelo de RNA .....	54
4.3.4. Detección de falla.....	55
4.3.4.1. Distancia de Mahalanobis .....	56
4.3.4.2. Valor del umbral.....	58
CAPÍTULO V: APLICACIÓN DEL ALGORITMO PROPUESTO A CASOS DE ESTUDIO REALES .....	61
5.1. Introducción .....	61
5.2. Características principales de los transformadores de prueba .....	61
5.3. Selección de variables .....	62
5.3.1. Análisis de correlación .....	63
5.3.2. Selección de los datos de entrenamiento .....	63
5.4. Comparación y selección entre modelos de RNA NARX y <i>Feed-Forward</i> .....	65
5.5. Comparación y selección de estructuras para el modelo de RNA propuesto .....	67
5.5. Estimación de temperatura promedio diaria.....	75
5.6. Detección de fallas utilizando la distancia de Mahalanobis.....	77
5.8. Análisis comparativo contra otros algoritmos similares .....	80
CAPÍTULO VI: CONCLUSIONES Y TRABAJOS FUTUROS .....	84
6.1. Conclusiones .....	84
6.2. Trabajos futuros.....	84
BIBLIOGRAFÍA .....	86
APÉNDICES .....	94
Apéndice A: Matriz de correlaciones parciales para el Transformador T1 .....	94
Apéndice B: Teoría de redes neuronales .....	98
Apéndice C: Regla delta generalizada.....	108
Apéndice D: Método de Levenberg-Marquardt para el entrenamiento de RNA .....	112
Apéndice E: Datos de los transformadores utilizados como casos de estudio .....	116

## ÍNDICE DE FIGURAS

FIGURA 2.1. FUNCIONALIDADES DE LA RED ELÉCTRICA CLASIFICADA EN FUNCIÓN DEL TIEMPO EN LA GESTIÓN OPERATIVA, DE MANTENIMIENTO Y DE PLANIFICACIÓN [11].....	10
FIGURA 2.2. PROCESO GENERAL DE TOMA DE DECISIONES DE GESTIÓN DE RIESGOS [21] .....	15
FIGURA 2.3. TIPOS DE ESTRATEGIAS DE MANTENIMIENTO [9] .....	16
FIGURA 2.4. EVOLUCIÓN DEL MANTENIMIENTO EN LA IMPLEMENTACIÓN DE POLÍTICAS.....	20
FIGURA 2.5. PORCENTAJE DE FALLA POR ELEMENTO DE TRANSFORMADORES DE POTENCIA, SEGÚN A) CIGRE Y B) CFE MÉXICO RESPECTIVAMENTE .....	22
FIGURA 3.1. MODELO DE APRENDIZAJE SUPERVISADO .....	35
FIGURA 3.2. MODELO DE REGRESIÓN .....	36
FIGURA 3.3. MODELO DE CLASIFICACIÓN.....	36
FIGURA 3.4. MODELO DE SOBRE- Y BAJO-AJUSTE.....	37
FIGURA 3.5 MODELO DEL PERCEPTRÓN [60] .....	41
FIGURA 3.6. ESTRUCTURA DE RED NEURONAL .....	42
FIGURA 3.7. ESTRUCTURA DE RNA RECURRENTE .....	42
FIGURA 3.8. FUNCIÓN DE ACTIVACIÓN DE LA RNA. A) FUNCIÓN ESCALÓN, B) FUNCIÓN SIGMOIDEA.....	46
FIGURA 3.9. MODELO DE UNA SVM'S KERNEL [67].....	47
FIGURA 3.10. MODELO DE SEPARACIÓN ÓPTIMA [77].....	47
FIGURA 4.1. MODELO DE RNA CON ESTRUCTURA NARX [86] .....	53
FIGURA 4.2. COMPARACIÓN ENTRE LOS RESULTADOS DE LA RNA Y LOS DATOS MEDIDOS POR EL SCADA .....	55
FIGURA 4.3. A) COMPARACIÓN SCADA Y RESULTADOS DE LA RNA; B) DISTANCIA DE MAHALANOBIS .....	57
FIGURA 4.4. AMPLIACIÓN DE LA FIGURA 4.3 .....	58
FIGURA 4.5. ALGORITMO GENERAL DE MONITOREO DE LA CONDICIÓN UTILIZANDO RNA.....	60
FIGURA 5.1. ANÁLISIS EN COMPONENTES PRINCIPALES DEL CONJUNTO DE DATOS DE ENTRENAMIENTO .....	62
FIGURA 5.2. CONJUNTO DE DATOS DE ENTRENAMIENTO .....	64
FIGURA 5.3. COMPARACIÓN DE LA DISTRIBUCIÓN DE PROBABILIDAD DEL CONJUNTO DE DATOS DE ENTRENAMIENTO CON RESPECTO A DATOS HISTÓRICOS.....	64
FIGURA 5.4. COMPARACIÓN DE RESULTADOS ENTRE EL MODELO DE RNA NARX, <i>FEED-FORWARD</i> Y <i>SCADA</i> .....	65
FIGURA 5.5. ERRORES OBTENIDOS DE LA ESTIMACIÓN DE TEMPERATURA DE LOS MODELOS DE RNA NARX Y <i>FEED-FORWARD</i> .....	66
FIGURA 5.6. MODELO DE RNA NARX PROPUESTO.....	68
FIGURA 5.7. COMPARACIÓN ENTRE LA TEMPERATURA ESTIMADA POR LA RNA DE 5 NEURONAS Y LA TEMPERATURA MEDIDA POR EL SCADA .....	69
FIGURA 5.8. COMPARACIÓN ENTRE LA TEMPERATURA ESTIMADA POR LA RNA DE 10 NEURONAS Y LA TEMPERATURA MEDIDA POR EL SCADA .....	70
FIGURA 5.9. COMPARACIÓN ENTRE LA TEMPERATURA ESTIMADA POR LA RNA DE 20 NEURONAS Y LA TEMPERATURA MEDIDA POR EL SCADA .....	71
FIGURA 5.10. COMPARACIÓN ENTRE LA TEMPERATURA ESTIMADA POR LA RNA DE 30 NEURONAS Y LA TEMPERATURA MEDIDA POR EL SCADA .....	72
FIGURA 5.11. COMPARACIÓN ENTRE LA TEMPERATURA ESTIMADA POR LA RNA DE 40 NEURONAS Y LA TEMPERATURA MEDIDA POR EL SCADA .....	73
FIGURA 5.12. COMPARACIÓN DE RESULTADOS DE LOS MODELOS NARX.....	74
FIGURA 5.13. ERRORES OBTENIDOS PARA CADA MODELO DE RNA NARX.....	74
FIGURA 5.14.COMPARACIÓN ENTRE LA TEMPERATURA ESTIMADA POR LOS MODELOS DE RNA NARX Y LA TEMPERATURA MEDIDA POR EL SCADA .....	76
FIGURA 5.15. ERRORES OBTENIDOS DE LA ESTIMACIÓN DE TEMPERATURA DE LOS MODELOS DE RNA NARX .....	77
FIGURA 5.16. DISTANCIA DE MAHALANOBIS CALCULADA PARA LA TEMPERATURA DEL ACEITE EN EL TRANSFORMADOR T1 .....	78
FIGURA 5.17. DISTANCIA DE MAHALANOBIS CALCULADA PARA LA TEMPERATURA DEL ACEITE EN EL TRANSFORMADOR T2 .....	80

FIGURA 5.18. COMPARACIÓN DE LA SALIDA DE LA RNA Y LOS VALORES DE TEMPERATURA ALMACENADOS EN EL SCADA .....	81
FIGURA 5.19. VALOR RMSE DE MONITOREO DE LA CONDICIÓN PARA EL TRANSFORMADOR T2.....	82
FIGURA B.1. A) NEURONA BIOLÓGICA, B) NEURONA ARTIFICIAL .....	98
FIGURA B.3. ESTRUCTURA DE UNA RED NEURONAL MULTICAPA .....	99
FIGURA B.4. ESTRUCTURA GENERAL DE UNA NEURONA.....	99
FIGURA B.5. FUNCIÓN DE ACTIVACIÓN CONTINUA BIPOLAR.....	101
FIGURA B.6. FUNCIÓN DE ACTIVACIÓN CONTINUA UNIPOLAR .....	102
FIGURA B.7. FUNCIÓN DE ACTIVACIÓN RAMPA UNIPOLAR.....	102
FIGURA B.8. PERCEPTRÓN DISCRETO (BINARIO) .....	103
FIGURA B.9. PERCEPTRÓN CONTINUO.....	103
FIGURA B.10. DIAGRAMA A BLOQUES DEL APRENDIZAJE SUPERVISADO.....	105
FIGURA B.11. DIAGRAMA A BLOQUES DEL APRENDIZAJE NO SUPERVISADO .....	106
FIGURA C.1. RED MULTICAPA CON DOS CAPAS DE PERCEPTRONES CONTINUOS .....	109
FIGURA E.1. TRANSFORMADOR T1 .....	116
FIGURA E.2. DATOS DE PLACA DEL TRANSFORMADOR T1 .....	117
FIGURA E.3. RELÉ UTILIZADO PARA EL MONITOREO DE TRANSFORMADORES .....	117

## ÍNDICE DE TABLAS

TABLA 2.1. MATRIZ DE RIESGO EJEMPLIFICADA A PARTIR DEL MÉTODO DE GESTIÓN DE MANTENIMIENTO A GENERADORES DE ENERGÍA HIDROELÉCTRICA [20] .....	13
TABLA 2.2. SISTEMAS Y COMPONENTES CRÍTICOS EN UN TRANSFORMADOR .....	22
TABLA 2.3. ANÁLISIS DE SISTEMAS DE TRANSFORMADORES DE POTENCIA.....	24
TABLA 4.1. DETALLES DEL MODELO DE RNA PROPUESTO.....	54
TABLA 5.1. DATOS NOMINALES DE LOS TRANSFORMADORES DE PRUEBA .....	61
TABLA 5.2. CORRELACIÓN ENTRE VARIABLES DE ENTRADA Y SALIDA DE LOS DATOS DE ENTRENAMIENTO.....	63
TABLA 5.3. COMPARACIÓN DE RESULTADOS DE LOS MODELOS PROPUESTOS .....	75
TABLA 5.4. COMPARACIÓN RESUMIDA DE ALGORITMOS SIMILARES PROPUESTOS PREVIAMENTE VS EL ALGORITMO PROPUESTO .....	83
TABLA A. 1 ANÁLISIS DE CORRELACIÓN DE VARIABLES DEL CONJUNTO DE DATOS ENTRENAMIENTO PARA TRANSFORMADOR T1.....	94

## ACRÓNIMOS

<b>CBM</b>	Mantenimiento Basado en la Condición (Condition-Based Maintenance)
<b>CFE</b>	Comisión Federal de Electricidad
<b>CIGRE</b>	Consejo Internacional de Grandes Sistemas Eléctricos
<b>CM</b>	Mantenimiento Correctivo
<b>CONACyT</b>	Consejo Nacional de Ciencia y Tecnología
<b>DGA</b>	Análisis de Gases Disueltos
<b>HI</b>	Índice de Salud
<b>IEEE</b>	Institute of Electrical and Electronics Engineers
<b>LCC</b>	Costo del Ciclo de Vida
<b>LMA</b>	Algoritmo de Aprendizaje Levenberg-Marquardt
<b>MD</b>	Distancia de Mahalanobis
<b>ML</b>	Machine Learning
<b>MM</b>	Modelo de Márkov
<b>MPB</b>	Máquina de puntos Bayes
<b>NARX</b>	Red Autorregresiva no Lineal con Entradas Exógenas
<b>pdf</b>	Función de densidad de probabilidad
<b>PM</b>	Mantenimiento Preventivo
<b>RCM</b>	Confiabilidad Centrada en Mantenimiento
<b>REI</b>	Redes Eléctricas Inteligentes
<b>REP</b>	Redes Eléctricas de Potencia
<b>RNA</b>	Red Neuronal Artificial
<b>RTU</b>	Unidades Terminales Remotas
<b>SCADA</b>	Sistema de Supervisión, Control y Adquisición de Datos
<b>SIMAP</b>	Sistema Inteligente para el Mantenimiento Predictivo
<b>SIG</b>	Sistemas de Información Geográfica
<b>SVM</b>	Máquinas de Soporte Vectorial
<b>TCO</b>	Costo Total de Propiedad
<b>VDM</b>	Mantenimiento Impulsado por el Valor
<b>TPM</b>	Mantenimiento Productivo Total

## VARIABLES

$F$	Función de aproximación no lineal
$H$	Matriz de Hessiana
$g$	Vector de gradiente
$\lambda$	Parámetro escalar
$w$	Vector de ponderación
$S$	Matriz de covarianzas
$D$	Matriz diagonal de orden $k$
$R_p$	Matriz de correlaciones parciales
$U_t$	Vector de entrada de la RNA
$y_t$	Salida de la RNA
$P_{(a,b,c)}$	Potencia activa por fase
$I_{(an,bn,cn)}$	Corrientes de fase
$T_{amb}$	Temperatura ambiente
$MD_i$	Distancia de Mahalanobis
$Error_{C_{Metapa}}$	Diferencia entre las cantidades estimadas
$C$	Matriz de covarianza del modelo de RNA y reales



## CAPÍTULO I: INTRODUCCIÓN

### 1.1. Introducción

Las Redes Eléctricas de Potencia de todo el mundo se enfrentan a diversos cambios con la intención de hacerlas inteligentes. Estos desafíos van desde la integración masiva de sistemas de generación distribuida, principalmente basada en renovables, hasta la interoperabilidad de equipos de última generación con equipos cerca del final de su vida útil. Esto ha obligado a buscar soluciones de control y comunicaciones, cada país o región enfrenta sus propios retos. En los países desarrollados, el envejecimiento de los activos es un problema a vencer, mientras que en economías en desarrollo se busca identificar ejemplos de buenas prácticas para modelar sus operaciones [1].

El problema expuesto hace que las empresas eléctricas de servicios deban operar en un ambiente que plantea nuevos retos, tales como el aumento de los costos operativos, el crecimiento de la demanda, el requerimiento de elevados niveles de confiabilidad y calidad de servicio, además, de la necesidad de afrontar y gestionar el envejecimiento del equipamiento instalado.

La liberalización del mercado de la energía y la presión creciente orientada hacia la reducción de costos, obligan a las compañías a buscar permanentemente nuevas formas de reducir los costos de explotación de sus activos instalados y de mejorar la rentabilidad de sus inversiones. En el entorno indicado, la búsqueda de mayor rentabilidad de los activos físicos pasa a ser parte fundamental del trabajo diario de las compañías eléctricas, empleando así, métodos de gestión de activos más proactivos que les permitan obtener sistemas de distribución de energía más rentables [2].

México no es la excepción; a partir del año 2013 la Comisión Federal de Electricidad (CFE) cambió la metodología del mantenimiento de los activos, pasando de un mantenimiento basado en el tiempo, a un mantenimiento basado en condición. Esto ha llevado a la necesidad de desarrollar aplicaciones de gestión de activos, tiempo de vida útil y fuerza de trabajo móvil [3].

### 1.2. Antecedentes y planteamiento del problema

Luego de la Primera y Segunda Guerra Mundial se dio una rápida expansión económica en los períodos posteriores a la guerra dentro de sociedades desarrolladas. Esto incentivó la inversión gubernamental para la construcción de redes eléctricas y se promovió el uso de la energía eléctrica en la esfera civil [1]. Las técnicas de ingeniería tradicionales durante ese tiempo han dado lugar a un gran volumen de activos de larga duración dentro de la industria de distribución eléctrica, como transformadores, interruptores y postes de potencia [4]. Esto a su vez ha dado lugar a infraestructuras de distribución eléctrica en todo el mundo con un

gran número de activos que han venido envejeciéndose [1,5]. Estos activos sobreviven bajo regímenes de mantenimiento de rutina hasta que alcanzan el final de su vida, en cuyo punto a menudo se produce un corte de suministro. Este modo de operación, conocido como el modelo de ejecución a falla (mantenimiento correctivo), ha sido el método de operación tradicional para esta industria [1]. Evitar tales interrupciones, se ha convertido en una necesidad para muchas empresas de distribución eléctrica en todo el mundo desarrollado [1, 5, 6].

En la actualidad, las Redes Eléctricas de Potencia representan una importante área de oportunidades, dadas las posibilidades que representa para el desarrollo económico y social de un país. Por otro lado, el creciente número de actores y participantes del sector eléctrico, los avances tecnológicos desarrollándose rápidamente en los campos de la informática y la competencia como premisa entre las empresas, se conjugan para impulsar una constante búsqueda de REP cada vez más competitivas y eficientes. Para estas nuevas redes, es de vital importancia que funcionen de forma óptima, confiable y segura [7].

El aumento exponencial de la demanda eléctrica y los requisitos de calidad del suministro, cada vez más exigentes, han obligado a avanzar hacia un nuevo concepto de red eléctrica, las “Redes Eléctricas Inteligentes”. La principal característica de una red inteligente, es que permite la distribución de electricidad desde los proveedores hasta los consumidores, utilizando tecnología digital con el objetivo de ahorrar energía, reducir costos y aumentar la confiabilidad. Las “Redes Eléctricas Inteligentes” utilizarán el desarrollo tecnológico para optimizar el uso de sus activos en tres horizontes de tiempo diferentes: el corto plazo, se centrará en las operaciones diarias, el mediano plazo, se encargará de la gestión de mantenimiento y el largo plazo, se enfocará en mejorar los procesos de administración de activos [8].

Una de las características principales del concepto de redes inteligentes es la capacidad que tienen las mismas de generar un gran número de datos utilizando los sistemas de monitoreo instalados en éstas, los cuales no los encontramos en las redes eléctricas convencionales. Hay un gran número de posibilidades para utilizar estos datos con el objetivo de mejorar la gestión del sistema eléctrico en su conjunto y los componentes individuales. Los datos obtenidos en las mediciones de “Redes Eléctricas Inteligentes” pueden ser utilizados para la gestión de activos en las Redes Generales de Distribución [9].

La gestión de activos puede, en términos generales, denominarse como: “el proceso de gestión efectiva del costo de un activo durante su vida útil”. La gestión de activos incluye la gestión del costo durante la fase de planificación inicial, construcción, operación, mantenimiento y finalmente la disposición del activo [10]. Entre los costos mencionados anteriormente, el costo de mantenimiento a los activos se considera un costo crítico debido a

que es uno de los gastos más altos durante la vida útil del activo, este puede ser optimizado usando diferentes métodos y herramientas [9].

La gestión de mantenimiento se ha vuelto un área de interés, dando lugar a diferentes investigaciones alrededor del mundo, debido al avance tecnológico que existe en la actualidad, hoy en día se tienen una gran diversidad de sensores, sistemas de comunicación y grandes capacidades de almacenamiento y procesamiento de datos [9]. Estos sistemas de monitoreo proporcionan una gran cantidad de datos, que en muchos casos no se analizan en profundidad debido a que no existen algoritmos de minería de datos implementados en las empresas.

Por otro lado, los métodos de mantenimiento predictivo existentes no evalúan la condición de los equipos en tiempo real, sino que realizan mediciones periódicas para evaluar la condición operativa de los equipos.

Actualmente se desarrollan algoritmos que permiten determinar la condición de los equipos utilizando los sistemas de monitoreo que están instalados, pero haciendo un uso más eficiente de la información que éstos generan. Las investigaciones más significativas en esta área las encabezan el Instituto Tecnológico KTH en Estocolmo, Suecia y la Universidad Tecnológica de Chalmers en Gotemburgo, Suecia.

Por lo anterior, es importante utilizar métodos predictivos para evaluar la condición de transformadores en subestaciones de distribución. El presente trabajo hace una aportación científica en el mantenimiento predictivo de transformadores a través del uso eficiente de la infraestructura digital de las empresas eléctricas y más concretamente. Con esta infraestructura, se busca desarrollar una estrategia para mejorar el mantenimiento de los activos en la red. En particular, este trabajo se enfocará en mejorar el mantenimiento de los transformadores en Subestaciones de Distribución, mediante el procesamiento de los datos provistos por sensores y sistemas de comunicación. En esta tesis se utilizarán métodos predictivos basado en algoritmos de Aprendizaje Automático, en particular RNA, por la capacidad que tienen de procesar grandes cantidades de datos y modelar sistemas cuando estos son muy complejos y no hay ecuaciones que los gobiernen.

### **1.3. Hipótesis**

Conocer los datos históricos eléctricos y físicos de un conjunto de activos permitirá establecer programas de mantenimiento óptimos. Esto nos lleva a la hipótesis general en la que se basa este trabajo, la cual establece que: Contar con una herramienta computacional para realizar mantenimiento “inteligente”, contribuirá a reducir los riesgos de disturbios y además ayudará a reducir los gastos globales de mantenimiento de los activos en una REI.

## 1.4. Objetivo general y objetivos específicos

Desarrollar y programar un algoritmo basado en el análisis de datos históricos de parámetros eléctricos y físicos, que permita establecer rutinas de mantenimiento óptimo en activos de subestaciones eléctricas.

A continuación, se desglosan los principales objetivos específicos para dar cumplimiento al objetivo general del trabajo de tesis:

- Estudio del estado del arte sobre gestión de activos, análisis de riesgo y gestión de mantenimiento. Establecer definiciones y generalidades.
- Estudio del estado del arte de los principales métodos aplicados en confiabilidad centrada en mantenimiento y modelos similares.
- Investigar, desarrollar e implementar un método de minería de datos para la recolección, análisis e interpretación de los mismos.
- Estudio del estado del arte sobre modelos predictivos usados para la optimización del mantenimiento en activos de sistemas eléctricos.
- Seleccionar y desarrollar el modelo idóneo para el análisis de mantenimiento predictivo.
- Proponer un algoritmo para gestión de mantenimiento que permita optimizar las rutinas de mantenimiento a los activos que así lo requieran.
- Desarrollar e implementar el algoritmo propuesto para la gestión de mantenimiento.
- Validar el algoritmo con casos de estudio con datos proporcionados por CFE distribución.
- Documentar la información para escribir el trabajo de tesis.

## 1.5. Justificación

Hoy en día se han abierto las puertas a la descentralización de funciones y al establecimiento del libre mercado en las REP, situación que se ha expandido a numerosos países [1, 4]. Esto ha provocado la aparición de nuevas estructuras y organizaciones que a fin de cuentas están marcando un nuevo paradigma para operar las redes eléctricas en el país. Dentro de este entorno, aspectos como la confiabilidad que conlleva a la suficiencia, la seguridad y la calidad del suministro eléctrico han cobrado vital importancia.

Debido a la reforma energética en México, la CFE Distribución tiene que enfrentarse a problemas cada vez más complejos por la situación de crecientes negocios y los marcos regulatorios. Entre estas cuestiones hay dos que destacan especialmente: la previsión de que las redes eléctricas funcionarán más cerca de su capacidad máxima y la necesidad, cada vez mayor, de monitorear y operar de forma segura y confiable estas redes.

Actualmente, en operadores de redes de distribución como CFE y la Red de Media Tensión de Ciudad Universitaria, existe un interés creciente en el manejo de sus activos, así como la reducción de sus gastos de mantenimiento.

Un primer paso inherente ha sido el monitorear las Redes Eléctricas de Potencia. Esto abre muchas posibilidades para mejorar la operación y gestión de las mismas, una de ellas es la gestión de activos. Tener conocimiento sobre la condición de los activos de la REP, permitirá a los operadores de la misma tomar rápidas y eficaces acciones sobre los elementos que la conforman, cuando se necesite.

En México la nueva Ley de la Comisión Federal de Electricidad penaliza a los operadores por la desconexión de usuarios. Esto ha llevado a la CFE Distribución a buscar herramientas que le permitan establecer programas de mantenimiento que brinden mínimas o nula desconexión de clientes y que a su vez garanticen la reducción de riesgos de fallas de los activos. Esto tiene importantes implicaciones de confiabilidad.

## **1.6. Metodología**

En el trabajo de tesis se analiza el estado del arte sobre la gestión de activos y su relación con gestión de mantenimiento y el análisis de riesgo. Se describen las principales acciones de mantenimiento predictivo en transformadores descritas en la literatura.

Para obtener el conjunto de datos necesario para desarrollar este trabajo, se utilizó el sistema de monitoreo instalado en la subestación en la que se encuentran instalados los transformadores analizados. En el Apéndice E se ofrecen detalles de éste sistema; además, se muestran los datos de placa de los transformadores y el relé utilizado por el sistema de monitoreo.

Se propone el uso de redes neuronales artificiales para el desarrollo del algoritmo utilizado para el monitoreo de la condición en transformadores. Se analizan todas las estructuras descritas en la literatura y se propone el uso de RNA NARX para el desarrollo del algoritmo para la estimación de la temperatura del aceite en tiempo real debido a la fortaleza de éstas en el análisis de series de tiempo, los intervalos de cada muestra son de 30 minutos debido a la inercia de temperatura que tiene el aceite en los transformadores. Se selecciona la estructura que mejor estima esta variable y para la detección de falla se propone el uso de la distancia de Mahalanobis, la cual es una medida de distancia estadística que correlaciona la variable estimada y la medida. Utilizando el conjunto de datos de entrenamiento se establece un valor umbral que indica cuando el transformador entra en un estado de posible falla, a partir de ese momento se propone la revisión del equipo por parte de los operadores y personal de mantenimiento para la identificación del problema.

## 1.7. Organización de la tesis

A continuación, se describen de forma general los principales aspectos que se desarrollan en cada uno de los capítulos, que conforman el trabajo de tesis:

En el **Capítulo I** de este trabajo de investigación está dedicado a introducir el tema de estudio.

En el **Capítulo II** se presentan los aspectos generales sobre la gestión de activos, análisis de riesgos y gestión de mantenimiento.

En el **Capítulo III** se presenta una introducción a los métodos predictivos y su aplicación en Redes Eléctricas de Potencia.

En el **Capítulo IV** se presenta el método estudiado para su aplicación al mantenimiento predictivo, así como el desarrollo del algoritmo.

En el **Capítulo V** se presentan los resultados obtenidos y se valida el algoritmo con los casos de estudio utilizados.

En el **Capítulo VI** se presentan las conclusiones más relevantes sobre el trabajo de tesis realizado y trabajos futuros.

## 1.8. Metas

Las metas alcanzadas en este trabajo de tesis fueron:

- La implementación de un algoritmo basado en redes neuronales artificiales para el monitoreo de la condición en transformadores de distribución.
- Desarrollar bases sólidas en el uso de redes neuronales y sus aplicaciones al sector eléctrico.
- La aplicación del algoritmo desarrollado a transformadores de distribución reales.
- Se logró la detección de funcionamiento anómalo en uno de los transformadores analizados antes de que ocurriera la falla.

### **CAPÍTULO II: ASPECTOS GENERALES DE TEORÍA SOBRE GESTIÓN DE ACTIVOS, RIESGO Y MANTENIMIENTO**

#### **2.1. Introducción**

En la actualidad, la gestión de activos, riesgos y mantenimiento son tres tipos de gestión estrechamente relacionados entre sí. Los tres se encuentran con frecuencia en los servicios públicos de redes eléctricas y otras industrias con un gran número de activos en todo el mundo. Por lo tanto, deben ser estudiados y discutidos de manera integrada. Por ejemplo, los aspectos cubiertos por la gestión de activos pueden variar desde el mantenimiento y la renovación de activos específicos o grupos de activos hasta la gestión o el equilibrio de aspectos financieros con aspectos de ingeniería y riesgo para una gran población de activos geográficamente extendidos. El objetivo de este capítulo es investigar la gestión de mantenimiento de manera integrada con la gestión de activos y riesgos. Por lo tanto, en este capítulo las literaturas que cubren estos aspectos se discuten brevemente, tanto por separado como su relación entre sí. Para esto se realiza el análisis de las principales estrategias de mantenimiento predictivo utilizadas en transformadores de potencia en la actualidad. También se describen en el capítulo los principales componentes en los transformadores de potencia, su función y posibles fallas, así como las posibles consecuencias ante la ocurrencia de éstas. El caso de estudio es enfocado a los transformadores de potencia, como principal componente de las subestaciones eléctricas

#### **2.2. Gestión de activos**

##### **2.2.1. ¿Qué es un activo?**

Un activo es un elemento, cosa o entidad que tiene un valor potencial o real para una organización. El valor variará entre las diferentes organizaciones y sus partes interesadas y puede ser tangible o intangible, financiero o no financiero [11].

El período desde la creación de un activo hasta el final de su vida útil es la vida del activo. La vida de un activo no necesariamente coincide con el período durante el cual una organización es responsable de ello; en su lugar, un activo puede proporcionar valor potencial o real a una o más organizaciones a lo largo de su vida útil, y el valor del activo a una organización puede cambiar a lo largo de su vida útil.

Una organización puede optar por administrar sus activos como un grupo, en lugar de individualmente, de acuerdo a sus necesidades, y para obtener beneficios adicionales. Dichos agrupamientos de activos pueden ser por tipos de activos, sistemas de activos o carteras de activos [11].

Los factores que influyen en el tipo de activos que una organización requiere para alcanzar sus objetivos y cómo se gestionan los activos son los siguientes:

- naturaleza y finalidad de la organización;
- su contexto operativo;
- sus limitaciones financieras y requisitos reglamentarios;
- las necesidades y expectativas de la organización y de sus grupos de interés.

Estos factores de influencia deben considerarse al establecer, implementar, mantener y mejorar continuamente la gestión de activos.

El control y la gobernabilidad efectivos de los activos por parte de las organizaciones es esencial para lograr el valor a través de la gestión del riesgo y la oportunidad, con el fin de lograr el equilibrio deseado de costo, riesgo y rendimiento. El entorno normativo y legislativo en el que las organizaciones operan es cada vez más difícil y los riesgos inherentes que muchos activos presentan están en constante evolución.

La gestión de activos traduce los objetivos de la organización en decisiones, planes y actividades relacionados con los activos, utilizando un enfoque basado en el riesgo [11].

### **2.2.2. Beneficios de la gestión de activos**

La gestión de activos permite a una organización obtener valor de los activos en el logro de sus objetivos organizacionales. Lo que constituye valor dependerá de estos objetivos, la naturaleza y el propósito de la organización y las necesidades y expectativas de sus partes interesadas. La gestión de activos apoya la realización de valor y equilibra los costos financieros, ambientales y sociales, el riesgo, la calidad de servicio y el desempeño relacionado con los activos.

Los beneficios pueden incluir, pero no están limitados a lo siguiente:

- a) **mejora del rendimiento financiero:** se puede mejorar el rendimiento de las inversiones y reducir los costos, preservando al mismo tiempo el valor de los activos y sin sacrificar la realización a corto o largo plazo de los objetivos de la organización;
- b) **decisiones informadas de inversión de activos:** permitir a la organización mejorar su toma de decisiones y equilibrar efectivamente los costos, riesgos, oportunidades y desempeño;



- c) **riesgo gestionado:** reducir las pérdidas financieras, mejorar la salud y la seguridad, la buena voluntad y la reputación, minimizar el impacto ambiental y social, puede resultar en una reducción de pasivos como primas de seguros, multas y sanciones;
- d) **servicios y productos mejorados:** asegurar el desempeño de los activos puede conducir a mejores servicios o productos que consistentemente cumplan o exceden las expectativas de los clientes y las partes interesadas;
- e) **responsabilidad social demostrada:** mejorar la capacidad de la organización para, por ejemplo, reducir las emisiones, conservar los recursos y adaptarse al cambio climático, le permite demostrar prácticas de negocios socialmente responsables y éticas y administración responsable;
- f) **conformidad demostrada:** conformidad transparente con los requisitos legales, estatutarios y reglamentarios, así como adherirse a las normas, políticas y procesos de gestión de activos, puede permitir demostración de cumplimiento;
- g) **mayor reputación:** a través de la mejora de la satisfacción del cliente, el conocimiento de las partes interesadas y la confianza;
- h) **mejora de la sostenibilidad organizativa:** la gestión eficaz de los efectos a corto y largo plazo, los gastos y el desempeño, pueden mejorar la sostenibilidad de las operaciones y la organización;
- i) **mejora de la eficiencia y la eficacia:** revisar y mejorar los procesos, los procedimientos y el desempeño de los activos puede mejorar la eficiencia y la eficacia y el logro de los objetivos de la organización [11].

### 2.3. Gestión de activos en redes eléctricas. Clasificación.

Gestión de activos es una “expresión” que ha sido ampliamente usada en los sectores público y privado, con una amplia gama de interpretaciones diferentes, esto a menudo se confunde con el mantenimiento, que es sólo una de las partes del proceso completo.

El sistema eléctrico está formado por activos esenciales para que la producción, el transporte y la distribución de energía eléctrica se realicen de forma confiable y con calidad. En todos los segmentos del sistema eléctrico, los activos son vitales para lograr el suministro de electricidad. Actualmente, la desregulación del sector hace que la rentabilidad de la empresa sea indirectamente una función de su desempeño, lo que motiva a las empresas a concentrar los esfuerzos en obtener una mejor utilización de sus activos.

En el contexto de una empresa de distribución de electricidad, la gestión de activos puede definirse como un proceso sistemático de costo-efectividad, operación, mantenimiento y mejora de los activos eléctricos mediante la combinación de prácticas de ingeniería y análisis económico con una sólida práctica empresarial. La eficacia de estos procesos depende en gran medida de la información de la que dependen. Por ejemplo, las decisiones de planificación se hacen utilizando previsiones de carga. Cuanta más información disponga, más precisas serán las previsiones y más eficaz será la planificación. Los ingenieros de mantenimiento que deseen implementar un mantenimiento basado en la condición se beneficiarán de tener información sobre la condición del activo. El monitoreo del estado del equipo en tiempo real es limitado sin información en tiempo real.

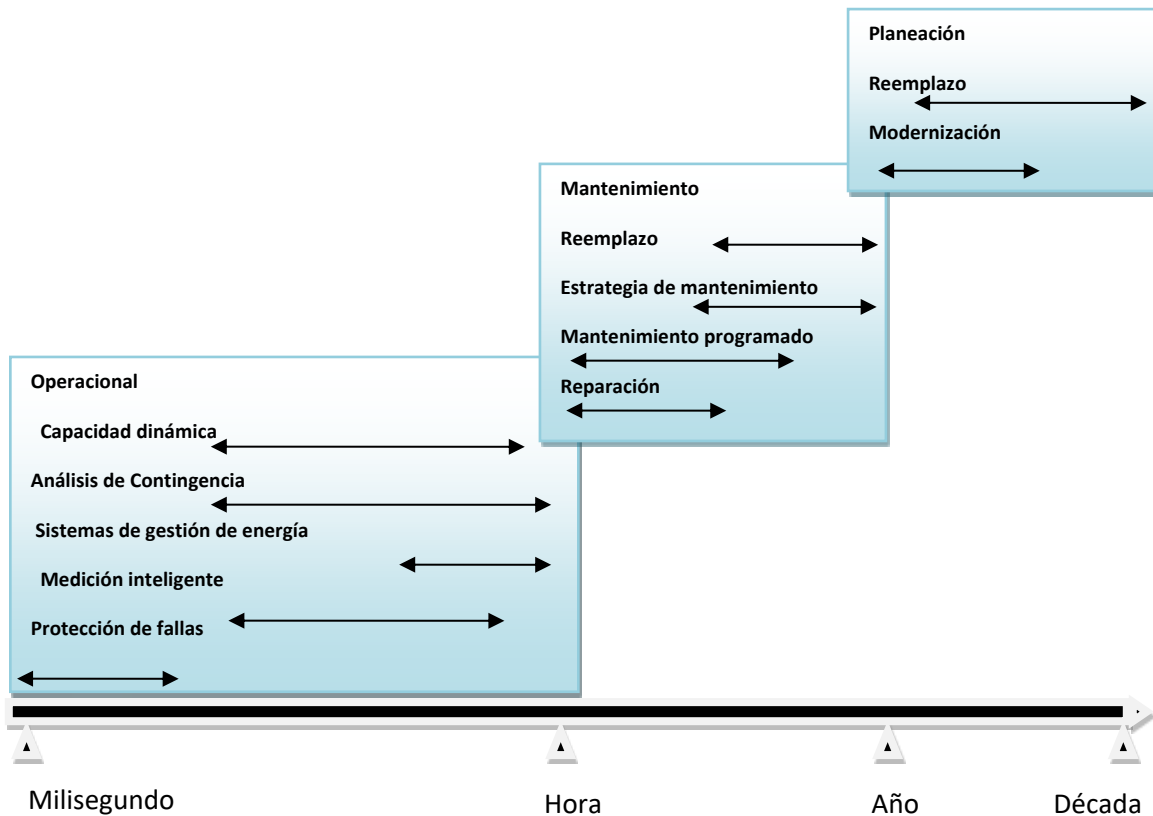
De hecho, la optimización de la utilización de la capacidad de los activos y de las operaciones, dará lugar a importantes reducciones de costos y mejoras de rendimiento. Hoy en día, mejorar la eficiencia y la utilización de los activos es una prioridad importante. La penetración de nuevas tecnologías en la red permite lograr un mayor grado de rendimiento y eficiencia. La gestión dinámica de activos, permite una mayor interacción entre los principales procesos de la red, combinado con la automatización de la red, optimizaría la red de distribución [8].

El objetivo principal de las redes eléctricas es proporcionar energía eléctrica confiable con una calidad de energía aceptable para los consumidores, con un coste óptimo.

Para cumplir con este objetivo, la red debe ser gestionada de manera efectiva. Las funcionalidades que hay que gestionar pueden clasificarse en tres divisiones: gestión operativa, de mantenimiento y de planificación. La escala de tiempo de estas tres administraciones varía, con las administraciones operativas que tienen la escala de tiempo más corta y la gestión de planificación con la escala de tiempo más larga [11].

La gestión de activos también se define como el proceso de maximizar el retorno de la inversión de los equipos a lo largo de todo su ciclo de vida, al maximizar el rendimiento y minimizar los costos relacionados con los equipos.

Las categorías y escalas de tiempo típicas se indican en la Figura 2.1. Algunas de las funcionalidades clave de la gestión operativa, de mantenimiento y de planificación también se enumeran en la figura. La escala de tiempo de la gestión varía de milisegundos a horas para la gestión operativa. Para la gestión de la planificación esta escala de tiempo se extiende de años a décadas.



**Figura 2.1. Funcionalidades de la red eléctrica clasificada en función del tiempo en la gestión operativa, de mantenimiento y de planificación [11]**

Como se puede apreciar en la Figura 2.1, la gestión de activos se puede clasificar en función del dominio del tiempo y las actividades que se realizan. Éstas se dividen en planificación de corto, mediano y largo plazo. A continuación, se realiza la descripción de cada una de estas clasificaciones, las cuales se relacionan con el esquema mostrado en la Figura 2.1.

- i. **Gestión de Planeación:** Abarca la planificación futura, es decir, la inversión en nuevos activos, como los transformadores desfasadores, los dispositivos compensadores de reactivos, la expansión de líneas o corredores de transmisión, o la mejora de los equipos de las subestaciones. Esto implica mayores riesgos financieros y por lo tanto, una planificación adecuada puede evitar los riesgos que conllevan a retrasos en el tiempo, tasas de interés y la diversidad de la carga a largo plazo. El margen de tiempo va desde un año y más, y apunta a la mejora de los activos de generación y transmisión existentes [13].
- ii. **Gestión de Mantenimiento:** El objetivo principal es prolongar la vida útil de las instalaciones existentes mediante un mantenimiento adecuado y asignar de forma óptima, los recursos energéticos renovables. El costo de mantenimiento es el factor más importante o decisivo, ya que incide directamente en las interrupciones. Éstas

pueden ser, planificadas y no planificadas, las no planificadas pueden reducirse considerablemente cuando se programan interrupciones planeadas según la disponibilidad de recursos y las distribuciones de carga. Por lo tanto, la programación de mantenimiento óptima reduce en gran medida la posibilidad de interrupciones imprevistas. En [13] se explica la gestión de activos a medio plazo como:

- Minimizar los riesgos corporativos financieros y físicos basados en interrupciones planeadas y forzadas de activos.
- Reducción de los costos de operación para el suministro de clientes en una era competitiva
- Optimización de la asignación de recursos naturales volátiles y limitados para la utilización de activos corporativos.
- Ampliación de la vida útil de los activos a través de programas operativos y de mantenimiento adecuados.
- Prolongación de los costos de inversión para la adquisición de nuevos activos.

El mediano plazo varía de un mes a un año e implica la programación óptima del mantenimiento del equipo y la asignación de los recursos disponibles.

- iii. **Gestión Operacional:** La gestión de activos a corto plazo se clasifica en gestión de activos operativos (diaria y semanal) y gestión de activos en tiempo real (gestión de interrupciones). La gestión de activos operativos tiene como objetivo minimizar los riesgos relacionados con los activos, tanto físicos como financieros, debido a la demanda de carga y a los precios por hora. La gestión de activos en tiempo real también se conoce como gestión de interrupciones o análisis de contingencias. El análisis de contingencias constituye una parte vital, pues ayuda a evaluar el efecto de las interrupciones inesperadas debido al cambio en las condiciones climáticas, cualquier descomposición repentina o fluctuaciones de la carga en la condición y el funcionamiento del activo. En los últimos años, debido a los avances tecnológicos, el monitoreo en tiempo real de los activos es posible debido a sistemas de Supervisión, Control y Adquisición de Datos (SCADA), Unidades Terminales Remotas (RTU) y Sistema de Información Geográfica (SIG) [14].

### 2.4. Gestión de riesgo

La gestión de activos toma cuenta todas las limitaciones de rendimiento, costo y riesgo de los activos [15]. La fiabilidad de la red se puede mejorar monitoreando el estado del equipo y tomando acciones preventivas previniendo riesgos [16], por lo tanto, el análisis de riesgos

constituye una parte importante de la gestión de activos.

El riesgo se define como el producto de la probabilidad de un evento y sus consecuencias. En términos generales, la existencia de riesgos implica que existe al menos la posibilidad de consecuencias negativas o desviaciones de los valores esperados cuando un evento tiene probabilidad de ocurrir. Es esencial para la gestión de riesgos que no se limite solo a los riesgos técnicos de la red, sino que también se extienda a la gestión de los riesgos que implican el análisis de eventos para un sistema completo (es decir, organización, red y activos) que podría tener un efecto negativo en el sistema [17]. Además, la gestión del riesgo implica el proceso de toma de decisiones de las medidas para evaluar los riesgos y la valoración de estas medidas.

El riesgo se evalúa de forma tal que se pueda cuantificar el efecto en un valor comercial y la probabilidad de que ocurra el efecto. Para garantizar la objetividad del análisis de riesgos, la probabilidad y las consecuencias se evalúan por separado y luego se combinan para formar una evaluación del riesgo. Para calificar los riesgos, se agrega un conjunto de posibles efectos a cada valor comercial [18].

### **2.4.1. Análisis de riesgo**

El análisis de riesgos es de vital importancia para la gestión proactiva de los activos. El objetivo principal es entender la causa, el efecto y la probabilidad de eventos adversos para administrar de manera óptima los riesgos y reducirlos a un nivel aceptable.

El análisis de riesgo podría realizarse utilizando tres preguntas [19]:

1. “¿Qué puede salir mal?” (Evento).
2. “¿Cuán probable es que suceda?” (Probabilidad).
3. “¿Cuáles son las consecuencias?” (Consecuencia).

Un enfoque común para evaluar un riesgo o comparar y priorizar los riesgos es el uso de “matrices de riesgo”. Este enfoque es de común uso por el sector eléctrico [20]. La probabilidad y la consecuencia se estiman primero y se dividen en categorías establecidas. La categorización podría hacerse por métodos cualitativos o cuantitativos. Las categorías de probabilidad y consecuencia dan dos ejes con una matriz bidimensional resultante con valores de riesgo estimados o medidas de acción propuestas. Una definición simple, pero comúnmente utilizada, del valor de riesgo se ejemplifica en la ecuación (2.1), donde  $P$  es un número establecido asociado con una categoría de probabilidad, mientras que  $C$  es un número establecido asociado con una categoría de consecuencia. El valor del riesgo podría usarse

como entrada para decidir si se debe tratar el riesgo y cómo (a menudo se combina con un análisis de costos).

$$\text{valor de riesgo} = P \times C \tag{2.1}$$

Sin embargo, una matriz de riesgos podría diseñarse de manera diferente, por ejemplo, como se ilustra en la Tabla 2.1. La matriz de riesgos ejemplificada se usa como parte de un método RCM aplicado a los generadores hidroeléctricos en Suecia. Si el valor de riesgo es 3 o más, se recomendará una estrategia de mantenimiento preventivo [20].

**Tabla 2.1. Matriz de riesgo ejemplificada a partir del método de gestión de mantenimiento a generadores de energía hidroeléctrica [20]**

Consecuencia \ Probabilidad	A Menor	B Promedio/Serio	C Severo/Catastrófico
Fallas que ocurren más de una vez al año	2	3 Seguridad, pérdida de eficiencia, ambiental	4
Fallas que ocurren en intervalos de 1 a 20 años	2	3	4
Fallas que ocurren en intervalos de 20 a 50 años	1	2 Seguridad	3
Fallas que ocurren en intervalos de más de 50 años	0	1	2 3 Interrumpido, producción, inundación

En general, la gestión de riesgos como un proceso comercial regular, proporciona un método para identificar, evaluar y tratar el riesgo de una organización para dar cuenta de eventos futuros con un impacto negativo en la organización. Además, crea una conciencia de los riesgos socio-técnicos evaluados y las posibilidades de tratarlos. Sumado a esto, la aplicación

de un enfoque de gestión de riesgos objetivos, permite a las empresas de servicios de distribución de electricidad tener control sobre los riesgos y tener un lenguaje común para compartir y discutir los riesgos entre la administración y los ingenieros. También se usa para cuantificar y respaldar argumentos para presupuestos y planificación de cartera. Los servicios de distribución de energía eléctrica utilizan la gestión de riesgos para varios aspectos, por ejemplo, nuevos proyectos de inversión, inversiones de expansión, etc.

No siempre es obvio que dos riesgos con la misma ubicación en la matriz deben manejarse de la misma manera. Esto se ejemplifica en la Tabla 2.1, donde, por ejemplo, los riesgos de seguridad a veces se priorizan más y se manejan de forma diferente en comparación con otros riesgos dentro de la misma categoría de probabilidad y consecuencia.

### **2.4.2. Estándares de gestión de riesgos**

Los estándares generales de gestión de riesgos que se publican son [21]:

- ISO 31000: 2009 - Gestión de riesgos: proporciona directrices genéricas sobre la gestión de riesgos.
- ISO Guide 73 - Vocabulario de gestión de riesgos.
- ISO 31010: 2009 - Técnicas de evaluación de riesgos.

En la norma ISO 31000: 2009, se ofrece una visión general de un proceso general de toma de decisiones de gestión de riesgos, como se muestra en la Figura 2.2 [21].

En [21], la gestión de riesgos se describe como una evaluación que proporciona a los responsables de la toma de decisiones y partes responsables una mejor comprensión de los riesgos que podrían afectar el logro de los objetivos, y la adecuación y eficacia de los controles ya existentes. Esto proporciona una base para las decisiones sobre el enfoque más apropiado que se utilizará para tratar los riesgos. El resultado de la evaluación de riesgos es un insumo para los procesos de toma de decisiones de la organización. La evaluación de riesgos es el proceso general de identificación de riesgos, análisis de riesgos y evaluación de riesgos (ver Figura 2.2). La forma en que se aplica este proceso depende no sólo del contexto del proceso de gestión de riesgos sino también de los métodos y técnicas utilizados para llevar a cabo la evaluación de riesgos.

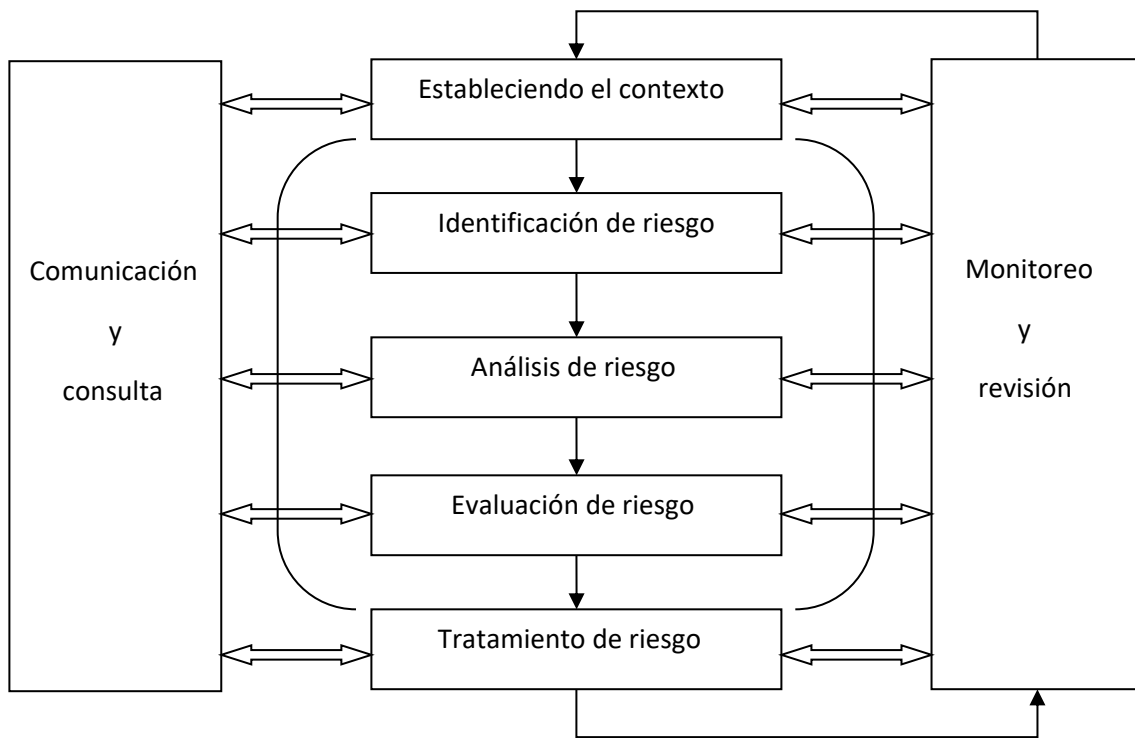


Figura 2.2. Proceso general de toma de decisiones de gestión de riesgos [21]

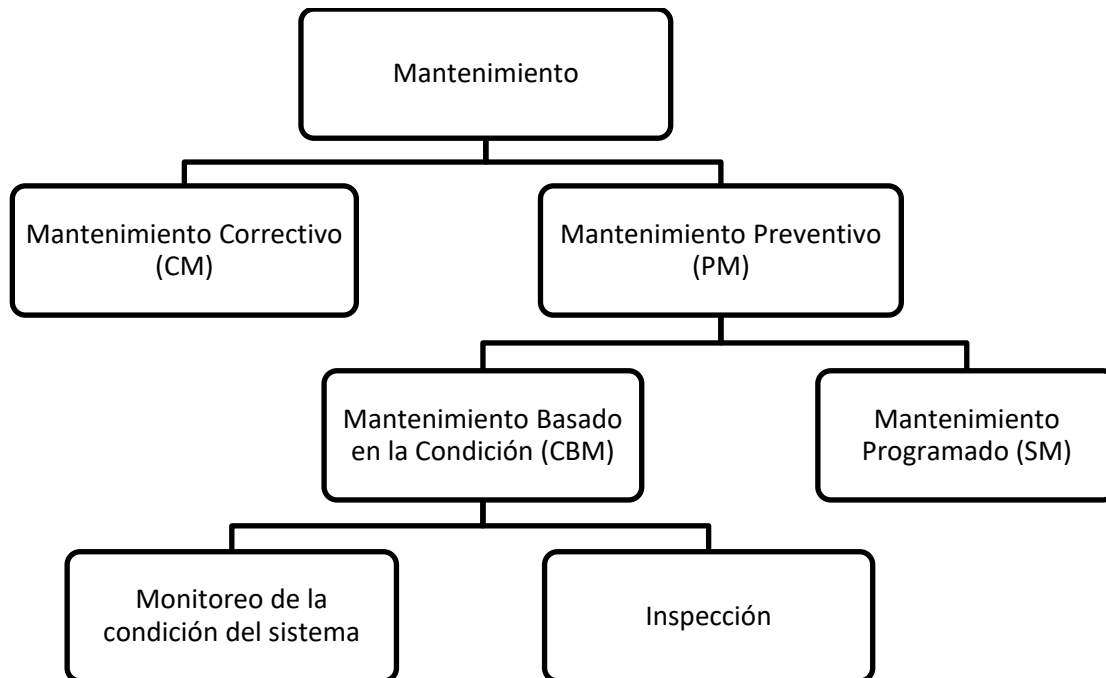
## 2.5. Gestión de mantenimiento

La gestión de mantenimiento se ha vuelto mucho más atractiva debido al avance tecnológico que existe; esta se puede definir como: “todas las actividades de la administración que determinan los objetivos o prioridades de mantenimiento, estrategias y responsabilidades, estas se implementan por medios tales como planificación de mantenimiento, control y supervisión de mantenimiento, y varios métodos de mejora que incluyen aspectos económicos” [22].

### 2.5.1. Conceptos y definiciones

El mantenimiento se puede definir como: “La combinación de todas las acciones técnicas, administrativas y de gestión durante el ciclo de vida de un activo destinado a mantenerlo, o restaurarlo a un estado en el que puede realizar la función requerida” [22].





**Figura 2.3. Tipos de estrategias de mantenimiento [9]**

Como se muestra en la Figura 2.3, el mantenimiento puede dividirse en dos categorías principales: correctivo (CM) y preventivo (PM).

- i) **Mantenimiento correctivo** se produce después de un fallo, y se reemplazan parte de los activos que han fallado. Durante el tiempo en que el mantenimiento correctivo se programa y se realiza, el activo está fuera de servicio. Como regla general, en un desglose a menudo es diez veces más caro en comparación con la situación en que la falla se puede identificar y corregir (o prevenir) de una manera planificada y programada. Las acciones correctivas son difíciles de predecir porque el comportamiento de falla es estocástico y las interrupciones en una red que causan interrupciones generalmente son imprevistas.
- ii) **Mantenimiento preventivo** se concentra en el cuidado y mantenimiento por parte de personal calificado con el fin de mantener el equipo y las instalaciones en condiciones de funcionamiento satisfactorias al permitir la inspección sistemática, la detección y la corrección de fallas incipientes antes de que ocurran o antes de que se conviertan en defectos importantes, así como la realización de pruebas, medición, ajustes y reemplazo de piezas, realizado específicamente para evitarlas [23].

El mantenimiento preventivo tiene los siguientes beneficios:

1. Aumenta la vida del equipo.

2. Reduce fallas.
3. Reduce el tiempo de inactividad.
4. Disminuye el costo de reemplazo.

El mantenimiento preventivo a su vez se puede clasificar como mantenimiento programado el cual es el mantenimiento realizado en un equipo basado en un calendario. Esto significa que el tiempo es el desencadenante de mantenimiento. El mantenimiento basado en el tiempo es mantenimiento planificado, ya que debe programarse con anticipación. Por otro lado, el mantenimiento basado en la condición difiere del mantenimiento basado en el tiempo en el sentido de que se realiza un cambio en los métodos de programación; es decir, de un método periódico a un método predictivo “totalmente”. Por lo tanto, CBM es una política predictiva. Ser predictivo se refiere a estimar la probabilidad de fallas en los activos. Con el mantenimiento basado en la condición, se puede detectar una indicación temprana de una falla inminente (mediante la aplicación de monitoreo de condición, métodos de diagnóstico o inspección) y se pueden evitar las consecuencias de una falla inesperada. CBM puede usar métodos de bajo nivel como los sentidos humanos para inspecciones (denominado mantenimiento basado en inspección) o implementar sofisticadas herramientas de monitoreo y diagnóstico (denotadas como mantenimiento basado en predicción). Más avanzado es el uso de varios parámetros de monitoreo y condición de la red para predecir los tiempos de vida restantes del componente (denotados como pronósticos o mantenimiento basado en la condición) [24].

El mantenimiento basado en la condición puede estar basado en la inspección o el monitoreo. La inspección se puede definir como la determinación de la condición física y tiene como objetivo identificar los signos de problemas pasados, presentes o futuros. Busca identificar cualquier corrosión, deterioro o daño notable (evidencia de que un automóvil golpeó un poste, o se ha disparado un aislador), o signos externos de mal funcionamiento interno (fuga de aceite alrededor de los sellos). La inspección física del dispositivo generalmente incluye un examen visual o infrarrojo del exterior de la unidad mientras está en servicio, pero puede incluir un examen detallado de sus partes internas (como por ejemplo al desmontar un interruptor de circuito). También puede incluir la evaluación de ruido e incluso olor (“¿Huele algo ardiendo?”), usando los sentidos, o un equipo especial para aumentarlos (escáneres infrarrojos, analizadores de señal acústica, llamados “inhaladores” hechos para detectar traza de gases en la atmósfera). Por otro lado, el monitoreo es la medición remota y el reporte de datos relevantes para su función y operación. La corriente, el voltaje y el factor de potencia medidos en un transformador de potencia o a través de un punto de medición proporcionan datos necesarios para observar cómo está funcionando el equipo y tomar buenas decisiones en cuanto a la operación continua de ese equipo. Por ejemplo, la temperatura a menudo se monitorea en un transformador y se usa para decisiones operativas de carga. De manera

similar, los interruptores automáticos y los interruptores podrían monitorearse para verificar su posición (abierto, cerrado, interruptor recargado y listo para operar, etc.) y reguladores de voltaje, cambiadores de tomas de carga o condensadores de bancos conmutados monitoreados en cuanto a su posición [25].

En la práctica, existe un conjunto de políticas y acciones de mantenimiento de diversos tipos y requiere una estructura general de toma de decisiones para seleccionar entre ellos. Esto es conocido en la literatura como conceptos de mantenimiento. La literatura proporciona una amplia gama de conceptos de mantenimiento que se han desarrollado a través de una combinación de conocimiento teórico y experiencia práctica. En [23,26], un resumen de los conceptos de mantenimiento más populares y sus características se dan e ilustran en este trabajo.

- i) **Costo del ciclo de vida:** Esta es una metodología para calcular y hacer un seguimiento del costo total de un sistema, desde su inicio hasta su eliminación. Se tienen en cuenta todos los gastos de compra, operación y disposición de una inversión, incluidos los costos del proyecto. Este método se origina a partir de los años 60 y está ganando interés nuevamente. Tal vez debido al enfoque en los ciclos de vida en los estándares PAS-55 e ISO 55000. El método sigue una serie de pasos según una estructura de desglose detallada del costo del sistema bajo investigación a lo largo de la vida. Este método se basa en una filosofía sólida, sin embargo, es intensivo tanto en recursos como en datos. Otro enfoque comúnmente utilizado es el método del Costo Total de Propiedad, que puede verse como una expansión del LCC. En los cálculos de TCO, se agregan los gastos o los costos indirectos, como el uso improductivo de los equipos y los costos totales de la cadena de suministro del negocio.
- ii) **Mantenimiento Productivo Total (TPM):** TPM tiene como objetivo obtener el uso más eficiente / efectivo de los equipos. En este método se necesita la participación total (en toda la organización). TPM promueve principalmente la implementación de tareas de mantenimiento preventivo basadas en tareas de grupos pequeños. Este método ha tenido éxito en la industria manufacturera, considera aspectos humanos y técnicos, sin embargo, lleva mucho tiempo implementarlo.
- iii) **Mantenimiento Centrado en la Confiabilidad:** RCM es un método estructurado enfocado en la confiabilidad y fue desarrollado inicialmente para alta tecnología y componentes de alto riesgo. Es un enfoque poderoso basado en un procedimiento paso a paso, sin embargo, consume muchos recursos y consume mucho tiempo.
- iv) **Basado en RCM:** En la literatura, se pueden encontrar varios métodos / conceptos inspirados en los principios de RCM. Por ejemplo, se desarrolló un concepto similar a RCM en el que el enfoque se centra en aspectos técnicos y organizativos más que en aspectos económicos. RCM es una metodología central para garantizar que una

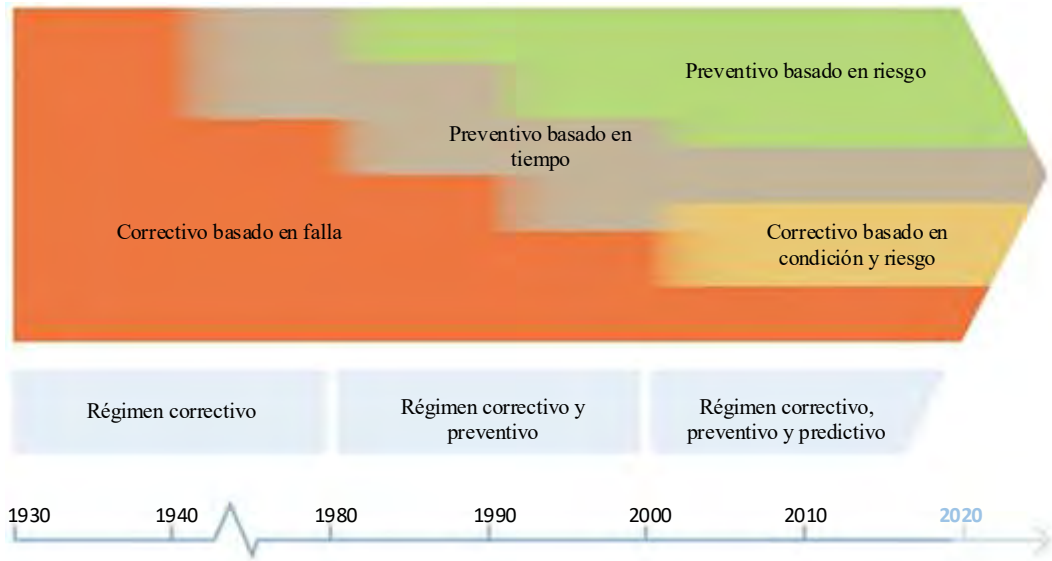
organización pueda lograr resultados de clase mundial. Sin embargo, se propone un nuevo RCM en este concepto donde las tareas de mejora de la calidad, se centran en los modos de falla más importantes en la empresa y la elaboración de paquetes de tareas son nuevas características. Mantenimiento Basado en el Riesgo (RBM) [32] es básicamente RCM, sin embargo, con una sólida base estadística. En la literatura, la gestión basada en los resultados se considera a veces como una política de mantenimiento o un concepto de mantenimiento. El RCM optimizado se ve como una versión simplificada o abreviada del método RCM tradicional, generalmente promovido por líderes industriales. Sin embargo, la RCM simplificada debe aplicarse cuidadosamente para no perder los beneficios de RCM.

- v) **Personalizado:** Estos conceptos generalmente se desarrollan en la empresa utilizando los beneficios de los conceptos existentes. Algunos ejemplos son el Mantenimiento Impulsado por el Valor, en el que la gestión de los valores de los accionistas está vinculada a las filosofías de mantenimiento tradicionales. Las empresas generalmente tienen su propio método único y prioritario y desean utilizar los beneficios de múltiples conceptos de mantenimiento existentes.

### 2.5.2. Gestión de mantenimiento en redes eléctricas

En la descripción general y la descripción de las tareas de mantenimiento, existe una gran cantidad de literatura disponible. En este contexto, se ha encontrado que la gestión del mantenimiento ha cambiado drásticamente en las últimas décadas. En el sector de la red eléctrica, se han producido desarrollos similares como resultado del cambiante entorno de mantenimiento. En la actualidad, existen diferentes niveles de tareas de mantenimiento mixto, políticas y conceptos implementados en las redes eléctricas de transmisión y distribución.

Si bien la administración de mantenimiento está evolucionando de regímenes reactivos y/o correctivos a más y más preventivos y/o proactivos, una migración completa a una sola tarea, política o concepto es poco probable. En la Figura 2.4, se muestra en un orden cronológico como se ha desarrollado la gestión del mantenimiento desde el pasado hasta la actualidad [27-30]. Sin embargo, esta es una representación general de los regímenes de gestión del mantenimiento y no refleja necesariamente la situación de las redes. En éstas el desarrollo ha seguido un ritmo más lento, principalmente, como resultado de la vida útil de los activos de las Redes Eléctricas de Potencia.



**Figura 2.4. Evolución del mantenimiento en la implementación de políticas.**

En este sentido, el mantenimiento podría posponerse por períodos más largos en comparación con otros sectores. Debido al hecho de que estos activos están llegando gradualmente al final de sus vidas técnicas, las estrategias de administración y reemplazo de mantenimiento son aspectos cruciales en la gestión de las poblaciones de activos. Sin embargo, estas tendencias también están sucediendo para las redes de transmisión y más recientemente, a un ritmo más rápido en las redes de distribución. La evolución en la gestión de mantenimiento, de acuerdo con la mayoría de la revisión de la literatura, se puede resumir y representar como se muestra en la Figura 2.4.

## 2.6. Gestión de mantenimiento en transformadores de potencia

La confiabilidad de un sistema de energía depende del adecuado funcionamiento de cada uno de los elementos que lo componen, entre los que se encuentra el transformador de potencia. Es por esta razón que en las empresas de energía eléctrica surge la necesidad de implementar estrategias de mantenimiento predictivo en los transformadores de potencia, para poder garantizar que los parámetros de operación de dichos equipos se mantengan en valores aceptables, que aseguren la prestación del servicio y a su vez, que éstos alcancen su vida útil. Los transformadores de potencia conforman la columna vertebral de los sistemas de transmisión y distribución de energía eléctrica de un país, son sin lugar a dudas los dispositivos más comunes en las redes eléctricas, de ahí su importancia estratégica ya que una salida inesperada de ellos puede conducir a grandes pérdidas por indisponibilidad y afectar grandes regiones de un país cualquiera. Su salida implica pérdidas incalculables a la economía y al medioambiente, de ahí la importancia de mantener su disponibilidad, confiabilidad y seguridad.

### **2.6.1. Diagnóstico y mantenimiento en los transformadores de potencia**

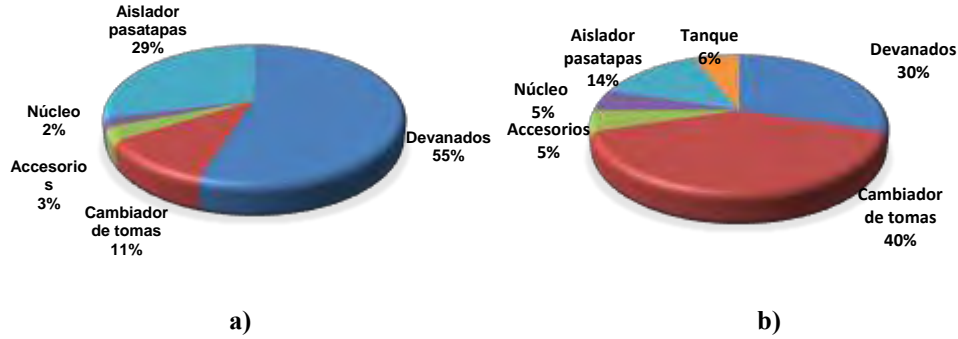
El diagnóstico y mantenimiento preventivo de transformadores de potencia es una tarea sistemática que deben realizar las diferentes empresas y especialmente las de generación, transmisión y distribución de energía eléctrica, con la finalidad de mantener en condiciones de operación óptimas los equipos y de identificar posibles condiciones de operación críticas, que puedan dar lugar a la aparición de incipientes. Esto permite una detección oportuna de las mismas para su corrección mediante un mantenimiento preventivo previo a una posible falla en el equipo, logrando así, ahorros importantes en los costos operativos [31].

De acuerdo [32], las pruebas que permiten diagnosticar el estado operativo de los transformadores de potencia se pueden clasificar en pruebas eléctricas, pruebas de aceites dieléctricos y pruebas de funcionamiento de sus protecciones eléctricas y mecánicas. Estas pruebas de diagnóstico están descritas con referencia a las categorías de sistemas y componentes que constituyen el transformador tales como: devanados o bobinas, aislador pasatapas o bushing, aceite aislante, conmutador de carga o cambiador de tomas, el núcleo, tanque o cuba, y los dispositivos asociados.

En cada uno de estos sistemas y subsistemas en [32] recomienda una serie de pruebas y mediciones que deben ser realizadas para registrar y evaluar su comportamiento. Sin embargo, la implementación de estas pruebas varía de acuerdo con la práctica regular de las empresas y puede depender de la historia operativa de los transformadores, así como de la disponibilidad de los equipos necesarios y recursos humanos para su implementación, ya que esto requiere de una gran inversión económica sobre todo en el área de equipamiento y entrenamiento del personal calificado. No siempre es suficiente con lo que indican las normas y fabricantes de los transformadores, sino que además se deben tomar en cuenta las estadísticas de fallas presentados por algunas instituciones de reconocido prestigio, para completar la información requerida y lograr un buen diseño de la estrategia de mantenimiento predictivo aplicable a transformadores de potencia [33].

### **2.6.2. Estadísticas de fallas en transformadores de potencia**

En la Figura 2.5, se pueden apreciar los resultados de las estadísticas de fallas en transformadores reportadas en [34].



**Figura 2.5. Porcentaje de falla por elemento de transformadores de potencia, según a) CIGRE y b) CFE México respectivamente**

En [35] se pueden encontrar las estadísticas de la CFE de México hasta 1999 y por otro lado, [36] muestra las estadísticas de la CIGRE, donde especifican los principales componentes susceptibles de fallas y su porcentaje en función del número de transformadores fallados.

De acuerdo a la Figura 2.5, los elementos mayormente fallados según los autores son: el cambiador de tomas, los devanados, tanque/aceite y los bushings que en su conjunto representan un promedio del 88 %, por lo que se debe hacer énfasis en estos elementos y realizar la identificación de los componentes críticos y evaluar en cada uno de ellos los modos de falla [31].

### 2.6.3. Identificación de los sistemas y componentes críticos

Basado en la información mostrada en la estadística anteriormente presentada y la documentación de trabajos realizados sobre las fallas más comunes en transformadores de potencia [37,38], se pueden identificar los sistemas y componentes más críticos que en el caso de fallar, representarían un riesgo inminente de daño del transformador e incluso un riesgo altamente potencial para el medio ambiente y las personas que se encuentran dentro del ámbito donde están instalados dichos equipos [37]. A continuación, en la Tabla 2.2, se muestran los sistemas y componentes críticos.

**Tabla 2.2. Sistemas y componentes críticos en un transformador**

<i>Sistema</i>	<i>Componente</i>
<i>Parte activa</i>	Devanados, núcleo
<i>Tanque/aceite</i>	Caja principal o cuba, aceite dieléctrico, tuberías y válvulas, tanque de expansión, radiadores, control de nivel de aceite de la cuba principal
<i>Cambiador de tomas o conmutador bajo carga</i>	Aceite dieléctrico, cuba, selector, pre-selector, mando motor, control de flujo del conmutador
<i>Bushing o aislador pasatapas</i>	Bushing del primario, bushing del secundario, bushing del terciario

### **2.6.4. Selección de las variables para fundamentar el mantenimiento predictivo basado en la condición.**

La medición periódica y sistemática de las variables eléctricas y la evaluación de las propiedades físico-químicas y análisis de gases disueltos en el aceite aislante del transformador, son parámetros suficientes, aunque no absolutos para establecer un monitoreo de la condición. El comportamiento de una gran cantidad de variables del transformador influye en la mayoría de los componentes críticos establecidos en la estadística de fallas anteriormente presentadas [32, 36, 38, 39, 40]. Esto representa una fortaleza en la gestión del mantenimiento ya que sirve de base referencial para la implementación de esta metodología.

Entre las pruebas eléctricas recomendadas por las normas [32], se encuentran: resistencia del aislamiento, prueba de aterramiento (*ground test*), temperatura (infrarrojo), resistencia del devanado, capacitancia, factor de potencia/factor de disipación, relación de transformación, ruidos, descargas parciales, entre otras. En las mismas normas [32, 40] se obtiene que las pruebas para determinar los niveles de las propiedades físico-químicas que indican el comportamiento del aceite dieléctrico son: contenido de agua, gases disueltos, esfuerzo dieléctrico, conteo de partículas, pérdidas dieléctricas, factor de potencia/factor de disipación, tensión interfacial, acidez, visual, color y estabilidad a la oxidación.

Las pruebas eléctricas y las propiedades físico-químicas del aceite indicadas anteriormente son variables que se controlan con técnicas de mantenimiento predictivo ya que el patrón de falla es desgaste progresivo (fallas periódicas). No aplica para patrones de falla de desgaste aleatorio [41].

### **2.6.5. Definición de los modos de fallas de los componentes críticos**

A continuación, se hace una descripción de los modos de fallas de los sistemas y componentes críticos del transformador y para efectos de este trabajo sólo se presenta para la parte activa; sin embargo, para obtener la totalidad del estudio de las definiciones de los modos de falla consultar [42].

#### **i) Sistema parte activa**

La mayoría de las fallas en la parte activa ocurren en los devanados por condiciones de cortocircuitos los cuales ejercen fuerzas axiales a través de una presión para desplazar de forma telescópica las bobinas del primario y del secundario, las que se repelen una de la otra debido a que las líneas eléctricas centrales no están alineadas. También existen fuerzas radiales en la bobina que tratan de desplazar las bobinas del primario y del secundario ocasionando que se fracture el aislamiento y falle el transformador. Esta ocurrencia de fallas es creciente y están asociadas a la antigüedad del aislamiento [38], ya que en estas condiciones de cortocircuitos se originan esfuerzos electrodinámicos que debido a su antigüedad ya no son



soportados por los enrollados. [43] plantea que: “en la operación de los transformadores de potencia éstos están sujetos a esfuerzos térmicos, eléctricos y mecánicos, los cuales provocan cierta degradación en el sistema de aislamiento. Las causas principales del fenómeno de degradación del sistema aislante son la temperatura excesiva, oxígeno, y humedad combinadas con los esfuerzos eléctricos, los cuales actúan como acelerador del proceso de degradación”. Las causas secundarias que aceleran el envejecimiento del sistema aislante son los esfuerzos mecánicos, los ácidos y lodos. Un exceso de cualquiera de estos esfuerzos puede acelerar el proceso de degradación. El proceso de degradación del sistema aislante puede evolucionar gradualmente hasta que se presenta una, la cual puede ser muy dañina. Estas fallas algunas veces ocurren sin ninguna alarma o señal de que un problema se está presentando. Pero en otras ocasiones existen pequeños indicios que indican la presencia de agentes dañinos o de deterioro y son detectables en los análisis periódicos al aceite aislante, en las mediciones de las pruebas eléctricas y en las mediciones de las descargas parciales [44]. La detección oportuna de estos indicios puede ser la diferencia entre someter el transformador a una reparación, sustituir una pieza dañada o tener un equipo fallado con todos los graves problemas que esto ocasiona.

En la Tabla 2.3 se muestra el análisis de fallas y sus efectos en cada uno de los sistemas principales desarrollados por los autores en [45].

**Tabla 2.3. Análisis de sistemas de transformadores de potencia**

<i>Sistema</i>	<i>Función</i>	<i>Modo de falla</i>	<i>Causas</i>	<i>Efecto</i>	<i>Modo de control</i>
<i>Parte activa-Bobina</i>	Constituye el circuito eléctrico que transfiere la energía de un circuito al otro y su función principal es la de crear el flujo magnético	Sobretensión Sobrecarga Cortocircuitos Cortocircuitos entre espiras Descargas parciales Corrosión	Sobretensiones producidas por operación del sistema o descargas atmosféricas Sobrecargas no admitidas Circulación de altas corrientes producto de fallas externas al transformados	Perdidas de energía Aumento de las corrientes de fuga (superficiales y volumétricas) que desencadenan y aceleran los procesos de envejecimiento Formación de gases disueltos Descomposición del aislamiento solido Chisporroteo a lo largo de grandes superficies. Pueden iniciarse descargas o una falla Deterioro del papel en la zona superficial en contacto con el aceite Cortocircuito entre espiras	Evitar los sobrecalentamientos en las bobinas Monitorear las temperaturas Realizar análisis físico químico y cromatografía de gases al aceite dieléctrico Realizar mediciones de la resistencia del devanado Realizar mediciones de la relación de transformación
<i>Parte activa -Núcleo</i>	Constituye el circuito magnético que transfiere la energía de un circuito a otro y	Aislamiento deficiente de los tornillos de afianzamiento del núcleo.	Carga excesiva. Calentamiento excesivo del núcleo.	Deterioro del aceite y de los materiales de aislamiento. Rompimiento de los terminales de los devanados.	Realizar pruebas periódicas de ruido teniendo en cuenta la hora y la carga en el momento de la prueba

	su función principal es la de conducir el flujo magnético	Canal de enfriamiento de aceite obstruido. Contacto a tierra deficiente. Aumento de las pérdidas en vacío. Aumento del ruido o sus decibeles.	Corriente de cortocircuito parcial. Pérdida del asilamiento entre láminas. Aflojamiento de los yugos.	Si se presenta una falla aguas abajo del transformador y esto provoca movimientos internos del mismo, estos movimientos son los que pueden en determinado caso que el núcleo se desplace o se rompa un perno de sujeción provocando que el núcleo sea afectado	
<i>Sistema Cuba-Aceite</i>	Refrigerante, aislante eléctrico y protector de las partes internas del transformador, teniendo a la vez buena estabilidad a la oxidación y una aceptable tendencia a la gasificación	Oxidación. Mal estado del radiador	Naturaleza o composición del aceite. Cantidad de oxígeno disponible para la reacción de oxidación. Presencia de agua y otros catalizadores de oxidación. Incremento excesivo de la temperatura. Agotamiento del inhibidor de oxidación Acumulación de sedimentos en la obleas o en el tubo, el flujo del aceite se dificulta y la temperatura desciende	Formación de lodo el cual limita las capacidades de refrigeración, aislamiento y protección. Presencia de fugas de aceite en las cabeceras del radiador y de las partes soldadas del panel o del tubo	En las pruebas del aceite se evidencia el deterioro del mismo y si algunos valores están por debajo de los valores permitidos se deben realizar mantenimiento al aceite (en caliente o en frío). Esto con el objetivo de mejorar las condiciones del transformador mejorando el aceite en sus características de refrigerante y dieléctrico.
<i>Conmutador</i>	Permitir la regulación de voltaje en el devanado secundario de un transformador a través de un cambiador de tomas que actúa en los puntos de conexión a lo largo de un devanado que permite seleccionar el número de espiras.	Existencia de alta fuente de contaminación y esfuerzos eléctricos, descargas y extinción de arco eléctrico producto de la disipación de energía al interior del interruptor.	Generación de gran cantidad de carbón, agua y gases. Reducción de la capacidad de transferencia de corriente del contacto e incremento de la temperatura.	Aparición de sobrecalentamientos localizados con temperaturas mayores a 200 grados centígrados. Degradación del aceite generando subproductos que se acumulan sobre la superficie de los componentes y contactos ubicados en el interior	Realizar pruebas al aceite con el número de maniobras realizadas se debe programar el mantenimiento del conmutador para realizar según sea el caso cambio de contactos, cambio de resistencia de transición, cambio total de aceite, lavado y limpieza general

<i>Bushing</i>	Permite el paso de la corriente a través de las bobinas del transformador y evita que haya un cortocircuito entre las fases de los devanados y entre los devanados y la cuba.	Son elementos capacitivos y es donde hay mayor esfuerzo dieléctrico concentrado en muy poco volumen, por lo que su condición debe evaluarse y asegurarse de una forma más intensiva que la de otros elementos.	Ingreso de contaminación externa, principalmente humedad, a causa del deterioro del empaque. Sobretensiones	Aparición de puntos calientes que aceleran el envejecimiento de los empaques y sellos. Depósito de impurezas en el canal de separación entre el cuerpo capacitivo y la porcelana como subproducto del deterioro del aceite del mismo. Generación de gases, especialmente nitrógeno, en el interior del bushing.	Realizar termografías periódicas o cada vez que se quiera aumentar la carga del transformador. Realizar pruebas de asilamiento y factor de potencia a los bushing en los periodos de mantenimiento programados.
----------------	---	--	---	---	---

En la Tabla 2.3 se describe cada sistema del transformador; además, se describen las funciones, las posibles fallas, las causas que pueden provocar estas fallas, las consecuencias que pueden tener sobre el transformador, así como las acciones de mantenimiento predictivo para minimizar la ocurrencia de averías en cada uno de estos sistemas. Para garantizar la operación óptima en los transformadores es necesario implementar un programa de inspección y medición de las variables eléctricas que permita en un corto plazo llevar los registros históricos de cada una de las pruebas realizadas y esto permite evaluar la variación de éstas en el tiempo y determinar las condiciones operativas.

Además, el monitoreo de las propiedades físico-químicas y análisis de los gases disueltos en los aceites dieléctricos reflejan la degradación y el deterioro del aislamiento del transformador, de tal manera que se pueden planificar las actividades de mantenimiento e inferir en cuanto a su periodicidad y muestreo en el tiempo.

Del análisis anterior se puede obtener una conclusión importante y es que el aceite del transformador además de ser un refrigerante y dieléctrico, proporciona una información importante acerca del estado técnico del transformador y de su estado operativo, siguiendo esta idea se propone desarrollar un algoritmo que permita desarrollar una estrategia de mantenimiento predictivo para transformadores utilizando el sistema de supervisión, control, y adquisición de datos (SCADA), lo que va a permitir el monitoreo de la condición utilizando como variable a monitorear la temperatura del aceite para obtener información de funcionamiento anómalo en el transformador.

El presente trabajo pretende usar las mediciones en línea de los parámetros eléctricos en conjunto con la temperatura ambiente y del transformador para desarrollar un algoritmo que evalúa la condición del transformador a partir de la estimación de la temperatura del aceite en el interior del transformador. Este método va a servir de complemento a los métodos utilizados en la literatura investigada, ayudando así, a mejorar la evaluación de la condición de los transformadores.

---

## CAPÍTULO III: INTRODUCCIÓN AL ESTUDIO DE MÉTODOS PREDICTIVOS Y SU APLICACIÓN EN REDES ELÉCTRICAS DE POTENCIA

### 3.1. Introducción

Para lograr una estrategia de mantenimiento predictiva es necesario contar con métodos matemáticos que a partir de un conjunto de variables sea capaz de modelar una o varias variables que describan el comportamiento físico y eléctrico de uno a varios activos de las redes eléctricas. En la literatura se pueden encontrar varios métodos predictivos utilizados para hacer mantenimiento predictivo, siendo los más utilizados: las cadenas de Márkov, método de Montecarlo y los métodos de Aprendizaje Automático (*Machine Learning*).

En este capítulo, se realiza una breve descripción de cada uno de los métodos mencionados anteriormente haciendo énfasis en el Aprendizaje Automático y específicamente las Redes Neuronales Artificiales, así como, algunas de sus aplicaciones en Redes Eléctricas de Potencia, destacando los beneficios que pueden aportar a un mantenimiento basado en la condición de los activos. También, se presentan las principales estructuras utilizadas en RNA; así como, algunos de los algoritmos de aprendizaje y funciones de activación utilizados por éstas.

### 3.2. Cadenas de Márkov

#### 3.2.1. Procesos Estocásticos

Un proceso estocástico se puede definir como una colección indexada de variables aleatorias  $X_t$ , donde el índice  $t$  toma valores de un conjunto  $T$  dado. Un proceso estocástico busca describir el comportamiento de un sistema en operación, durante periodos de tiempo. Un proceso estocástico tiene la siguiente estructura:

El estado actual del sistema puede estar en una de las  $M + 1$  categorías mutuamente excluyentes llamadas estados. Por conveniencia en la notación, estos estados se etiquetan  $0, 1, 2, \dots, M$ . La variable aleatoria  $X_t$  representa el estado del sistema en el tiempo  $t$ , de manera que sus únicos valores posibles son  $0, 1, \dots, M$ . El sistema se observa en puntos del tiempo dados, etiquetados  $t = 0, 1, 2, \dots$  así, el proceso estocástico  $X_t = X_0, X_1, X_2, \dots$  proporciona una representación matemática de cómo evoluciona el estado del sistema físico [45].

Lo anterior hace referencia a procesos estocásticos de *tiempo discreto* con *espacio de estados finito*.

### 3.2.2. Cadenas de Márkov

Según [46] todo proceso estocástico  $\{X_t\}$  que cumple con la siguiente propiedad markoviana esencial se puede denominar como una cadena de Márkov:

$$P(X_{t+1} = j | X_0 = k_0, X_1 = k_1, \dots, X_{t-1} = k_{t-1}, X_t = i) = P(X_{t+1} = j | X_t = i), \quad (3.1)$$

Para  $t = 0, 1, 2, \dots$ , y toda sucesión  $i, j, k_0, k_1, \dots, k_{t-1}$ .

Lo que quiere decir que esta propiedad markoviana es “*que la probabilidad condicional de cualquier evento futuro dados cualquier evento pasado y el estado actual  $X_t = i$  es independiente del evento pasado y sólo depende del estado actual del proceso*” [45].

Las probabilidades condicionales  $P(X_{t+1} = j | X_t = i)$  para una cadena de Márkov se llaman probabilidades de transición (de un paso). Si para cada  $i$  y  $j$ ,

$$P(X_{t+1} = j | X_t = i) = P(X_1 = j | X_0 = i), \text{ para toda } t = 1, 2, \dots, \quad (3.2)$$

Entonces, se dice que las probabilidades de transición (de un paso) son estacionarias. Así, tener **probabilidades de transición estacionarias** implica que las probabilidades de transición no cambian con el tiempo. La existencia de probabilidades de transición (de un paso) estacionarias también implica que, para cada  $i, j$  y  $n$  ( $n = 0, 1, 2, \dots$ ),

$$P(X_{t+n} = j | X_t = i) = P(X_n = j | X_0 = i), \quad (3.3)$$

Para toda  $t = 0, 1, \dots, n$ . Estas probabilidades condicionales se llaman **probabilidades de transición de  $n$  pasos**.

Así, las probabilidades de transición de  $n$  pasos  $p_{ij}^{(n)}$  son simplemente la probabilidad condicional de que el sistema se encuentre en el estado  $j$  justo después de  $n$  pasos (unidades de tiempo), dado que comenzó en el estado  $i$  en cualquier tiempo  $t$ . Cuando  $n = 1$ , observe que  $p_{ij}^{(n)} = p_{ij}^1$ .

Como las  $p_{ij}^{(n)}$  son probabilidades condicionales, deben ser no negativas y como el proceso debe hacer una transición a algún estado, deben satisfacer las propiedades:

$$p_{ij}^{(n)} \geq 0, \text{ para toda } i \text{ y } j; n = 0, 1, 2, \dots \quad (3.4)$$

y

---

<sup>1</sup>Para  $n = 0$ ,  $p_{ij}^{(0)}$  es  $P\{X_0 = j | X_0 = i\}$ , que es igual a 1 cuando  $i = j$  y es 0 cuando  $i \neq j$

$$\sum_{j=0}^M p_{ij}^{(n)} = 1, \text{ para toda } i; n = 0,1,2, \dots \quad (3.5)$$

Una notación conveniente para representar las probabilidades de transición de  $n$  pasos es la forma matricial:

$$P^{(n)} = \begin{pmatrix} p_{00}^n & p_{01}^n & \dots & p_{0M}^n \\ p_{10}^n & p_{11}^n & \dots & p_{1M}^n \\ \vdots & \vdots & \ddots & \vdots \\ p_{M0}^n & p_{M1}^n & \dots & p_{MM}^n \end{pmatrix} \quad (3.6)$$

Se puede observar en (3.6) que la probabilidad de transición en un renglón y columna dados es para la transición del estado en ese renglón al estado en la columna. Cuando  $n = 1$ , el superíndice  $n$  no se escribe y se hace referencia a ésta como la matriz de transición [46].

### 3.2.3. Aplicación de las Cadenas de Márkov en transformadores

Cadenas de Márkov, es una técnica que se ha venido utilizando en el análisis de redes eléctricas. En particular y por ser referente al presente trabajo a continuación se presentan dos trabajos sobre mantenimiento predictivo en transformadores

En [47], se desarrolla un modelo de mantenimiento de Márkov para predecir la vida útil restante del aislamiento del transformador. Se incorporan varias etapas de deterioro del aislamiento y un estado de mantenimiento menor, debido a que se alcanza un estado actual de envejecimiento del aislamiento, a partir de una prueba e inspección de diagnóstico. El modelo es capaz de calcular el tiempo promedio previsto antes de que ocurra una falla de aislamiento. El algoritmo desarrollado en este modelo se implementó en un programa de computadora utilizando Matlab. También se discuten los datos requeridos para este programa y el ejemplo numérico basado en el algoritmo desarrollado.

Por otro lado, en [48] se presenta un estudio sobre la aplicación de un Modelo de Márkov (MM) para determinar el índice de salud (HI) de un grupo de transformadores. En total se analizan 3195 muestras de aceite de 373 transformadores con edades comprendidas entre 1 y 25 años. En primer lugar, el índice de salud de los transformadores se calculó en base a los datos anuales de monitoreo de la condición del aceite que consistían en la calidad del aceite, gases disueltos y compuestos furanos. Con estos datos, se calcula el HI promedio para cada edad y se obtienen las probabilidades de transición basadas en una técnica de optimización no lineal. Finalmente, la curva de rendimiento de deterioro futuro de los transformadores se determina con base en el algoritmo de la cadena MM. Se encontró que el MM se puede utilizar para predecir los estados futuros de la condición de los transformadores. El análisis de sensibilidad de ajuste de chi cuadrado reveló que el HI pronosticado para el grupo de

transformadores obtenida en base al MM está de acuerdo con el HI promedio calculado a lo largo de los años, y el error promedio es del 3,59%.

### 3.3. Método de Montecarlo

El método de Montecarlo fue nombrado así por la ciudad de Montecarlo en Mónaco donde se juega “la ruleta”, el juego de azar que genera resultados aleatorios. Este método surge formalmente en el año 1944, sin embargo, ya existían prototipos y procesos anteriores que se basaban en los mismos principios.

El empleo del método de Montecarlo para fines de investigación comenzó con el desarrollo de la bomba atómica en la Segunda Guerra Mundial en el Laboratorio Nacional de Los Álamos. Durante el desarrollo de este proyecto, los científicos Von Neumann y Ulam perfeccionaron la técnica y la aplicaron a problemas de cálculo de difusión de neutrones en un material. Alrededor de 1970, los desarrollos teóricos en complejidad computacional comienzan a proveer mayor precisión y relación para el empleo del método Montecarlo [49].

Actualmente el método de Montecarlo a veces es usado para analizar problemas que no tienen un componente aleatorio explícito; en estos casos un parámetro determinista del problema se expresa como una distribución aleatoria y se simula dicha distribución. La simulación de Montecarlo también fue creada para resolver integrales que no se pueden resolver por métodos analíticos, para solucionar estas integrales se usaron números aleatorios. Posteriormente fue utilizado para cualquier esquema que emplee números aleatorios, usando variables aleatorias con distribuciones de probabilidad conocidas.

En el método de Montecarlo se combinan conceptos estadísticos como lo es el muestreo aleatorio, con la generación de números aleatorios y la automatización de los cálculos. Es un procedimiento matemático que consiste en la generación numérica de series mediante un muestreo aleatorio de las distribuciones de probabilidad.

#### 3.3.1. Descripción del algoritmo

El algoritmo de Simulación Montecarlo fundamentado en la generación de números aleatorios, se basa en las distribuciones acumuladas de frecuencias:

- Determinar la o las variables aleatorias y sus funciones de distribución de probabilidad.
- Generar un número aleatorio uniformemente distribuido en  $(0,1)$ .
- Determinar el valor de la variable aleatoria para el número aleatorio generado de acuerdo a la función de densidad de probabilidades (pdf) que tengamos.
- Iterar los dos pasos anteriores tantas veces como muestras necesitamos.
- Calcular la media, desviación estándar, error y realizar el histograma.

- Analizar resultados para distintos tamaños de muestra.

Otra opción para trabajar con el método Montecarlo, cuando la variable aleatoria no es directamente el resultado de la simulación o tenemos relaciones entre variables es la siguiente:

- Diseñar el modelo lógico de decisión.
- Especificar las distribuciones de probabilidad para las variables aleatorias relevantes.
- Incluir posibles dependencias entre variables.
- Muestrear valores de las variables aleatorias.
- Calcular el resultado del modelo según los valores del muestreo (iteración) y registrar el resultado.
- Repetir el proceso hasta tener una muestra estadísticamente representativa.
- Obtener la distribución de frecuencias del resultado de las iteraciones.
- Calcular media, desviación.
- Analizar los resultados.

Las principales características a tener en cuenta para la implementación o utilización del algoritmo son:

- El sistema debe ser descrito por una o más funciones de (pdf).
- El proceso de generación de los números aleatorios es importante para evitar que se produzca correlación entre los valores de muestra.
- Establecer límites y reglas de muestreo para las pdf; conocemos qué valores pueden adoptar las variables.
- Definir cuando un valor aleatorio tiene o no sentido para el modelo a simular.
- Estimar con qué error trabajamos, cuánto error podemos aceptar para que una corrida sea válida.
- Paralelización y vectorización: En aplicaciones con muchas variables se estudia trabajar con varios procesadores paralelos para realizar la simulación [50].

### 3.3.2. Aplicación del método de Montecarlo en REP

En [51] se realiza un estudio que evalúa la eficiencia económica de sistemas de generación distribuida. Este estudio se basa en el costo de la capacidad de generación de energía distribuida. Los cálculos de las mejores ganancias obtenibles se hicieron mediante una simulación de Montecarlo. Las simulaciones de Montecarlo que se basan en el muestreo aleatorio repetido para calcular sus resultados tienen en cuenta el costo de la producción de electricidad, las cargas diarias y el costo de ventas y generan un resultado más rápido que los cálculos matemáticos. También se sugiere el diseño óptimo, que considera que la pérdida de



distribución asociada con los sistemas de distribución de energía eléctrica se centra en el aspecto de detección y la generación de energía distribuida.

Por otro lado, técnicas de monitoreo de la condición de transformadores de la industria son ampliamente utilizadas por las empresas eléctricas para evaluar las condiciones de los sistemas de aislamiento de papel y aceite en los transformadores de la industria. Entre los métodos de monitoreo existentes, el análisis de gases disueltos (DGA) es una de las técnicas más comúnmente utilizadas en la industria energética. En [52], se realiza la predicción de la salud de transformadores industriales basados en la técnica de DGA. La simulación de Montecarlo se realiza con base en el método DGA para estimar la confiabilidad de series de tiempo de los transformadores de la industria. Se han desarrollado varios modelos de diagnóstico basados en DGA para identificar los tipos de falla de los transformadores de la industria. Sin embargo, la predicción de salud del transformador también es significativa en la industria.

### 3.4. Aprendizaje Automático

El Aprendizaje Automático (*Machine Learning (ML)*) es el tópico principal de la inteligencia computacional que se ocupa de la cuestión de cómo construir programas informáticos que mejoren automáticamente con la experiencia. [53]

En otras palabras, se trata de construir máquinas que adapten y modifiquen sus acciones o predicciones de tal forma que sean más precisas. Para comprender correctamente el Aprendizaje Automático, es útil entender y definir primero el aprendizaje. El aprendizaje es una función que permite, a las máquinas, modificar y reforzar o adquirir nuevos conocimientos, habilidades, preferencias y comportamientos [54].

Con algunos conocimientos básicos de aprendizaje, ahora podemos avanzar hacia el Aprendizaje Automático. ML es simplemente la implementación de capacidades de aprendizaje en máquinas. En [55] se define que el Aprendizaje Automático es lo que da a las computadoras la capacidad de aprender sin estar explícitamente programadas.

El Aprendizaje Automático es un campo de investigación dentro de la informática. En la mayoría de las referencias consultadas se busca la construcción de algoritmos que pueden aprender y luego predecir, una predicción es una declaración de que se espera algún resultado definido. Los algoritmos de ML se basan en un modelo a partir de datos de ejemplo. El Aprendizaje Automático difiere de la computadora tradicional en ciencias. La informática tradicional escribe códigos específicamente para el dominio del problema y utiliza un conocimiento a priori de reglas de código rígido. Sin embargo, los algoritmos de Aprendizaje Automático no están diseñados para un solo problema, el algoritmo mismo aprende un

modelo que se ajusta a muestras en los datos y por extensión crea sus propias reglas autodidactas.

ML no es una tarea trivial. Después de décadas de investigación, no se puede decir que alguna máquina tenga capacidades de aprendizaje similares a las de los humanos. En el Aprendizaje Automático: un enfoque de inteligencia artificial [54], los autores afirman que el problema de implantar capacidades de aprendizaje en las computadoras es muy exigente. Y algunos investigadores afirman que hacer que las máquinas aprendan, es una de las metas más desafiantes en la inteligencia artificial.

Sin embargo, se han obtenido éxitos en muchos tipos de datos y problemas. Entre ellos, estos algoritmos se han utilizado con éxito para modelar series de tiempo financieras [56], así como, perfiles de carga en sistemas eléctricos de potencia entre otros. Además, se han utilizado con éxito en el monitoreo de la condición en turbinas eólicas.

### **3.4.1. Datos de muestra y modelos de construcción**

Un requisito previo para el aprendizaje es que haya algo de lo que aprender, es decir, se necesitan muestras. Para las máquinas, estas muestras son conjuntos de datos. Para que un algoritmo de ML derive un modelo estable y duradero, es necesario que los datos sean adecuados. También se necesita una correlación o un contexto entre los datos de muestra. Puede haber varios problemas con los datos de muestra que hacen que sea más difícil derivar un modelo sostenible para la predicción. Los datos deben ser estacionarios, lo que significa que la distribución no cambia con el tiempo. El grado de ruido aleatorio en los datos de muestra no debe ser demasiado alto y las entradas iguales deben producir salidas iguales.

Cuando se recopilan datos, generalmente se divide en dos grupos. Estos grupos de conjuntos de datos se llaman conjunto de entrenamiento y conjunto de prueba. El conjunto de entrenamiento es la información que aplica para construir el modelo. Este conjunto se utiliza para ajustar el modelo de ML, emparejando la entrada con la salida esperada y tratando de encontrar alguna relación entre las variables y de ese modo, minimizar el error. El proceso de ajuste se repite hasta que el modelo haya minimizado los errores hasta el punto en que haya alcanzado un umbral mínimo. Cuando uno finaliza el proceso, a menudo se decide con la validación cruzada, que se analiza con más detalle en la sección 3.5.2. Después de ajustar el modelo, el conjunto de prueba se pone en uso. Como el modelo no puede ver el conjunto de pruebas, podemos usarlo para obtener una medida final sobre qué tan bien el modelo de Aprendizaje Automático adapta los datos; esta medida debe indicar qué tan bien el modelo interpretará los datos del mundo real [57].

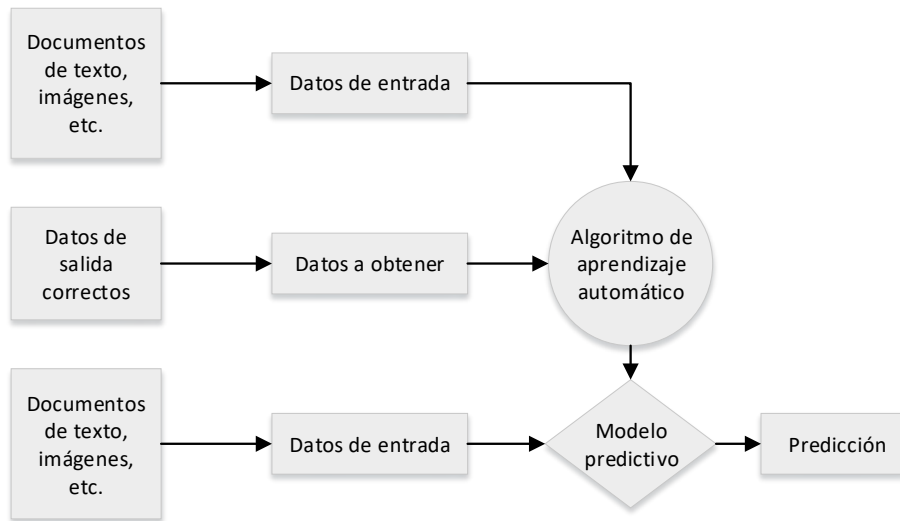
### 3.4.2. Aprendizaje no supervisado

¿Puede aprender algo útil la red si no se le proporciona un conjunto de entrenamiento con la salida deseada de cada entrada? La respuesta es positiva [58]. Esta forma de aprendizaje se denomina aprendizaje no supervisado (*unsupervised learning*) ya que no se necesita ningún “profesor” que muestre la salida deseada a la red con cada entrada. Dado un conjunto de entradas, a la red se le permite “jugar” con ellos para intentar descubrir regularidades y relaciones entre las diferentes partes de las entradas.

El aprendizaje se produce frecuentemente a través de alguna noción sobre qué características del conjunto de entrada son importantes. Sin embargo, con frecuencia esto no es conocido de antemano. El aprendizaje no supervisado se puede utilizar como un módulo descubridor de características, o agrupador de entradas similares, que preceda a un aprendizaje supervisado [59].

### 3.4.3. Aprendizaje supervisado

Para el aprendizaje supervisado hay un conjunto de datos de entrada y de salida, siendo los datos de salida la respuesta que el algoritmo debe producir a partir de la entrada. Estos dos conjuntos de datos combinados generalmente se denominan datos de entrenamiento. Los datos objetivos pueden ser preparados por expertos o personas normales. Si tuviéramos un conjunto de entrenamiento infinitamente grande, con cada entrada y resultado posible de un problema, podríamos haber hecho una tabla de búsqueda, donde simplemente buscaríamos la entrada y luego la salida correspondiente. Desafortunadamente, no sería factible desde el punto de vista computacional, ya que ninguna computadora puede almacenar una cantidad infinita de datos, ni es posible generar todas las entradas y salidas posibles para la mayoría de los problemas del mundo real. Esta es la razón por la cual el aprendizaje supervisado es útil, porque el aprendizaje supervisado puede hacer que la computadora sea capaz de generalizar. La Figura 3.1 muestra el modelo estándar para el aprendizaje supervisado.



**Figura 3.1. Modelo de Aprendizaje Supervisado**

El algoritmo de Aprendizaje Automático debería producir una salida sensible para la entrada. Al generalizar, los algoritmos separan el ruido, o pequeños errores, en los datos deseados [60].

Dentro del aprendizaje supervisado podemos dividir los tipos de problemas en dos subgrupos: i) Regresión, predicción de valores continuos y ii) Clasificación, prediciendo categorías discretas.

### **i) Regresión**

Como se dijo, la regresión se trata de predecir valores. La regresión predice un valor continuo, se realiza estimando las relaciones entre las variables de entrada y el resultado, siendo la relación cómo un cierto cambio en la entrada que afectaría el resultado. Un ejemplo de regresión en el Aprendizaje Automático es la predicción de la actividad cerebral utilizada en [60] donde predijeron con éxito ciertos valores de actividad cerebral en áreas del cerebro después de algunos estímulos. Otro ejemplo podría ser predecir el precio futuro de una acción basada en los precios actuales. Un gráfico típico de regresión se muestra en la Figura 3.2.

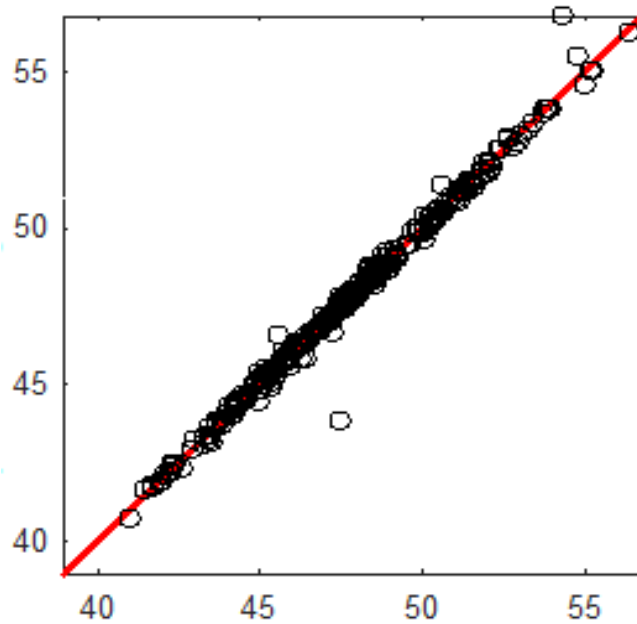


Figura 3.2. Modelo de Regresión

## ii) Clasificación

La clasificación es la predicción de valores discretos. El problema es tomar una nueva entrada u observación y decidir a cuál, de varias clases, pertenece la entrada, sabiendo que cada entrada pertenece exactamente a una clase. Se realiza en base a los ejemplos de formación de cada clase y, por lo tanto, es un algoritmo de aprendizaje supervisado. Un ejemplo de clasificación sería asignar un diagnóstico a un paciente basándose en las observaciones, las observaciones podrían ser proporción de glóbulos blancos, los síntomas, la frecuencia cardíaca, el género, etc. En la Figura 3.3 vemos un ejemplo de clasificación, un separador lineal divide las dos clases.

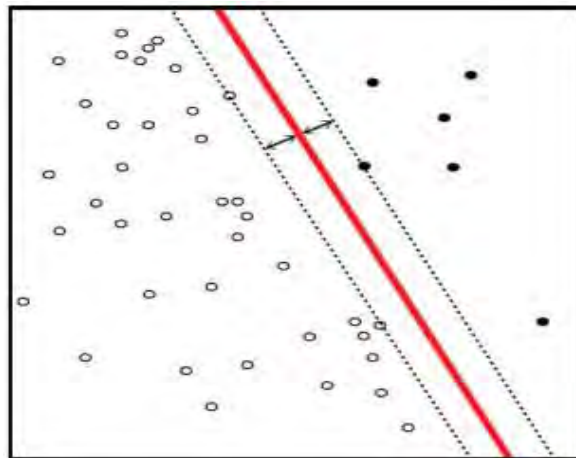


Figura 3.3. Modelo de Clasificación

La naturaleza discreta de la clasificación, sin embargo, no siempre es realista en problemas del mundo real, algunos ejemplos pueden pertenecer, o estar precisamente en el borde de dos o más clases diferentes. También hay problemas cuando es imposible categorizar cada entrada posible. Otro ejemplo de clasificación es la predicción del nivel de rendimiento del índice bursátil realizado por [61] donde los buscadores probaron varios modelos de clasificación y las existencias pronosticadas regresan con diversos éxitos. En el estudio usaron las direcciones del movimiento de las acciones como categorías.

### 3.5. Problemas inherentes al Aprendizaje Automático

#### 3.5.1. Sobre- y bajo-ajuste

Uno de los problemas más importantes para realizar el Aprendizaje Automático es el sobre- y bajo-ajuste. Los problemas ocurren debido al ruido. La Figura 3.4 intenta mostrar los problemas de sobre- y bajo-ajuste. Para explicar mejor el sobre- y bajo-ajuste, podemos usar el principio de la “Navaja de Occam” para deducir cómo y por qué ocurre este problema. Como sabemos, el aprendizaje automático intenta construir un modelo para que, dado un conjunto de entradas, pueda proporcionar el resultado deseado. El principio de la “Navaja de Occam” afirma que un modelo debe contener todo lo que es necesario para el modelado, pero nada más.

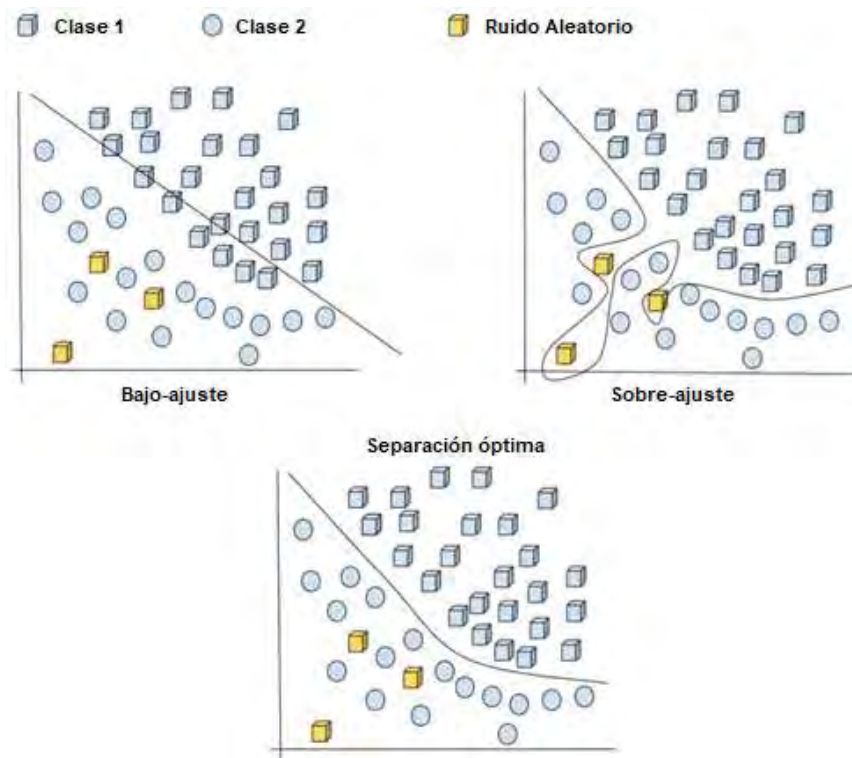


Figura 3.4. Modelo de sobre- y bajo-ajuste

En términos de Aprendizaje Automático, esto significa que el conjunto de entradas debe contener exactamente lo que se necesita para una buena predicción, lo que significa que el modelo incluye más términos de los necesarios o usa enfoques más complicados de lo necesario [62]. Cuando el modelo hace hincapié en que tiene muy poco error, el modelo crea un límite de decisión que es demasiado complicado e incluye el ruido. Cuando el modelo permite un error demasiado grande, no puede dividir correctamente las clases. Estos problemas pueden ser difíciles de manejar. Y desafortunadamente, casi todos los conjuntos de datos del mundo real contienen algún tipo de ruido. Lo que sucede es que se intenta predecir el ruido y crea un modelo que requiere demasiadas entradas o es demasiado complejo.

### 3.5.2. Validación cruzada

Para evitar los problemas de sobre- y bajo- ajuste y mantener el poder de generalización del algoritmo de aprendizaje automático, se utiliza la validación cruzada (*cross validation*). La validación cruzada es un método de evaluación modelo. El problema con simplemente dividir el conjunto de datos en una parte para las pruebas y una parte para construir el modelo es que no hay forma de saber cuándo detener la adaptación del modelo. No hay forma de saber cuándo el modelo se está sobre-ajustando y cuándo el modelo está con bajo-ajuste. Para contrarrestar este problema, la validación cruzada elimina parte de la información antes de que comience el entrenamiento. Cuando finaliza el entrenamiento, los datos que se eliminaron se utilizan para probar el rendimiento del modelo ajustado con datos no vistos. Hay varias formas diferentes de validación cruzada. Una de las más comunes se conoce como *K – Fold*. Los datos se dividen en  $k$  conjuntos, y uno de los conjuntos se utiliza para la prueba, mientras que los otros conjuntos  $k - 1$  se utilizan para el entrenamiento. El proceso se repite  $k$  veces. *K – Fold* es una técnica completa de validación cruzada, pero tiene algunos problemas con el tiempo y el consumo computacional. Es común establecer  $k$  como 10 [63].

### 3.5.3. Análisis de series temporales

Las series temporales son observaciones ordenadas en secuencias, generalmente ordenadas por tiempo. El análisis de series de tiempo difiere de otros análisis porque las observaciones dependen del tiempo. Esto requiere algunas consideraciones al realizar el Aprendizaje Automático, ya que no se puede aleatorizar el orden de las entradas ni se puede afirmar con certeza de cuántas de las observaciones anteriores depende una observación actual. Por esta última razón, se debe realizar una prueba preliminar sobre qué tan atrás en el tiempo uno se ve al intentar hacer una predicción. Los estudios han demostrado que los algoritmos de ML se ajustan a las series de tiempo de pronósticos [64] y que el Aprendizaje Automático se puede utilizar completamente para uno o dos pasos de predicción [65].

### 3.5.4. Normalización

Un problema común con los conjuntos de datos es que la escala está desactivada, lo que significa esencialmente que algunos de los puntos de datos son simplemente mucho más grandes que otros sin representar necesariamente un cambio grande. Un ejemplo de los mercados de valores es el volumen de acciones vendidas, que puede variar de cero a millones, mientras que el cambio de un precio de una acción varía con un cierto porcentaje. Sin embargo, un pequeño porcentaje del cambio en el precio podría ser un mejor indicador del precio de las acciones de mañana que un cambio de miles en el volumen. Estas diferencias de escalado pueden causar problemas para los algoritmos de aprendizaje automático y, por lo tanto, puede ser conveniente aplicar algunas escalas de características al conjunto de datos antes de usarlo para el aprendizaje automático. Para ilustrar mejor el problema, podemos ver los algoritmos de ML que utilizan la distancia euclidiana. Si una de las características tiene una amplia gama de valores, como el volumen de almacenamiento, la distancia se regirá por esa característica particular. Una forma de manejar estos problemas se conoce como normalización, que es el proceso de cambiar uniformemente la amplitud de un conjunto de datos, de modo que ninguna característica varía en un rango mayor que el otro [57].

### 3.5.5. Maldición de dimensionalidad y selección de características

Existen muchos factores que pueden influir en una variable a predecir. Este es un problema común en muchas aplicaciones de ML y puede llevar a los diseñadores de algoritmos de Aprendizaje Automático a la tentación de simplemente incluir grandes cantidades de datos para evitar el sesgo de la variable omitida. Sin embargo, incluir demasiados datos viene con muchos otros problemas, uno de los cuales se conoce como la maldición de la dimensionalidad. El término fue introducido por Bellman para describir los problemas que ocurren cuando uno aumenta las dimensiones de una entrada. El problema surge debido al aumento exponencial del volumen causado al agregar más dimensiones [66, 67]. Cuando aumenta la complejidad de la información, la complejidad de los patrones subyacentes también puede aumentar. Para extraer estos patrones más complejos, podemos necesitar más muestras. En pocas palabras, una entrada de gran dimensión puede requerir una gran cantidad de muestras para construir un modelo significativo.

Para contrarrestar problemas como la maldición de la dimensionalidad, es común usar algún tipo de selección de características. La selección de características es el proceso de selección de un subconjunto de todas las características para su uso. Los beneficios de la selección de características son muchos, ya que puede reducir el tiempo de entrenamiento, las necesidades



de almacenamiento y reducir las implicaciones de la maldición de la dimensionalidad [68]. Hay tres formas principales de realizar la selección de características; métodos de filtrado, envoltura e integrado. Los métodos de filtrado analizan las propiedades del conjunto de datos para encontrar el conjunto óptimo de características. Uno de estos métodos de filtrado se conoce como la correlación de Pearson y utiliza una medida de la correlación lineal entre dos variables para encontrar un buen conjunto de parámetros. Los métodos de envoltura usan un clasificador para encontrar las características, y los métodos integrados son métodos más complejos donde la selección de características es parte del algoritmo de Aprendizaje Automático.

### 3.6. Modelos de Aprendizaje Automático

Hay una gran cantidad de algoritmos de Aprendizaje Automático; esta sección brindará una descripción general de algunos de los algoritmos más comunes y una breve explicación de cómo funcionan. Primero, sin embargo, una explicación rápida de por qué hay tantos algoritmos de aprendizaje automático.

#### Teorema de “*No Free Lunch*” (NFL):

El teorema de NFL para el aprendizaje supervisado establece que todos los algoritmos de Aprendizaje Automático supervisados son equivalentes cuando se promedia su rendimiento en todos los problemas posibles [69]. Las implicaciones del teorema son que no existe un algoritmo óptimo de ML para todos los problemas y que algunos son más apropiados para ciertos problemas que otros. Para el aprendizaje supervisado, la causa de esto es que los diferentes algoritmos manejan problemas como el ruido y el sobre-ajuste de forma diferente. No hay forma de saber con certeza cómo se comportará un algoritmo de Aprendizaje Automático en un determinado problema. Y por lo tanto, tenemos que probar varios algoritmos para saber con certeza si el algoritmo puede resolver un problema [70].

#### 3.6.1. Redes neuronales

La relación entrada-salida para una red neuronal depende estrictamente de la estructura de la red. Se pueden realizar diferentes arquitecturas de redes neuronales por la forma en que las neuronas se conectan entre sí. En esta sección se discuten tres tipos principales de estructuras de red. Además, en el Apéndice B y C se explica en detalle el funcionamiento de las redes neuronales y la regla delta de generalizada para el aprendizaje en este tipo de redes.

##### 3.6.1.1. El Perceptrón

El perceptrón es un algoritmo que puede resolver problemas limitados y simples. Está modelado a partir de las neuronas en el cerebro humano. Funciona simplemente al ponderar varias entradas; si la suma de las entradas por la ponderación está por encima de un umbral

predefinido, la salida es verdadera; de lo contrario, la salida es falsa. Un modelo del perceptrón se puede ver en la Figura 3.5. Comenzando con pesos aleatorios, el algoritmo se entrena alimentándose de los datos de entrada y los objetivos correspondientes. Los pesos se adoptan con una función matemática hasta que el algoritmo haya cumplido algunos criterios mínimos de error.

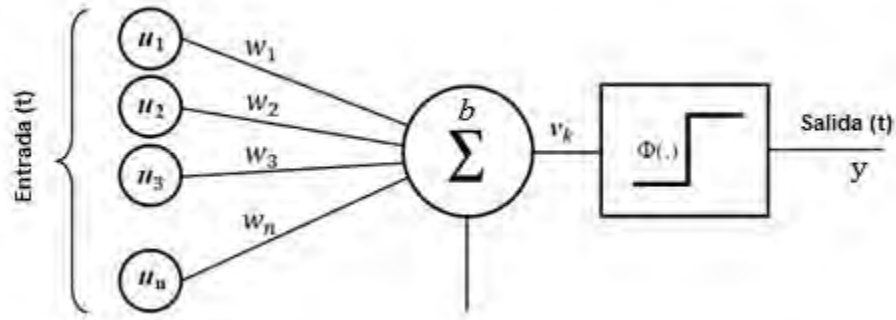


Figura 3.5 Modelo del Perceptrón [62]

La representación matemática de una neurona representada en la Figura 3.7 se puede lograr de la siguiente manera:

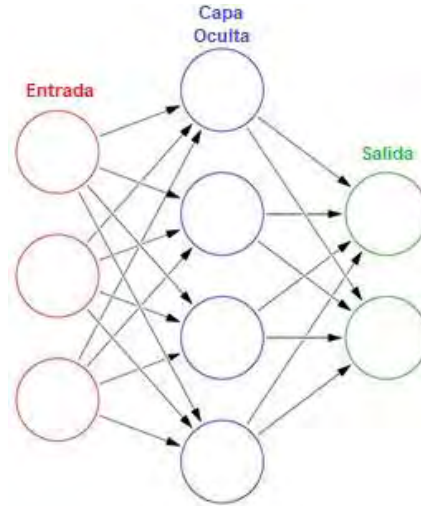
$$v = \sum_{j=1}^n w_j u_j \quad (3.7)$$

$$y = \Phi(v + b) \quad (3.8)$$

Como el perceptrón solo puede resolver problemas lineales y por lo tanto, tiene poderes de clasificación bastante limitados, es más comúnmente utilizado en el aprendizaje automático como base de otros algoritmos más complicados [71].

### 3.6.1.2. Red neuronal multicapa con conexiones hacia adelante

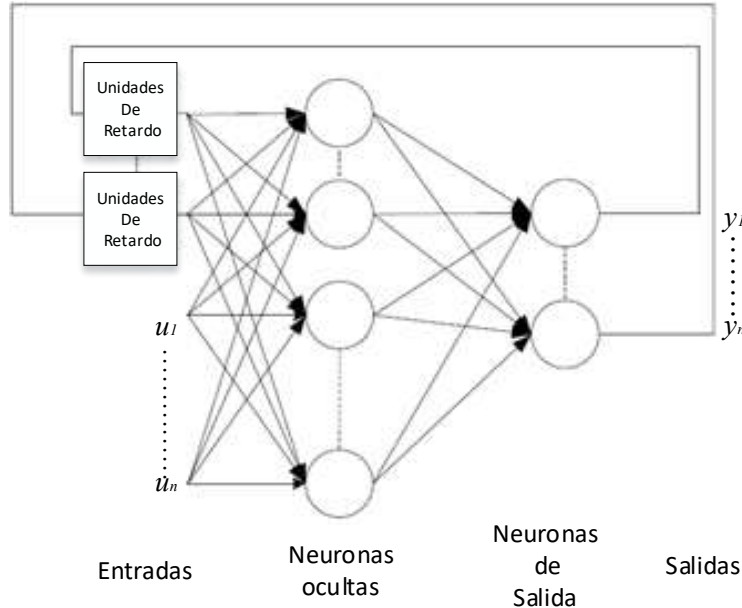
Uno de los algoritmos más comunes que utilizan el perceptrón como base son las redes neuronales multicapa, a veces también denominadas perceptrones multicapa. Combinando varios perceptrones en capas, como se muestra en la Figura 3.6, el algoritmo puede crear clases más complejas y límites de decisión que no son lineales. Cuantas más capas de perceptrones se agreguen, más complejas pueden ser las clases y los límites que separan las clases. El algoritmo está entrenado de forma similar al perceptrón, pero como contiene varios perceptrones en capas, la actualización de los pesos comienza desde la parte posterior, es decir, los nodos de salida y se actualiza desde atrás para comenzar con una función matemática conocida como función de retropropagación [72].



**Figura 3.6. Estructura de Red Neuronal**

**3.6.1.3. Red recurrente multicapa**

En contraste con las redes neuronales con conexiones hacia adelante (*Feed-Forward*), las redes neuronales recurrentes se caracterizan por al menos un ciclo de retroalimentación. La Figura 3.7 muestra una representación esquemática de una red neuronal recurrente.



**Figura 3.7. Estructura de RNA recurrente**

Se puede ver en la Figura 3.7 que la red neuronal exhibe una estructura de avance a través de la capa oculta de neuronas. Además, las unidades de retardo hacen que el comportamiento de la red neuronal no sea lineal. Esta clase de redes neuronales ha demostrado un mejor

rendimiento en términos de precisión para diferentes aplicaciones, en comparación con las redes neuronales *Feed-Forward* tradicionales [73-75].

### 3.6.2. Métodos de aprendizaje en RNA

#### 3.6.2.1. Aprendizaje por lotes

En el método de aprendizaje supervisado por lotes, todas las muestras del conjunto de datos de entrenamiento se presentan a la RNA al mismo tiempo. Los pesos sinápticos definen la fuerza de una conexión sináptica entre dos neuronas y se deciden en base a  $N$  muestras en el conjunto de datos de entrenamiento. El proceso de presentar todas las muestras se llama una época. Las ponderaciones sinápticas finales se deciden sobre la base de un enfoque de época por época, donde las muestras en el conjunto de datos de entrenamiento se barajan aleatoriamente y se presentan de nuevo a la RNA en cada época. El rendimiento de la red neuronal inicializada aleatoriamente se minimiza usando un algoritmo de minimización. La curva de aprendizaje se construye promediando el rendimiento de dicha RNA inicializada al azar en un número suficientemente grande de épocas. El entrenamiento se detiene cuando la curva de aprendizaje no muestra ninguna mejora o esta es muy pequeña.

Las dos ventajas del método de aprendizaje por lotes son garantizar la convergencia a un mínimo local y la paralelización del proceso de aprendizaje. La paralelización del aprendizaje hace que el aprendizaje sea más rápido. Sin embargo, la desventaja del método de aprendizaje por lotes es que no se puede lograr el mínimo global, por lo que no se garantiza el mejor rendimiento posible. Los requisitos de almacenamiento para el aprendizaje por lotes también son más altos en comparación con el método de aprendizaje en línea [76].

#### 3.6.2.2. Aprendizaje en línea

Contrariamente al método de aprendizaje por lotes, en el aprendizaje en línea las ponderaciones sinápticas se ajustan en función de un enfoque muestra por muestra. Se alcanza una época cuando todas las  $N$  muestras en el conjunto de datos de entrenamiento se presentan a la RNA. De forma similar al método de aprendizaje por lotes, las redes neuronales inicializadas aleatoriamente se entrenan para diferentes épocas en las que las muestras en el conjunto de datos de entrenamiento se barajan aleatoriamente. La curva de aprendizaje se traza como la función de rendimiento promedio para cada época.

El aprendizaje en línea tiene la ventaja de ser más receptivo a las redundancias en el conjunto de datos de entrenamiento; es decir, si se repiten las muestras de entrenamiento, el método de aprendizaje en línea aprovecha este hecho ya que las muestras se presentan a la RNA una a una. El aprendizaje en línea es relativamente más fácil de implementar y proporciona una mejor solución para los problemas de reconocimiento de patrones a gran escala. Sin embargo,

como la paralelización del proceso no es posible, es más lento que el método de aprendizaje por lotes [76].

El entrenamiento supervisado de RNA se puede resumir en los siguientes pasos:

1. Decidir la arquitectura la RNA.
2. Decidir el conjunto de datos de entrenamiento de modo que las muestras representen el entorno o el sistema que se está modelando.
3. Decidir qué parámetro de rendimiento se debe utilizar.
4. Decidir el método de entrenamiento (aprendizaje por lotes supervisado o aprendizaje en línea).
5. Entrenar a la RNA.

### 3.6.2.3. Algoritmo de aprendizaje de Levenberg-Marquardt

Los pesos sinápticos ' $w$ ' se actualizan para una estructura dada de RNA basada en el algoritmo de entrenamiento adoptado. El algoritmo de entrenamiento Levenberg-Marquardt (LMA) es uno de los más comunes utilizados para el entrenamiento de las RNA. Combina el método de descenso de gradiente, que garantiza la convergencia y el método de Newton, que converge rápidamente. LMA ofrece un mejor rendimiento en términos de precisión para redes neuronales con menos de 100 neuronas [77]. Por lo tanto, en esta tesis, el LMA se ha utilizado para el entrenamiento de la RNA ya que el número de neuronas requeridas para el modelado es inferior a 100.

La relación entrada / salida para una RNA se puede representar por la ecuación siguiente ecuación:

$$y = F(u_1, u_2, \dots, u_m) \quad (3.9)$$

$F$  es la función de aproximación no lineal que modela la red para emular la relación entre las entradas  $u$  y la salida. El sufijo  $m$  en la ecuación (3.42) representa la cantidad total de parámetros de entrada  $u$ , que se utilizan para modelar un parámetro de salida.

Considere que  $\{u(i), d(i)\}_{i=1}^N$  está entrenando el conjunto de datos con  $N$  puntos de muestra y  $F(u(i); w)$  es la función no lineal emulada por la RNA, donde  $w$  es el vector de ponderación. La capacitación en red se logra al minimizar la función de costo presentada en la siguiente ecuación:

$$\zeta(w) = \frac{1}{2N} \sum_{i=1}^N [d(i) - F(u(i); w)]^2 \quad (3.10)$$

Según LMA, el vector de peso se actualiza usando la siguiente ecuación:

$$\Delta w = [H + \mu I]^{-1} g \quad (3.11)$$

$H$  es la matriz de Hessiana aproximada y  $g$  es el vector de gradiente.  $I$  es una matriz de identidad con dimensiones iguales a  $H$  y  $\lambda$  es un parámetro escalar utilizado para cambiar entre el método de Newton y el método de gradiente descendente. Éstas se calculan utilizando las siguientes expresiones:

$$H = \frac{1}{N} \sum_{i=1}^N \left[ \frac{\partial F(x(i); w)}{\partial w} \right] \left[ \frac{\partial F(x(i); w)}{\partial w} \right]^T \quad (3.12)$$

$$g = \frac{\partial \zeta(w)}{\partial w} \quad (3.13)$$

Si el valor de  $\lambda$  es cero, la ecuación (3.44) se reduce al método de Newton y si  $\lambda$  es lo suficientemente grande como para sobrealimentar  $H$ , el método es similar al método de gradiente descendente. El objetivo del método es pasar al método de Newton, que es rápido cerca del valor mínimo y por lo tanto, el valor de  $\lambda$  se reduce en cada paso consecutivo, siempre que la función de rendimiento de la red definida por la ecuación (3.43) sea reducida. El valor de  $\lambda$  aumenta si la función de rendimiento aumenta en un paso consecutivo. El valor aumenta o disminuye en un factor de 10 y el valor inicial de  $\lambda$  es considerado como 1. En el Apéndice D se realiza la descripción de este algoritmo de aprendizaje con más detalle para una mejor comprensión del mismo.

#### 3.6.2.4. Funciones de activación

La salida de la neurona depende de la función de activación. Dos tipos de funciones de activación, que se usan comúnmente en redes neuronales son:

##### i) Función Escalón

El tipo de umbral de la función de activación es básicamente una función escalón definida por la siguiente ecuación.

$$\Phi(v) = \begin{cases} 1 & \text{si } v \geq 0 \\ 0 & \text{si } v < 0 \end{cases} \quad (3.14)$$

La función de umbral puede tener salida 1 o 0 dependiendo del campo inducido. Las funciones de umbral a menudo se usan en la capa de salida de la RNA donde podría ser necesaria una clasificación binaria de la entrada.

##### ii) Función sigmoidea

La función sigmoidea es una función de activación no lineal definida por la siguiente ecuación:

$$\Phi(v) = \frac{1}{1+e^{-av}} \quad (3.15)$$

La función sigmoidea es una de las funciones de activación más comunes utilizadas en las redes neuronales. La pendiente de la función sigmoidea puede variarse mediante el parámetro de pendiente 'a', ya que tiende a infinito, la función sigmoidea se convierte en la función de umbral. A diferencia de la función de umbral, que puede asumir un valor de 0 o 1, la función sigmoidea puede asumir cualquier valor entre 0 y 1. La Figura 3.8 muestra el comportamiento de las funciones escalón y sigmoidea.

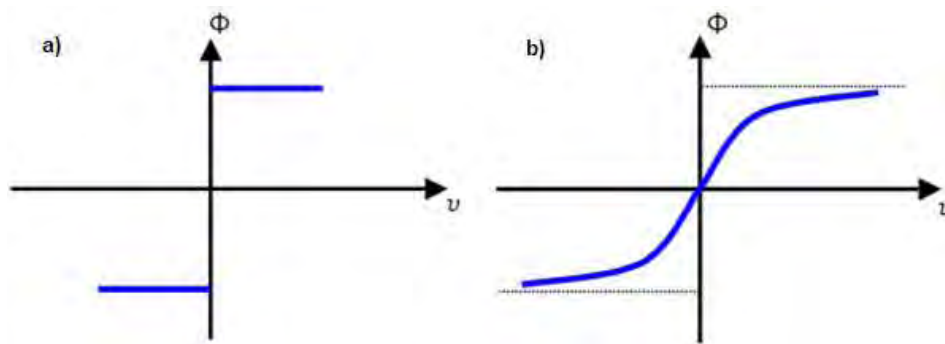


Figura 3.8. Función de activación de la RNA. a) Función escalón, b) Función sigmoidea

### 3.6.3. Máquinas de soporte vectorial

Máquinas de soporte vectorial es un esquema de clasificación que, al construir el modelo, utiliza una función matemática para aumentar las dimensiones de las muestras hasta que pueda separar linealmente las clases en el conjunto de prueba. La función matemática que aumenta las dimensiones se conoce como función kernel; la función kernel transforma los datos de tal manera que hay una mayor posibilidad de clases separables [78]. La Figura 3.9 muestra cómo un kernel  $\phi$  transforma los datos en clases separables.

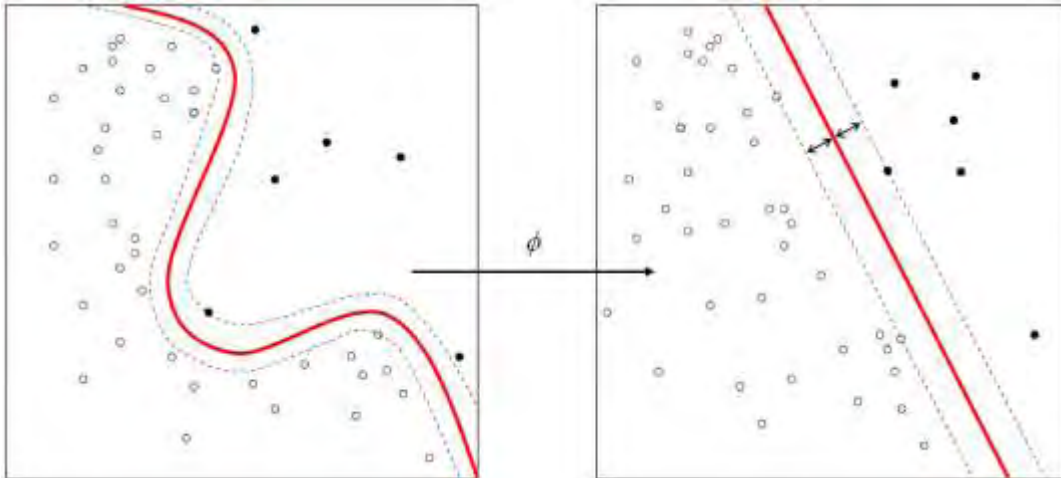


Figura 3.9. Modelo de una SVM Kernel [68]

Cuando la SVM ha alcanzado un estado donde puede separar linealmente las clases, intenta encontrar la separación óptima. La separación óptima es cuando el separador está igualmente lejos de la muestra más cercana en cada clase, como se muestra en la Figura 3.10.

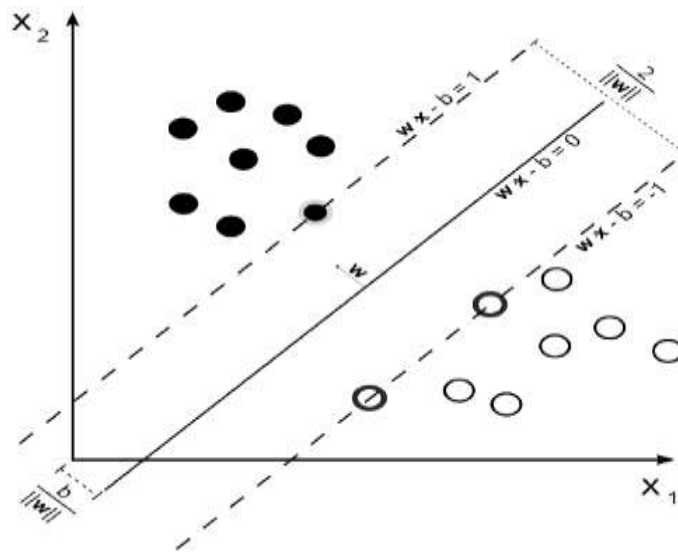


Figura 3.10. Modelo de separación óptima [78]

Cuando la SVM ha creado su modelo, puede predecir con nuevos datos realizando la misma transformación de kernel en los nuevos datos y luego observa a qué clase debe pertenecer. Las SVM son conocidas por funcionar bien en conjuntos de muestras de tamaño razonable, pero tienen un rendimiento inferior en conjuntos grandes [79]. Los SVM pueden usar funciones kernel tanto lineales como no lineales.

### 3.6.4. Máquina de puntos Bayes



Al igual que las SVM, la máquina de puntos Bayes también utiliza una función kernel. En lugar de encontrar una dimensión donde hay separabilidad lineal y usar solo este límite de decisión, la MPB intenta separar las clases en varias dimensiones y almacena cada uno de los separadores intentados. Posteriormente, se aproxima a un error mínimo, la intersección de los separadores intentados y lo utiliza para hacer sus predicciones. La intersección es un punto medio de la región y divide el espacio en dos mitades de igual volumen y se conoce como el punto de Bayes [80].

### 3.6.5. Regresión logística

La regresión logística, a pesar de su nombre, no es un algoritmo de regresión, sino un clasificador binario. El modelo estima la probabilidad de un resultado binario basado en algunas características. Funciona al medir la relación de las variables y las probabilidades de estimación utilizando una distribución logística acumulativa. El modelo es rápido de entrenar, pero se limita a modelos lineales. [57]

### 3.6.6. Conjunto de aprendizaje

El conjunto de aprendizaje (*ensemble learning*), utiliza múltiples algoritmos de aprendizaje para obtener un mejor rendimiento predictivo. La idea es que varias cabezas piensan mejor que una. Una característica común de los problemas de aprendizaje automático es que es imposible obtener clasificadores perfectos, lo que abre la posibilidad de que diferentes algoritmos, con diferentes sesgos, fortalezas y debilidades, en conjunto produzcan mejores resultados que uno solo. Los “estudiantes” se entrenan de forma independiente y las predicciones se combinan de alguna manera para hacer la predicción general. Hay varias formas de combinar a los “alumnos” independientes y garantizar que no arrojen predicciones que sean demasiado similares una de otra.

Quizás el algoritmo más común para el conjunto de aprendizaje se conoce como impulso. El impulso incremental construye el modelo de conjunto formando cada nueva instancia de modelo para enfatizar las muestras en las que los modelos anteriores no se clasificaron correctamente [81].

La agregación de empaquetamiento o *Bootstrap*, a menudo abreviado como embolsado, usa múltiples “estudiantes” que tienen el mismo peso en el comité de conjunto. La varianza se logra al entrenar cada modelo en un subconjunto de los datos extraídos al azar [82].

El apilamiento entrena varios algoritmos diferentes de aprendizaje automático en todos los datos disponibles y luego usa un algoritmo combinado que está entrenado para hacer una predicción final usando tanto las predicciones de los datos disponibles. Cualquier algoritmo de Aprendizaje Automático se puede usar como un combinador, sin embargo, comúnmente se usa la regresión logística de una sola capa. [83]

### 3.6.7. Árboles de decisión

Los árboles de decisión sirven como base para varios algoritmos de aprendizaje en conjunto. El árbol de decisión es una estructura similar a los diagramas de flujo, donde cada nodo contiene una prueba en una característica, cada rama representa el resultado de la prueba y cada nodo final contiene una etiqueta de clase. Cuando los árboles de decisión se usan como algoritmo de clasificación, los árboles se crean dividiendo el conjunto de datos en subconjuntos según los resultados de una característica y creando sub-árboles correspondientes. Esto se lleva a cabo hasta que cada sub-árbol derivado conduce a un nodo final [84].

### 3.6.8. Bosques aleatorios

Bosques aleatorios (*decision forest*) es un algoritmo que agrupa los árboles de decisión. Crea una multitud de árboles de decisión y el resultado es el modo de las predicciones de los árboles individuales. Puede usar varias de las técnicas de conjunto presentadas, pero lo más común es utilizar el “*bagging*” para la selección aleatoria de características para construir diferentes árboles de decisión [85]. El “*bagging*” es un algoritmo de conjunto de Aprendizaje Automático diseñado para mejorar la estabilidad y la precisión de los algoritmos utilizados en la clasificación estadística y la regresión. También reduce la varianza y ayuda a evitar el sobreajuste.

### 3.6.9. Jungla de decisión

Jungla de decisión (*decision jungle*) es una modificación de bosques aleatorios. En lugar de permitir solo una ruta a cada nodo en los árboles, jungla de decisión permite varias rutas desde el nodo raíz hasta los nodos finales. Esto cambia la estructura de los árboles y complica el proceso de aprendizaje, pero, sin embargo, permite que los árboles sean más pequeños y consuman menos memoria y poder de cómputo. La clase se encuentra utilizando la búsqueda de gráficos acíclicos dirigidos [86].

### 3.6.10. Árbol rápido

Árbol rápido (*fast tree*) es otra extensión de bosques aleatorios. Lo que diferencia al árbol rápido de los algoritmos antes mencionados es que usa el esquema de un vecino más cercano para reducir la longitud de los árboles individuales. Es, como su nombre lo indica, una implementación más rápida de bosques aleatorios, pero en ciertos casos tiene un mejor desempeño que bosques aleatorios [87].

### **CAPÍTULO IV: APLICACIÓN DE REDES NEURONALES ARTIFICIALES (RNA) UTILIZANDO EL SISTEMA DE MONITOREO DE TRANSFORMADORES DE DISTRIBUCIÓN**

#### **4.1. Introducción**

El propósito de este capítulo es desarrollar un modelo predictivo para gestionar el mantenimiento de transformadores a partir del perfil de carga y la temperatura ambiente. La metodología propuesta tiene como base la utilización de algoritmos de aprendizaje automático, en este caso se utilizan modelos de RNA para estimar la temperatura de operación del aceite en el transformador. Con la obtención de esta variable se presenta un enfoque estadístico utilizado en [88]. Este enfoque se utiliza para analizar el resultado del modelo de RNA y detectar fallas utilizando la distancia Mahalanobis. Además, se discute la metodología para decidir el umbral. Para entrenar los modelos RNA, se aplica el procedimiento supervisado de aprendizaje descrito en la sección 3.6.2.3 y el algoritmo de entrenamiento Levenberg-Marquardt se usa para decidir los pesos sinápticos.

#### **4.2. Filosofía del desarrollo**

De los métodos analizados en el Capítulo III, se selecciona el método de Aprendizaje Automático, en su modalidad de aprendizaje supervisado, utilizando las redes neuronales artificiales, esto, debido a que es el método que más se ajusta a los requerimientos de este trabajo de tesis. La clave del éxito para este trabajo va a ser seleccionar correctamente el modelo de red neuronal que más se ajuste a las características del sistema analizado.

En este trabajo de tesis se analizan transformadores de distribución, los cuales operan en un ambiente dinámico. Estas variaciones pueden ser producidas por: cambios en la demanda y en las condiciones ambientales, éstos pueden ser de manera general, por hora, días y meses. Estos frecuentes cambios en las características de funcionamiento producen cambios en el comportamiento de los componentes en los transformadores, por ejemplo, la temperatura del aceite que depende de manera directa de la demanda del transformador y las condiciones climatológicas.

Los proveedores de mantenimiento y operación de transformadores de distribución, en general siguen un régimen de mantenimiento preventivo de seis meses para algunos componentes y un año o más para mantenimiento general. Como se ve en el Capítulo II, las actividades de mantenimiento programado se realizan con el objetivo de mantener el rendimiento de los transformadores en un nivel óptimo. Sin embargo, después de una falla de un componente principal, que conlleve a un cambio de aceite en el transformador, este cambio puede provocar que las características operativas de este componente varíen

drásticamente. El modelo de RNA diseñado para modelar el estado de equilibrio con el componente anterior no se puede aplicar para el monitoreo de la condición del nuevo componente. Por lo tanto, se vuelve necesario actualizar el modelo de RNA con nuevos datos, para evitar falsas alarmas.

Para desarrollar esta metodología, el primer paso va a ser seleccionar la variable a ser monitoreada. El siguiente paso es seleccionar datos históricos acumulados por el sistema de monitoreo que representen la operación óptima del transformador, estas variables deben ser parámetros operativos relevantes para modelar el comportamiento normal de la variable a monitorear. En la siguiente sección se describe el enfoque utilizado para seleccionar estos parámetros, que son utilizados por el conjunto de variables de entrenamiento, además, el modelo de RNA entrenado se utiliza para el monitoreo del estado de la variable seleccionada.

Si se realiza una actividad de mantenimiento importante sobre el equipo analizado que conlleve al reemplazo del componente monitoreado, se inicia un nuevo entrenamiento de la RNA. De acuerdo con el enfoque descrito anteriormente, será necesario recolectar nuevamente el conjunto de datos para el entrenamiento de la RNA, por lo que durante el tiempo en que ocurre esta recolección no estará disponible el sistema implementado. Sin embargo, la probabilidad de que el equipo recientemente reemplazado falle dentro del tiempo en que se recolectan los nuevos datos de entrenamiento es baja.

### 4.3. Método propuesto para el monitoreo de la condición en transformadores

#### 4.3.1. Matriz de correlación y correlaciones parciales

Para seleccionar las variables de entrada del conjunto de datos de entrenamiento de la RNA se realiza un análisis de correlación de variables, con el objetivo de que las variables seleccionadas sean las más representativas de la variable de salida. El cálculo se realiza a partir del análisis matemático descrito a continuación

Llamaremos matriz de correlación a la matriz cuadrada y simétrica que tiene la unidad en la diagonal y fuera de ella los coeficientes de correlación entre las variables.

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \dots & 1 \end{bmatrix} \quad (4.1)$$

Esta matriz es semidefinida positiva.

Para verlo, llamemos  $D$  a la matriz diagonal de orden  $k$  construida colocando en la diagonal principal las desviaciones típicas de las variables. La matriz  $R$  está relacionada con la matriz de covarianzas  $S$  mediante:

$$R = D^{-1}SD^{-1} \quad (4.2)$$

que implica

$$S = DRD \quad (4.3)$$

la condición  $w'Sw \geq 0$  equivale a:

$$w'DRDw = Z'RZ \geq 0 \quad (4.4)$$

Llamando  $Z = Dw$ . Por tanto, la matriz  $R$  es también semidefinida positiva.

Se define la matriz de correlaciones parciales como la matriz que mide la relación entre pares de variables eliminando el efecto de las restantes. Por ejemplo, para cuatro variables:

$$R_p = \begin{bmatrix} 1 & r_{12,34} & r_{13,24} & r_{14,23} \\ r_{21,34} & 1 & r_{13,14} & r_{24,12} \\ r_{31,24} & r_{32,14} & 1 & r_{34,12} \\ r_{41,23} & r_{42,13} & r_{43,12} & 1 \end{bmatrix} \quad (4.5)$$

donde,  $r_{12,34}$  es la correlación entre las variables 1 y 2 cuando eliminamos el efecto de la 3 y la 4, es decir, cuando las variables 3 y 4 permanecen constantes.

Puede demostrarse que el coeficiente de correlación parcial entre dos variables es proporcional al coeficiente de una regresión entre las dos variables que incluye también al resto de las variables. En concreto, por ejemplo:

$$r_{12,34} = \hat{\beta}_{12,34} \left( \sqrt{\hat{\beta}_{12,34}^2 + \hat{s}_{12,34}^2(n-k-1)} \right) \quad (4.6)$$

donde  $k$  es aquí igual a 4 y  $\hat{\beta}_{12,34}$  se obtiene a partir de la recta de regresión

$$\hat{x}_1 = \hat{\beta}_0 + \hat{\beta}_{12,34}x_2 + \hat{\beta}_{13,34}x_3 + \hat{\beta}_{14,34}x_4 \quad (4.7)$$

siendo  $\hat{s}_{12,34}^2$  la varianza estimada del coeficiente  $\hat{\beta}_{12,34}$  en esta ecuación [89].

4.3.2. Modelado de RNA para el monitoreo de la condición

En la Sección 3.6.1, se presentaron diferentes arquitecturas de RNA. Las redes neuronales recurrentes han demostrado tener un mejor rendimiento que la estructura de *Feed-Forward* [86]. Todas las redes dinámicas específicas discutidas hasta ahora han sido redes enfocadas, con la dinámica sólo en la capa de entrada, o redes *Feed-Forward*. La red autorregresiva no lineal con entradas exógenas (NARX) es una red dinámica recurrente, con conexiones de retroalimentación que encierran varias capas de la red. El modelo NARX se basa en el modelo ARX lineal, que se usa comúnmente en el modelado de series de tiempo. Los modelos NARX se han aplicado con éxito al modelado de sistemas no lineales, o sea, sistemas físicos como intercambiadores de calor y tratamiento de aguas residuales [90, 91]. A continuación, se presenta el modelo de RNA NARX propuesto en este trabajo de tesis:

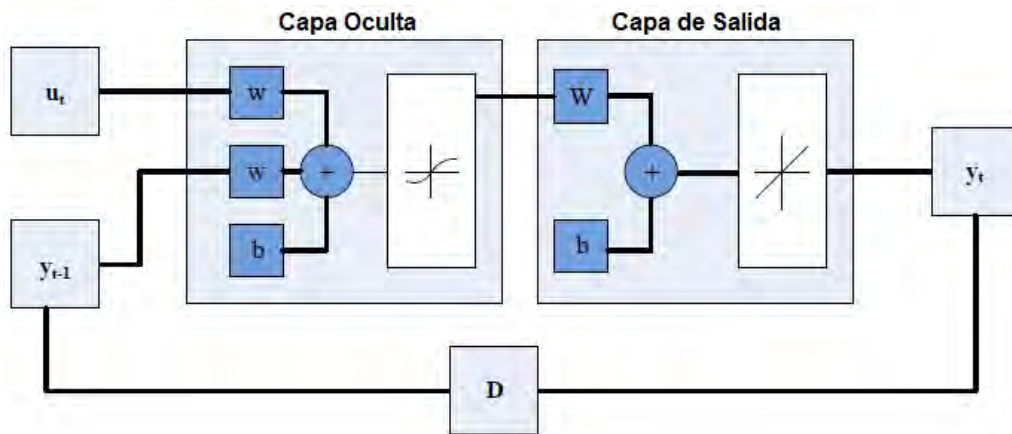


Figura 4.1. Modelo de RNA con estructura NARX [87]

Donde  $u_t$  representa el vector de entrada en el instante ' $t$ ' y  $y_t$  es la salida correspondiente; ' $w$ ' es el vector de ponderación en la capa oculta y la salida.  $D$  es el vector de retardo, que decide la demora en la retroalimentación entre la salida y la entrada.

La entrada  $u_t$  representa las entradas exógenas a medida que se originan fuera de la red neuronal. La salida del modelo RNA  $y_t$  se regresa en el valor retardado de la salida  $y_{t-1}$ . El comportamiento generalizado del modelo NARX de la Figura 4.1 se expresa como:

$$y_t = F(y_{t-1}, \dots, y_{t-D}; u_t) \quad (4.8)$$

Para decidir el valor del vector de retardo  $D$ , es importante considerar el comportamiento dinámico del sistema que se está modelando. En esta tesis, el enfoque de monitoreo de condición basado en RNA se aplica para el monitoreo de la condición del aceite del transformador. Por lo tanto, es deseable modelar el comportamiento dinámico del aceite del

transformador. Se supone que este aceite tiene una cierta inercia térmica y el valor de la temperatura para aceite en el instante  $(t - 1)$  influirá en el valor de la temperatura en el instante  $t$ . Sin embargo, como las mediciones son medidas promedio de 30 minutos, no se incluyeron más señales de temperatura retardada en el modelado de la temperatura del aceite en el instante  $t$  de tiempo. Por lo tanto, se considera que el vector de retardo  $D$  solo tiene un valor de retardo, lo que da como resultado que solo  $y_{t-1}$  se devuelva al lado de entrada. Por lo tanto, la formulación específica del modelo de RNA se describe como:

$$y_t = F(y_{t-1}; u_t) \tag{4.9}$$

### 4.3.3. Estructura del modelo de RNA

No existe una metodología estándar para decidir la cantidad óptima de neuronas en la capa oculta de la RNA. Para determinar éstas se adoptó un proceso iterativo donde se compararon los resultados de la red neuronal con diferentes números de neuronas en la capa oculta. Se encontró que la red neuronal con 30 neuronas en la capa oculta proporciona el mejor rendimiento para los casos con diferente número de neuronas investigadas.

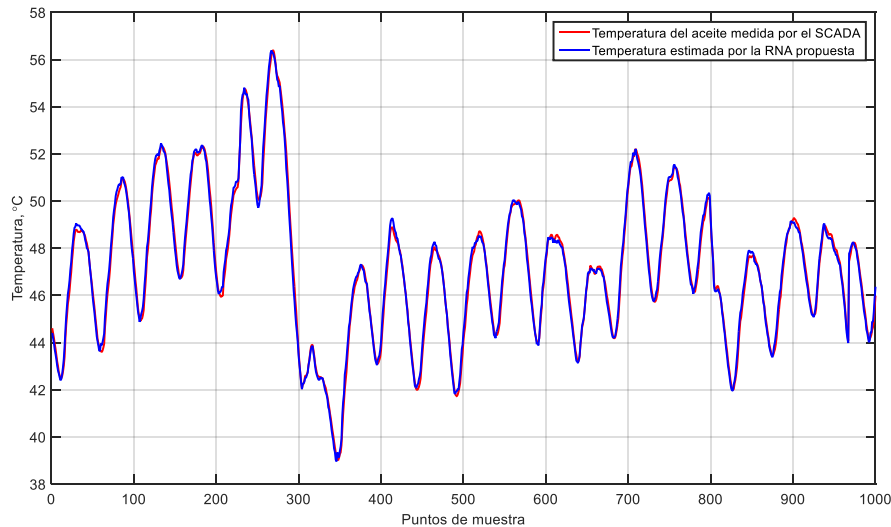
Además, es importante verificar la influencia de cada entrada en la salida. Si se encuentra que una entrada no tiene influencia en la salida, puede ser ignorada. Sin embargo, para darle un significado físico al modelo, se han seleccionado las entradas que tienen un efecto físico en la temperatura del aceite del transformador. Las entradas y salidas del modelo RNA se analizan con más detalle en el Capítulo V. Sin embargo, para fines de demostración, se presenta un modelo RNA y su resultado en esta sección. La Tabla 4.1 muestra los detalles del modelo utilizado para fines de demostración.

**Tabla 4.1. Detalles del modelo de RNA propuesto**

<i>Modelo de RNA</i>	<i>Detalles del Modelo</i>
<i>Número de neuronas en la capa oculta</i>	30
<i>Número de neuronas en la capa de salida</i>	1
<i>Función de activación para las neuronas de la capa oculta</i>	<i>Función sigmoidea</i>
<i>Función de activación para las neuronas de la capa de salida</i>	<i>Función umbral</i>
<i>Entradas</i>	$u(t) = [P_{a,b,c}, I_{an,bn,cn}, T_{amb}]$
<i>Salida</i>	$y(t) = [T_{aceite}]$

El modelo de RNA seleccionado está diseñado para estimar la temperatura del aceite del transformador. Las variables de entrada que se utilizan para el modelado son: la demanda de potencia activa por fase del transformador, así como la corriente y la temperatura ambiente.

La Figura 4.3 muestra la comparación de la salida prevista de la RNA y la temperatura real medida, extraída del SCADA, para 1000 puntos de muestra, tomados en intervalos de 30 minutos.



**Figura 4.2. Comparación entre los resultados de la RNA y los datos medidos por el SCADA**

Es importante destacar que estos 1000 puntos de muestra son diferentes de los utilizados para el entrenamiento del modelo de RNA. Por lo tanto, el resultado del modelo de RNA se basa en las entradas que el modelo de RNA entrenado no ha visto antes.

De la figura anterior se puede observar que el modelo de RNA, es capaz de predecir la temperatura del aceite en el transformador de manera muy precisa. Para el desarrollo del caso de estudio de este trabajo se utilizó la estructura descrita en esta sección.

#### 4.3.4. Detección de falla

Un error en el modelo de RNA se define como la diferencia entre el valor del parámetro estimado de la red neuronal y el valor real del parámetro medido almacenado en el SCADA. En cualquier momento dado, existe algún error en el modelo de RNA debido a imprecisiones durante la etapa de modelado. La detección de fallas es el proceso de diferenciar cualquier error normal de los que podrían ser importantes desde la perspectiva del monitoreo de la condición. Por lo tanto, la detección de fallas es un paso importante en el proceso de monitoreo de la condición usando RNA.

Un inconveniente principal de los métodos de RNA es la falta de comprensión física del sistema que se modela. Como no existe un modelo analítico involucrado, los métodos de



redes neuronales son modelos de caja negra, en donde la relación entre la entrada y la salida se deriva únicamente de la comprensión de los datos representativos proporcionados al modelo de RNA. A pesar de que se puede tener cuidado para asegurar que los datos de entrenamiento estén libres de errores, no se puede asegurar que la red neuronal pueda modelar el sistema sin ningún error. Esto también puede atribuirse al hecho de que los modelos de RNA tienen la tendencia a no converger en el mínimo local durante el entrenamiento, dando lugar a errores inadvertidos.

Para superar este inconveniente del método de modelado de RNA, se utiliza un enfoque estadístico para la detección de fallas, desarrollado en [88], donde utilizan una medida de distancia llamada distancia de Mahalanobis.

#### 4.3.4.1. Distancia de Mahalanobis

En estadística, la distancia de Mahalanobis (MD) es una medida de distancia introducida por Mahalanobis en 1936. Su utilidad radica en que es una forma de determinar la similitud entre dos variables aleatorias multidimensionales [91]. Se diferencia de la distancia euclídea en que tiene en cuenta la correlación entre las variables aleatorias [92]. La distancia de Mahalanobis se ha aplicado con éxito para capturar valores atípicos en diferentes campos de aplicación; ver por ejemplo [92, 93]. Esta fuerza de detección de valores atípicos de la medición de distancia de Mahalanobis se utiliza para la detección de fallas.

La distancia de Mahalanobis es adimensional y se puede calcular utilizando la siguiente ecuación:

$$MD_i = \sqrt{(X_i - \mu)C^{-1}(X_i - \mu)^T} \quad (4.10)$$

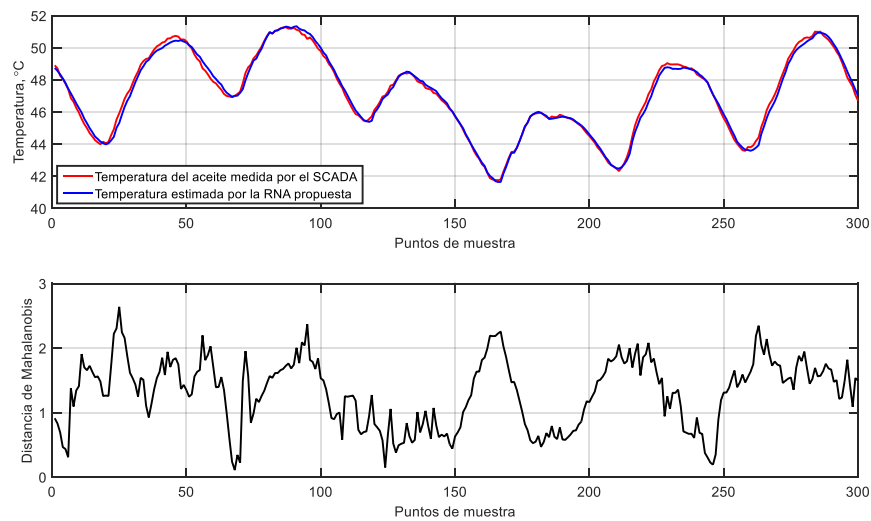
Donde  $X_i = [X_1, X_2, \dots, X_m]$  es el *i*ésimo vector de observación,  $\mu = [\mu_1, \mu_2, \dots, \mu_m]$  es el vector de los valores medios obtenidos del conjunto de datos de entrenamiento “saludables”,  $C$  es la matriz de covarianza para el conjunto de datos “saludables” y  $m$  es el número total de parámetros.

Un nuevo valor de  $MD$ , obtenido durante la etapa de monitorización de la condición se calcula utilizando las siguientes ecuaciones.

$$X_{C_{Metapa}} = [\text{Error}_{C_{Metapa}}, \text{SCADA datos}] \quad (4.11)$$

$$(MD_{C_{Metapa}})_i = \sqrt{\left( (X_{C_{Metapa}})_i - \mu_{ref} \right) C_{ref}^{-1} \left( (X_{C_{Metapa}})_i - \mu_{ref} \right)^T} \quad (4.12)$$

El *SCADA datos* en la ecuación (4.11) es la temperatura real registrada en el SCADA durante la etapa de monitoreo de la condición y  $Error_{C_{Metapa}}$  es la diferencia entre las cantidades estimadas del modelo de RNA y reales durante la etapa de monitoreo de la condición. La Figura 4.3 muestra la aplicación de la medición de distancia de Mahalanobis a la monitorización del estado basado en RNA para la temperatura de aceite en el transformador. La aplicación de ejemplo muestra 300 puntos de muestra. Además, se puede observar que los valores estimados y reales de la temperatura del modelo de RNA para el aceite, son similares presentando poca diferencia. Sin embargo, la distancia de Mahalanobis es alta para algunos puntos de muestra.



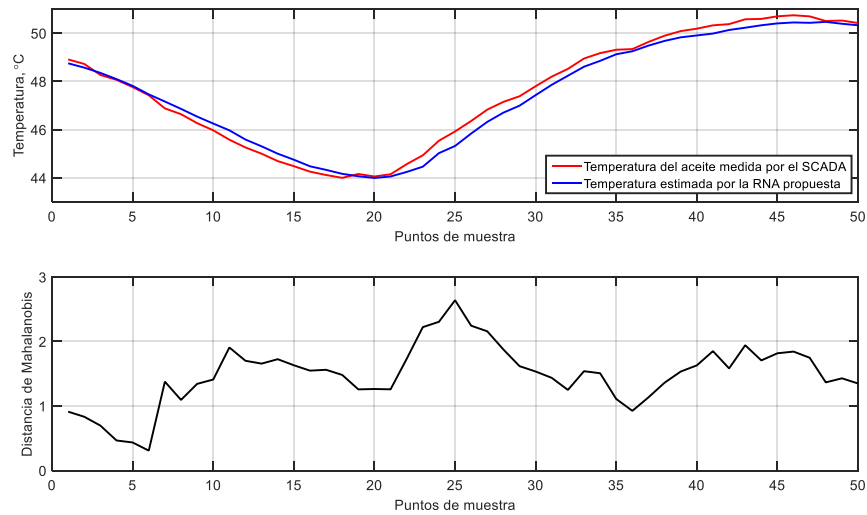
**Figura 4.3. a) Comparación SCADA y resultados de la RNA; b) Distancia de Mahalanobis**

La Figura 4.4 muestra la imagen ampliada en las primeras 50 muestras de la Figura 4.3. Se puede observar que una pequeña desviación entre la temperatura estimada de la RNA y la real se acompaña de un aumento observable en el valor de la distancia de Mahalanobis.

El modelo de RNA es capaz de modelar la temperatura del aceite con precisión; por lo tanto, el método estándar para usar RMSE se considera insuficiente. En este sentido, la medición de distancia Mahalanobis proporciona un método robusto para la detección de valores atípicos.

Sin embargo, es importante notar que pequeñas variaciones con respecto al rango en que varía la temperatura del aceite en el transformador, como se muestra en la Figura 4.3 y 4.4, podrían ocurrir y esto puede no apuntar a daño en el componente. Con el fin de evitar falsas alarmas durante la monitorización del estado, se calcula un promedio de la distancia de

Mahalanobis durante un período de tiempo predefinido para garantizar que la falla haya existido por un período de tiempo en lugar de solo un instante.



**Figura 4.4. Ampliación de la Figura 4.3**

#### 4.3.4.2. Valor del umbral

Se ha señalado que uno de los inconvenientes del método de RNA es la falta de modelos analíticos. Por lo tanto, para asegurar que la anomalía se debe al deterioro en el componente que se está monitoreando y no a inexactitudes en el modelo de RNA, se adopta un enfoque estadístico para decidir un valor umbral. Para una detección de fallas exitosa, es importante decidir un valor de umbral que establecer.

El entrenamiento del modelo de RNA se realiza en un conjunto de datos específicamente seleccionados de un intervalo de tiempo, donde no se han informado fallas en el transformador. Por lo tanto, es natural considerar que los errores entre la estimación de la RNA y el valor de temperatura real obtenidos durante la etapa de capacitación, se deben a las inexactitudes en el modelo de RNA. Esta información se usa más para decidir el valor umbral que para la detección de anomalías.

El umbral para la detección de anomalías se calcula en función de los valores de distancia de Mahalanobis para los errores obtenidos durante la etapa de entrenamiento. La MD para el conjunto de datos de entrenamiento se calcula utilizando las siguientes ecuaciones:

$$X_{ref} = [Error, Valor_{objetivo}] \quad (4.13)$$

$$(MD_{ref})_i = \sqrt{\left( (X_{ref})_i - \mu_{ref} \right) C_{ref}^{-1} \left( (X_{ref})_i - \mu_{ref} \right)^T} \quad (4.14)$$

El vector de referencia se usa para calcular la matriz de covarianza  $C_{ref}^{-1}$  y el vector de valor medio  $\mu_{ref}$ .  $(MD_{ref})_i$  representa la distancia de Mahalanobis para la  $i$ ésima fila del vector de referencia. El *Error* en la ecuación (4.13) es la diferencia entre la temperatura estimada por el modelo de RNA y la medida registrada en el SCADA y el *Valor\_objetivo* es la medida SCADA registrada utilizada durante el entrenamiento del modelo de RNA.

Los valores de *MD* obtenidos durante la etapa de entrenamiento se pueden representar con precisión mediante una función de distribución de probabilidad de Weibull de dos parámetros [91], definida en la siguiente ecuación:

$$f(x) = \beta \eta^{-\beta} (x)^{\beta-1} e^{\left(-\frac{x}{\eta}\right)^\beta} \quad (4.15)$$

Donde  $x$  es el valor de distancia de Mahalanobis calculado durante la etapa de entrenamiento. Los parámetros de forma y escala  $\beta$  y  $\eta$  para la distribución de Weibull se pueden estimar utilizando el método de máxima probabilidad de campana [93]. Sin embargo, en esta tesis, la funcionalidad incorporada de Matlab se ha utilizado para estimar los parámetros.

El umbral se decide de tal manera que la probabilidad de ocurrencia de la distancia Mahalanobis tiene baja probabilidad en comparación con la difusión de los datos obtenidos de la etapa de entrenamiento. Conceptualmente, esto se hace usando la lógica de que el valor *MD* es una anomalía si  $f(MD_{CMstage})_i < tol$ . Por lo tanto, cualquier valor *MD* obtenido durante la etapa de monitoreo, que tiene una probabilidad de ocurrencia menor que *tol* se considera que indica una anomalía en el componente [93]. La Figura 4.5 muestra el algoritmo propuesto para monitorear la condición utilizando RNA. Este algoritmo ha sido codificado y será usado para la gestión de mantenimiento en subestaciones eléctricas. En el Capítulo V se muestran las variables utilizadas por las RNA y se establece un análisis de correlación para definir cuáles son las variables más importantes para el monitoreo de la condición en los transformadores analizados.

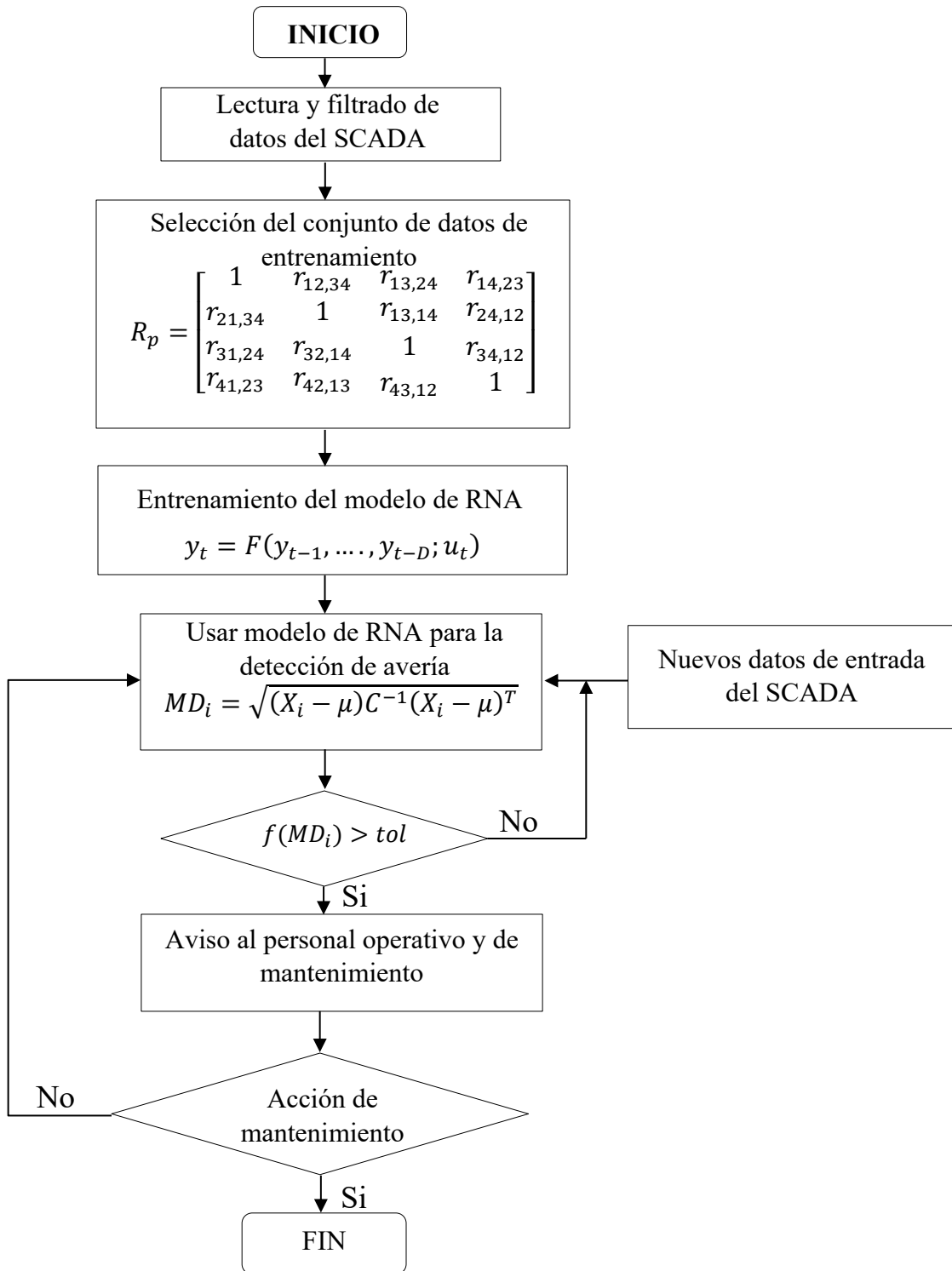


Figura 4.5. Algoritmo general de monitoreo de la condición utilizando RNA

**CAPÍTULO V: APLICACIÓN DEL ALGORITMO PROPUESTO A CASOS DE ESTUDIO REALES**

**5.1. Introducción**

En este capítulo se presentan los resultados del algoritmo propuesto para el monitoreo de la condición en transformadores de distribución como herramienta para la gestión de mantenimiento en subestaciones. Inicialmente se proporcionan algunas características del transformador analizado, luego se selecciona el conjunto de datos de entrenamiento con el objetivo de que sean los más representativos de la variable a predecir, a continuación, se selecciona el modelo de RNA más preciso con el fin de obtener un proceso correcto de monitoreo de la condición en el transformador. Luego de seleccionar el modelo de RNA se calcula la distancia de Mahalanobis y se establece el valor umbral para la detección de fallas utilizando nuevos datos obtenidos en la etapa de monitoreo de la condición.

Además, se establece un análisis comparativo de trabajos similares desarrollados por otros investigadores con el fin de evaluar las ventajas del modelo desarrollado en este trabajo.

**5.2. Características principales de los transformadores de prueba**

Los transformadores que se utilizan en los casos de estudio pertenecen a instalaciones de CFE Distribución. Debido al acuerdo de confidencialidad en el proyecto, no está permitido revelar más información sobre los mismos, por lo que se les denominará sólo como transformador de prueba T1 y T2. En la Tabla 5.1 se muestran algunas de las características de los transformadores de prueba. En el Apéndice E se muestra más información sobre los transformadores.

**Tabla 5.1. Datos nominales de los transformadores de prueba**

<i>Parámetros</i>	<i>Valores nominales(T1/T2)</i>
<i>Frecuencia</i>	60 (Hz)
<i>Relación de Voltaje</i>	85000-23000Y/13280 (V)
<i>Año de fabricación</i>	07/2014
<i>Elevación de temperatura</i>	55/65 (°C)
<i>Fases</i>	3
<i>kVA</i>	18000/24000/30000/33600 (kVA)

Debido a la cantidad de variables que se tienen es necesario realizar una reducción de dimensiones en los datos de entrenamiento para la RNA con el fin de evitar sesgos por información duplicada. Otro factor importante es evitar la utilización de variables que no estén altamente correlacionadas con la variable de salida y puedan provocar un coste computacional más elevado del requerido.



### 5.3.1. Análisis de correlación

En esta etapa de reducción de características, se calcula la matriz de correlación entre todas las variables del conjunto de datos de entrada utilizando las ecuaciones de la sección 4.3. Debido a las dimensiones de la matriz de correlación se seleccionan para mostrar en esta sección solo las variables que tengan una correlación mayor a 0.9 con respecto a la temperatura del aceite ( $T1OILM$ ). La Tabla 5.2 muestra la correlación entre las variables de entrada y salida utilizadas para los casos de estudio. Éstas se obtienen utilizando R.

**Tabla 5.2. Correlación entre variables de entrada y salida de los datos de entrenamiento**

<i>Variable</i>	$I_{an}$	$I_{bn}$	$I_{cn}$	$P_{an}$	$P_{bn}$	$P_{cn}$	$t_{amb}$
<b><math>T1OILM</math></b>	0.932645	0.934596	0.933518	0.913251	0.920146	0.916493	0.908423

Como se aprecia en la Tabla 5.2 las variables de  $t_{amb}, I_{an, bn, cn}$ , y  $kw_{an, bn, cn}$ , son las que tienen mayor correlación con respecto a la variable  $T1OILM$ , resultado que es similar al de la Figura 5.1 por lo que se seleccionan estas variables para modelar la temperatura del aceite en el transformador debido a que son las más representativas de la variable objetivo.

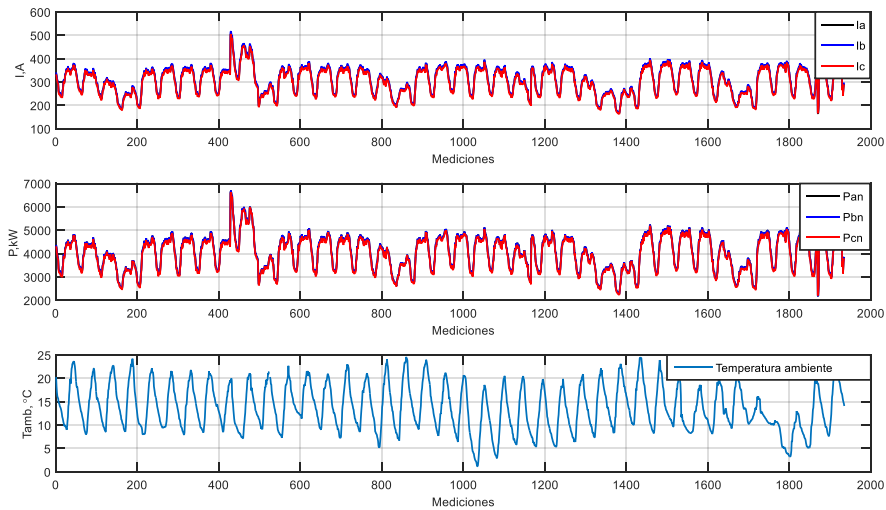
### 5.3.2. Selección de los datos de entrenamiento

El modelo RNA será preciso si el conjunto de datos de entrenamiento contiene todas las condiciones de funcionamiento normales para el transformador. El monitoreo de la condición basado en el método RNA usando datos SCADA ha sido desarrollado por varios investigadores [94-99]. Sin embargo, la selección de datos de entrenamiento, que es un paso importante en la construcción de buenos modelos de RNA, rara vez se discute en la literatura. Para la selección de datos de forma manual se utiliza el enfoque utilizado en [96]. Por lo tanto, este procedimiento se utiliza para la selección del conjunto de datos de entrenamiento.

La selección del conjunto de datos se realizó en función del valor de demanda de potencia activa en el transformador. Todos los valores repetidos en el conjunto de datos de entrenamiento fueron despreciados de tal manera que sólo se obtiene un vector de entradas para cada valor de salida. El conjunto de datos de entrenamiento final resultante seleccionado contiene 1935 puntos de muestra. La Figura 5.2 muestra el conjunto de datos de entrenamiento seleccionado de forma convencional.

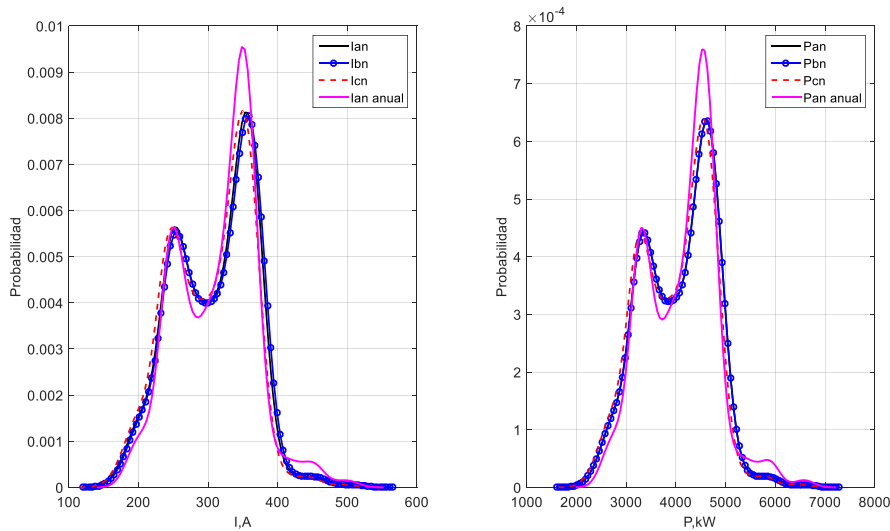
En este capítulo se utilizan datos obtenidos del transformador T1. Este transformador no presentó fallas durante la etapa de monitoreo, por lo que se selecciona para mostrar cómo se desarrolla el algoritmo y se seleccionan los modelos utilizados para la detección de fallas.





**Figura 5.2. Conjunto de datos de entrenamiento**

La Figura 5.3 presenta la distribución de probabilidad bimodal de los parámetros de entrada, es decir, los valores de corriente y potencia en el transformador para el conjunto de datos históricos y el conjunto de datos de entrenamiento seleccionado.



**Figura 5.3. Comparación de la distribución de probabilidad del conjunto de datos de entrenamiento con respecto a datos históricos**

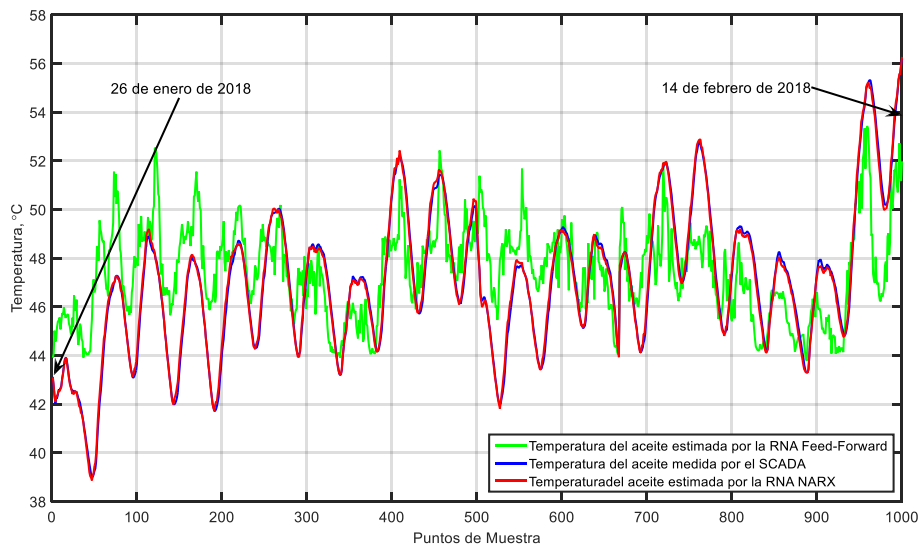
Se puede concluir que los datos de entrenamiento cubren toda la gama de operación del transformador en comparación con el conjunto de datos que abarca un año de funcionamiento del transformador. En el caso de estudio se seleccionan como se ve en la Figura 5.2, 1935 mediciones que van desde el 24 de diciembre hasta el 25 de enero con intervalos de medición

de 30 minutos, de modo que el conjunto de datos de entrenamiento tenga puntos de muestra de al menos un mes de funcionamiento del transformador. Es necesario destacar que no se muestran los valores de temperatura ambiente debido a que no se cuenta con los datos históricos y no es posible realizar la comparación.

#### 5.4. Comparación y selección entre modelos de RNA NARX y *Feed-Forward*

Como se muestra en el Capítulo III las RNA NARX son más precisas cuando se quiere realizar una predicción dónde existen problemas de series de tiempo. Su principal ventaja sobre una red de *Feed-Forward* es que permite ganar grados de libertad al incluir la predicción de un período como dato en períodos siguientes, resumiendo información de las variables exógenas, permitiendo incluir menos rezagos de las mismas y resumiendo el número de parámetros a estimar.

En esta sección se realiza una comparación de ambas estructuras. Para este análisis se toman 1000 muestras y se utiliza la estructura que se muestra en la Tabla 4.1. En la Figura 5.4 se observa la estimación de ambos modelos de RNA y se compara con las mediciones realizadas por el SCADA, se puede apreciar que el modelo de RNA NARX es mucho más preciso para este tipo de análisis en donde la variable a predecir es una variable física.



**Figura 5.4. Comparación de resultados entre el modelo de RNA NARX, *Feed-Forward* y SCADA**

Para comparar los resultados entre estos modelos de redes neuronales, los errores de ambos modelos se consideran durante el período descrito anteriormente. Es importante resaltar que el entrenamiento del modelo de RNA se realiza en base a datos seleccionados en el mes de enero de 2018, se presenta el modelo de RNA con datos del mes de febrero. Para hacer que

los resultados sean claros y fáciles de entender, se calcula el error entre ambas estimaciones y se comparan entre ellos.

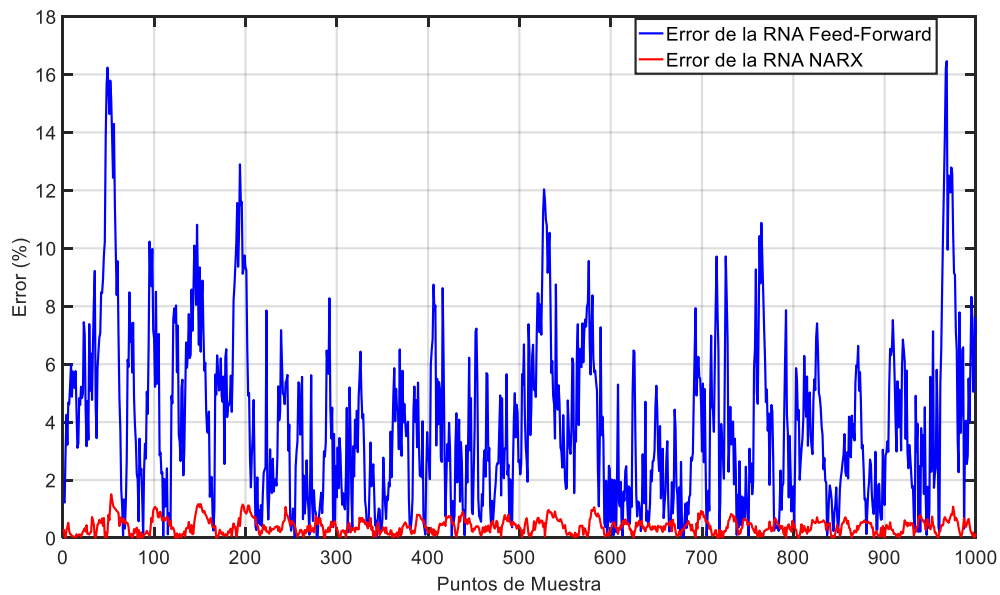
**i) Error de los modelos de RNA**

Un error se define como la diferencia entre el resultado del parámetro estimado del modelo de RNA y el valor real medido del parámetro almacenado en el SCADA. Para evaluar el error de ambos modelos se utiliza la siguiente ecuación:

$$E_i = abs \left[ \frac{y_i^{medida} - y_i^{RNA}}{y_i^{medida}} \right] * 100 \tag{5.1}$$

$E_i$  es el  $i^{ésimo}$  valor de error porcentual para cada muestra,  $y_i^{medida}$  es el  $i^{ésimo}$  valor de parámetro medido registrado en SCADA y  $y_i^{RNA}$  es el  $i^{ésimo}$  valor de la salida estimada del modelo de RNA e  $i$  es el  $i^{ésimo}$  punto de muestra sobre los que se aplica el modelo de RNA.

La Figura 5.5 presenta el resultado de los errores para los modelos de RNA NARX y *Feed-Forward* durante el período de monitoreo de la condición durante 20 días. Se puede observar que la salida  $E_i$  de ambos modelos RNA es muy diferente durante todo el período.



**Figura 5.5. Errores obtenidos de la estimación de temperatura de los modelos de RNA NARX y *Feed-Forward***

De la Figura 5.5 se puede concluir que el modelo de RNA más apropiado para el monitoreo de la condición del transformador es el NARX, debido a que es mucho más preciso. Para obtener resultados precisos es necesario ajustar el modelo, buscando reducir el error a los

valores esperados. En la siguiente sección se ajusta el modelo propuesto para el monitoreo de la condición del transformador.

### **5.5. Comparación y selección de estructuras para el modelo de RNA propuesto**

En esta sección se comparan los resultados obtenidos utilizando diferentes estructuras de redes neuronales NARX. Se realizan cambios en la cantidad de neuronas en la capa oculta con el objetivo de reducir el error máximo y promedio entre, la temperatura estimada y la medida. Los datos utilizados para comparar los diferentes modelos son los medidos por el SCADA durante la etapa de monitoreo de la condición; éstos van desde el 26 de enero al 14 de febrero de 2018.

La selección de los datos de entrenamiento en esta red neuronal se realizó siguiendo el mismo concepto que se ha desarrollado a lo largo del capítulo, donde las variables de entradas son: la temperatura ambiente, la potencia entregada por fase y la corriente por fase. Como variable de salida se utiliza en el entrenamiento de la RNA la temperatura del aceite en el transformador. Luego de entrenar la red neuronal se procede a validar el modelo entrenado, para validar cada estructura se utilizan datos que este modelo de red neuronal no haya visto, o sea nuevos datos que son tomados durante la etapa de monitoreo y se comparan con la predicción obtenida por cada estructura de RNA.

El algoritmo de entrenamiento utilizado para decidir los pesos sinápticos es el Levenberg-Marquardt descrito en el Capítulo III y las funciones de activación de la capa oculta y salida son la función sigmoidea y umbral respectivamente.

En este análisis se utilizan redes neuronales artificiales NARX con 5, 10, 20, 30, y 40 neuronas en la capa oculta. En la Figura 5.6 se muestra la estructura propuesta, la variable  $NN$  representa el número de neuronas en la capa oculta,  $x(t)$  y  $y(t)$  son las variables de entrada y salida respectivamente. La capa de salida tiene una neurona.

Es importante destacar, que a medida que se aumenta el número de neuronas en la capa oculta, la velocidad procesamiento de datos es más lenta, esto, debido a que la matriz de pesos es de mayor dimensión y la red neuronal debe realizar un mayor número de cálculos.

La cantidad de datos analizados y la velocidad de procesamiento son directamente proporcionales. En la medida que la cantidad de datos aumenta, el procesamiento de datos es más lento, por lo que se debe buscar un equilibrio entre la cantidad de neuronas en la capa oculta y la velocidad de procesamiento.

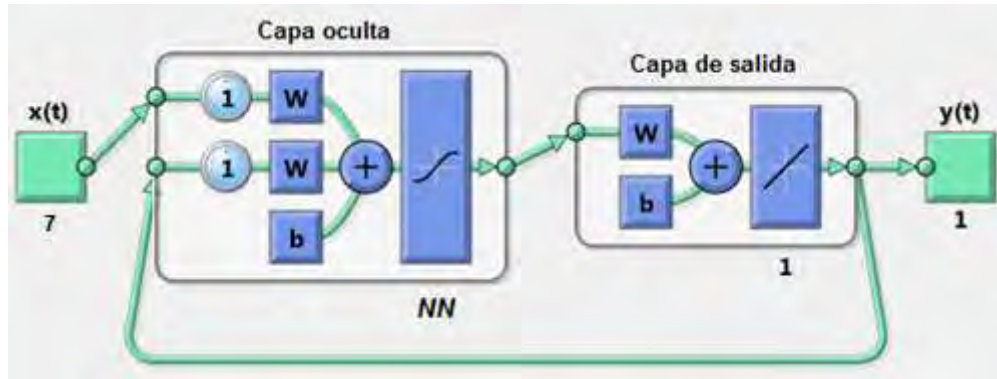


Figura 5.6. Modelo de RNA NARX propuesto

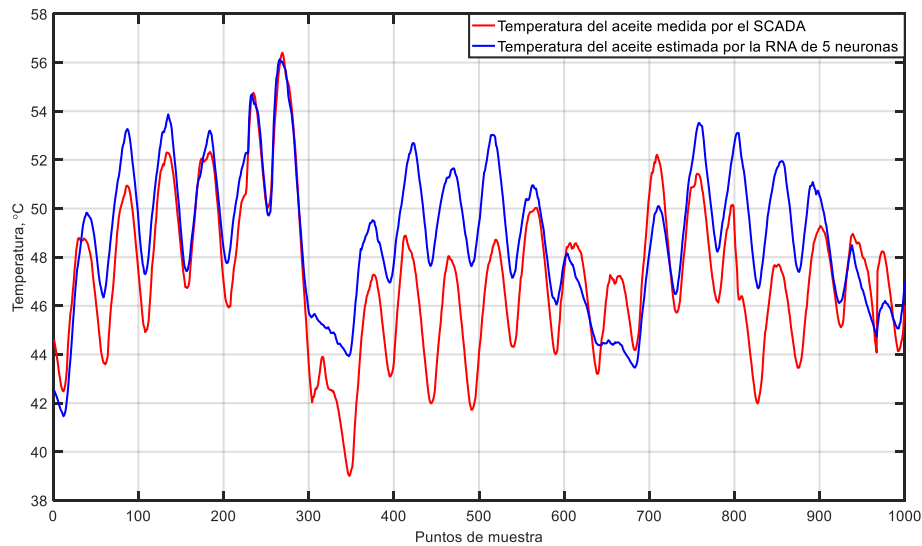
Para simplificar y comprender mejor este análisis se calcula promedio diario de la temperatura del aceite medida y se compara con cada resultado de las estructuras utilizadas. Esto es necesario debido a que se busca con este trabajo, desarrollar un algoritmo para evaluar la condición de transformadores y planificar mantenimiento predictivo a partir del estado técnico del transformador.

A continuación, se analizan los resultados obtenidos de cada estructura descrita anteriormente:

#### i) RNA NARX de 5 capas ocultas

Para modelar la red neuronal con 5 neuronas en la capa oculta se utiliza el conjunto de datos de entrenamiento descrito en la sección 5.3. En este caso la velocidad de procesamiento es rápida, pero la precisión en la predicción no cumple los requisitos esperados.

En la Figura 5.7 se muestran los resultados obtenidos durante en la etapa de monitoreo de la condición del transformador y se compara con la estimación obtenida con la RNA de 5 neuronas.



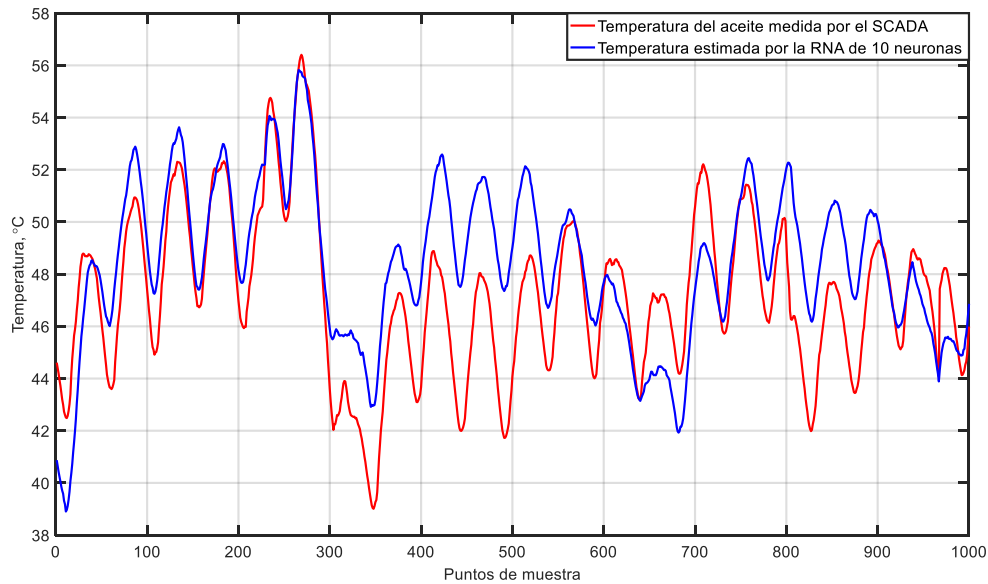
**Figura 5.7. Comparación entre la temperatura estimada por la RNA de 5 neuronas y la temperatura medida por el SCADA**

De la figura anterior se puede concluir, que esta topología de red neuronal es capaz de predecir la temperatura del aceite en el transformador. En este caso el error alcanza hasta los 5 grados centígrados de temperatura, superando el error admisible para lograr el objetivo de este trabajo de tesis. Si consideramos el error porcentual absoluto medio para evaluar la estimación de este modelo, se obtiene que este error es 4.2249%, este error se mejora aumentando el número de neuronas en la capa, por lo que se decide aumentar el número de neuronas a 10.

#### ii) RNA NARX de 10 capas ocultas

En este caso se utiliza una estructura de red neuronal con 10 neuronas en la capa oculta, en este caso la velocidad de procesamiento sigue siendo rápida. Al igual que en el caso anterior de 5 capas el resultado no es el óptimo debido a que el error sigue siendo alto.

En la Figura 5.8 se muestra la estimación de temperatura obtenida por la red neuronal de 10 neuronas y se compara con las mediciones tomadas por el SCADA durante la etapa de monitoreo de la condición del transformador.



**Figura 5.8. Comparación entre la temperatura estimada por la RNA de 10 neuronas y la temperatura medida por el SCADA**

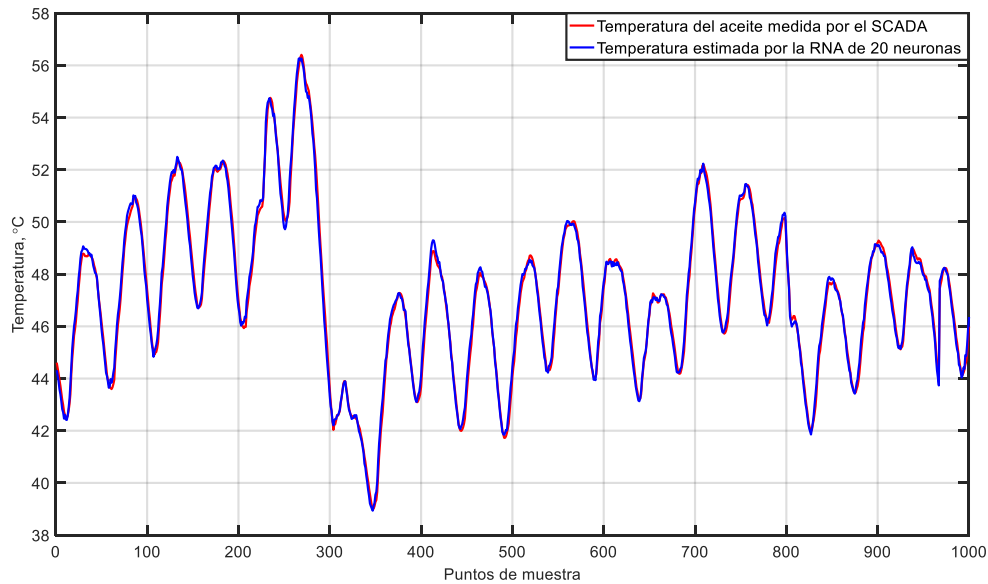
En este caso se puede observar, que al igual que en el caso de la RNA de 5 neuronas en la capa oculta, el error porcentual absoluto medio entre la temperatura estimada y la medida es 3.7374% durante la etapa de monitoreo. Este resultado no cumple con el error esperado para la implementación de este algoritmo, por lo que es necesario aumentar el número de neuronas en la capa oculta de la RNA a 20 neuronas.

### iii) RNA NARX de 20 capas ocultas

Para el caso de la red neuronal de 20 neuronas en la capa oculta la velocidad de procesamiento más lenta que en los casos anteriores, debido a que es necesario que la red calcule los pesos sinápticos para cada neurona y en este caso la matriz de pesos va a ser de mayor dimensión que las anteriores.

En la Figura 5.9 se puede apreciar una notable mejora en la estimación de la red neuronal en comparación con los casos de 5 y 10 neuronas en la capa oculta, en este caso el error porcentual absoluto medio en la predicción es del 0.4377% con respecto al rango en que varía la temperatura durante la etapa de monitoreo de la condición.

Este resultado es considerado satisfactorio para desarrollar la aplicación pues el error en la predicción se encuentra por debajo del 1%.



**Figura 5.9. Comparación entre la temperatura estimada por la RNA de 20 neuronas y la temperatura medida por el SCADA**

Con el objetivo de mejorar este resultado se aumenta el número de neuronas en la capa oculta de la RNA a 30 y se realiza el mismo procedimiento.

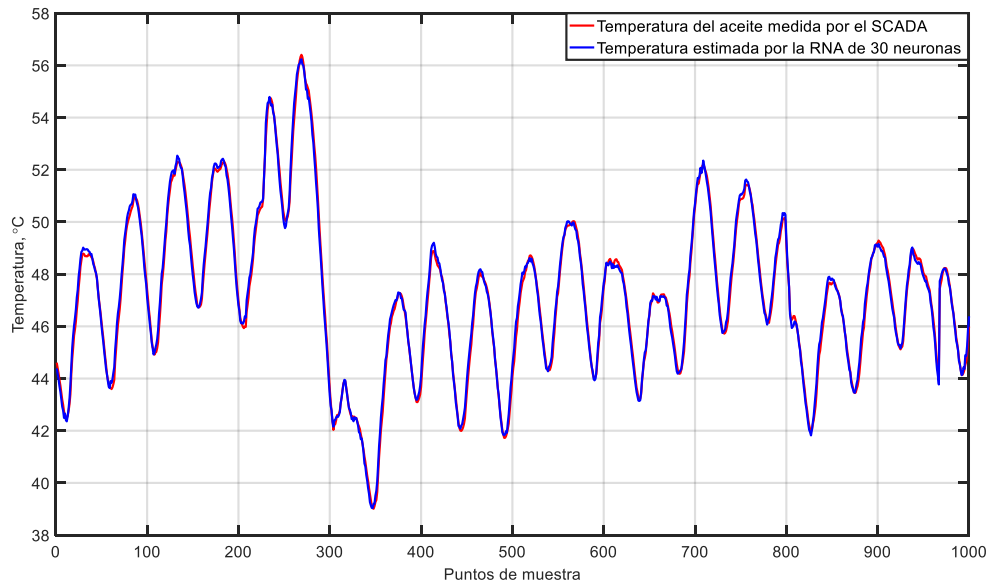
**iv) RNA NARX de 30 capas ocultas**

La red neuronal de 30 neuronas en la capa oculta presenta una velocidad de procesamiento más lenta que el caso anterior, aunque este tiempo es aceptable para el desarrollo de la aplicación que se ha desarrollado.

En la Figura 5.10 se comparan los resultados obtenidos de la estimación de la red neuronal de 30 neuronas en la capa oculta y las mediciones realizadas durante la etapa de monitoreo.

Como se observa en la figura anterior los resultados obtenidos por la RNA son similares a los obtenidos en el caso anterior, en este caso se mejora ligeramente el rendimiento de la RNA y la velocidad de procesamiento de los datos es similar a la de la red neuronal de 20 neuronas.





**Figura 5.10. Comparación entre la temperatura estimada por la RNA de 30 neuronas y la temperatura medida por el SCADA**

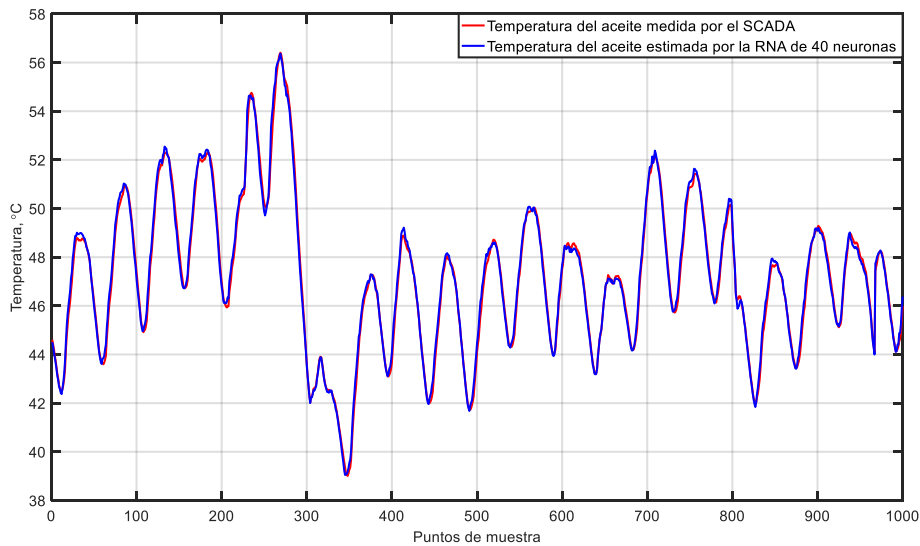
En este caso el error porcentual absoluto medio que se obtiene es 0.4093% con respecto al rango de temperatura del aceite en la etapa de monitoreo. A continuación, se realiza el procedimiento nuevamente utilizando 40 neuronas en la capa oculta de la RNA.

#### v) RNA NARX de 40 capas ocultas

La red neuronal con 40 neuronas en la capa oculta se modela con el objetivo de buscar un mejor rendimiento que en los casos anteriores, debido a que una mayor precisión en la estimación de la temperatura del aceite, influye en una detección de falla en el transformador más efectiva. Con el incremento en la cantidad de neuronas en la capa oculta de la RNA, se observa que el rendimiento de estas ha ido en aumento y el tiempo en que procesa los datos y se entrena la red neuronal también ha aumentado de manera significativa, sin embargo, en ocasiones, no se logra mejorar este rendimiento con el aumento de neuronas en la capa oculta de la RNA.

Para evaluar el rendimiento de esta RNA se utilizan los datos obtenidos en la etapa de monitoreo de la condición y se comparan con la estimación, los resultados se muestran en la Figura 5.11.

Siguiendo la secuencia desarrollada en esta sección se aumentó el número de neuronas en la capa oculta de la red neuronal a 40 neuronas con el objetivo de mejorar el rendimiento de la RNA.



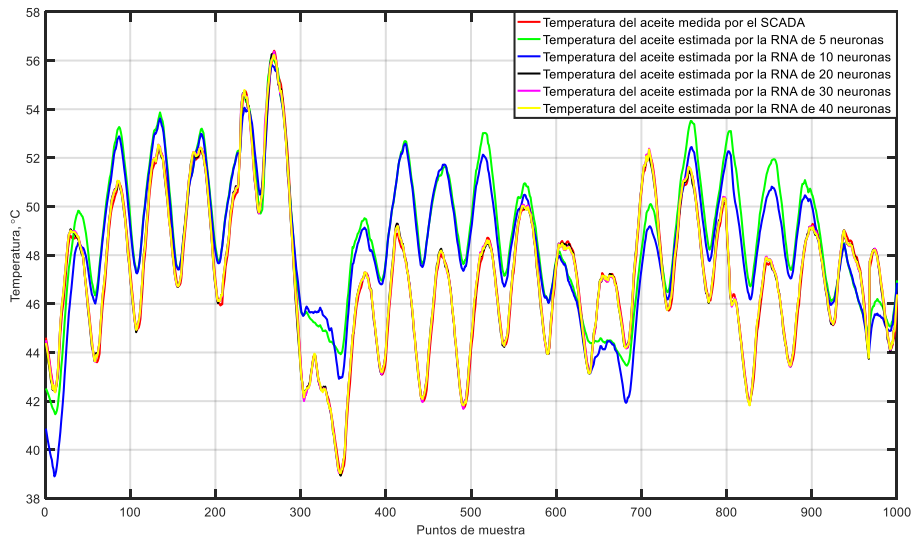
**Figura 5.11. Comparación entre la temperatura estimada por la RNA de 40 neuronas y la temperatura medida por el SCADA**

Sin embargo, los resultados son similares a los obtenidos en los casos de 20 y 30 neuronas, con un incremento en la velocidad de procesamiento de los datos de entrenamiento, en este caso el error porcentual absoluto medio es de 0.4542%, resultado que es mayor que en el caso de 30 neuronas, esto puede estar dado por varios factores que se mencionan en el Capítulo III.

#### vi) Comparación de resultados y selección de estructura para la implementación del algoritmo

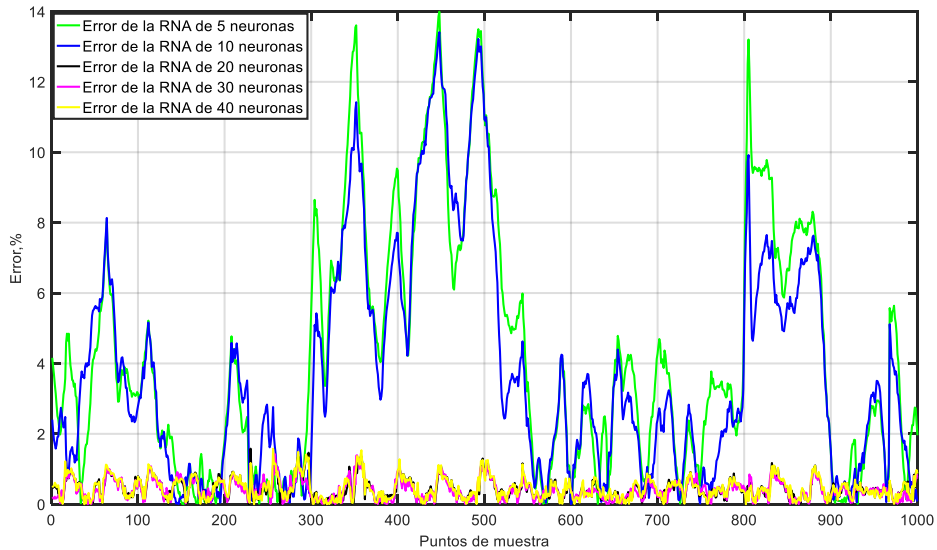
Durante la etapa de modelado se trabajó con RNA que presentaban más de 40 neuronas en la capa oculta. Los resultados obtenidos en todos los casos no mejoraron el rendimiento obtenido con la red neuronal de 30 neuronas en la capa oculta y por cuestiones de espacio no se muestran estos casos en este capítulo. En la Figura 5.12, se muestran los resultados obtenidos de cada modelo y se compara con la medición del SCADA.

Además, se aprecia que las estructuras de redes neuronales artificiales con 20, 30 y 40 neuronas son las que estiman con mayor exactitud y precisión la temperatura del aceite. Sin embargo, no es posible determinar qué modelo ofrece mejores resultados, por lo que es necesario calcular el error para cada estructura.



**Figura 5.12. Comparación de resultados de los modelos NARX**

En la Figura 5.13 se muestra el error de todos los casos de estudio y para cada punto de muestra.



**Figura 5.13. Errores obtenidos para cada modelo de RNA NARX**

Para comprender mejor la Figura 5.13 se calcula el error máximo y promedio de cada modelo. El objetivo de calcular estos errores es seleccionar el modelo de RNA para integrarlo al algoritmo de monitoreo de la condición en el transformador. En la Tabla 5.3 se muestran, los resultados obtenidos de cada modelo.

**Tabla 5.3. Comparación de resultados de los modelos propuestos**

<i>Cantidad de neuronas</i>	<i>Error Máximo (%)</i>	<i>MAPE (%)</i>	<i>Iteraciones</i>
5	13.9940	4.2249	13
10	13.4121	3.7374	43
20	1.5795	0.4377	87
30	1.4272	0.4093	125
40	1.5777	0.4542	193

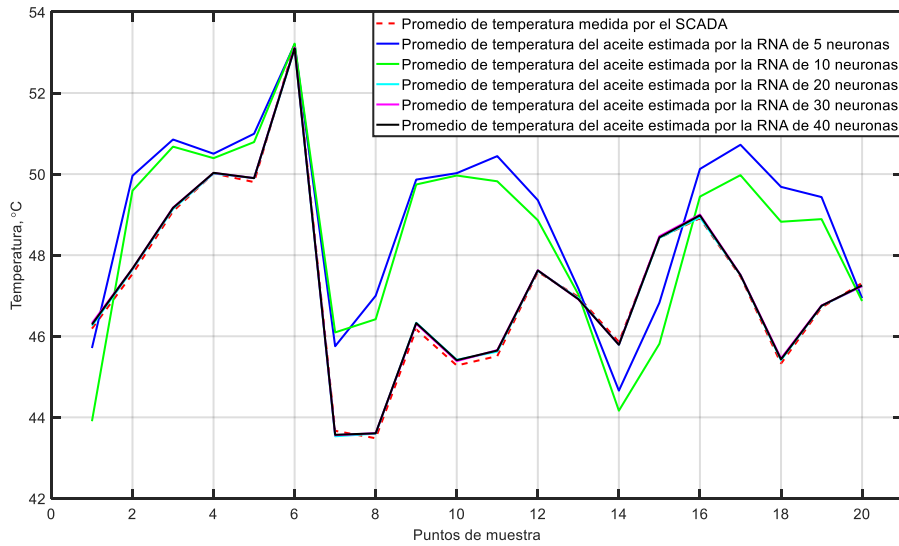
Se puede concluir que, el modelo de 30 neuronas en la capa oculta arroja mejores resultados en la estimación de la temperatura del aceite en el transformador. Sin embargo, los tiempos de procesamiento de datos y la cantidad de iteraciones son mayores al caso de 20 neuronas. Considerando que este tipo de análisis se debe realizar de manera diaria, debido a que la planificación de mantenimiento se realiza en escalas de tiempo de horas a un año según el concepto descrito en el Capítulo II; se selecciona el modelo de 30 neuronas para desarrollar la siguiente etapa del algoritmo.

### 5.5. Estimación de temperatura promedio diaria

Para comprender mejor qué ocurre con los equipos que se analizan, es necesario que los datos que se procesan ofrezcan información de manera más sencilla. Este trabajo de tesis analiza solo dos meses de operación de estos transformadores. Sin embargo, cuando el tiempo de análisis supera los 12 meses o más, la cantidad de datos aumenta considerablemente. Esto provoca la necesidad de promediar de manera diaria los datos para trazar un perfil de funcionamiento del equipo, que sea más sencillo de analizar para el personal técnico y de operación en las subestaciones. Además, se evitan errores que puedan estar provocados por los equipos de monitoreo. Por este motivo, se promedian los resultados obtenidos con los modelos vistos en la sección anterior y se analiza el resultado.

En la Figura 5.14 se observan los resultados obtenidos luego de promediar diariamente los valores de temperatura del aceite de las mediciones realizadas durante la etapa de monitoreo y la estimación de todas las RNA presentadas en la sección anterior.

También observa que los resultados de las estimaciones de temperatura comparadas con las mediciones realizadas durante la etapa de monitoreo se comportan de manera similar a los casos anteriores.



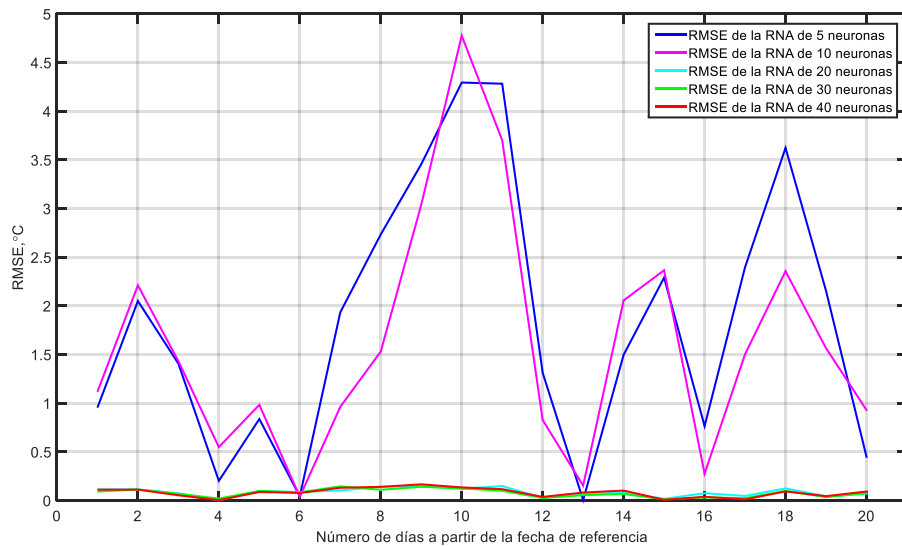
**Figura 5.14. Comparación entre la temperatura estimada por los modelos de RNA NARX y la temperatura medida por el SCADA**

Además se puede determinar qué estructura de RNA ofrece mejor rendimiento calculando el error medio cuadrático (*RMSE*), para hacer que los resultados sean claros y fáciles de entender y se calcula en un período de 20 días que abarcan el total de datos disponibles durante la etapa de monitoreo, y se define por la ecuación (5.2).

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{n_j} (y_{ij}^{medida} - y_{ij}^{RNA})^2}{n_j}} \quad (5.2)$$

$RMSE_j$  es el valor del error medio cuadrático de la  $j$  –ésima semana, donde  $n$  es el número total de puntos de datos en la semana  $j$ . La Figura 5.15 presenta los resultados de los modelos de RNA de 5, 10, 20, 30 y 40 neuronas en la capa oculta durante el período de monitoreo de condición. Se puede observar que la salida  $RMSE_j$  de todos los modelos de RNA es muy similar en los casos de 20, 30 y 40 neuronas.

En la Figura 5.15 se aprecia como varía el valor del error medio cuadrático de cada modelo de red neuronal, al igual que en los casos donde se modelan las estructuras de 5, 10, 20, 30 y 40, las dos primeras son las que tienen mayor error, por lo que se descartan desde el inicio para la selección del modelo para el monitoreo de la condición. Los resultados en este caso son consistentes y arrojan la misma conclusión que la sección anterior.



**Figura 5.15. Errores obtenidos de la estimación de temperatura de los modelos de RNA NARX**

La selección de este modelo está dada, debido a que para determinar la ocurrencia de un disturbio es necesario que el error, entre los valores de temperatura estimados por el modelo de RNA y los valores registrados en el SCADA, sea mínimo en operación estable del transformador. En caso de la ocurrencia de una falla en el transformador, este error debe aumentar debido a que el modelo de red neuronal estima los valores en estado estable.

### 5.6. Detección de fallas utilizando la distancia de Mahalanobis

La detección de fallas, como se ve en el Capítulo IV es el proceso de diferenciar cualquier error normal de los que podrían indicar una falla en el equipo monitoreado. Por lo tanto, lograr una correcta detección de fallas es la parte mas importante en el proceso de monitoreo de la condición.

Con el objetivo de resolver este problema se propone un enfoque estadístico en el Capítulo IV con el cual se detectan las fallas que ocurren en el transformador que incrementan la temperatura en el aceite durante la etapa de monitoreo, este utiliza la medida de distancia nombrada distancia de Mahalanobis. Las averías que pueden ser detectadas con este método pueden ser problemas en el sistema de enfriamiento en el transformador así como, algún problema interno que genere un incremento de esta temperatura.

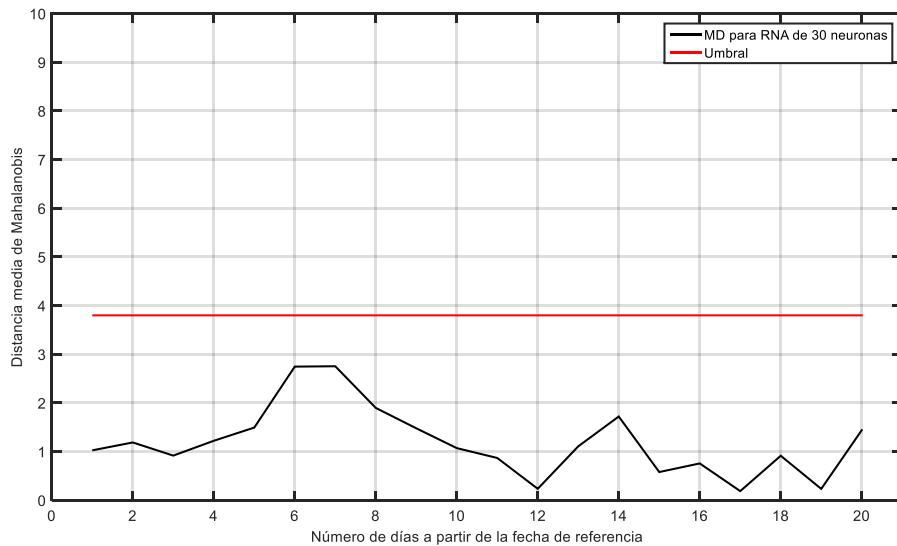
La distancia de Mahalanobis ( $MD$ ) es una unidad de medición de distancia adimensional, que tiene la capacidad de capturar la correlación de variables en un proceso o sistema. La distancia de Mahalanobis proporciona un valor de distancia univariante para datos multivariados.

Para la detección de fallas en el transformador monitoreado es necesario definir un valor umbral para determinar qué valores de  $MD$  indican una falla, el procedimiento para decidir en umbral se expone en el Capítulo IV.

**i) Distancia de Mahalanobis para el caso de estudio sin falla (Transformador T1)**

Se presenta la salida de la detección de fallas usando el enfoque de monitoreo de la condición basado en RNA para el aceite del transformador. El valor umbral se ha calculado para el modelo de RNA de 30 neuronas en la capa oculta, en función de sus propios datos de entrenamiento.

La Figura 5.16 muestra el valor medio de la distancia de Mahalanobis para la temperatura del aceite en el transformador T1 ( $T1OILM$ ) durante el período de monitoreo de la condición. Durante esta etapa de monitoreo de 20 días que corresponde a la fecha del 26 de enero de 2018 al 14 de febrero de 2018 no se reportaron fallas internas en el transformador por el personal de operación y mantenimiento.



**Figura 5.16. Distancia de Mahalanobis calculada para la temperatura del aceite en el transformador T1**

Se puede observar que durante todo el período de monitoreo de la condición, la distancia media de Mahalanobis no ha cruzado el umbral y no ha habido falsas alarmas. Por lo tanto, el enfoque de monitoreo de la condición basado en RNA propuesto ha sido validado para un caso sin fallas.

### ii) Distancia de Mahalanobis para el caso de estudio con falla (Transformador T2)

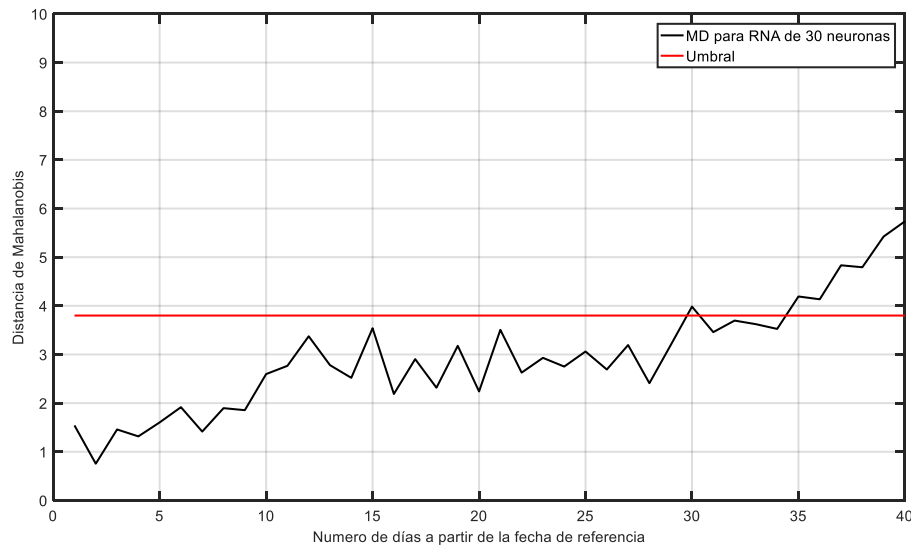
Al igual que en el caso en que no se presentan fallas en el transformador T1, la detección de fallas para el transformador T2, se presenta utilizando el procedimiento desarrollado en este trabajo de tesis. El valor umbral calculado para este caso es igual al transformador T1 debido a que son transformadores similares y se encuentran ubicados en la misma subestación. Para el caso en que el transformador presenta una falla se utiliza el modelo de red neuronal de 30 neuronas en la capa oculta.

La Figura 5.17 muestra el valor promedio diario de la distancia de Mahalanobis para la temperatura del aceite en el transformador T2 (*T2OILM*) durante la etapa de monitoreo de condición que en este caso es de 40 días. Se puede observar que la distancia media de Mahalanobis ha cruzado el valor umbral en el día 30 de la etapa de monitoreo de la condición. Sin embargo, en esta fecha el sistema SCADA no emitió ninguna alarma, en este caso se puede observar una tendencia al aumento de la distancia de Mahalanobis desde que comienza el monitoreo, utilizando el algoritmo desarrollado en este trabajo. En esta situación, se recomienda que el proveedor del servicio de mantenimiento o el operador del transformador tome una decisión basada en la experiencia adquirida mediante la aplicación continua del algoritmo propuesto. Además, debe tenerse en cuenta que una inspección en una etapa temprana podría haber revelado el deterioro en etapas muy iniciales. La información sobre el deterioro en etapas muy tempranas puede ser útil desde la perspectiva de la administración del mantenimiento.

En la Figura 5.17, se puede observar que el valor promedio de la distancia de Mahalanobis ha cruzado el día 34 de la etapa de monitoreo. A partir de este día la diferencia entre el umbral y el valor medio de la distancia de Mahalanobis se incrementa notablemente para el día 40. La cantidad de alarmas y advertencias producida por el SCADA que se han producido a partir del día 37 de la etapa de monitoreo es más alta de lo normal. Estas dos condiciones juntas proporcionan una fuerte indicación de deterioro en el funcionamiento del transformador.

El monitoreo de la condición basado en RNA muestra que a medida que pasan los días la *MD* aumenta significativamente hasta que cruza el umbral calculado para este caso. Para determinar que está causando este funcionamiento anómalo es necesario realizar una inspección en el equipo monitoreado.





**Figura 5.17. Distancia de Mahalanobis calculada para la temperatura del aceite en el transformador T2**

En este caso el transformador presentó problemas en una bomba de aceite, específicamente en un rodamiento de la misma lo que evitaba un adecuado bombeo de aceite en el transformador, alterando así la temperatura del aceite en la parte interna del transformador, el sistema SCADA emitió varios avisos antes de que se presentara el fallo.

Por lo tanto, se puede concluir que, aunque la monitorización del estado basado en RNA está diseñada para monitorear directamente solo la temperatura del aceite, cualquier daño en otro componente del transformador, se va a ver reflejado en el aceite.

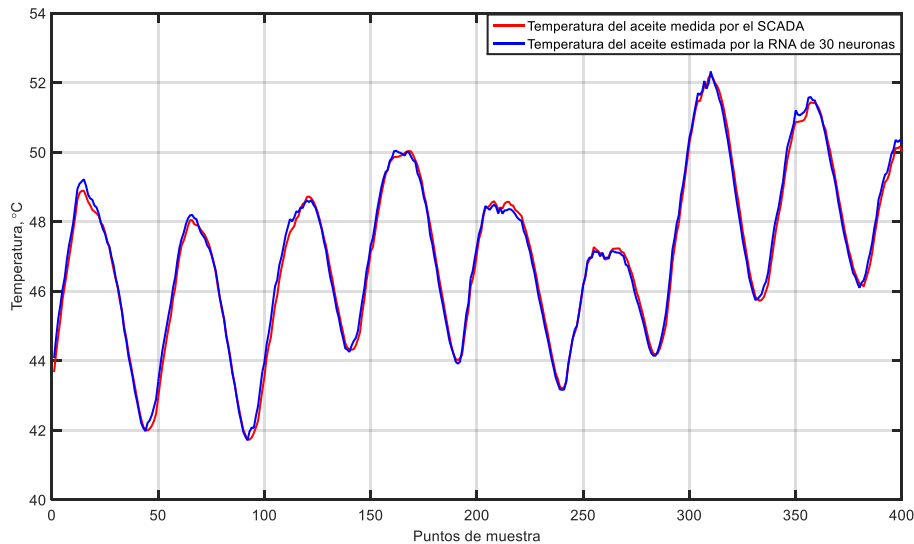
### 5.8. Análisis comparativo contra otros algoritmos similares

Varios investigadores han presentado métodos de monitoreo de condición basados en RNA utilizando datos almacenados en sistemas SCADA. La mayoría de los métodos presentados anteriormente se basan en una filosofía similar de usar la RNA para modelar las características de comportamiento normal del componente que se está monitoreando, con el fin de detectar fallas en tiempo real. En esta sección, el enfoque de monitoreo de condición basado en RNA propuesto en esta tesis se compara con dos enfoques de monitoreo de la condición basados en RNA utilizados en [95, 96].

En [95] se ha introducido SIMAP (Sistema Inteligente para el Mantenimiento Predictivo). El enfoque propuesto utiliza RNA para modelar el comportamiento normal de la temperatura del cojinete de la caja de engranes. Un modelo similar se presenta en [96], donde los datos de temperatura promedio de 10 minutos se utilizan para modelar el comportamiento normal de las características de temperatura de la caja de engranes utilizando redes neuronales. En

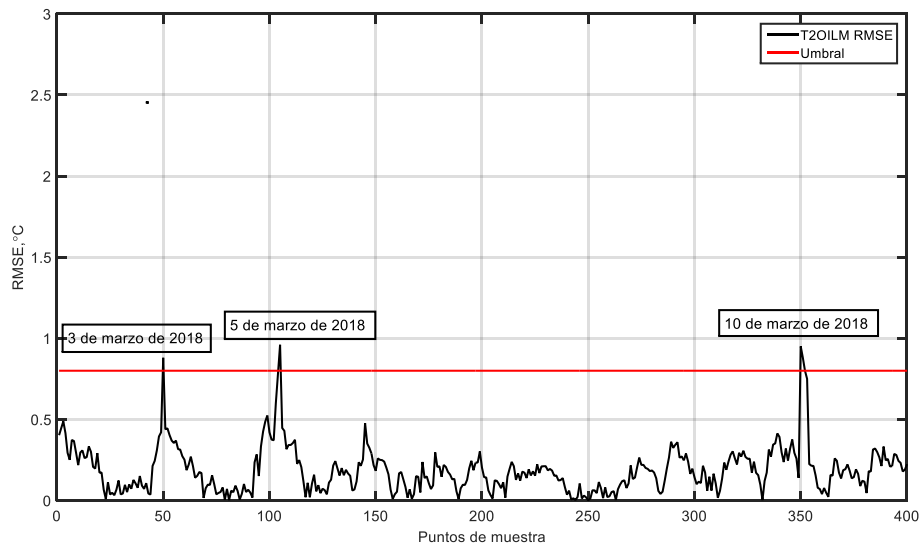
ambos enfoques propuestos, la detección de fallas se realiza en función del error entre la temperatura estimada de la RNA y la temperatura medida real. En [95], la falla se detecta en función del aumento del error en relación con un intervalo de confianza predefinido y en [96] la falla se detecta en función del aumento de la duración y la frecuencia de los errores. Sin embargo, con respecto al caso presentado anteriormente, un valor de error simple; es decir, la diferencia en la temperatura estimada y real, se encontró insuficiente para la detección de fallas.

En este trabajo se propone un algoritmo para la selección de datos de entrenamiento para la predicción de temperatura del aceite en transformadores. Este enfoque ofrece un resultado muy preciso en la predicción de la temperatura del aceite, y se aprecian en la Figura 5.18.



**Figura 5.18. Comparación de la salida de la RNA y los valores de temperatura almacenados en el SCADA**

Como se muestra, la temperatura estimada del aceite en T2 es muy cercana al valor de temperatura real registrado en el sistema SCADA. Un análisis del caso en que ocurre una falla en uno de los transformadores analizados (T2), muestra que la temperatura de operación del aceite durante una etapa de operación normal cruzo el umbral establecido para este caso. Por lo tanto, el RMSE de temperatura no es un buen indicador de funcionamiento anómalo. La Figura 5.19 muestra el valor de RMSE para el período de monitoreo de condición durante varios días de la etapa de monitoreo.



**Figura 5.19. Valor RMSE de monitoreo de la condición para el Transformador T2**

En la Figura 5.19 para un valor de umbral bajo de 0,8 grados centígrados, no es posible definir la ocurrencia de una falla, ya que la diferencia entre la temperatura estimada y la real del aceite en el transformador es muy pequeña y en ocasiones algunas mediciones pueden arrojar falsas alarmas. Por lo tanto, para superar este inconveniente, se propone la utilización del método de distancia de Mahalanobis, introducido en la sección 4.3.4 de este trabajo.

Además, el enfoque propuesto en [96] no define un valor de umbral simple para la detección de fallas. El enfoque propuesto en esta tesis proporciona una comprensión fácil de la salida del modelo RNA al proporcionar un valor umbral, que se calcula en base a los datos de entrenamiento. La provisión de dicho valor umbral simplifica la aplicación del monitoreo de condición basado en RNA, principalmente debido a las siguientes ventajas:

- No se requiere un juicio subjetivo sobre la falla por parte de un experto
- No es necesario almacenar los detalles sobre el valor de MD promediado del punto de inspección anterior para la detección de fallas

La Tabla 5.4 muestra el resumen de la comparación de los enfoques basados en RNA propuestos previamente con el enfoque presentado en esta tesis.

**Tabla 5.4. Comparación resumida de algoritmos similares propuestos previamente vs el algoritmo propuesto**

<i>Puntos de comparación</i>	<i>Modelos en [95] y [96]</i>	<i>RNA basada en el enfoque propuesto en este trabajo de tesis</i>
<i>Selección del conjunto de datos de entrenamiento</i>	No se discute en [95] y manual en [96]	Manual
<i>Detección de falla</i>	Usan la diferencia entre la temperatura estimada y la real	Distancia de Mahalanobis
<i>Valor umbral</i>	Bandas de confianza del 95% en [95], no definidas en [96]	Un enfoque estadístico para decidir el umbral basado en datos de operación de estado estable del conjunto de datos de entrenamiento

### CAPÍTULO VI: CONCLUSIONES Y TRABAJOS FUTUROS

Este trabajo se enfocó en el desarrollo de un algoritmo con el que se utiliza la información almacenada en SCADA para estimar el estado técnico de transformadores a partir del monitoreo de la temperatura del aceite.

#### 6.1. Conclusiones

Se propone un algoritmo para el diagnóstico de fallos en máquinas eléctricas con un enfoque eminentemente práctico, orientado hacia el mantenimiento predictivo basado en el monitoreo de la condición en transformadores de potencia. El algoritmo propuesto se implementó con un objetivo claro: la detección de fallas o funcionamiento anómalos en el transformador. El procedimiento propuesto es aplicable a otros transformadores. A continuación, se describen las principales conclusiones y aportaciones hechas en la tesis:

- La presente tesis propone un algoritmo basado en Redes Neuronales para el diagnóstico de fallos en máquinas eléctricas cuyas principales características son:
  - La primera etapa del algoritmo consiste en un proceso de dos etapas, filtrado de datos y análisis de correlaciones parciales del conjunto de datos. Esto reduce el número de variables indicativas de fallo para disminuir el esfuerzo computacional y emplear sólo las características necesarias para el diagnóstico.
  - La metodología permite combinar variables indicativas de fallo de diferentes orígenes. De hecho, en el presente trabajo se emplearon variables eléctricas y de temperatura para el conjunto de datos de entrenamiento.
  - El algoritmo permite detectar cuando el transformador comienza a operar de forma errónea o con parámetros fuera de lo normal.
  - El algoritmo es capaz de detectar la tendencia hacia un funcionamiento anómalo y establece un umbral el cual se establece a partir de la probabilidad de falla que va a tener el equipo en caso de no aplicar mantenimiento o inspección para solucionar el problema existente.
  - La metodología ha demostrado ser robusta al obtener buenos resultados empleando variables eléctricas y físicas de un sistema real. Cabe resaltar que el comportamiento complejo de algunas de las características y la gran cantidad de variables consideradas representan un reto para la metodología de diagnóstico.
- De la investigación realizada en este trabajo se puede concluir que el aceite del transformador es sin dudas uno de los componentes esenciales para el correcto funcionamiento de todo transformador. Sus funciones principales son: aislar y refrigerar el equipo; protegiéndolo de sobrecalentamiento y formación de gases que pudieran afectar su funcionamiento.
- El enfoque de monitoreo de la condición basado en RNA se aplica a dos casos de estudio y los resultados validan el rendimiento del algoritmo propuesto.

- La realización de la presente tesis de maestría y la forma en la que se desarrolló, abre un amplio panorama para el desarrollo de nuevas estrategias en el mantenimiento predictivo utilizando RNA.
- La estructura utilizada en este trabajo puede ser utilizada para otros transformadores, aunque éstos sean de diferente potencia y nivel de voltaje.
- El algoritmo de aprendizaje Levenberg-Marquardt es el que mejor reduce el error de entrenamiento en este tipo de problemas. Sin embargo, otros algoritmos de aprendizaje pueden ser utilizados.
- La utilización de variables de entrada que tengan una alta correlación con la variable de salida permite estimar con mayor exactitud la variable objetivo.

### 6.2. Trabajos futuros

La tesis se ha centrado en el desarrollo del método basado en inteligencia artificial para extraer signos de daño de los datos almacenados en el SCADA. Se presentan los resultados para el monitoreo de la condición del transformador usando el enfoque de monitoreo de la condición basado en RNA. Los trabajos futuros de esta tesis se pueden enumerar de la siguiente manera:

- Realizar el análisis de manera similar a lo visto en este trabajo de tesis, pero utilizando datos históricos con un año o más de datos almacenados, y evaluar el rendimiento de la RNA debido a que este tiempo cubre el perfil histórico de temperatura ambiente y de carga de todo el año por lo que se tiene mayor cantidad de información para que la red neuronal sea entrenada.
- Desarrollo de un modelo similar basado en ML para detectar que componentes causan el funcionamiento anómalo en el transformador.
- Aplicar el algoritmo desarrollado en este trabajo de tesis a otras subestaciones para evaluar el rendimiento del mismo en otros equipos
- Desarrollar un esquema de mantenimiento predictivo autónomo que tenga en cuenta la importancia de los equipos y realice un análisis de riesgo para evitar grandes pérdidas, producto de una salida del equipo no deseada.

## BIBLIOGRAFÍA

- [1] IEC White Paper. 2016. “Strategic asset management of power networks”. Geneva, Switzerland: International Electrotechnical Commission.
- [2] J. Martínez. 2006. “Diseño de estrategias de mantenimiento para la gestión de activos físicos en empresas distribuidoras de energía eléctrica en argentina”. Tesis de maestría. Universidad Tecnológica Nacional Facultad Regional Buenos Aires.
- [3] SENER. “Programa de Redes Eléctricas Inteligentes”. Mayo 2016.
- [4] R. Brown. 2004. “Reliability and Distribution Asset Management”, Electric Energy T&D Magazine. Vol 8, No.6.
- [5] S. Di Matteo. 2015. “Big Data, Predictive Analytics and Asset Health Indexing”, Industry Update Manufacturing Magazine and Directory. Issue 85.
- [6] R. Otal & T. Hjartarson. 2016. “Transmission & Distribution World, Sustainment Actions Take a New Direction” retrieved from <http://tdworld.com/smart-grid/sustainment-actions-take-new-direction>.
- [7] J. Millán, 2006. “Entre el mercado y el Estado: Tres décadas de reformas en el sector eléctrico de América Latina: Departamento de Desarrollo Sostenible del BID”, Banco Interamericano de Desarrollo Washington D. C., Estados Unidos ISBN:1-59782-028-8
- [8] S. García, 2016. “Improved maintenance and management of distribution assets due to the digitalization of the electrical grid”, Official Master's Degree in the Electric Power Industry, Universidad Pontificia Comillas, Madrid, España.
- [9] C. Liu; S. McArthur & S. Lee. 2016. “Smart Grid Handbook”, 1st. Ed. John Wiley & Sons Ltd. ISBN: 9781118755471.
- [10] T. Kostic. 2003. “Asset management in electrical utilities: how many facets it actually has”, Power Engineering Society General Meeting, 2003, IEEE, vol.1, pp.275- 281.
- [11] ISO 55000-2014: Asset Management - Overview, principles and terminology.
- [12] M. Shahidehpour & R. Ferrero. 2005. “Time management for assets: chronological strategies for power system asset management”, IEEE Power and Energy Magazine, vol. 3, no. 3, pp. 32–38.
- [13] O. Tor & M. Shahidehpour, 2006. “Power Distribution Asset Management”, IEEE Power Engineering Society General Meeting, pp. 2-4.

- [14] S. R. Khuntia<sup>a</sup>; J. L. Rueda<sup>a</sup> & S. Bouwman<sup>b</sup>. 2015. “Classification, Domains and Risk Assessment in Asset Management: A Literature Study,” 50th International Universities Power Engineering Conference (UPEC). Stoke on Trent. pp. 1-5.
- [15] R. Brown & B. Humphrey. 2005. “Asset management for transmission and distribution,” IEEE Power and Energy Magazine, vol. 3, no. 3, pp. 39–45.
- [16] J. Smith. 2005. “Trends in emerging technologies in power systems”. International Conference on Future Power Systems, Amsterdam, 2005, pp. 7 pp.-7.
- [17] R. Jones. 1995. “Risk-based management: A reliability centered approach. In: Risk-based management: A reliability centered approach.” Gulf Publishing Company. New York, USA. ISBN-13:978-0-88415-785-4
- [18] J. Nilsson & L. Bertling. 2007. “Maintenance management of wind power systems using condition monitoring systems-Life cycle cost analysis for two case studies,” IEEE Trans. Energy Conversion, vol. 22, no. 1, pp. 223-229.
- [19] D. Nordgård, et al. 2007. “A Risk Based Approach to Distribution System Asset Management and a Survey of Perceived Risk Exposure among Distribution Companies”, the 19th International Conference on Electricity Distribution (CIRED), Vienna, pp. 21-24.
- [20] O. Wilhelmsson. 2005. “Evaluation of the introduction of RCM for hydro power generators at Vattenfall Vattenkraft”, Master of Science Thesis, KTH, Stockholm, Sweden.
- [21] ISO 31000:2009, Risk management - Principles and guidelines.
- [22] R. Mehairjan, 2016. “Risk-Based Maintenance for Electricity Networks Organizations” Master of Science in Electrical Engineering, University of Technology, Paramaribo, Suriname,
- [23] L. Pintelon & F. Van Puyvelde. 2013. “Asset Management - The Maintenance Perspective”. Ed. Acco, Leuven, Belgium, ISBN- 9789033493447.
- [24] G. Bajracharya. 2014. “Multi-Agent Model-Based Optimization for Future Electrical Grids”. PhD Thesis, University of Technology, Paramaribo, Suriname.
- [25] H. Willis & R. Schrieber. 2017. “Aging Power Delivery Infrastructures”, CRC Press, Boca Raton, FL, USA. ISBN-1439896917
- [26] G. Balzer & C. Schorn, 2015. “Asset Management for Infrastructure Systems”. Springer International Publishing. Switzerland. ISBN 978-3-319-17879-0.



- [27] J. Moubray. 1997. “Reliability- Centered Maintenance,” Industrial Press Inc. New York, USA. 2<sup>nd</sup> Edition. ISBN-0-8311-3078-4.
- [28] B. S. Hauge & D. Johnston. 2001. “Reliability and risk treatment centered maintenance”. In Proceedings of IEEE Annual Reliability and Maintainability Symposium, Cape Canaveral, FL, USA.
- [29] A. Pharmatrisanti. 2012. “Long Term Performance of Gas-Insulated Switchgear Operating under Tropical Conditions”, PhD Thesis, University of Technology, Paramaribo, Suriname.
- [30] Q. Zhuang, 2015. “Managing Risks in Electrical Infrastructure Assets from a Strategic Perspective”. PhD Thesis, University of Technology, Paramaribo, Suriname.
- [31] S. Fernández, 1999. “Diagnostico táctica y estrategia”. Revista Ingeniería Energética del ISPJAE XX, Habana, Cuba, vol. 4, pp. 718.
- [32] IEEE Standard Conformance Test Procedure for Instrument Transformers. 2005. In: IEEE Std C57.13.2-2005, vol., no., and pp.0\_1-6.
- [33] R. Pérez, H. Torres & S. Fernández. 2012. “Sistema de Monitoreo en Tiempo Real para el Diagnóstico”. 10th Latin American and Caribbean Conference for Engineering and Technology. Panamá City, Panamá.
- [34] O. Ramírez, 2001. “Modelación fisicomatemática para el diagnóstico de grandes transformadores en tiempo real”. Phd Tesis. CIPEL-CUJAE, Habana, Cuba.
- [35] R. Liñán. 2001. “Experiencias en el desarrollo de sistemas de monitoreo y diagnóstico para transformadores de potencia”. Instituto de Investigaciones Eléctricas. Biental CIGRÉ-México, paper 12-02.
- [36] C. Neumann. 2006. “The impact of insulation monitoring and diagnostics on reliability and exploitation of service life”. Memorias Congreso de la CIGRE, Paris, Francia, paper C4-201.
- [37] R. Pérez. 2009. “Modelación térmica para el diagnóstico del aislamiento principal en transformadores de potencia”. Phd Tesis. CIPEL-CUJAE, Habana, Cuba.
- [38] R. Ferreli & F. Milani. 2009. “Estadística De Fallas En Transformadores De Potencia. Implementación De Bobinas Limitadoras De Corriente De Cortocircuito”. Décimo Tercer Encuentro Regional Iberoamericano de CIGRÉ, Argentina. 2009.

- [39] S. Fernández. 2006. “Diagnóstico Integral Curso de la Especialidad de Diagnóstico Integral de Centrales y Subestaciones Eléctricas”. CIPEL-CUJAE, Habana, Cuba. 3<sup>er</sup> Edición.
- [40] IEC 60599:2015. Mineral oil-filled electrical equipment in service - Guidance on the interpretation of dissolved and free gases analysis.
- [41] A. Castro & L. Cárdenas. 2008. “Mantenimiento preventivo basado en la confiabilidad de los elementos que componen la red de distribución”, Congreso de Las Américas de Distribución Eléctrica, Mar de Plata, Argentina.
- [42] E. Acevedo & F. Saldivia. 2012. “Estrategias de mantenimiento predictivo basado en análisis de aceites dieléctricos en transformadores de potencia de Centrales Eléctricas de Norte de Santander S.A. E.S.P. Área Cúcuta”. Tesis de Maestría, Táchira, Venezuela.
- [43] P. Chambilla. “Técnicas de alta tensión, medición de las descargas parciales”. Escuela Profesional de Ingeniería Eléctrica. disponible en: <http://es.scribd.com/doc/62939759/Descargas-Parciales>. Consultado el 1 de febrero de 2018.
- [44] L. Pettersson. 1990. “Estimation of the remaining service life of power transformers and their insulation”, *Electra*, no.133, pp.65-71.
- [45] M. Castillo. 2010. “Análisis modal y sus efectos en los generadores de vapor de centrales azucareros”, Ediciones UNEXPO Barquisimeto, Venezuela.
- [46] S, Frederick & G. Lieberman, “Investigación de Operaciones,” McGraw-Hill. 1997.
- [47] A. Adoghe, & J.C. Ekeh. 2010. “A Márkov model for estimating there maining life of electrical insulation in distribution transformer”, *American journal of scientific and industrial research* ©, Science Huß, <http://www.scihub.org/AJSIR> ISSN: 2153-649X doi:10.5251/ajsir.2010.1.3.539.548
- [48] M. Yahaya, et al. 2017. “Estimation of Transformers Health Index Based on the Markov Chain”, *Energies* 2017, vol. 10, 1824.
- [49] M. Sóbol, “Método de Montecarlo,” Editorial MIR Moscú 1987
- [50] R. Ramírez. 2010. “Desarrollo de modelos para el diseño del combustible de un reactor nuclear rápido enfriado por gas”. PhD Tesis. Universidad Nacional Autónoma de México, CDMX, México.

- [51] H. Moon, et al. 2011. “A Study Using a Montecarlo Method of the Optimal Configuration of a Distribution Network in Terms of Power Loss Sensing, Sensors”. Published by MDPI AG, Basel, Switzerland. 11, 7823-7834.
- [52] H. Liu & Y. Gu. 2013. “A Monte-Carlo Simulation Method for Industry Transformer Health Prediction Based on Dissolved Gas Analysis,” International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering, School of Chemistry, Physics and Mechanical Engineering Queensland University of Technology Brisbane, Australia.
- [53] T. M. Mitchell, 1997. “Machine learning”. WCB/McGraw-Hill, vol. 8.
- [54] S. Ryszard & T. Mitchell. 2013. “Machine learning: An artificial intelligence approach,” Springer Science & Business Media. ISBN-13: 978-1493303489.
- [55] P. Simon. 2013. Too Big to Ignore: “The Business Case for Big Data,” John Wiley & Sons, SAS Institute Inc., Cary, North Carolina, USA.
- [56] A. Refenes. 1997. “Financial time series modelling with discounted least squares back propagation”. Elsevier, in: Neurocomputing, vol. 14.2, pp. 123-138.
- [57] M. Olden. 2016 “Predicting Stocks with Machine Learning”. Master’s Thesis, Department of Informatics, University of Oslo. Noruega.
- [58] E. Rich & K. Knight. 1994. “Inteligencia Artificial”, 2<sup>nd</sup> Edición, Editorial McGraw- Hill, ISBN-0-07-052263-4.
- [59] J. T. Palma & R. Marín. 2008. “Inteligencia Artificial. Técnicas, métodos y aplicaciones”. Editorial McGraw- Hill/Interamericana de España.
- [60] E. Formisano & G. Valente. 2008. “Multivariate analysis of fMRI time series: classification and regression of brain responses using machine learning”. In: Magnetic resonance imaging, vol. 26, pp.345-354
- [61] M. Leung & A. Chen. 2000. “Forecasting stock indices: a comparison of classification and level estimation models”. In: International Journal of Forecasting vol. 16, pp. 173-190.
- [62] D. M. Hawkins. 2004. “The problem of overfitting,” In: Journal of chemical information and computer sciences, vol. 44, pp. 1-12.
- [63] Y. Bengio & Y. Grandvalet. 2004. “No unbiased estimator of the variance of k-fold cross-validation”. In: The Journal of Machine Learning Research 5, pp. 1089-1105.

- [64] N. K. Ahmed. 2010. “An Empirical Comparison of Machine Learning Models for Time Series Forecasting Econometric Reviews”. Taylor & Francis Journals, vol. 29(5-6), pp. 594-621.
- [65] G. Bontempi & Y. L. Borgne. 2013. “Machine learning strategies for time series forecasting,” In: Business Intelligence. Springer, pp. 62-77.
- [66] R. Bellman. 1957. “A Markovian decision process”. In: Indiana Univ. Math. J., vol. 6, pp. 679-684
- [67] E. Keogh & A. Mueen. 2011. “Curse of dimensionality”. In: Encyclopedia of Machine Learning. Springer, pp. 257-258.
- [68] I. Guyon & A. Elisseeff. 2003. “An introduction to variable and feature selection,” In: The Journal of Machine Learning Research 3, pp. 1157-1182.
- [69] D. H. Wolpert & W. G. Macready, “No free lunch theorems for optimization,” In: Evolutionary Computation, IEEE Transactions on 1.1 (1997), pp. 67-82.
- [70] D. H. Wolpert, “The supervised learning no-free-lunch theorems”. In: Soft Computing and Industry. Springer, 2002, pp. 25-42.
- [71] cdipaolo. Perceptron. URL: <https://github.com/cdipaolo/goml/tree/master/perceptron>.
- [72] J. Schmidhuber. 2015. “Deep Learning in Neural Networks: An Overview”. In: Neural Networks, vol: 61, pp. 85-117.
- [73] D. Brezak, et al. 2012. “A comparison of *Feed-Forward* and recurrent neural networks in time series forecasting”. IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr), New York, NY, pp. 1-6.
- [74] M. Salmasi, H. Mahdavi-Nasab and H. Pourghassem. 2011. “Comparison of *Feed-Forward* and Recurrent Neural Networks in Active Cancellation of Sound Noise”. 2011 International Conference on Multimedia and Signal Processing, Guilin, Guangxi, 2011, pp. 25-29.
- [75] M. Sundermeyer, et al. 2013. “Comparison of *Feed-Forward* and recurrent neural network language models”. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, 2013, pp. 8430-8434.
- [76] S. S. Haykin. 2009. “Neural networks and learning machines”. 3. Ed. Upper Saddle River: Pearson Education.

- [77] M. T. Hagan & M. B. Menhaj. 1994. “Training *Feed-Forward* networks with the Marquardt algorithm”. In: *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989-993.
- [78] H. William & Brian P. 2011. “Section 16.5. Support Vector Machines”. *Numerical Recipes: The Art of Scientific Computing* (3rd Ed.). New York: Cambridge University Press. ISBN 978-0-521-88068-8. Archived from the original on 2011-08-11.
- [79] S. Marsland. 2009. “Machine Learning: An Algorithmic Perspective”. (1st Ed.). Chapman & Hall/CRC.
- [80] R. Herbrich & C. Campbell. 1999. “Bayes point machines: Estimating the bayes point in kernel space”. In: 2nd International Joint Conference on Artificial Intelligence (*IJCAI*) Workshop SVMs, pp. 23-27.
- [81] Y. Freund & N. Abe. 1999. “A short introduction to boosting”. In: *Journal-Japanese Society for Artificial Intelligence*, vol. 14, pp. 771-780.
- [82] L. Breiman. 1996. “Bagging predictors”. In: *Machine learning* 24.2, pp. 123-140.
- [83] D. H. Wolpert. 1992. “Stacked generalization”. In: *Neural networks* 5.2, pp. 241-259.
- [84] L. Rokach & O. Maimon. 2014. “Data mining with decision trees: theory and applications”. World Scientific Publishing Co Inc. 2<sup>nd</sup> Ed. ISBN 978-9812771711.
- [85] J. R. Quinlan. 1986. “Induction of decision trees”. In: *Machine learning* 1.1, pp. 81-106.
- [86] J. Shotton. 2013. “Decision jungles: Compact and rich models for classification”, in: *Advances in Neural Information Processing Systems*, pp. 234-242.
- [87] M. Price & A. P. Arkin. 2009. “Fast Tree: computing large minimum evolution trees with profiles instead of a distance matrix”. In: *Molecular biology and evolution*, pp. 1641-1650.
- [88] P. Bangalore & L. Bertling. 2015. “An Artificial Neural Network Approach for Early Fault Detection of Gearbox Bearings”, *IEEE transactions on smart grid*, vol. 6, No. 2.
- [89] R. Álvarez. 2011. “Estudio de correlación de las variables significativas analizadas en el censo económico 2010”, *Retos* 2. Julio/diciembre. Pp.19-38.
- [90] S. Chen & S. A. Billings, P. M. Grant. 1990. “Non-linear system identification using neural networks”. In: *International Journal of Control*, vol. 51, pp. 1191-1214.

- [91] H. T. Su & T. J. McAvoy. 1991. "Identification of chemical processes using recurrent networks". In: Proceedings of Proceedings of the 1991 American Control Conference, Boston, pp. 2314-2319.
- [92] P.C. Mahalanobis. 1936. "On the generalised distance in statistics", Proceedings of the National Institute of Science of India, vol. 12, pp. 49-55.
- [93] G. Niu & M. Pecht. 2011. "Health monitoring of electronic products based on Mahalanobis distance and Weibull decision metrics". In: Microelectronics Reliability, vol. 51, pp. 279-284,
- [94] Y. Wang & M. G. Pecht. 2013. "Online anomaly detection for hard disk drives based on Mahalanobis distance". In: IEEE Transactions on Reliability, vol. 62, pp. 136-145,
- [95] M. C. Garcia & J. del Pico. 2006. "SIMAP: Intelligent System for Predictive Maintenance. Application to the health condition monitoring of a wind turbine gearbox". In: Computers in Industry, vol. 57, pp. 552-568.
- [96] A. Zaher, & Y. Patel. 2009. "Online wind turbine fault detection through automated SCADA data analysis". In: Wind Energy, vol. 12, pp. 574-593.
- [97] K. Kim, et al. 2011 "Use of SCADA Data for Failure Detection in Wind Turbines". In: Proceedings of 5th International Conference on Energy Sustainability, Washington D.C, pp. 2071-2079.
- [98] J. Xiang, S. Watson & Y. Liu. 2009. "Smart monitoring of wind turbines using neural networks", in Proceedings of International Conference on Sustainability in Energy and Buildings, Brighton, pp. 1-8.
- [99] R. F. Mesquita & F. MacIel. 2012. "Forecast of faults in a wind turbine gearbox", in Proceeding of 9th International Conference on ELEKTRO 2012, Zilina - Rajecke Teplice; Slovakia, pp. 170-173.
- [100] A. Kusiak & A. Verma. 2012. "Analyzing bearing faults in wind turbines: A data-mining approach", Renewable Energy, vol. 48, pp. 110-116.

## APÉNDICES

### Apéndice A: Matriz de correlaciones parciales para el Transformador T1

**Tabla A.1. Análisis de correlación de variables del conjunto de datos entrenamiento para transformador T1.**

	<i>desb</i>	<i>desbv</i>	<i>F</i>	<i>fp</i>	<i>ia</i>
<i>desb</i>	1	-0.015219522	-0.012495073	0.012844543	-0.513270114
<i>desbv</i>	-0.01521952	1	-0.988495569	-0.007519025	0.003995801
<i>f</i>	-0.01249507	-0.988495569	1	0.002105656	0.022939785
<i>fp</i>	0.01284454	-0.007519025	0.002105656	1	0.075242267
<i>ia</i>	-0.51327011	0.003995801	0.022939785	0.075242267	1
<i>iaprom</i>	-0.51486533	0.00405828	0.023698126	0.074110767	0.999909483
<i>ib</i>	-0.50864791	0.004826118	0.023382172	0.07430414	0.9997216
<i>ic</i>	-0.52267048	0.003339958	0.024776416	0.072760162	0.999783359
<i>kvaran</i>	-0.48770203	-0.017311722	-0.013711006	0.044161122	0.804666535
<i>kvarbn</i>	-0.44941269	-0.020165738	-0.011744311	0.053045324	0.810205889
<i>kvarcn</i>	-0.47192823	-0.015342052	-0.012418149	0.047123209	0.828396309
<i>kvartot</i>	-0.46981389	-0.017647675	-0.012624808	0.048196901	0.81496466
<i>kwan</i>	-0.5080954	0.00311567	0.026863401	0.079821043	0.998425733
<i>kwbm</i>	-0.5080954	0.00311567	0.026863401	0.079821043	0.998425733
<i>kwen</i>	-0.51966664	0.004495748	0.026940818	0.077336304	0.997837073
<i>kwtot</i>	-0.5108685	0.004320714	0.026926156	0.078638258	0.997951941
<i>TIOILM</i>	-0.21688832	0.066554174	-0.047417412	0.009087656	0.932644946
<i>thdia</i>	-0.42332795	0.015507954	0.056219346	-0.094820054	0.315869528
<i>thdib</i>	-0.3931165	0.008430584	0.062013931	-0.096435558	0.24341548
<i>thdic</i>	-0.41564541	0.015782244	0.057078634	-0.103288889	0.31357546
<i>thdvan</i>	-0.53151661	0.039269477	0.01275978	-0.036510975	0.675666104
<i>thdvbn</i>	-0.50750829	0.006458989	0.039388787	-0.042829582	0.620775407
<i>thdvcn</i>	-0.50424351	0.033764463	0.014421015	-0.050752176	0.641708376
<i>van</i>	0.29220561	-0.466298084	0.476233974	-0.013823186	-0.564915409
<i>vbn</i>	0.27838831	-0.455844944	0.475011279	-0.014728134	-0.559909721
<i>vcn</i>	0.28882637	-0.460702061	0.471465952	-0.011558954	-0.581375476
<i>vprom</i>	0.28651775	-0.461706699	0.475021578	-0.013372488	-0.568962773
	<i>iaprom</i>	<i>ib</i>	<i>ic</i>	<i>kvaran</i>	<i>kvarbn</i>
<i>desb</i>	-0.51486533	-0.508647915	-0.522670475	-0.48770203	-0.44941269
<i>desbv</i>	0.00405828	0.004826118	0.003339958	-0.01731172	-0.02016574
<i>f</i>	0.02369813	0.023382172	0.024776416	-0.01371101	-0.01174431
<i>fp</i>	0.07411077	0.07430414	0.072760162	0.04416112	0.05304532
<i>ia</i>	0.99990948	0.9997216	0.999783359	0.80466653	0.81020589



<i>iaprom</i>	1	0.999923313	0.99994292	0.8009004	0.80626467
<i>ib</i>	0.99992331	1	0.999822228	0.79741695	0.80313852
<i>ic</i>	0.99994292	0.999822228	1	0.8004713	0.80529522
<i>kvaran</i>	0.8009004	0.79741695	0.8004713	1	0.99842265
<i>kvarbn</i>	0.80626467	0.803138523	0.805295224	0.99842265	1
<i>kvarcn</i>	0.8252984	0.822630365	0.824707213	0.9973475	0.99762846
<i>kvartot</i>	0.81135636	0.808260853	0.810689403	0.99932079	0.99942969
<i>kwan</i>	0.99858478	0.998601096	0.998502465	0.7796154	0.7854443
<i>kwbn</i>	0.99858478	0.998601096	0.998502465	0.7796154	0.7854443
<i>kwcen</i>	0.99825607	0.998321678	0.998385166	0.7742153	0.77919961
<i>kwtot</i>	0.99832099	0.998454399	0.998330952	0.77388511	0.77945341
<i>TIOILM</i>	0.92706194	0.934595846	0.933517921	0.53210733	0.53583331
<i>thdia</i>	0.31622129	0.314247014	0.318509497	0.09297207	0.08308018
<i>thdib</i>	0.24359236	0.241414187	0.245928271	0.03165777	0.02109508
<i>thdic</i>	0.31374504	0.311718644	0.315903523	0.09555585	0.08594828
<i>thdvan</i>	0.67501451	0.672869079	0.676386423	0.46254723	0.45730395
<i>thdvbn</i>	0.61984422	0.617451598	0.621198521	0.4314637	0.42594528
<i>thdvcn</i>	0.64103979	0.638922538	0.642373609	0.4390218	0.43347621
<i>van</i>	-0.56234945	-0.561140415	-0.560869964	-0.62307706	-0.62730987
<i>vbn</i>	-0.55697824	-0.555595133	-0.555309301	-0.63909633	-0.64380757
<i>vcn</i>	-0.57868477	-0.577464601	-0.577087468	-0.63744345	-0.64198327
<i>vprom</i>	-0.56623213	-0.564961806	-0.56464844	-0.63344218	-0.63794027
	<i>kvarcn</i>	<i>kvartot</i>	<i>kwan</i>	<i>kwbn</i>	<i>kwcen</i>
<i>desb</i>	-0.47192823	-0.46981389	-0.5080954	-0.5080954	-0.519666636
<i>desbv</i>	-0.01534205	-0.01764767	0.00311567	0.00311567	0.004495748
<i>f</i>	-0.01241815	-0.01262481	0.0268634	0.0268634	0.026940818
<i>fp</i>	0.04712321	0.0481969	0.07982104	0.07982104	0.077336304
<i>ia</i>	0.82839631	0.81496466	0.99842573	0.99842573	0.997837073
<i>iaprom</i>	0.8252984	0.81135636	0.99858478	0.99858478	0.998256071
<i>ib</i>	0.82263036	0.80826085	0.9986011	0.9986011	0.998321678
<i>ic</i>	0.82470721	0.8106894	0.99850246	0.99850246	0.998385166
<i>kvaran</i>	0.9973475	0.99932079	0.7796154	0.7796154	0.774215301
<i>kvarbn</i>	0.99762846	0.99942969	0.7854443	0.7854443	0.779199611
<i>kvarcn</i>	1	0.99904804	0.80472738	0.80472738	0.799742336
<i>kvartot</i>	0.99904804	1	0.7904494	0.7904494	0.784894372
<i>kwan</i>	0.80472738	0.7904494	1	1	0.999731275
<i>kwbn</i>	0.80472738	0.7904494	1	1	0.999731275
<i>kwcen</i>	0.79974234	0.78489437	0.99973128	0.99973128	1
<i>kwtot</i>	0.79970921	0.78486083	0.99987174	0.99987174	0.999933474



<i>TIOILM</i>	0.54719492	0.53873957	0.91325083	0.92014627	0.916493163
<i>thdia</i>	0.09183674	0.08929548	0.32227482	0.32227482	0.326918534
<i>thdib</i>	0.02773277	0.02678824	0.25043998	0.25043998	0.254992579
<i>thdic</i>	0.09386567	0.09179515	0.31951421	0.31951421	0.323795839
<i>thdvan</i>	0.46851014	0.46306488	0.67692329	0.67692329	0.678681382
<i>thdvbn</i>	0.43475219	0.43098335	0.62226465	0.62226465	0.623624666
<i>thdvcn</i>	0.44428358	0.43918813	0.64298617	0.64298617	0.644554825
<i>van</i>	-0.62674681	-0.62618521	-0.5295482	-0.5295482	-0.526103801
<i>vbn</i>	-0.64162461	-0.64200302	-0.5238207	-0.5238207	-0.519604483
<i>vcn</i>	-0.64138758	-0.64075717	-0.54657316	-0.54657316	-0.542674083
<i>vprom</i>	-0.6368283	-0.63655483	-0.53353171	-0.53353171	-0.529676119
	<i>kwtot</i>	<i>TIOILM</i>	<i>thdia</i>	<i>thdib</i>	<i>thdic</i>
<i>desb</i>	-0.510868504	-0.216888324	-0.42332795	-0.393116498	-0.41564541
<i>desbv</i>	0.004320714	0.066554174	0.01550795	0.008430584	0.01578224
<i>f</i>	0.026926156	-0.047417412	0.05621935	0.062013931	0.05707863
<i>fp</i>	0.078638258	0.009087656	-0.09482005	-0.096435558	-0.10328889
<i>ia</i>	0.997951941	0.932645321	0.31586953	0.24341548	0.31357546
<i>iaprom</i>	0.998320992	0.932812367	0.31622129	0.243592355	0.31374504
<i>ib</i>	0.998454399	0.934596216	0.31424701	0.241414187	0.31171864
<i>ic</i>	0.998330952	0.933518791	0.3185095	0.245928271	0.31590352
<i>kvaran</i>	0.773885106	0.532107334	0.09297207	0.031657773	0.09555585
<i>kvarbn</i>	0.779453406	0.535833314	0.08308018	0.021095078	0.08594828
<i>kvarcn</i>	0.799709208	0.54719492	0.09183674	0.027732766	0.09386567
<i>kvartot</i>	0.784860831	0.538739575	0.08929548	0.026788236	0.09179515
<i>kwan</i>	0.99987174	0.913251326	0.32227482	0.250439978	0.31951421
<i>kwbvn</i>	0.99987174	0.920146127	0.32227482	0.250439978	0.31951421
<i>kvcn</i>	0.999933474	0.916493179	0.32691853	0.254992579	0.32379584
<i>kwtot</i>	1	0.916945233	0.32431777	0.25234617	0.32134958
<i>TIOILM</i>	0.915455357	1	-0.09893806	-0.148002279	-0.097781
<i>thdia</i>	0.324317772	-0.098938061	1	0.995101627	0.99652161
<i>thdib</i>	0.25234617	-0.148002279	0.99510163	1	0.99476222
<i>thdic</i>	0.321349577	-0.097780999	0.99652161	0.994762218	1
<i>thdvan</i>	0.677040156	0.129969496	0.84010682	0.811465856	0.85021507
<i>thdvbn</i>	0.621997831	0.083190906	0.83255463	0.812785426	0.84649133
<i>thdvcn</i>	0.643012295	0.109281211	0.80739593	0.784906743	0.82443409
<i>van</i>	-0.526442431	-0.265663612	-0.1254569	-0.072913884	-0.11973238
<i>vbn</i>	-0.520117858	-0.265105666	-0.09351459	-0.041410938	-0.08763322
<i>vcn</i>	-0.543220351	-0.282361939	-0.11828537	-0.06509498	-0.11308752
<i>vprom</i>	-0.530143136	-0.271241593	-0.11232626	-0.059687115	-0.10672008

	thdvan	thdvbn	thdvcn	van	vbn
<i>desb</i>	-0.53151661	-0.507508289	-0.50424351	0.29220561	0.27838831
<i>desbv</i>	0.03926948	0.006458989	0.03376446	-0.46629808	-0.45584494
<i>f</i>	0.01275978	0.039388787	0.01442101	0.47623397	0.47501128
<i>fp</i>	-0.03651098	-0.042829582	-0.05075218	-0.01382319	-0.01472813
<i>ia</i>	0.6756661	0.620775407	0.64170838	-0.56491541	-0.55990972
<i>iaprom</i>	0.67501451	0.619844223	0.64103979	-0.56234945	-0.55697824
<i>ib</i>	0.67286908	0.617451598	0.63892254	-0.56114042	-0.55559513
<i>ic</i>	0.67638642	0.621198521	0.64237361	-0.56086996	-0.5553093
<i>kvaran</i>	0.46254723	0.431463701	0.4390218	-0.62307706	-0.63909633
<i>kvarbn</i>	0.45730395	0.425945282	0.43347621	-0.62730987	-0.64380757
<i>kvarcn</i>	0.46851014	0.434752194	0.44428358	-0.62674681	-0.64162461
<i>kvartot</i>	0.46306488	0.430983351	0.43918813	-0.62618521	-0.64200302
<i>kwan</i>	0.67692329	0.622264647	0.64298617	-0.5295482	-0.5238207
<i>kwbvn</i>	0.67692329	0.622264647	0.64298617	-0.5295482	-0.5238207
<i>kvcn</i>	0.67868138	0.623624666	0.64455482	-0.5261038	-0.51960448
<i>kwtot</i>	0.67704016	0.621997831	0.6430123	-0.52644243	-0.52011786
<i>TIOILM</i>	0.1299695	0.083190906	0.10928121	-0.26566361	-0.26510567
<i>thdia</i>	0.84010682	0.832554633	0.80739593	-0.1254569	-0.09351459
<i>thdib</i>	0.81146586	0.812785426	0.78490674	-0.07291388	-0.04141094
<i>thdic</i>	0.85021507	0.846491333	0.82443409	-0.11973238	-0.08763322
<i>thdvan</i>	1	0.991792061	0.99176401	-0.35012159	-0.33157238
<i>thdvbn</i>	0.99179206	1	0.99458118	-0.2902395	-0.27448102
<i>thdvcn</i>	0.99176401	0.994581181	1	-0.31397534	-0.29779436
<i>van</i>	-0.35012159	-0.290239499	-0.31397534	1	0.99772837
<i>vbn</i>	-0.33157238	-0.274481017	-0.29779436	0.99772837	1
<i>vcn</i>	-0.35674202	-0.296990395	-0.32199373	0.99838856	0.99804527
<i>vprom</i>	-0.34621006	-0.287262475	-0.311313	0.99934957	0.9992436

## Apéndice B: Teoría de redes neuronales

### Definición de una red neuronal artificial

Las RNA se definen como sistemas de mapeos no lineales cuya estructura se basa en principios observados en los sistemas nerviosos humanos y animales. Constan de un número grande de procesadores simples ligados por conexiones con pesos. Las unidades de procesamiento se denominan neuronas. Cada unidad recibe entradas de otros nodos y genera una salida escalar simple que depende de la información local disponible guardada internamente o que llega a través de las conexiones con pesos. Pueden realizarse muchas funciones complejas dependiendo de las conexiones. En la Figura B.1 se muestra la analogía de la neurona biológica con la neurona artificial.

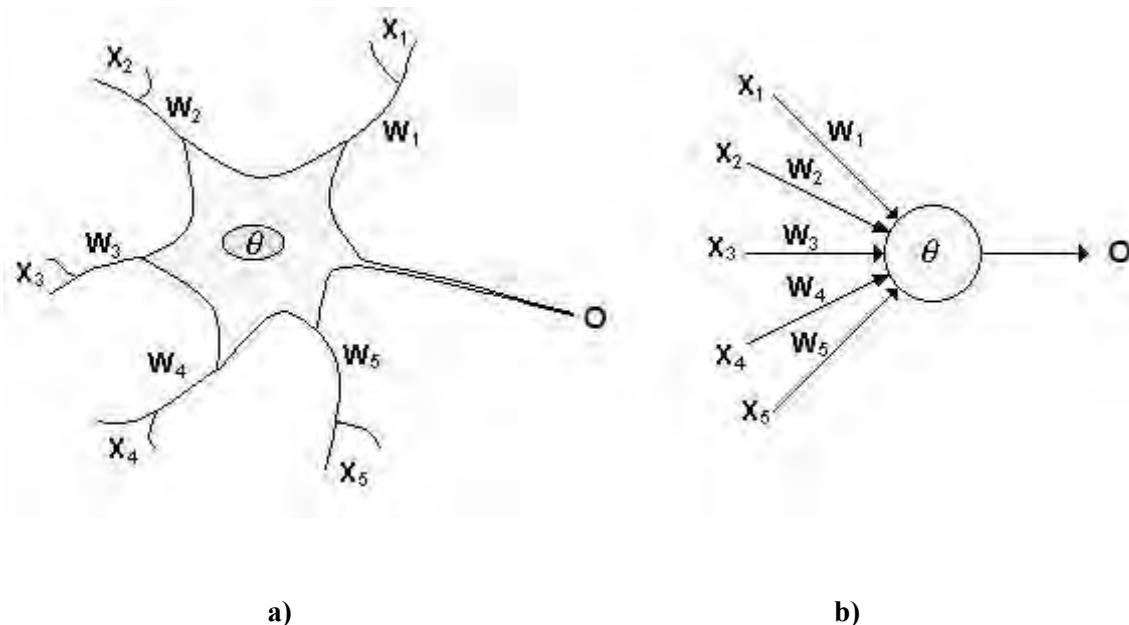


Figura B.1. a) Neurona biológica, b) Neurona artificial

De la observación detallada del proceso biológico se han hallado las siguientes analogías con el sistema artificial:

- Las entradas  $X_i$  representan las señales que provienen de otras neuronas y que son capturadas por las dendritas.
- Los pesos  $w_i$  son la intensidad de la sinapsis que conecta dos neuronas; tanto  $X_i$  como  $w_i$  son valores reales.
- $\theta$  es la función umbral que la neurona debe sobrepasar para activarse; este proceso ocurre biológicamente en el cuerpo de la célula.
- $O$  es la salida.

## Elementos que componen una red neuronal artificial

Las redes neuronales están constituidas por neuronas interconectadas y arregladas en capas. Los datos ingresan por medio de la “*capa de entrada*”, pasan a través de la “*capa oculta*” y salen por la “*capa de salida*”. Cabe mencionar que la capa oculta puede estar constituida por varias capas como lo podemos ver en la Figura B.2.

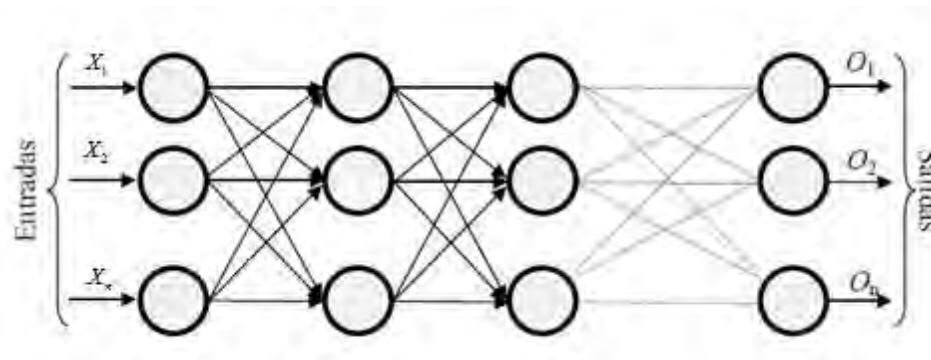


Figura B.2. Estructura de una red neuronal multicapa

## La neurona

Cada modelo de neurona consta de un elemento de procesamiento con conexiones sinápticas de entrada y una salida simple. El flujo de las señales de entrada de las neuronas se considera unidireccional. El símbolo general de una neurona se muestra en la Figura B.3. En esta representación se muestra un conjunto de pesos y el nodo.

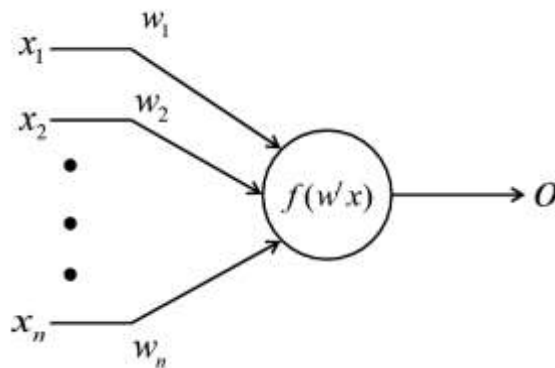


Figura B.3. Estructura general de una neurona

La señal de salida de la neurona está dada por la siguiente relación:

$$o = f(w^t x) \quad (B.1)$$

O bien

$$o = f\left(\sum_{i=1}^n w_i x_i\right) \quad (\text{B.2})$$

Donde  $w$  es el vector de pesos definido como:

$$w = [w_1 w_2 \dots w_n]^T \quad (\text{B.3})$$

Y  $x$  es el vector de entradas:

$$x = [x_1 x_2 \dots x_n]^T \quad (\text{B.4})$$

### La función de activación

La función  $f(w^t x)$  se refiere a una función de activación. La variable  $net$  se define como el producto escalar de los vectores de pesos y entradas:

$$net = w^t x \quad (\text{B.5})$$

La estructura de neurona mostrada en la Figura B.3 y descrita por las ecuaciones (B.1) y (B.5) es la más conocida en la literatura de redes neuronales, sin embargo, diferentes clases de redes neuronales artificiales hacen uso de diferentes definiciones de  $f(net)$ . Se nota en (B.1) que la neurona es como un nodo de procesamiento que realiza la operación de suma de sus entradas, o el producto escalar para obtener  $net$ . Posteriormente, se realiza la operación no lineal  $f(net)$  a través de su función de activación. Las funciones típicas de activación utilizadas son:

$$f(net) = \frac{2}{1 + e^{(-\lambda net)}} - 1 \quad (\text{B.6})$$

Y también,

$$f(net) = \text{sgn}(net) = \begin{cases} +1, & net > 0 \\ -1, & net < 0 \end{cases} \quad (\text{B.7})$$

Donde  $\lambda > 0$  en (B.6) es proporcional a la ganancia de la neurona determinado la pendiente de la función continua  $f(net)$  cercana a  $net = 0$ . La función de activación continua es mostrada en la Figura B.4 para varios  $\lambda$ . Se puede notar que mientras  $\lambda \rightarrow \infty$ , el límite de la función continua se convierte en la  $\text{sgn}(net)$ , la cual está definida en (B.7).

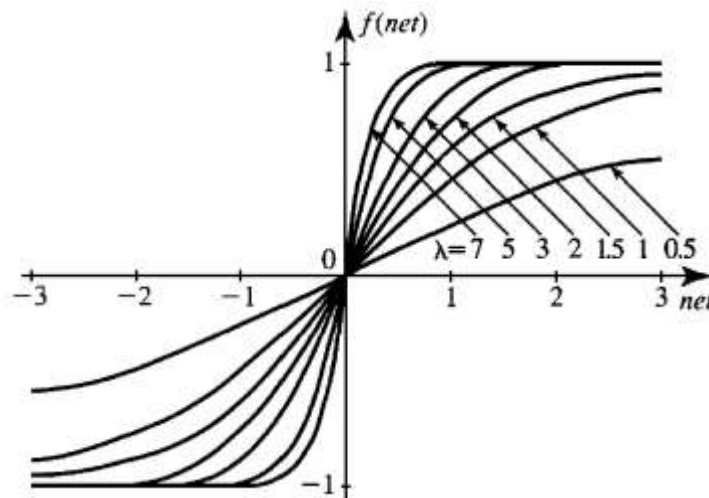
Estas funciones de activación son llamadas, función continua bipolar (B.6) y función binaria bipolar (B.7). La palabra "bipolar" se utiliza para indicar que las respuestas positivas y negativas de las neuronas son producidas por esta definición de la función de activación.

Al desplazar y ampliar las funciones de activación bipolares definidas por (B.6) y (B.7), la función de activación continua unipolar y la función de activación binaria unipolar pueden ser obtenidas respectivamente de la siguiente manera:

$$f(net) = \frac{1}{1+e^{(-\lambda net)}} \quad (B.8)$$

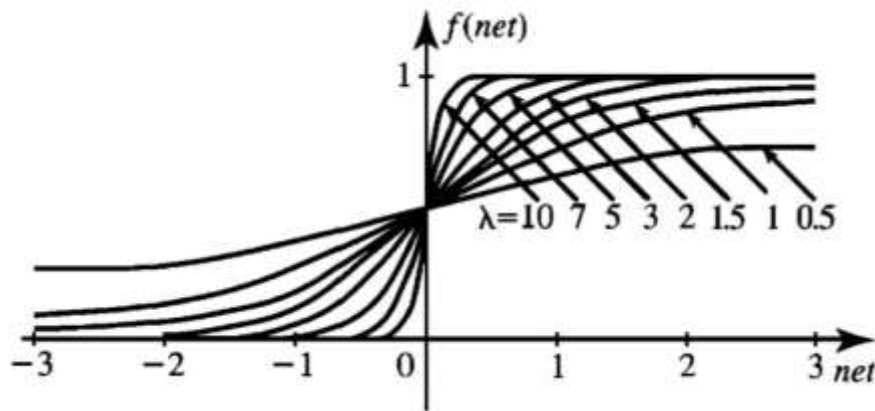
Y también,

$$f(net) = \begin{cases} 1, & net > 0 \\ 0, & net < 0 \end{cases} \quad (B.9)$$



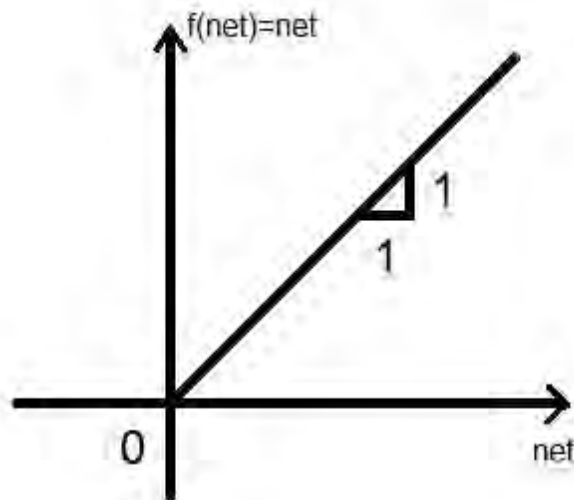
**Figura B.4. Función de activación continua bipolar**

La función (B.8) se muestra en la Figura B.5. Una vez más, la función binaria unipolar es el límite de la  $f(net)$  en (B.8) cuando  $\lambda \rightarrow \infty$ . A las funciones de activación de limitación suave (B.6) y (B.8) se les conoce como características sigmoidales, en contraparte de las funciones de activación de limitación dura dadas en (B.7) y (B.9). Las funciones de activación de limitación dura describen el modelo de neuronas discreto.



**Figura B.5. Función de activación continua unipolar**

La mayoría de las neuronas emplean funciones de activación bipolares. Algunas arquitecturas de redes neuronales o aplicaciones, sin embargo, requieren específicamente las respuestas de neuronas unipolares. Esencialmente, cualquier función  $f(net)$  que es monótonamente creciente y continua tal que  $net \in R$  y  $f(net) \in (-1,1)$  se puede utilizar en lugar de (B.6) para el modelado neuronal. Algunos modelos neuronales que a menudo implican alguna forma de realimentación requieren el uso de otro tipo de no linealidad que la definida en (B.6) a (B.9). Un ejemplo de función de activación rampa unipolar se muestra en la Figura B.6.



**Figura B.6. Función de activación rampa unipolar**

Si la función de activación de la neurona tiene la forma binaria bipolar de (B.7), el símbolo de la Figura B.3 puede ser sustituido por el diagrama mostrado en la Figura B.7, que en realidad es un diagrama a bloques de una neurona con función discreta que muestra la suma

realizada por el nodo de suma y el umbral de limitación dura, realizado por la unidad lógica de umbral (TLU). Este modelo consta de los pesos sinápticos, un nodo de suma, y el elemento TLU.

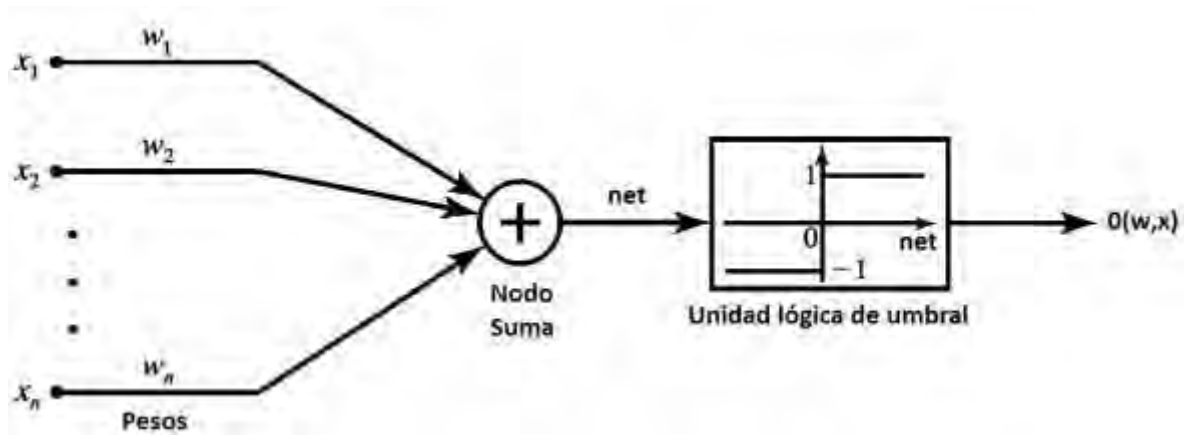


Figura B. 7. Perceptrón discreto (binario)

En el caso de la función de activación continua, como en (B.6), el modelo usado se muestra en la Figura. B.8, la neurona es representada como amplificador sumador de saturación de alta ganancia que amplía su señal de entrada  $w^T x$ . Los modelos de las Figuras B.7 y B.8 pueden ser llamados perceptrones discretos (binario) y perceptrones continuos, respectivamente. El perceptrón discreto, introducido por Rosenblatt (1958), fue la primera máquina de aprendizaje y puede ser considerado como un precursor de muchos de los modelos de red neuronal que se utilizan hoy. Además, su estudio proporciona una visión considerable en la naturaleza de los sistemas neurales artificiales.

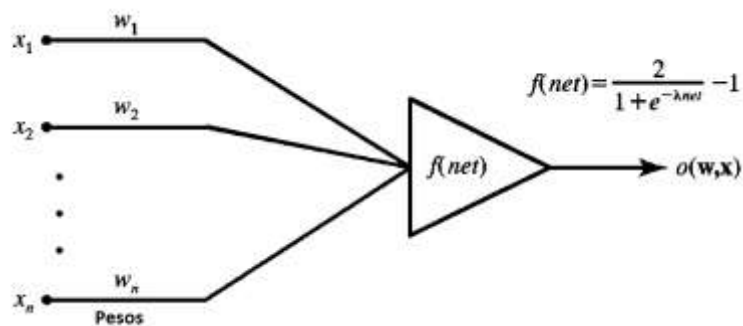


Figura B.8. Perceptrón continuo

## La salida



Como se vio anteriormente, las salidas de las neuronas son discretas o continuas. Dada una capa de  $m$  neuronas, sus valores de salida pueden estar dispuestos en un vector de salida  $o_1, o_2, \dots, o_n$  pueden estar dispuestos en un vector de salida de una capa:

$$o = [o_1 o_2 \dots o_n]^T \quad (\text{B.10})$$

Donde  $o_i$  es la señal de salida de la  $i^{\text{ésima}}$  neurona. El dominio de vectores  $o$  es definido en un espacio de dimensión  $m$  como el siguiente para  $i = 1, 2, \dots, m$ :

$$(-1,1)^m \equiv \{o \in R^m, o_i \in (-1,1)\} \quad (\text{B.11})$$

o bien

$$(0,1)^m \equiv \{o \in R^m, o_i \in (0,1)\} \quad (\text{B.12})$$

Para las funciones de activación continuas unipolares y bipolares definidas como en (B.6) y (B.8) respectivamente.

### **Aprendizaje y adaptación de redes neuronales artificiales**

El aprendizaje en los seres humanos y los animales es un proceso inferido, no es posible ver qué pasa directamente y se supone que se ha producido mediante la observación de los cambios en el rendimiento.

En redes neuronales artificiales se denomina aprendizaje al proceso de configuración de la red para que las entradas produzcan las salidas deseadas a través del fortalecimiento de las conexiones.

Una forma de llevar esto a cabo, es a partir del establecimiento de pesos conocidos con anterioridad, y el otro método implica el uso de técnicas de retroalimentación y patrones de aprendizaje que cambian los pesos hasta encontrar los adecuados.

Una red neuronal debe aprender a calcular la salida correcta para cada vector de entrada en el conjunto de datos. Este proceso de aprendizaje se denomina: proceso de entrenamiento o acondicionamiento. El conjunto de datos sobre el cual se basa este proceso es, por ende, llamado: conjunto de datos de entrenamiento.

La topología de la red y las diferentes funciones de cada neurona no pueden cambiar durante el aprendizaje, mientras que los pesos sobre cada una de las conexiones si pueden hacerlo; el aprendizaje de una red neuronal significa: adaptación de los pesos.

En otras palabras, el aprendizaje es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el mismo se reducen a la destrucción, modificación y creación de conexiones entre las neuronas. En los sistemas biológicos existe una continua destrucción y creación de conexiones entre las neuronas. En los modelos de redes neuronales artificiales, la creación de una nueva conexión implica que el peso de la misma pasa a tener un valor distinto de cero. De la misma manera, una conexión se destruye cuando su peso pasa a ser cero.

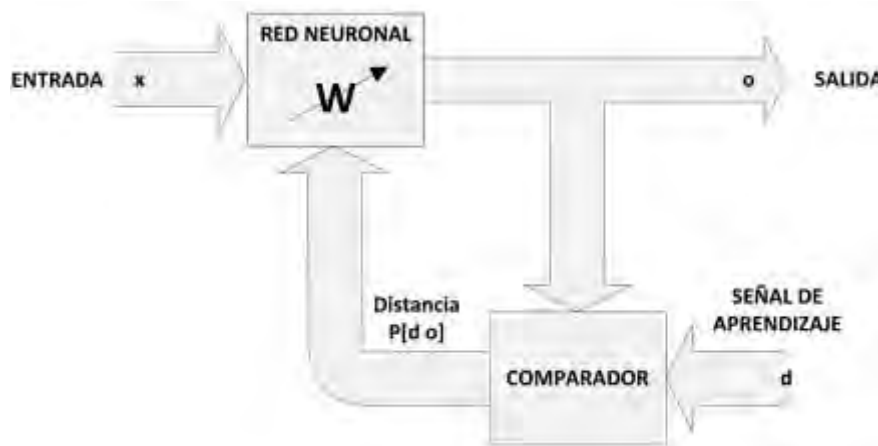
En general, existen dos tipos de aprendizaje en redes neuronales artificiales:

- Aprendizaje supervisado.
- Aprendizaje no supervisado.

### Aprendizaje supervisado

En el aprendizaje supervisado se asume que en cada instante de tiempo cuando la entrada es aplicada, la respuesta deseada del sistema es proporcionada por el maestro. Esto se ilustra en la Figura B.9, la distancia  $p[d, o]$  entre la respuesta obtenida y la respuesta deseada sirve como una medida de error y se utiliza para corregir los parámetros externos de la red.

Debido a que se asume que los pesos son ajustables, el maestro puede implementar un esquema de recompensa y castigo para adaptar los pesos de la matriz  $W$  en la red. Además, se utiliza en muchas situaciones de aprendizaje natural. Un conjunto de entrada y patrones de salida llamados un conjunto de entrenamiento es necesario para este modo de aprendizaje.



**Figura B.9. Diagrama a bloques del aprendizaje supervisado**

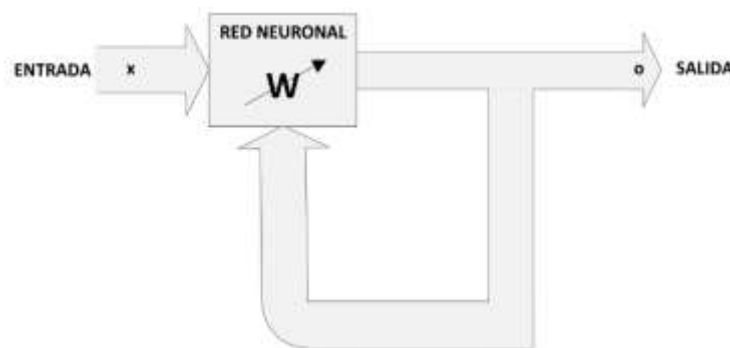
Por lo general, en el aprendizaje supervisado se premian las clasificaciones o asociaciones exactas y se castiga a aquellas que producen respuestas inexactas. El maestro estima el error

negativo de la dirección del gradiente y reduce el error correspondiente. En muchas situaciones, las entradas, las salidas y el cálculo del gradiente son deterministas, sin embargo, la minimización del error se realiza de manera aleatoria. Como resultado, la mayoría de los algoritmos de aprendizaje supervisado reducen a minimización estocástica del error en espacios multidimensionales de pesos.

### Aprendizaje no supervisado

En el aprendizaje sin supervisión, la respuesta deseada no se conoce, por lo tanto, la información de error explícito no se puede utilizar para mejorar el comportamiento de la red, dicho aprendizaje se muestra en la Figura B.10, dado que no se dispone de información en cuanto a la corrección o incorrección de las respuestas, el aprendizaje de alguna manera debe llevarse a cabo sobre la base de las observaciones de las respuestas a las entradas que se tienen.

En este tipo de aprendizaje, las redes deben encontrar las características, regularidades, correlaciones o categorías que se puedan establecer entre los datos que se presenten en su entrada. Existen varias posibilidades en cuanto a la interpretación de la salida de estas redes, que dependen de su estructura y del algoritmo de aprendizaje empleado.



**Figura B.10. Diagrama a bloques del aprendizaje no supervisado**

La técnica de aprendizaje no supervisado se utiliza a menudo para llevar a cabo la agrupación como la clasificación no supervisada de los objetos, sin proporcionar información acerca de las clases reales. Este tipo de aprendizaje responde con un mínimo de información *a priori* disponible.

### Modelos de redes neuronales artificiales

El modelo o arquitectura de una red neuronal consiste en la organización y disposición de las neuronas en la misma, formando capas o agrupaciones de neuronas alejadas de la entrada y salida de dicha red. En este sentido, los parámetros fundamentales de la red son: el número

de capas, el número de neuronas por capa, el grado de conectividad y el tipo de conexiones entre neuronas.

En las redes monocapa, se establecen conexiones entre las neuronas que pertenecen a la única capa que constituye la red. Las redes monocapas se utilizan generalmente en tareas relacionadas con lo que se conoce como auto-asociación (regenerar información de entrada que se presenta a la red de forma incompleta o distorsionada).

### **Redes multicapa**

Las redes multicapas son aquellas que disponen de un conjunto de neuronas agrupadas en varios (2, 3, etc.) niveles o capas. En estos casos, una forma para distinguir la capa a la que pertenece una neurona, consistiría en fijarse en el origen de las señales que recibe a la entrada y el destino de la señal de salida. Normalmente, todas las neuronas de una capa reciben señales de entrada desde otra capa anterior (la cual está más cerca a la entrada de la red), y envían señales de salida a una capa posterior (que está más cerca a la salida de la red). A estas conexiones se las denomina *conexiones hacia adelante o Feed-Forward*.

Sin embargo, en un gran número de estas redes también existe la posibilidad de conectar la salida de las neuronas de capas posteriores a la entrada de capas anteriores; a estas conexiones se las denomina *conexiones hacia atrás o Feed-Back*.

Estas dos posibilidades permiten distinguir entre dos tipos de redes con múltiples capas: las redes con conexiones hacia adelante o *redes Feed-Forward*, y las redes que disponen de conexiones tanto hacia adelante como hacia atrás o *redes Feed-Forward/Feed-Back*.

## Apéndice C: Regla delta generalizada

La arquitectura de la red de dos capas considerada a continuación se muestra en la Figura C.1, estrictamente hablando, dos capas del procesamiento de neuronas. Sin embargo, si se cuenta por las capas de nodos, la red también puede ser considerada como una red de tres capas. La  $i$ ésima columna de las señales se entiende en como una respuesta de una capa de salida inexistente (entrada) de la capa de neuronas. Por lo tanto, se puede referir a la arquitectura de la Figura C.1 como una red de dos capas de neuronas, o red de tres capas de nodos.

El término "capa", hace referencia al número real de capas neuronales existentes y procesadas. Por lo tanto, no se contará a las terminales de entrada como capas. Así, la red de la Figura C.1 es una red de dos capas. También debemos observar que una red de  $N$  capas tiene  $N - 1$  capas de neuronas cuyas salidas no son accesibles.

Las capas con neuronas cuyas salidas o resultados no son directamente accesibles se denominan capas internas u ocultas. Por lo tanto, todas excepto la capa de salida son las capas ocultas. Puesto que la salida de la capa  $j$  de las neuronas no es accesible desde la entrada y la salida la red de la Figura C.1 puede ser nombrada como la única red de capa oculta.

Los pesos calculados conducen ahora desde el nodo  $i$  hacia el nodo  $j$ , como se muestra en la Figura C.1. La fórmula del gradiente negativo para la capa oculta ahora se lee:

$$\Delta v_{ji} = \eta \frac{\partial E}{\partial v_{ji}}, \quad \text{Para } j = 1, 2, \dots, J \text{ y } i = 1, 2, \dots, I \quad (\text{C.1})$$

Y la fórmula de la componente del gradiente del error:

$$\frac{\partial E}{\partial v_{ji}} = \frac{\partial E}{\partial(\text{net}_j)} \frac{\partial(\text{net}_j)}{\partial v_{ji}} \quad (\text{C.2})$$

Se toma en cuenta que las entradas a la capa son  $z_i$  para  $i = 1, 2, \dots, I$ . El segundo término en el producto (C.2) es igual a  $z_i$  y podemos expresar el ajuste de peso como:

$$\Delta v_{ji} = \eta \delta_{yj} z_i \quad (\text{C.3})$$

Donde  $\delta_{yj}$  es el término para la señal de error de la capa oculta con salida  $y$ . Este término para la señal de error es producido por la  $j$ ésima neurona de la capa oculta, donde  $j = 1, 2, \dots, J$ . El término para la señal de error es igual a:

$$\delta_{yj} = -\frac{\partial E}{\partial(\text{net}_j)}, \quad \text{para } j = 1, 2, \dots, J \quad (\text{C.4})$$

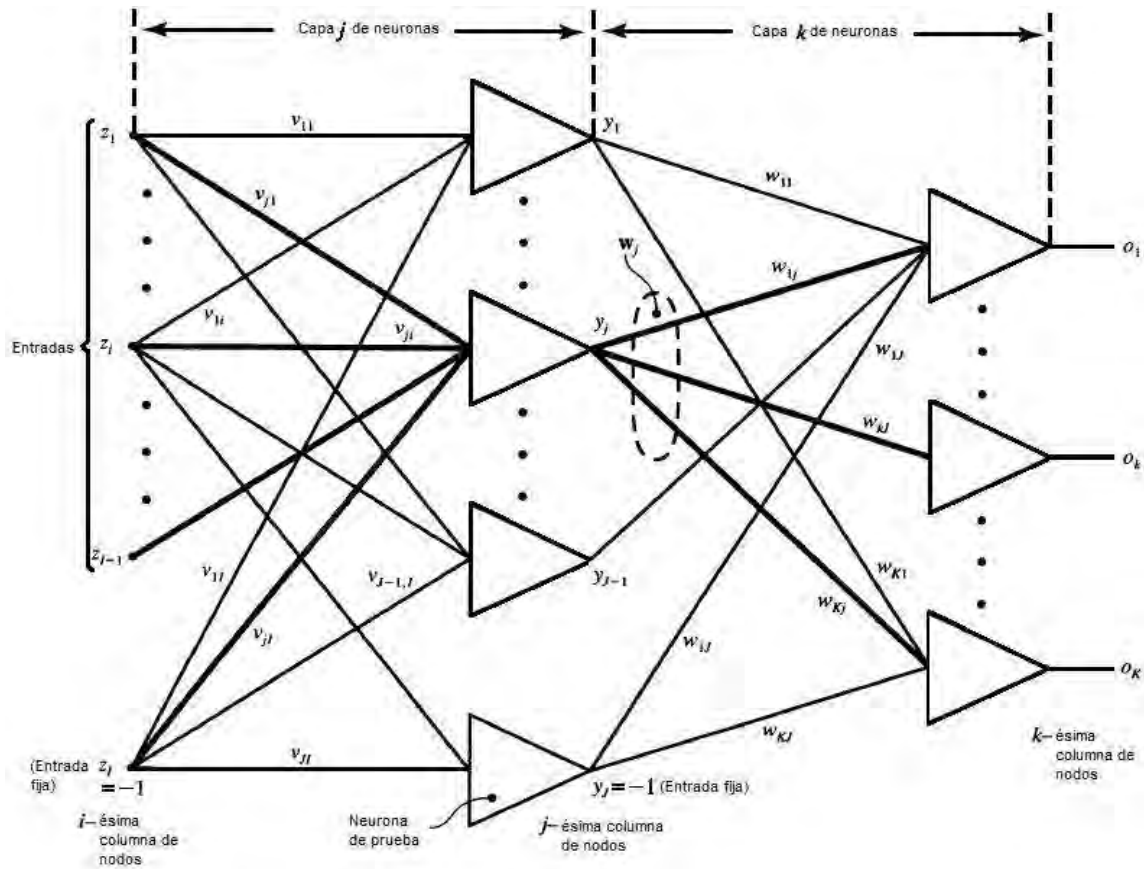


Figura C.1. Red multicapa con dos capas de perceptrones continuos

En contraste con la excitación de la capa de salida de las neuronas  $net_k$ , que afecta únicamente a la  $k$ -ésima neurona de salida, la  $net_j$ , contribuye ahora a cada componente de error en la suma de error que contiene los términos especificados en la capa  $k$ . El término para la señal de error  $\delta_{yj}$ , en el nodo  $j$  puede ser computado como sigue:

$$\delta_{yj} = -\frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial (net_j)} \quad (C.5)$$

Donde

$$\frac{\partial E}{\partial y_j} = \frac{\partial}{\partial y_j} \left( \frac{1}{2} \sum_{k=1}^K \{d_k - f[net_k(y)]\}^2 \right) \quad (C.6)$$

Y el segundo término de (C.5) es igual a

$$\frac{\partial y_j}{\partial (net_j)} = f_j'(net_j) \quad (C.7)$$

Cálculos de rutina de (C.6) dan como resultado

$$\frac{\partial E}{\partial y_j} = \sum_{k=1}^K (d_k - o_k) \frac{\partial}{\partial y_j} \{f[net_k(y)]\} \quad (C.8)$$

Cálculo de la derivada de la expresión entre llaves (C.8) se obtiene

$$\frac{\partial E}{\partial y_j} = \sum_{k=1}^K (d_k - o_k) f'(net_k) \frac{\partial (net_k)}{\partial y_j} \quad (C.9)$$

Podemos simplificar la expresión de arriba con la forma compacta que se presenta a continuación:

$$\frac{\partial E}{\partial y_j} = \sum_{k=1}^K \delta_{ok} w_{kj} \quad (C.10)$$

Combinando (C.7) y (C.10) se obtiene como resultado el reordenamiento  $\delta_{yj}$  expresado en (C.5) de la forma

$$\delta_{yj} = f_j'(net_j) \sum_{k=1}^K \delta_{ok} w_{kj}, \quad \text{para } j = 1, 2, \dots, J \quad (C.11)$$

El ajuste del peso (C.3) en la capa oculta ahora se convierte en

$$\Delta v_{ji} = \eta f_j'(net_j) z_i \sum_{k=1}^K \delta_{ok} w_{kj}, \quad \text{para } \begin{matrix} j = 1, 2, \dots, J \\ i = 1, 2, \dots, I \end{matrix} \quad (C.12)$$

Donde los términos  $f_j'(net_j)$  se van a calcular una simple regla delta de entrenamiento. La fórmula (C.12) expresa la llamada regla delta generalizada.

El ajuste de los pesos que conducen a la neurona  $j$  en la capa oculta es proporcional a la suma ponderada de todos los valores  $\delta$  en la capa adyacente siguiente de nodos conectando la neurona  $j$  con la salida. Los pesos que se abren en abanico desde el nodo  $j$  son por sí mismos los factores de ponderación. Los pesos que afectan  $\delta_{yj}$  de la  $j$ ésima neurona oculta se han destacado en la capa de salida de la Figura C.1. Todos los errores de las capas de salida  $\delta_{ok} w_{kj}$ , para  $k = 1, 2, \dots, K$  contribuyen al ajuste de pesos resaltado  $v_{ji}$  para  $i = 1, 2, \dots, I$  de la capa oculta. Los pesos modificados de la capa oculta se pueden expresar ahora como:

$$v'_{ji} = v_{ji} + \eta f_j'(net_j) z_i \sum_{k=1}^K \delta_{ok} w_{kj}, \quad \text{para } \begin{matrix} j = 1, 2, \dots, J \\ i = 1, 2, \dots, I \end{matrix} \quad (C.13)$$

El ajuste de peso de la capa oculta basado en el principio de la regla delta generalizada para la red de la Figura C.1 puede sucintamente ser declarado en notación vectorial como:

$$V' = V + \eta \delta_y z^T \quad (C.14)$$

Donde

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_I \end{bmatrix} \quad (C.15)$$

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1I} \\ v_{21} & v_{22} & \cdots & v_{2I} \\ \vdots & \vdots & \cdots & \vdots \\ v_{J1} & v_{J2} & \cdots & v_{JI} \end{bmatrix} \quad (C.16)$$

Y  $\delta_y$  es el vector de columna con entradas  $\delta_{yj}$  dada por (C.11). Definiendo ahora la columna  $j$ ésima de la matriz  $W$  como  $w_j$ , vector  $\delta_y$ , se puede expresar de la siguiente forma compacta

$$\delta_y = w_j^T \delta_o f'_{yj} \quad (C.17)$$

Donde  $f'_{yj}$  es el vector de columna con entradas  $f'_{yj}$ , expresado para cada una de las neuronas de capa la oculta  $1, 2, \dots, J$ , para funciones de activación unipolares y bipolares, respectivamente, como

$$f'_{yj} = y_j(1 - y_j) \quad (C.18)$$

$$f'_{yj} = \frac{1}{2} y_j(1 - y_j^2) \quad (C.19)$$

La regla de aprendizaje delta generalizada propaga el error hacia atrás por una capa, permitiendo que el mismo proceso se repita para cada capa anterior a la capa  $j$ . En la siguiente sección vamos a formalizar el método de entrenamiento para capas de redes neuronales.



## Apéndice D: Método de Levenberg-Marquardt para el entrenamiento de RNA

El algoritmo de retropropagación (EBP) ha sido una mejora significativa en la investigación de redes neuronales, pero tiene una tasa de convergencia débil.

Muchos esfuerzos se han hecho para acelerar el algoritmo EBP. Todos estos métodos conducen a resultados poco aceptables. El algoritmo Levenberg-Marquardt (LMA) se produjo por el desarrollo de métodos dependientes del algoritmo EBP.

Éste da un buen intercambio entre el algoritmo de velocidad de Newton y la estabilidad del método de pendiente descendente, éstos son dos teoremas básicos del algoritmo LMA. Se ha hecho un intento para acelerar el algoritmo LMA con la modificación del índice de desempeño modificado y el cálculo del gradiente, aunque es incapaz de reducir el error de oscilación. Se ha producido otro esfuerzo con la tasa de atenuación variable para reducir el error de oscilación, pero el algoritmo que se propuso tenía poca velocidad en comparación con el algoritmo LMA estándar.

En el algoritmo de EBP, el índice de rendimiento de  $F(w)$  al ser minimizado se define como la suma de errores cuadráticos entre los resultados previstos y los resultados simulados de la red, como a continuación se describe:

$$F(w) = e^T e \quad (D.1)$$

Donde  $w = [w_1 w_2 \dots w_N]$  consta de todos los pesos de la red,  $e$  es el vector de error que comprende el error de para todos los ejemplos de entrenamiento. Al entrenar con el método LMA, el incremento de pesos  $\Delta w$  se puede obtener de la siguiente forma:

$$\Delta w = [J^T J + \mu I]^{-1} J^T e \quad (D.2)$$

Donde  $J$  es la matriz Jacobiana,  $\mu$  es la tasa de aprendizaje que será actualizada utilizando  $\beta$  en función del resultado. En particular,  $\mu$  se multiplica por la tasa de descomposición  $\beta$  ( $0 < \beta < 1$ ) cuando  $F(w)$  disminuye, mientras que  $\mu$  se divide por  $\beta$  siempre que  $F(w)$  aumenta en un nuevo paso. El proceso de entrenamiento LMA estándar puede ser ilustrado en los siguientes pseudo-códigos:

1. Inicializar los pesos y el parámetro  $\mu$  ( $\mu = 0.01$  es apropiado).
2. Calcular la suma de los errores cuadráticos sobre todas entradas  $F(w)$ .
3. Resolver (2) para obtener el incremento de pesos  $\Delta w$
4. Recalcular la suma de errores cuadráticos  $F(w)$ .

Usando  $w + \Delta w$  como el nuevo  $w$  y seguir con:

Si el nuevo  $F(w) < F(w)$  en el paso 2, entonces,

$$w = w + \Delta w \quad (D.3)$$

$$\mu = \mu\beta (\beta = 0.1) \quad (D.4)$$

Volver al paso 2, de lo contrario:

$$\mu = \frac{\mu}{\beta} \quad (D.5)$$

Volver al paso 4

Considerando que el índice de rendimiento es  $F(w) = e^T e$  utilizando el método Newton tenemos como:

$$W_{K+1} = W_K - A_K^{-1} g_K \quad (D.6)$$

$$A_k = \nabla^2 F(w)|_{w=w_k} \quad (D.7)$$

$$g_k = \nabla F(w)|_{w=w_k} \quad (D.8)$$

$$[\nabla F(w)]_j = \frac{\partial F(w)}{\partial w_j} = 2 \sum_{i=1}^N e_i(w) \frac{\partial e_i(w)}{\partial w_j} \quad (D.9)$$

El gradiente se puede escribir como:

$$\nabla F(w) = 2J^T e(w) \quad (D.10)$$

Donde

$$J(w) = \begin{bmatrix} \frac{\partial e_{11}}{\partial w_1} & \frac{\partial e_{11}}{\partial w_2} & \dots & \frac{\partial e_{11}}{\partial w_N} \\ \frac{\partial e_{21}}{\partial w_1} & \frac{\partial e_{21}}{\partial w_2} & \dots & \frac{\partial e_{21}}{\partial w_N} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial e_{KP}}{\partial w_1} & \frac{\partial e_{KP}}{\partial w_2} & \dots & \frac{\partial e_{KP}}{\partial w_N} \end{bmatrix} \quad (D.11)$$

$J(w)$  es la matriz Jacobiana. A continuación, se debe hallar la matriz Hessiana. Los elementos  $k, j$  del rendimiento de la matriz Hessiana se encuentran con la siguiente fórmula:

$$[\nabla^2 F(w)]_{k,j} = \frac{\partial^2 F(w)}{\partial w_k \partial w_j} = 2 \sum_{i=1}^N \sum_{i=1}^n X_i^2 \left\{ \frac{\partial e_i}{\partial w_k} \frac{\partial e_i}{\partial w_j} + e_i(w) \frac{\partial^2 e_i(w)}{\partial w_k \partial w_j} \right\} \quad (D.12)$$

La matriz Hessiana puede entonces expresarse como sigue:

$$\nabla^2 F(W) = 2J^T(W)J(W) + S(W) \quad (D.13)$$

donde

$$S(W) = 2 \sum_{i=1}^N e_i(w) \nabla^2 e_i(w) \quad (D.14)$$

Si suponemos que  $S(W)$  es pequeña, se puede aproximar la matriz Hessiana como:

$$\nabla^2 F(w) \cong 2J^T(w)J(w) \quad (D.15)$$

Usando (D.7) y (D.15) se obtiene el método de Gauss-Newton como:

$$W_{k+1} = W_k - [2J^T(w_k)J(w_k)]^{-1} 2J^T(w_k)e(w_k) \cong W_k - [J^T(w_k)J(w_k)]^{-1} J^T(w_k)e(w_k) \quad (D.16)$$

La ventaja de Gauss-Newton es que no requiere cálculo de segundas derivadas. Hay un problema en el método de Gauss-Newton, se trata de la matriz  $H = J^T J$  que puede no ser invertible. Esto se puede superar mediante el uso la siguiente modificación. La matriz Hessiana se puede escribir como:

$$G = H + \mu I \quad (C.17)$$

Suponemos que los valores propios y los vectores propios de  $H$  son  $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$  y  $\{z_1, z_2, \dots, z_n\}$ . Entonces:

$$Gz_i = [H + \mu I]z_i = Hz_i + \mu z_i = \lambda_i z_i + \mu z_i = (\lambda_i + \mu)z_i \quad (C.18)$$

Por lo tanto, los vectores propios de  $G$  son los mismos que los vectores propios de  $H$  y los valores propios de  $G$  son  $(\lambda_i + \mu)$ . La matriz  $G$  es definida positiva, aumentando  $\mu$  hasta  $(\lambda_i + \mu) < 0$  para todo  $i$  por lo tanto, la matriz será invertible. Esto conduce al algoritmo Levenberg-Marquardt:

$$W_{k+1} = W_k - [J^T(w_k)J(w_k) + \mu I]^{-1} J^T(w_k)e(w_k) \quad (C.19)$$

$$\nabla W_k = [J^T(w_k)J(w_k) + \mu I]^{-1} J^T(w_k)e(w_k) \quad (C.20)$$

Como es conocido, el parámetro de aprendizaje,  $\mu$  es ilustrador de pasos de movimientos de salida real a la salida deseada. En el método LMA estándar,  $\mu$  es un número constante. Sin embargo, para mejores resultados se modifica  $\mu$  como:

$$\mu = 0.01e^T e \quad (C.21)$$

Donde  $e$  es a la matriz  $k \times 1$  por lo tanto,  $e^T e$  es a  $1 \times 1$  tanto  $[J^T J + \mu I]$  es invertible. Por lo tanto, si la salida real está lejos de salida deseada o similar, los errores son grandes, así, converge a la salida deseada con pasos grandes. Del mismo modo, cuando la medida de error es pequeña, entonces, la salida real se aproxima a la salida deseada con pasos suaves. Por lo tanto, el error oscilación se reduce grandemente.

## Apéndice E: Datos de los transformadores utilizados como casos de estudio

Para complementar la información del Capítulo V sobre los transformadores utilizados en los casos de estudio se muestran a continuación los datos de placa, así como el relé utilizado para obtener los datos utilizados en este trabajo. Además, algunas características de los dispositivos que se utilizan para la medición.



**Figura E.1. Transformador T1**

En la Figura E.2 se muestra la placa del Transformador T1 y T2.



**Figura E.2. Datos de placa del Transformador T1 y T2**

En la Figura E.3 se muestra el relé utilizado para el monitoreo de los transformadores analizados en este trabajo.



**Figura E.3. Relé utilizado para el monitoreo de transformadores**

Esta subestación se encuentra ubicada en Av. de las Granjas 280C, Libertad, CDMX. Los transformadores se alimentan 85 kV por el lado de alta con conexión delta y por el lado de baja está conectado a una red de 23 kV con conexión Y.

Estos transformadores de potencia se diseñan para contar con las características de operar a su capacidad plena siempre que la temperatura ambiente no exceda de 40 °C y bajo un promedio de temperatura durante cualquier periodo de 24 horas no exceda de 30 °C [1].

Así mismo, la elevación de temperatura de los devanados y punto más caliente a tensión nominal, frecuencia nominal y a capacidad plena puede ser hasta 55 °C y 65 °C respectivamente, considerando la medición realizada por el método de resistencia sobre una temperatura ambiente máxima de 40 °C.

El transformador debe contar con un termómetro indicador de aceite tipo carátula y debe contar con 2 contactos que pueden usarse para la implementación de esquemas de monitoreo y acciones de control.

De igual manera, el equipo debe contar con un indicador de temperatura de devanados conocido también como imagen térmica, para ello debe incluir un termo pozo que tenga la capacidad de alojar un elemento resistivo por el cual podrá circular una corriente en baja tensión (corriente secundaria) y un transformador de corriente con el objeto de obtener una indicación de la temperatura de los devanados. Este indicador debe contar con cuatro (4) contactos.

Ambos transformadores se protegen mediante un esquema diferencial (87T) como protección primaria. Este relevador proporciona una protección selectiva de alta velocidad ante fallas internas al equipo y para aquella región de protección delimitada por los transformadores de corriente (TCs).

Así mismo, el equipo cuenta con elementos de protección que vienen integrados en su construcción, los cuales pueden especificarse de acuerdo a la capacidad del transformador, tensiones de operación entre otras características. Los principales dispositivos que se usan en los transformadores son:

- Relevador Buchholz (63G).
- Válvula de sobrepresión (63P).
- Indicador de temperatura del aceite (26Q).
- Indicador de temperatura de devanados (49T).



- Indicador de nivel de aceite (T1Q).

A continuación, se describen los accesorios de protección que se usan para proteger al transformador contra alta temperatura.

### **Indicador de temperatura del aceite (26Q)**

Dispositivo que consiste en un accesorio con un bulbo sensor y que se instala en la parte superior del tanque principal del transformador con el objetivo de captar la temperatura del aceite. Este indicador es equipado con microinterruptores los cuales son operados por las levas en la flecha de su aguja indicadora.

El indicador 26Q está equipado para enviar señalización y medición de temperatura a los centros de control.

### **Indicador de temperatura de devanados (49T), imagen térmica**

Dispositivo que es operado por un elemento bimetálico y una bobina calefactora que se construye con el objetivo de calentar el elemento a la temperatura del punto más caliente del devanado del transformador. Este indicador es utilizado en transformadores en su esquema de control del sistema de ventilación forzada con el fin de evitar que éstos sean sometidos a sobrecargas que derivan en elevación de temperatura inadmisibles. Así mismo, puede habilitarse con esquemas de señalización y alarma.

Este indicador generalmente es equipado con tres o cuatro microinterruptores que se pueden calibrar a diferentes niveles de temperatura para funciones de controlar el sistema de ventilación forzada; señalización de alarma y activación de disparo o desconexión del transformador.

Este dispositivo cumple con las 3 funciones siguientes:

- i) Indicar la temperatura máxima alcanzada dentro del transformador,
- ii) Energizar o desenergizar el sistema de ventiladores y
- iii) Activar una señalización de alarma y/o disparo ante condiciones adversas por temperatura para el transformador.

Este dispositivo desempeña la función de obtener la medición de temperatura del cobre partiendo de la premisa que esta temperatura depende de la propia del aceite contenido en el transformador como el calor que se genera ante la circulación de la corriente en sus devanados. Por tal motivo, ambas mediciones se toman en cuenta en la determinación de la referencia medida.