

UNIVERSIDAD DE CARABOBO
ÁREA DE ESTUDIOS DE POSTGRADO
FACULTAD DE INGENIERÍA
MAESTRÍA EN
INGENIERÍA ELÉCTRICA

**COMPARACIÓN DE LOS MÉTODOS
EMERGENTES UTILIZADOS EN EL
ANÁLISIS ESPECTRAL DE CORRIENTE
ESTATÓRICA PARA DETERMINAR LAS
CONDICIONES DE COJINETES EN
MOTORES DE INDUCCIÓN**

AUTOR: RUBÉN TERÁN
TUTOR: CARLOS JIMÉNEZ

Naguanagua 28 de julio de 2019

UNIVERSIDAD DE CARABOBO
ÁREA DE ESTUDIOS DE POSTGRADO
FACULTAD DE INGENIERÍA
MAESTRÍA EN
INGENIERÍA ELÉCTRICA

**COMPARACIÓN DE LOS MÉTODOS
EMERGENTES UTILIZADOS EN EL
ANÁLISIS ESPECTRAL DE CORRIENTE
ESTATÓRICA PARA DETERMINAR LAS
CONDICIONES DE COJINETES EN
MOTORES DE INDUCCIÓN**

AUTOR: RUBÉN TERÁN

Trabajo presentado ante le Área de Estudios de
Postgrado de la Universidad de Carabobo para
optar al Título de Magister en Ingeniería Eléctrica

Naguanagua 28 de julio de 2019



UNIVERSIDAD DE CARABOBO
FACULTAD DE INGENIERIA
DIRECCIÓN DE ESTUDIOS PARA GRADUADOS
SECCIÓN DE GRADO



ACTA DE DISCUSIÓN DE TRABAJO DE GRADO

En atención a lo dispuesto en los Artículos 137, 138 y 139 del Reglamento de Estudios de Postgrado de la Universidad de Carabobo, quienes suscribimos como Jurado designado por el Consejo de Postgrado de la Facultad de Ingeniería, de acuerdo a lo previsto en el Artículo 135 del citado Reglamento, para estudiar el Trabajo de Grado titulado:

**“COMPARACIÓN DE LOS MÉTODOS EMERGENTES UTILIZADOS
EN EL ANÁLISIS ESPECTRAL DE CORRIENTE ESTATÓRICA
PARA DETERMINAR LAS CONDICIONES DE COJINETES EN
MOTORES DE INDUCCIÓN**

*Presentado para optar al grado de **MAGÍSTER EN INGENIERÍA ELÉCTRICA** por el (la) aspirante:*

RUBÉN TERÁN
V.- 13.552.117

*Habiendo examinado el Trabajo presentado, decidimos que el mismo está **APROBADO.***

En Valencia, al primer (01) día del mes de Julio del año dos mil diecinueve.

Prof. Carlos Jiménez
C.I.: 7523966
Fecha: 01/07/2019

Prof. Eva Monagas
C.I.: 6122881
Fecha: 01/07/2019
FE: 03/07/2019 /D6.-

Prof. Juan Arcila
C.I.: 3.947.614
Fecha: 01/07/2019

*Dedicado a
la memoria de mi madre
María Duarte de Terán.
A mi esposa Ledy Laura y
a mi hija María Laura.*

AGRADECIMIENTOS

A todos aquellos compañeros de trabajo que me ayudaron ya sea con asesorías o con palabras de apoyo a realizar este trabajo, especialmente a la profesora Irahís Rodríguez de quien se originó el tema, la profesora Marlenis Gonzáles por su asesoría metodológica, el profesor Carlos Jiménez por su orientación en el uso de las máquinas de vectores soporte y en el tratamiento de la data y al profesor César Seijas por su diligencia para aclararme diversas dudas de programación y en el estudio de los clasificadores.

Índice general

Agradecimientos	VI
Glosario	X
Lista de figuras	XI
Lista de tablas	XII
Resumen	1
Abstract	2
Introducción	3
1. EL PROBLEMA	5
1.1. Planteamiento del problema	5
1.2. Formulación del problema	7
1.3. Objetivos	8
1.3.1. General	8
1.3.2. Específicos	8
1.4. Justificación	8
1.5. Alcance y limitación	10
2. MARCO TEÓRICO	11

2.1. Antecedentes	11
2.2. Bases teóricas	14
2.2.1. El motor de inducción	14
2.2.2. Partes principales	15
2.2.3. Principio de funcionamiento	16
2.2.4. Fallas en motores de inducción	18
2.2.5. Fallas en cojinetes	19
2.2.6. Monitoreo de las condiciones del motor de inducción	20
2.2.7. Máquinas de Vectores de Soporte	22
2.2.8. Redes Neuronales Artificiales	26
2.2.9. Análisis discriminante	30
3. MARCO METODOLÓGICO	34
3.1. Tipo de estudio y de investigación	34
3.2. Población y muestra	34
3.3. Procedimiento metodológico a utilizar	36
3.4. Técnicas e instrumentos de recolección de datos	38
3.5. Técnicas de procesamiento y análisis de datos	39
3.5.1. Extracción de rasgos de la corriente de armadura	39
3.5.2. Normalización de la data	42
3.5.3. Reducción de atributos	42
3.5.4. Selección de características	44
3.5.5. Clasificación del estado del cojinete	45
4. RESULTADOS	46
4.1. Data adquirida	46
4.1.1. Extracción de rasgos en el dominio del tiempo	48
4.1.2. Extracción de componentes en el dominio de la frecuencia	52
4.2. Evaluación de los clasificadores	54
4.3. Clasificación usando Redes Neuronales Artificiales	56
4.4. Clasificación múltiple usando Redes Neuronales	59

4.5. Clasificación usando Máquinas de Vectores Soporte	66
4.6. Clasificación múltiple usando Máquinas de Vectores Soporte . . .	72
4.7. Clasificación partiendo del Análisis Discriminante	78
4.8. Clasificación Múltiple partiendo del Análisis] Lineal Discriminante	84
CONCLUSIONES	88
RECOMENDACIONES	90
Bibliografía	91

Glosario

ALD Análisis Lineal Discriminante.

ANN Redes Neuronales Artificiales.

ANOVA Análisis de la Varianza.

DM Distancia de Mahalanobis.

PCA Análisis de Componentes Principales.

RFE Eliminación Recursiva de Características.

SMT Sistema de Mahalanobis-Taguchi.

SVM Máquinas de Vectores de Soporte.

Índice de figuras

2.1. Dimensiones del cojinete	19
2.2. Plano clasificador	22
2.3. Función discriminante	23
2.4. Clasificación con hiperplano. El margen de clasificación viene dado por la distancia entre los puntos de cada clase que son más cercanos al hiperplano	24
2.5. Neurona artificial	27
2.6. Capa de S neuronas	29
3.1. Cojinete con falla en pista interna	36
3.2. Cojinete con falla en pista externa	36
3.3. Arquitectura de diagnóstico de falla en cojinetes	38
3.4. Esquema del montaje	39
4.1. Muestreo de corriente en el dominio del tiempo	47
4.2. Clasificación de los datos	48
4.3. Matriz de confusión, caso clasificación binaria.	55
4.4. Comparación de dos espectros de la corriente armadura.	86

Índice de tablas

3.1. Motor utilizado	35
4.1. Corriente medida	47
4.2. Descripción de los datos. Clasificación binaria	49
4.3. Descripción de los datos. Clasificación múltiple	49
4.4. Características extraídas en cada condición de operación	50
4.5. Matriz de covarianza	51
4.6. Selección de rasgos	52
4.7. Datos del cojinete	53
4.8. Frecuencias características de movimiento de la bola en el cojinete	53
4.9. Efecto de variar los rasgos en las Redes Neuronales Artificiales (ANN)	57
4.10. Efecto del número de neuronas ocultas	58
4.11. Resultados completos de la clasificación binaria, usando Redes Neu- ronales	58
4.12. Matriz de confusión. Clasificación con Redes Neuronales	59
4.13. Selección de arquitectura, clasificación múltiple con ANN, caso 1 .	60
4.14. Matriz de confusión. Clasificación con Redes Neuronales	61
4.15. Selección de arquitectura, clasificación múltiple con ANN, caso 2 .	61
4.16. Matriz de confusión. Clasificación con Redes Neuronales	62
4.17. Evaluación de la clasificación múltiple con las ANN. Caso 2.	63
4.18. Selección de arquitectura, clasificación múltiple con ANN, caso 3 .	63
4.19. Matriz de confusión. Clasificación con Redes Neuronales	64
4.20. Evaluación de la clasificación múltiple con las ANN. Caso 3.	64

4.21. Selección de arquitectura, clasificación múltiple con ANN, caso 4 .	65
4.22. Matriz de confusión. Clasificación con Redes Neuronales	65
4.23. Resumen de la clasificación múltiple con ANN.	66
4.24. Clasificación con Máquinas de Vectores de Soporte (SVM) caso 1	67
4.25. Matriz de confusión. Clasificación con SVM	67
4.26. Evaluación de la clasificación con SVM. Caso 1.	68
4.27. Clasificación con SVM.	69
4.28. Matriz de confusión. Clasificación con SVM	69
4.29. Evaluación de la clasificación con SVM. Caso 2.	70
4.30. Matriz de confusión. Clasificación con SVM	70
4.31. Evaluación de la clasificación con SVM. Caso 3.	71
4.32. Matriz de confusión. Clasificación con SVM	71
4.33. Clasificación múltiple con SVM.	73
4.34. Matriz de confusión. Clasificación múltiple con SVM caso 1	73
4.35. Evaluación de la clasificación múltiple con SVM. Caso 1.	74
4.36. Matriz de confusión. Clasificación múltiple con SVM caso 2	75
4.37. Evaluación de la clasificación múltiple con SVM. Caso 2.	75
4.38. Matriz de confusión. Clasificación múltiple con SVM caso 3	76
4.39. Evaluación de la clasificación múltiple con SVM. Caso 3.	76
4.40. Matriz de confusión. Clasificación múltiple con SVM caso 4	77
4.41. Evaluación de la clasificación múltiple con SVM. Caso 4.	78
4.42. Valores medios de cada características	79
4.43. Matriz de confusión. Clasificación con Redes Neuronales	81
4.44. Evaluación de la clasificación con el Análisis Lineal Discriminante (ALD). Caso 1.	81
4.45. Matriz de confusión. Clasificación con Redes Neuronales, caso 2 .	82
4.46. Evaluación de la clasificación con el ALD. Caso 2.	82
4.47. Matriz de confusión. Clasificación con Redes Neuronales, caso 3 .	83
4.48. Evaluación de la clasificación con el ALD. Caso 3.	83
4.49. Matriz de confusión. Clasificación con Redes Neuronales, caso 4 .	83

	XIV
4.50. Evaluación de la clasificación con el ALD. Caso 4.	84
4.51. Matriz de confusión. Clasificación múltiple con ALD con señal en el dominio del tiempo	85
4.52. Evaluación de la clasificación múltiple con el ALD y la señal en el dominio del tiempo.	85
4.53. Matriz de confusión. Clasificación múltiple con ALD con señal en el dominio de la frecuencia	87
4.54. Evaluación de la clasificación múltiple con el ALD y la señal en el dominio de la frecuencia.	87

RESUMEN

En este trabajo se compararon los métodos emergentes como son: el análisis lineal discriminante, las máquinas de vectores soporte y las redes neuronales, en la detección de fallas incipientes en motores de inducción. Para ello se hicieron mediciones en la corriente de armadura de un motor de inducción perteneciente al Laboratorio de Mantenimiento Predictivo de la Universidad de Carabobo.

Se hicieron distintas pruebas con cojinetes sin fallas y con fallas en bola, pista externa e interna. Se midieron señales de corriente en el dominio del tiempo, los datos de las mediciones fueron almacenados en una computadora con la ayuda de una tarjeta de adquisición de datos. La señal en el dominio del tiempo se llevó al dominio de la frecuencia por medio de la Transformada Rápida de Fourier.

De la señal en el dominio del tiempo se extrajeron 11 métricas estadísticas, que posteriormente se redujeron usando el Análisis de la Varianza y la Eliminación Recursiva de Características.

Tanto la señal en el dominio como en el espectro en frecuencia se utilizaron para analizar el comportamiento los métodos emergentes. Primero se estudió el grado con el que detectaban si hay falla y luego su capacidad para discriminar el tipo de falla.

Para comparar a los clasificadores entre sí, se observaron los tiempos de cálculos, y parámetros como exactitud, precisión, recall entre otros.

ABSTRACT

In this paper, the emerging methods were compared: linear discriminant analysis, support vector machines, and neural networks, in order to detect incipient failures in induction motors. For this purpose, it was made measurements to the armature current of an induction motor, which belongs to the Predictive Maintenance Laboratory of the University of Carabobo.

Different tests were made on bearings without failures and bearings with balls with faults, and with an external and internal track. It was measured current signals in the time-domain, and the measurements data were stored in a computer with the help of a data acquisition card. The signal in the time-domain was converted into the domain of the frequency by means of the Fast Fourier Transform.

From the signal in the time-domain, it was 11 extracted statistical metrics, which subsequently were reduced using the Analysis of Variance and Recursive Feature Elimination.

Both the signal in the domain and in the spectrum in frequency was used to analyze the emerging method behavior. First, it was studied the level to which it was detected the fault and then it was analyzed their ability to discriminate what kind of fault was.

It was observed the calculation times, and parameters such as accuracy, precision, recall, among others, to compare the classifiers with each other.

INTRODUCCIÓN

La falla en cojinete es la más común de las que se presentan en motores de inducción trifásicos, entre los síntomas de ésta se tienen ruidos, vibraciones e inclusive la detención del motor. Para disminuir la posibilidad de ocurrencia, generalmente una solución puede consistir en sustituirlo cada cierto tiempo. Otra opción es detectarla cuando se encuentre en su estado incipiente por medio de técnicas de mantenimiento predictivo.

Entre las técnicas más comunes se tienen estudio de vibraciones, análisis de la corriente de armadura, el voltaje del estator y la potencia eléctrica. En este trabajo se utiliza el análisis de la corriente de armadura y se comparan diversos métodos emergentes para la detección de fallas en cojinetes.

El trabajo se divide en cuatro capítulos. En el primero se plantea el problema y los objetivos de la investigación. En el segundo capítulo se muestra el marco teórico en donde se explica porqué ocurre la falla en rodamiento, cuál es su efecto en la corriente de armadura. Adicionalmente se estudian los métodos emergentes como son: el análisis discriminante, las máquinas de vectores soporte y las redes neuronales artificiales, con el propósito de ver que variables se deben ajustar para lograr la detección de la falla.

En el capítulo 3 se muestra el marco metodológico en donde se explica el proceso de adquisición de la data, su preparación y el entrenamiento de los clasificadores. En el capítulo 4 se muestran los resultados, para cada clasificador se muestra su desempeño cuando se quiere indicar si hay falla y cuando se indica el tipo de falla. Finalmente se finaliza con las conclusiones y futuros trabajos como

continuación de éste.

Capítulo 1

EL PROBLEMA

1.1. Planteamiento del problema

Los motores de inducción jaula de ardilla, para el desarrollo de los procesos y actividades, tradicionalmente se habían considerado como una de las máquinas eléctricas que presentaban pocas averías, las cuales requerían poca atención excepto cuando producían paradas de planta generales. En la actualidad la tendencia es construir máquinas de bajo costo, que trabajen más cerca de los límites de resistencia mecánica y electromagnética, y en especial cuando el motor trabaja en condiciones ambientales desfavorables como por ejemplo altas temperaturas, la posibilidad de un deterioro progresivo se ha incrementado.

Teniendo en cuenta que con frecuencia los motores eléctricos se encuentran en puntos vitales de los sistemas industriales muy complejos, y los criterios industriales actuales tienden hacia una producción de elevada calidad y sin interrupciones, se hace necesario el control de funcionamiento de estos para detectar y dar respuesta a posibles fallas de operación (vibraciones, aumento de corrientes) que garanticen el buen funcionamiento del proceso en general.

Dentro de las fallas que ocurren en estos motores y que es muy importante detectar se tienen: fallas en el devanado del estator, conexión incorrecta de devanados, barras del rotor rotas o daños en los anillos de cortocircuito, irregula-

ridades estáticas y dinámicas en el entrehierro, curvatura de eje, fallas en apoyos y engranajes, fallas en los cojinetes, entre otras.

Una de las fallas que se presenta comúnmente en los motores de inducción es la de cojinetes, tal como lo plantean en un estudio realizado en el Estado Carabobo (Mago y Monagas, 2004), siendo esta la que tiene el mayor índice de ocurrencia (71,2 %). Es de hacer notar que el estado de funcionamiento en estos motores originan otras fallas como: desalineaciones de ejes, excentricidad estática y dinámica, entre otras. El aumento de las vibraciones en el eje produce un incremento de la falla en el cojinete, lo que ocasiona una reacción en cadena que lleva al “estado catastrófico del cojinete”. Es importante destacar que la falla en cojinetes se puede detectar antes de ocasionar la parada del motor y, por consiguiente, del sistema de producción.

Para contrarrestar estas fallas que se presentan en los sistemas industriales, las empresas establecen planes de mantenimiento, los cuales pueden ser de tres tipos: correctivo, en el que la intervención se realiza una vez que se ha producido la falla; preventivo, en el cual se actúa antes de presentarse la falla, pero en la mayoría de los casos se reemplazan piezas que podrían seguir en funcionamiento por un tiempo prolongado; y predictivo, en el cual se realiza un monitoreo continuo del sistema y por consiguiente la intervención se realiza en el momento más óptimo donde se aprovecha al máximo la vida útil de los equipos.

El mantenimiento predictivo para motores de inducción tiene grandes ventajas económicas, entre las que se puede mencionar: aprovechamiento máximo de las piezas que conforman al motor de inducción, la reducción de staff de piezas para la realización del mantenimiento, menos pérdidas económicas ya que se pueden planificar intervenciones. Adicionalmente se puede destacar que, al planificar las intervenciones de mantenimiento, la calidad del producto final se mantiene con lo que se beneficia también el consumidor. Lo expuesto anteriormente muestra la importancia de tomar medidas anticipadas a la ocurrencia de las fallas que puedan presentarse en el motor de inducción.

Una de las técnicas utilizadas en el mantenimiento predictivo para la detección

de fallas en cojinetes es el análisis de corrientes estatóricas. Esta técnica considera que el hecho de fallas en cojinetes afecta la longitud del entrehierro en la máquina, viéndose esto reflejado en la magnitud de los armónicos de la corriente, razón por la cual se hace necesario comparar los distintos métodos emergentes como son: el ALD, las ANN y las SVM, utilizados en el análisis de la corriente estatórica para determinar la condición de cojinetes en motores de inducción jaula de ardilla.

El desarrollo de esta propuesta contempla el análisis de una de las técnicas de mantenimiento predictivo aplicada en la detección de fallas en cojinetes, esta técnica es el análisis de la corriente estatórica. Por otra parte será necesario analizar los datos que resultan de las pruebas realizadas a través de los tres métodos emergentes. Con esto se pretende establecer criterios de observación del espectro de corriente estatórica, para ubicar las variaciones y el nivel en que estas ocurren, y por último comparar dichos métodos para indicar cuál de estos es el mejor para indicar el grado de la falla.

1.2. Formulación del problema

Los cambios en la magnitud de la corriente se pueden detectar por medio del uso de técnicas o métodos emergentes como son: el ALD, las ANN y las SVM. En la actualidad se han realizado diversos trabajos en los cuales se obtienen resultados aceptables en cuanto a la detección de fallos en cojinetes usando estas técnicas.

Debido a lo anteriormente expuesto surge la siguiente interrogante ¿Será posible comparar los métodos emergentes para determinar cuál discrimina mejor el grado de la falla?

1.3. Objetivos

1.3.1. General

Comparar los métodos emergentes, utilizados en el análisis de la corriente estática para determinar las condiciones de cojinetes en motores de inducción jaula de ardilla.

1.3.2. Específicos

- Determinar las características de los diversos métodos emergentes (ANN, SVM y análisis discriminante) para determinar las condiciones de los cojinetes.
- Analizar la corriente estática en motores de inducción jaula de ardilla por medio de los métodos emergentes (ANN, SVM y análisis discriminante) para determinar las condiciones de los cojinetes.
- Comparar los métodos emergentes (ANN, SVM y análisis discriminante) para determinar cuál discrimina mejor las condiciones de los cojinetes de motores de inducción jaula de ardilla.

1.4. Justificación

En la actualidad para el sector industrial venezolano, ha cobrado gran importancia la calidad en la producción debido a que este necesita ser más competitivo en el ámbito nacional e internacional. Esto hace que se produzca gran interés en mantener el funcionamiento óptimo de los motores de inducción ya que, el buen desarrollo de los procesos industriales se ve afectado notablemente por estos motores y por lo tanto la calidad de producción también.

Los motores de inducción constituyen un recurso de vital importancia en los procesos productivo industriales, los cojinetes juegan un papel importante en la

confiabilidad y el rendimiento de los sistemas que usan motores. Debido a la estrecha relación entre su funcionamiento y los cojinetes, es difícil imaginar el progreso de las máquinas rotacionales modernas sin tomar en cuenta las aplicaciones de los cojinetes. Adicionalmente, la mayoría de las fallas en los motores están relacionadas a fallas en cojinetes. De allí la necesidad de detectar en que estados se encuentran estos y es aquí donde los métodos emergentes juegan un papel importante.

Esto ha incrementado el interés en el desarrollo de los planes de mantenimiento predictivo ya que, de acuerdo a (Giordanelli, 1997) existen tres ventajas de este sobre los otros planes de mantenimiento, como lo son: no se asignan recursos en reparar máquinas en buen estado; pueden tomarse decisiones estratégicas en programas de producción, ya que en los motores pueden prolongarse las paradas por mantenimiento; y por último, la decisión de intervenir un motor no depende del tiempo de instalación, por lo cual, las fallas pueden ser detectadas durante los períodos de garantía de los mismos lo que disminuye los costos de mantenimiento.

Por lo antes expuesto, se requiere detectar fallas en su estado incipiente para tomar las correcciones pertinentes y asegurar así la interrupción de la operación de los procesos productivos en el momento oportuno, lo cual redundaría en la reducción de los gastos de mantenimiento.

Al anticiparse a la ocurrencia de la falla y realizar paradas programadas del proceso se traduce en beneficios económicos, ya que no se tiene pérdidas de material por producto defectuoso ni es necesario mantener gran cantidad de piezas disponibles para realizar los reemplazos, las cuales pueden adquirirse en el momento más apropiado.

Por otro lado, se proporcionarían múltiples aportes teóricos que servirían como aporte para futuras investigaciones con motores de inducción jaula de ardilla. Una de las líneas de investigación de la Universidad de Carabobo es Mantenimiento Predictivo de Máquinas y Sistemas, por lo que los conocimientos adquiridos con el desarrollo de este trabajo de investigación, contribuirán al desarrollo de nuevos estudios que conlleven a que esta universidad pueda ofrecer servicios de asesorías

técnica a las industrias, lo que traería beneficios económicos a nuestra casa de estudios.

Este tema de investigación contribuirá a realizar planes de mantenimiento aplicados al sector industrial, lo que se traducirá en beneficios para los procesos productivos. Estos planes contribuyen con la calidad del producto, favoreciendo a la vez a los consumidores.

1.5. Alcance y limitación

El presente trabajo de investigación se realizará en el Laboratorio de Mantenimiento Predictivo de la Universidad de Carabobo.

El trabajo responde a necesidades específicas de la línea de investigación Mantenimiento Predictivo de Máquinas y Sistemas la cual pertenece al Departamento de Potencia de la Universidad de Carabobo. Abarca el análisis de la técnica de medición del análisis espectral de la corriente estática para determinar la forma de adquisición de los datos, los cuales serán producto de las mediciones que se realizarán en el Laboratorio de Mantenimiento Predictivo de la Universidad de Carabobo, la determinación de la red neuronal, la función discriminante y SVM para discriminar sobre el estado de los cojinetes de los motores de inducción objetos de estudio.

Capítulo 2

MARCO TEÓRICO

En este capítulo se muestran los antecedentes al tema de investigación. Seguidamente se dará una breve explicación del funcionamiento del motor de inducción y sus fallas típicas, en específico la de cojinete, y el proceso de monitoreo de sus variables en funcionamiento, con el fin de detectar estas fallas. Finalmente se comentarán los métodos emergentes objetos de estudio como son: Las Máquinas de Vectores Soporte, las Redes Neuronales Artificiales y el Análisis Discriminante.

2.1. Antecedentes

La determinación oportuna de las condiciones de los cojinetes de los motores de inducción jaula de ardilla reduce en gran medida los riesgos de interrupción de los procesos productivos y los costos que estos pudieran ocasionar. En relación a las investigaciones previas se puede mencionar al trabajo publicado por (Jin y Chow, 2013), en este se muestra que la Distancia de Mahalanobis (DM) se puede usar para indicar la condición del motor y de su ventilador basada en la medición de la señal de vibración. Los resultados muestran que una falla en estado incipiente en el ventilador de enfriamiento, causada por vibraciones debidas a fallas en cojinetes, puede ser detectada satisfactoriamente, y las diferentes fallas desbalanceadas en los motores de inducción pueden ser clasificadas con una exactitud alta usando el

Sistema de Mahalanobis-Taguchi (SMT).

El SMT es una tecnología de reconocimiento de patrones y es usada ampliamente en la clasificación de datos, combina los métodos de distancia de Mahalanobis y Taguchi. El procedimiento propuesto consiste en formar una data con las características extraídas, cuando el motor está sin falla, en donde las DM forman un espacio de referencia conocido como Espacio de Mahalanobis (EM); luego cuando el motor funcione en una condición anormal se comparará la DM con el EM para luego clasificar el estado de funcionamiento del motor. Los experimentos demostraron que las DM correspondientes al funcionamiento con falla se desvían grandemente del normal, condición esta deseable para la detección de fallas.

Las SVM son ampliamente usadas para monitorear las condiciones de las máquinas y para el diagnóstico de fallas como lo explican (Widodo y Yang, 2007). En este artículo se resumen los trabajos logrados en el uso de las SVM en el diagnóstico de fallas, se describen las ecuaciones de las SVM y se discuten los métodos utilizados para realizar tareas de clasificación múltiple. Entre los trabajos descritos para detectar fallas se nombra el de Pohyonen en el cual se estima la densidad de potencia usando el método de Welch, la cual se calculó de corrientes circulantes en ramas en paralelo del motor de inducción. Después del entrenamiento las SVM distinguían entre un patrón de señal con falla de uno sin la misma. En otros trabajos se emplean SVM y extracción de características vía análisis de componentes; las características en el dominio del tiempo y de la frecuencia se extraen de la señal de corriente y de la vibración.

Las ANN también son usadas como herramienta en la detección de fallas, en este sentido (Kowalski y Orłowska-Kowalska, 2003) demuestran problemas de diagnósticos de motores de inducción en el caso de fallas en el rotor, estator y cojinetes. Para el propósito de diagnóstico se propone el método de inteligencia artificial basado en ANN. Se proponen dos clases de ANN: perceptrón multicapa y redes de Kohonen. Estas redes fueron entrenadas y probadas usando las datos adquiridas de los espectros de corriente y vibración. Finalmente se evaluó la eficiencia de los detectores basados en ANN. Se obtuvieron cuatro tipos de salidas:

una salida para la determinación de una condición del cojinete (fallado o sano) basada en el espectro de la corriente de armadura, una salida para la determinación de una condición del cojinete (fallado o sano) basada en el espectro de vibración, una salida para la determinación de una condición del cojinete (fallado o sano) basada en el espectro de la corriente de armadura y vibración y una salida para la determinación de una condición del cojinete (fallado o sano) basada en el desbalance de la señal.

En el trabajo de (Filippetti, Franceschini, Tassoni, y Vas, 2000) se presenta una revisión de los desarrollos más recientes en el campo de diagnósticos de máquinas eléctricas y accionamientos basados en Inteligencia Artificial (IA). Este cubre la aplicación de sistemas expertos, ANN y sistemas con lógica difusa, los cuales pueden integrarse en conjunto así como también con técnicas más tradicionales. La metodología discutida es aplicada a un motor de inducción, utilizando como señales de entrada voltajes y corrientes en el dominio del tiempo. Adicionalmente, el convertidor de frecuencia es considerado para incorporarlo en el procedimiento de diagnóstico de la mayoría de las fallas típicas de los elementos electrónicos. Se propone calcular el deslizamiento del motor por medio del espectro de la señal obtenido después de aplicar la transformada rápida de Fourier. En los resultados obtenidos se muestra que las ventajas de combinar las ANN con lógica difusa.

Las fallas en los cojinetes afectan a la corriente de armadura del motor, esto se explica en el artículo de (Schoen, Habetler, Kamran, y Bartfield, 1995). En este trabajo se explica como el análisis espectral de la corriente de armadura se puede utilizar para determinar fallas en los elementos rodantes de motor. Se realiza una correlación entre la frecuencia de las vibraciones mecánicas, cuya medición es el método más usado para detectar fallas incipientes en cojinetes, y la frecuencia de las señales eléctricas. Se estudian los distintos tipos de fallas en rodamientos, y se muestra como las frecuencias características asociadas a estas fallas, dependen de las dimensiones de los cojinetes. Se describen los efectos en el espectro de la corriente de armadura y su relación con las frecuencias determinadas. Los resultados de los experimentos muestran que, al ocurrir fallas, ocurren cambios

en la densidad de flujo magnético, que causan cambios visibles en la corriente de armadura a frecuencias previsible. A pesar de que estos cambios son pequeños en comparación con las otras amplitudes de los armónicos, estos ocurren a frecuencias distintas a los relacionados con los armónicos de ranuras, por lo que se puede utilizar su observación para la detección de fallas.

2.2. Bases teóricas

El motor de inducción es una máquina que transforma la energía eléctrica en mecánica. Se encuentra presente en muchísimas aplicaciones industriales tales como ventiladores, compresores, bombas, prensas, cintas transportadoras, extrusoras, ascensores, molinos de cemento. Los fabricantes y usuarios de los motores en las primeras décadas del siglo XX usaban únicamente protecciones de sobrecorriente, sobrecarga y sobrevoltaje en la mayoría de los casos para asegurar su funcionamiento. Sin embargo con la automatización de los procesos industriales se ha reducido la interacción entre el hombre y la máquina para supervisar la operación del accionamiento eléctrico; la tendencia es al monitoreo de las condiciones del proceso, con el propósito de detectar fallas incipientes y reducir el tiempo de parada del accionamiento.

A continuación se estudiará el funcionamiento del motor de inducción, seguidamente de las fallas específicamente la de cojinetes. Luego se explicará el proceso de monitoreo de las condiciones del motor y finalmente de los métodos emergentes.

2.2.1. El motor de inducción

En el motor de inducción la corriente alterna es suplida directamente al estator, y al rotor por inducción o acción transformadora desde el estator (Fitzgerald, Kingsley, Umans, y James, 2003). Sus partes principales son el estator, el rotor y los devanados.

2.2.2. Partes principales

El motor de inducción tiene múltiples partes como son: estator, rotor, conductores, aislamientos, rodamientos, eje, bornera entre otras. Las principales son el estator y rotor, los que se describen a continuación:

Estator: parte que recibe a energía eléctrica. Está formado por un núcleo de chapas magnéticas apiladas y aisladas entre sí, que aloja un devanado trifásico en las ranuras situadas en su superficie interior. Estas ranuras están distribuidas de forma uniforme a lo largo de su periferia de forma que la distribución de conductores sea idéntica para las tres fases, con la única diferencia de que los conductores que ocupan posiciones homólogas en cada una de las fases están desplazados entre sí 120° eléctricos espaciales.

Rotor parte que transfiere la energía mecánica. También está formado a bases de chapas magnéticas pero tiene la forma de un cilindro con las ranuras que alojan al devanado situadas en su exterior, y distribuidas también de manera uniforme. En la mayoría de los casos las ranuras del rotor están inclinadas para mejorar la forma de onda de la fuerza electromotriz producida. Existen dos ejecuciones posibles para los devanados:

Jaula de ardilla: en este caso los huecos de las ranuras se rellenan de barras de aluminio (u otro material conductor) que se unen en ambos extremos del rotor mediante unos anillos de cortocircuito igualmente conductores, adoptando la forma de una jaula de ardilla, que da origen a la denominación de este tipo de motores. Existen múltiples diseños para estos devanados (simple o doble jaula, ranuras normales o profundas, etc.) lo que da lugar a diferencias funcionales importantes, sobre todo en su comportamiento en el arranque.

Rotor devanado: este se fabrica con conductores de cobre, los cuales constituyen un devanado trifásico que son imágenes de los devanados del estator. Se conectan usualmente en estrella y se le puede colocar una

terminación específica que permita extraerla a través de anillos colectores.

2.2.3. Principio de funcionamiento

El funcionamiento del motor de inducción se basa en la creación de un campo magnético giratorio en el entrehierro de distribución sinusoidal (Fitzgerald y otros, 2003). Si se desprecian los armónicos de ranura, se puede demostrar que al alimentar el devanado de armadura con tres corrientes equilibradas se crea un campo en el entrehierro sinusoidal que gira a una velocidad constante. Considere tres corrientes que alimenten el devanado de armadura dadas por:

$$\begin{aligned} i_a &= I_{max} \cos(\omega_e t) \\ i_b &= I_{max} \cos\left(\omega_e t - \frac{2\pi}{3}\right) \\ i_c &= I_{max} \cos\left(\omega_e t + \frac{2\pi}{3}\right) \end{aligned}$$

donde: I_{max} es el valor pico de la corriente de armadura y ω_e la velocidad angular de la misma.

Cada fase producirá una onda de fuerza magnetomotriz pulsante sobre su propio eje magnético. Para un ángulo espacial, las ecuaciones de fuerza magnetomotriz de cada fase será:

$$\mathcal{F}_a(\theta, t) = F i_a \cos(\theta) = F_{max} \cos(\omega_e t) \cos(\theta) \quad (2.1)$$

$$\mathcal{F}_b(\theta, t) = F i_b \cos\left(\theta - \frac{2\pi}{3}\right) = F_{max} \cos\left(\omega_e t - \frac{2\pi}{3}\right) \cos\left(\theta - \frac{2\pi}{3}\right) \quad (2.2)$$

$$\mathcal{F}_c(\theta, t) = F i_c \cos\left(\theta + \frac{2\pi}{3}\right) = F_{max} \cos\left(\omega_e t + \frac{2\pi}{3}\right) \cos\left(\theta + \frac{2\pi}{3}\right) \quad (2.3)$$

donde: F es una constante y depende del devanado y θ el ángulo de par, medido a partir del eje magnético de la fase a.

La fuerza magnetomotriz de armadura medida a un ángulo θ es dada por la suma de las fuerzas magnetomotrices por fase:

$$\mathcal{F}_{ar}(\theta, t) = \mathcal{F}_a(\theta, t) + \mathcal{F}_b(\theta, t) + \mathcal{F}_c(\theta, t) \quad (2.4)$$

Sustituyendo las ecuaciones (2.1),(2.2) y (2.3) en la ecuación (2.4) se obtiene:

$$\mathcal{F}_{ar}(\theta, t) = \frac{3}{2}F_{max} \cos(\theta - \omega_e t) \quad (2.5)$$

La ecuación (2.5) describe a una onda viajera de valor pico $\frac{3}{2}F_{max}$ y que se mueve en el entrehierro a una velocidad de $\omega_e \frac{2}{P}$ rad/s. Para una máquina de P polos, la velocidad expresada en rpm viene dada por:

$$N_s = \frac{120f}{P}$$

donde: N_s es la velocidad sincrónica en rpm, f la frecuencia de la corriente en hertz y P el número de polos.

En el rotor se inducirán tensiones cuyo valor es proporcional a la velocidad relativa con la que el campo magnético corta a los conductores $N_s - N_{rotor}$, luego al existir caminos cerrados circularan corrientes, en las barras (rotor jaula de ardilla) o en el devanado del rotor (rotor devanado), la cual interactuará con el campo produciendo un torque electromagnético. Nótese que para que pueda existir torque la velocidad del rotor no puede ser igual a la sincrónica, en otras palabras la velocidad relativa no debe ser cero, se define al deslizamiento como:

$$s = \frac{\omega_s - \omega_{rotor}}{\omega_s}$$

donde: s es el deslizamiento ω_s la velocidad angular del campo magnético en rad/s y ω_{rotor} la del rotor en rad/s.

La ecuación de torque electromagnético se puede formular con las variables del circuito eléctrico del rotor:

$$T_e = \frac{3}{\omega_s} I_r^2 \frac{R_2}{s}$$

donde: T_e es el torque eléctrico en N.m, I_r la corriente de una fase del rotor en amperios y R_2 la resistencia de una fase del rotor en ohmios.

2.2.4. Fallas en motores de inducción

El motor de inducción es una pieza vital en el proceso industrial y, generalmente, está sometido a ambientes y ciclos de trabajo severos. Muchas veces el motor se llega a deteriorar de manera tal de que la vida de sus partes es menor a la estimada por el fabricante.

Las fallas más importantes que aparecen en los motores de inducción se pueden dividir en (Neelam y Ratna, 2007):

- Fallas en cojinetes.
- Fallas en el rotor en donde se incluyen: fallas en barras, en devanados y roturas de anillos.
- Fallas relacionadas con excentricidad.
- Fallas en el estator.

La mayoría de las fallas que ocurren en los motores de inducción en la zona industrial del Estado Carabobo son en cojinetes (71, 2 %) (Mago y Monagas, 2004). A continuación se describirá esta falla la cual es de interés para este trabajo.

2.2.5. Fallas en cojinetes

La mayoría de los motores eléctricos usan cojinetes, estos consisten en dos anillos uno exterior y otro interior que contienen un conjunto de bolas o elementos rodantes, puestas en unas pistas dentro de los anillos como se muestra en la figura 2.1. Las fallas se clasifican de acuerdo a los elementos afectados (Hong, Yoon, y Kim, 2005):

- Falla en pista externa.
- Falla en pista interna.
- Falla en bola.

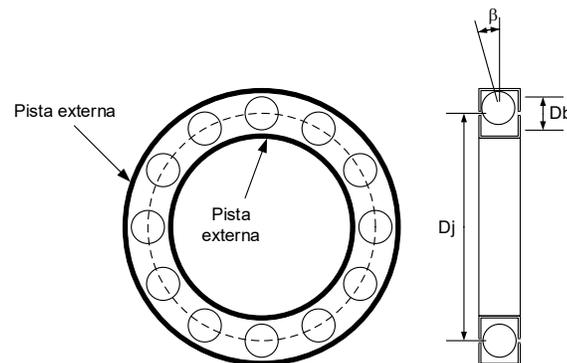


Figura 2.1: Dimensiones del rodamiento
(Schoen y otros, 1995)

Las fallas nombradas producen vibraciones causando que el entrehierro de la máquina no sea uniforme. Esta variación afecta el flujo magnético introduciendo componentes únicas de frecuencia en la corriente medida, las cuales, dependen de la geometría del cojinete y de su velocidad de giro dada por (Hong y otros, 2005),(Sin, Soong, y Ertugrul, 2003)(Benbouzid, 2000):

$$f_{cojinete} = |f_1 \pm m f_{1,0}| \quad (2.6)$$

donde: $m = 1, 2, \dots$ f_1 es la frecuencia de la corriente y $f_{1,0}$ es una frecuencia característica que depende de la geometría del rodamiento ver figura 2.1 y viene dada por la ecuación (2.7) para la pista interna:

$$f_{BPI} = \frac{N_b}{2} f_r \left(1 + \frac{D_b}{D_j} \cos \beta \right) \quad (2.7)$$

donde: N_b es el número de bolas del cojinete, f_r la frecuencia mecánica del rotor en hertz, D_b el diámetro de la bola, D_j el de la jaula y β el ángulo de contacto de las bolas con las pistas.

En la ecuación (2.7) f_{BPI} se refiere a la frecuencia con la que una bola toca un punto de la pista interna del cojinete, igualmente f_{BPE} se refiere a la frecuencia con que toca la pista externa y viene dada por la ecuación (2.8):

$$f_{BPE} = \frac{N_b}{2} f_r \left(1 - \frac{D_b}{D_j} \cos \beta \right) \quad (2.8)$$

Finalmente se determina la frecuencia de giro de la bola sobre su eje (f_B) según la ecuación (2.9):

$$f_B = \frac{D_j}{2D_b} f_r \left(1 - \frac{D_b^2}{D_j^2} \cos^2 \beta \right) \quad (2.9)$$

2.2.6. Monitoreo de las condiciones del motor de inducción

Los motores de inducción trifásicos son máquinas simétricas que funcionan en estado normal con campos magnéticos sinusoidales en el entrehierro. Cualquier falla que ocurra alterará de alguna manera estos campos y producirá uno o más de los siguientes síntomas (Wen, 2011):

- Desbalances de tensión y corrientes.
- Aumento en las pulsaciones de par.

- Disminución del par promedio.
- Disminución de la eficiencia.
- Calentamiento excesivo.

Con el monitoreo de las condiciones del motor se puede determinar el estado actual de la máquina. Este consiste en la medición de uno a varios parámetros relacionados con la condición mecánica del motor, estos se mencionan a continuación (Wen, 2011):

- Monitoreo eléctrico.
- Monitoreo de vibraciones.
- Monitoreo del par.
- Monitoreo de temperatura.
- Monitoreo químico.

En este trabajo las variables se obtuvieron con este tipo de monitoreo específicamente, por medio de medición de la corriente del estator. El análisis espectral de la corriente del motor es una de las técnicas más usadas para la detección de la falla del motor, este consiste en la observación de las componentes del espectro de corriente en donde al ocurrir alguna falla esta producirá cambios en las componentes. Este tipo de análisis requiere la presencia de un experto el cual distinguirá el tipo, la ubicación, la magnitud y el tiempo de la falla.

En los últimos años la detección de las fallas ha migrado hacia el uso de las técnicas de inteligencia artificial entre las más significativa se tienen: sistemas expertos, lógica difusa, ANN y SVM. Estas técnicas se usan en procesos en los que haya clasificaciones no lineales y usan asociación, razonamiento y decisiones, que son los procesos que usa el cerebro humano para resolver problemas de diagnóstico.

2.2.7. Máquinas de Vectores de Soporte

La máquina de vectores de soporte SVM (por sus siglas en inglés, Support Vector Machines) es un método general para la resolución de problemas de clasificación, regresión y estimación. La idea principal de las SVM es separar los datos por medio de una función, la cual es obtenida de una muestra de estos. A manera de ejemplo la figura 2.2 muestra como se separan datos de distinta clase por medio de una recta.

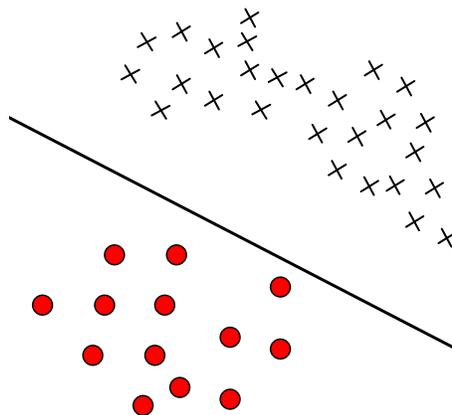


Figura 2.2: Los datos están separados por un plano clasificador
(Widodo y Yang, 2007)

El problema puede formularse de la siguiente manera se tiene un conjunto de datos los cuales pueden dividirse en dos grupos, a cada grupo se le asigna un valor o etiqueta; en la práctica se le asignan valores de $+1$ (para datos positivos) y -1 (para datos negativos). Para un conjunto de datos cada uno se representa por el vector (\mathbf{x}, y) en donde la componente \mathbf{x} tiene n atributos, y la componente y corresponderá a la etiqueta. Los valores de \mathbf{x} se les conocen como patrones o entradas. La clasificación se puede lograr usando la función $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ de la manera siguiente: a la entrada $\mathbf{x} = (x_1, x_2, \dots, x_n)$ se le asignará la etiqueta $+1$ si $f(\mathbf{x}) \geq 0$, en caso contrario se le asignará la etiqueta -1 (Widodo y Yang,

2007).

Considerando el caso en que $f(\mathbf{x})$ es una función lineal de $\mathbf{x} \in X$, se puede expresar de la siguiente forma:

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b$$

A \mathbf{w} se le conoce como el vector de pesos y a b como *bias*. Los puntos obtenidos de resolver la ecuación $\mathbf{w}^T \mathbf{x} = 0$ son todos aquellos perpendiculares a \mathbf{w} y que pasan por el origen, el valor de b lo que hace es trasladar estos puntos, esto se muestra en la figura 2.3 en donde se tiene un conjunto de datos en el plano para mejor visualización.

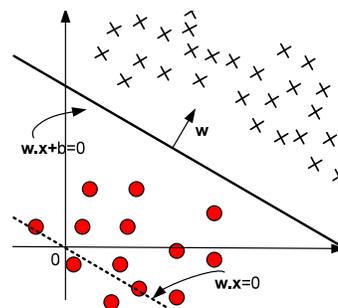


Figura 2.3: Función discriminante
(Widodo y Yang, 2007)

El hiperplano de separación óptimo es aquel que hace máxima la distancia entre el plano y las entradas más cercanas de las distintas clases como se muestra en la figura 2.4. Esta distancia se le llama margen de clasificación y viene dado por:

$$m_D(\mathbf{w}) = \frac{2}{\|\mathbf{w}\|}$$

Para determinar el mínimo margen de clasificación se debe resolver el siguiente problema de optimización:

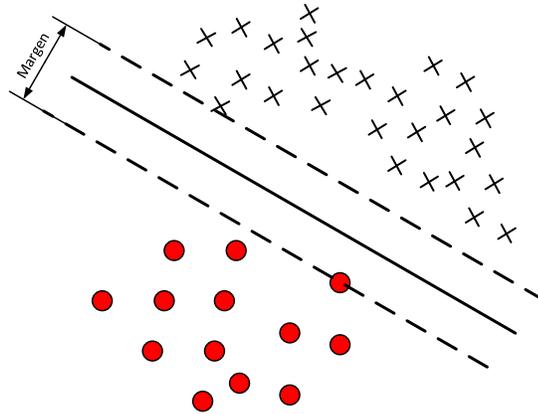


Figura 2.4: Clasificación con hiperplano. El margen de clasificación viene dado por la distancia entre los puntos de cada clase que son más cercanos al hiperplano (Widodo y Yang, 2007)

$$\begin{aligned} &\text{minimizar} && \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \\ &\text{condicionado a} && y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \quad i = 1, \dots, n \end{aligned}$$

donde: ξ_i es la distancia medida entre el margen y el punto \mathbf{x}_i que se encuentra en el lado incorrecto de este, C es un factor de penalización

Usando el método de los multiplicadores de Lagrange, podemos obtener la formulación *dual*, la cual es expresada en términos de la variable α_i (Benbouzid, 2000):

$$\begin{aligned} &\text{maximizar} && \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j \mathbf{x}_i^T \mathbf{x}_j \\ &\text{condicionado a} && \sum_{i=1}^n y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C \end{aligned}$$

La formulación dual conduce a una expansión del vector de pesos en función de los datos de entrada:

$$\mathbf{w} = \sum_{i=1}^n y_i \alpha_i \mathbf{x}_i$$

En la mayoría de las aplicaciones los datos no son linealmente separables, sin embargo se puede seguir usando un clasificador lineal el cual, es deseable porque, tiene la ventaja de algoritmos de entrenamiento rápidos. La clasificación se logra ahora transformando los datos del espacio de entrada X al espacio de características F de igual o mayor dimensión, en el que se supone una mayor separación entre las clases, usando la función no lineal ϕ (Gunn y otros, 1998).

Una vez que los datos están en el espacio F la función discriminante es:

$$f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$$

Con el propósito de reducir memoria y tiempo de cálculo el vector de pesos se expresa como una combinación lineal de los ejemplos de entrenamiento esto es $\mathbf{w} = \sum_{i=1}^l \alpha_i \phi(\mathbf{x}_i)$ entonces:

$$f(\mathbf{x}) = \sum_{i=1}^l \alpha_i k(\mathbf{x}, \mathbf{x}_i) + b$$

donde $k(\mathbf{x}, \mathbf{x}_i)$ se define como función kernel que para todo los $\mathbf{x}, \mathbf{z} \in X$ viene dada por:

$$k(\mathbf{x}, \mathbf{z}) = \phi(\mathbf{x})^T \phi(\mathbf{z})$$

Algunas funciones kernel comunes son:

1. kernel lineal:

$$k(\mathbf{x}, \mathbf{x}_i) = \mathbf{x}^T \cdot \mathbf{x}_i$$

2. kernel polinómico de grado d :

$$k(\mathbf{x}, \mathbf{x}_i) = (\gamma \mathbf{x}^T \cdot \mathbf{x}_i + r)^d \quad \text{con } \gamma > 0$$

3. Función básica radial (RBF):

$$k(\mathbf{x}, \mathbf{x}_i) = e^{-\gamma \|\mathbf{x} - \mathbf{x}_i\|^2} \quad \text{con } \gamma > 0$$

Como las SVM son clasificadores de dos clases existen varias estrategias para resolver problemas de clasificación múltiple, se explican las dos más utilizadas:

Todos contra uno: si hay k clases se crean k clasificadores de dos clases. Entonces el clasificador C_i se entrena con los datos de la clase i etiquetándolos positivos y los $i - 1$ restantes negativos. Una vez entrenados, un dato desconocido k quedará clasificado en la clase en la que se etiqüete positivo. Entre las desventajas de este método están: una entrada puede ser asignada a varias clases y que el entrenamiento es desequilibrado teniendo más datos en las clases negativas.

Uno contra uno: si hay k clases el número de clasificadores viene dado por la ecuación (2.10):

$$\text{Clasificadores} = k \left(\frac{k - 1}{2} \right) \quad (2.10)$$

Cada clasificador se entrena con dos clases cualquiera con la condición de que los otros no se pueden entrenar con las mismas. La clasificación de un dato k desconocido se logra por un proceso de votación en donde quedará en la clase que reciba más votos.

2.2.8. Redes Neuronales Artificiales

Las ANN son un sistema de procesamiento de datos consistente en un gran número de elementos simples, interconectados entre sí en una arquitectura similar a la corteza cerebral. Debido a la topología del sistema y a la manera en que la

información es almacenada y manipulada, las ANN realizan tareas que normalmente hacen los humanos y animales (Alguindigue, Loskiewicz-Buczak, y Uhrig, 1993).

Las siguientes propiedades permiten que las ANN sean una solución interesante para la vigilancia de algunos problemas en las industrias (Trigeassou, s.f.):

- Capacidad para resolver problemas relativos a la clasificación y reconocimiento de patrones.
- Alta habilidad de aprendizaje y generalización.
- Rapidez para el procesamiento de datos no lineales.

En la figura 2.5 se observa la arquitectura de una neurona artificial de R entradas, en donde cada una de estas es multiplicada por un peso para luego sumarlas y, el resultado, pasarla por una función de activación f .

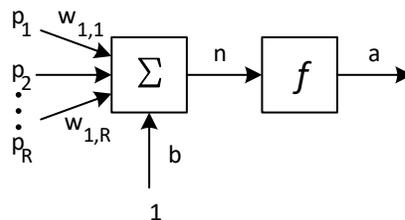


Figura 2.5: Neurona artificial

La neurona tiene un ajuste b (del inglés *bias*) el cual se le suma al resultado de la suma de las entradas multiplicadas por los pesos (Hagan, Demuth, Beale, y otros, 1996):

$$n = w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b$$

Este resultado se expresa en forma matricial de la manera siguiente:

$$n = \mathbf{W}\mathbf{p} + b$$

Finalmente la salida de una neurona se puede obtener con la función de activación:

$$a = f(\mathbf{W}\mathbf{p} + b)$$

Entre algunas funciones de activación más utilizadas se tienen:

Hard Limit: la salida solo puede tener los valores cero (0) y uno (1) bajo las siguientes condiciones:

$$\begin{aligned} a &= 0 & \text{si } n < 0 \\ a &= 1 & \text{si } n \geq 0 \end{aligned}$$

Symmetrical Hard Limit: igual al caso anterior pero con los valores de la salida menos uno (-1) y uno (1)

$$\begin{aligned} a &= -1 & \text{si } n < 0 \\ a &= 1 & \text{si } n \geq 0 \end{aligned}$$

Linear: la salida es igual a la entrada $a = n$

Log-Sigmoid la salida es una función suave entre cero (0) y uno (1)

$$a = \frac{1}{1 + e^{-n}}$$

Hyperbolic Tangent Sigmoid la salida es una función suave entre menos uno (-1) y uno (1)

$$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$$

Cuando se requieren varias salidas se pueden tener más de una neurona artificial según se observa en la figura 2.6, donde se tienen S neuronas. Una configuración como la mostrada se le conoce como capa de S neuronas.

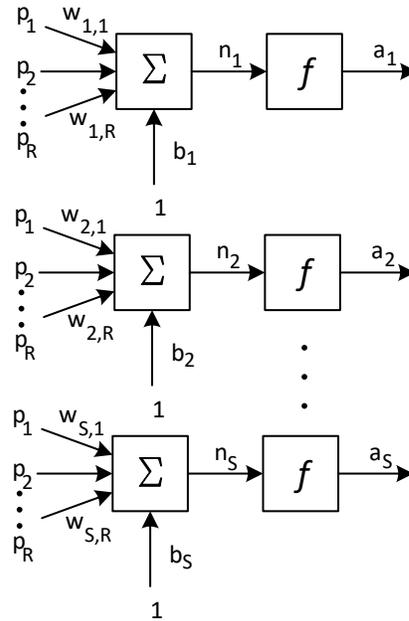


Figura 2.6: Capa de S neuronas

Cada elemento del vector de entradas \mathbf{p} es conectado a cada neurona después de multiplicarlo por el peso w . Como por cada neurona hay R entradas debe haber la misma cantidad de pesos por cada una, sin embargo, el valor de estos puede diferir entre una neurona y otra, por lo que el vector de pesos es una matriz de dimensiones $R \times S$:

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,R} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,R} \\ \vdots & \vdots & & \vdots \\ w_{S,1} & w_{S,2} & \cdots & w_{S,R} \end{bmatrix}$$

Las ANN pueden contener más de una capa, por ejemplo en el caso de dos capas las salidas de la primera son las entradas de la segunda. Aquella capa cuyas salidas sean las de la red se le conoce como *capa de salida*, toda las otras se les llama *capas ocultas* (Hagan y otros, 1996).

Para determinar los valores de los pesos y bias existen algoritmos de entrenamiento, uno de los más utilizados es el *Backpropagation*, el cual es una generalización de la regla de aprendizaje de Widrow-Hoff, y es usado para redes multicapas. Se pueden usar redes de dos capas, con una función de transferencia Log-Sigmoid para la capa oculta y una Lineal para la de salida, para aproximar cualquier función con un finito número de discontinuidades (Bagheri, Ahmadi, y Labbafi, 2010).

2.2.9. Análisis discriminante

El planteamiento estadístico de la función discriminante es el siguiente (Peña, 2002): se dispone de un conjunto amplio de elementos que pueden venir de dos o más poblaciones distintas, en cada elemento se ha observado una variable aleatoria p -dimensional \mathbf{Y} , cuya población se conoce en las poblaciones consideradas. Se desea clasificar un nuevo elemento, con valores de las variables conocidas, en una de las poblaciones. Hay dos grandes objetivos en la separación de grupos (Rencher, 2003):

1. Descripción de la separación de grupos, en el cual se usan funciones lineales (discriminantes) para describir las diferencias entre dos o más grupos.
2. Predicción o localización de observaciones a grupos, en el cual se emplean funciones lineales o cuadráticas (de clasificación) para asignar una medición individual a uno de los grupos.

Para determinar la función discriminante se debe conocer previamente la matriz de covarianza de las poblaciones, a continuación se explica como se obtiene esta matriz y la discriminación de los datos:

Sea \mathbf{y} un vector de p características asociado a una medición. Si hay n muestras individuales, los n vectores de observación se denotan por $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ donde

$$\mathbf{y}_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{ip} \end{pmatrix}$$

El vector de medias $\bar{\mathbf{y}}$ se puede determinar como el promedio de los vectores de observación.

$$\bar{\mathbf{y}} = \frac{1}{n} \sum_{i=1}^n \mathbf{y}_i$$

La matriz de covarianza tiene dimensiones p^2 y viene dada por:

$$\mathbf{S} = (s_{jk}) = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{pmatrix} \quad (2.11)$$

La covarianza se puede determinar con los vectores de observación y el vector de medias según la ecuación (2.12):

$$\mathbf{S} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}_i - \bar{\mathbf{y}})' \quad (2.12)$$

En el caso de dos poblaciones se requiere que las covarianzas sean iguales pero, si se obtiene de poblaciones de distintas cantidad de datos, se puede determinar una matriz de covarianza combinada:

$$\mathbf{S}_{pl} = \frac{\sum_{i=1}^k (n_i - 1) \mathbf{S}_i}{\sum_{i=1}^k (n_i - 1)} \quad (2.13)$$

donde: S_{pl} es la matrix de covarianza combinada, S_i la matrix de covarianza de la población i , n_i es la cantidad de mediciones realizadas en la población i .

La función discriminante es una combinación lineal de las p variables que maximiza la distancia de las dos poblaciones transformadas. Una combinación lineal $z = \mathbf{a}'\mathbf{y}$ transforma cada vector de observación a un escalar:

$$\begin{aligned} z_{1i} &= \mathbf{a}'\mathbf{y}_{1i} = a_1y_{1i1} + a_2y_{1i2} + \dots + a_py_{1ip}, & i = 1, 2, \dots, n_1 \\ z_{2i} &= \mathbf{a}'\mathbf{y}_{2i} = a_1y_{2i1} + a_2y_{2i2} + \dots + a_py_{2ip}, & i = 1, 2, \dots, n_2 \end{aligned}$$

Se busca un vector a que maximice la diferencia estandarizada $(\bar{z}_1 - \bar{z}_2)/s_z$, pero debido a que esta puede dar valores negativos se usa la distancia cuadrada $(\bar{z}_1 - \bar{z}_2)^2/s_z^2$ la cual puede ser expresada como:

$$\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2} = \frac{[\mathbf{a}'(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)]^2}{\mathbf{a}'\mathbf{S}_{pl}\mathbf{a}} \quad (2.14)$$

El máximo de (2.14) ocurre cuando

$$\mathbf{a} = \mathbf{S}_{pl}^{-1}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2) \quad (2.15)$$

Dado un nuevo dato \mathbf{y} este se clasificará en el grupo G_1 si:

$$\mathbf{a}'\mathbf{y} = (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{S}_{pl}^{-1}\mathbf{y} > \frac{1}{2}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{S}_{pl}^{-1}(\bar{\mathbf{y}}_1 + \bar{\mathbf{y}}_2) \quad (2.16)$$

Si no se cumple la inecuación (2.16) entonces la observación se clasificará en G_2 . Si se conocen las probabilidades previas p_1 y p_2 de las poblaciones, entonces la regla de clasificación se puede modificar de la siguiente manera para obtener mejores resultados:

$$\mathbf{a}'\mathbf{y} = (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{S}_{pl}^{-1}\mathbf{y} > \frac{1}{2}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{S}_{pl}^{-1}(\bar{\mathbf{y}}_1 + \bar{\mathbf{y}}_2) + \ln\left(\frac{p_2}{p_1}\right)$$

donde: p_1 es la proporción de datos que corresponden a G_1 y p_2 la proporción que corresponde a G_2

Si las matrices de covarianzas de las poblaciones no son iguales y se conocen las probabilidades previas p_1 y p_2 se usa la regla de clasificación cuadrática:

$$Q_i(\mathbf{y}) = \ln p_i - \frac{1}{2} \ln |\mathbf{S}_i| - \frac{1}{2} (\mathbf{y} - \bar{\mathbf{y}}_i)' \mathbf{S}_i^{-1} (\mathbf{y} - \bar{\mathbf{y}}_i)$$

donde: $i = 1, 2$ para el caso de dos poblaciones

Entonces \mathbf{y} se asignará a la población en el que se obtenga el máximo valor de $Q_i(\mathbf{y})$.

Capítulo 3

MARCO METODOLÓGICO

3.1. Tipo de estudio y de investigación

El presente trabajo consiste en una investigación del tipo interactiva, la cual es aquella cuyo objetivo consiste en modificar el evento estudiado, generando y aplicando sobre él una intervención especialmente diseñada (Hurtado, 2007)

Considerando que los datos serán tomados de motores, que se encuentran en el Laboratorio de Mantenimiento Predictivo de la UC, y se modificaran las variables independientes, el tipo de estudio es contemporáneo transeccional de laboratorio.

3.2. Población y muestra

Dos motores de inducción trifásicos constituyen la población de la cual se midió la corriente de armadura. Las características principales de estos motores se muestran en la tabla 3.1

La carga del motor consistió en un generador DC el cual alimentó un banco de resistencias. En este trabajo no se consideraron cargas cuyo par varíen en el tiempo.

Para la simulación de las fallas en las pistas interna y externa se tomaron dos cojinetes en buen estado y se realizó una perforación en la pista interna de uno

Tabla 3.1: Características del motor utilizado

Dato	Valor
Marca:	Siemens
Potencia:	2 hp – 1,5 kW
Tensión Nominal:	220 V YY – 440 V Y
Corriente Nominal:	7 A YY – 3,5 A Y
Frecuencia:	60 Hz
Velocidad:	1700 rpm
Factor de Servicio:	1,15
Factor de potencia:	0,8
Clase de Aislamiento:	<i>F</i>
Servicio:	<i>S1</i>
Rotor:	Jaula de ardilla
Cojinetes:	6205-2RSH

ver figura 3.1, y en la pista externa del otro cojinete figura 3.2, para ello se utilizó una mecha WIDEA de 3 mm de diámetro (Gonzales y J, 2009).



Figura 3.1: Cojinete con falla en pista interna



Figura 3.2: Cojinete con falla en pista externa

Por otro lado la falla en los elementos rodantes se obtuvo tomando un cojinete en buen estado y lubricándolo con una grasa especial que contiene virutas de hierro, lo cual genera roce y desgaste en el cojinete a medida que el motor aumenta su velocidad.

3.3. Procedimiento metodológico a utilizar

Los pasos seguidos para lograr el objetivo de esta investigación se describen a continuación:

Fase I Determinar las características de los diferentes métodos emergentes (ANN, SVM y análisis discriminante) que se utilizarán para la distinción de las fallas:

- Se analizaron las configuraciones de las ANN para seleccionar la más eficiente en la detección del estado del motor.

- Se estudió las SVM para establecer las características apropiadas a la investigación
- Se establecieron las características de las funciones discriminantes canónicas a utilizar en la diferenciación de las fallas en cojinetes

Fase II Analizar el espectro de la corriente estatórica en motores de inducción jaula de ardilla por medio de los métodos emergentes (ANN, SVM y análisis discriminante) para determinar las condiciones de los cojinetes.

- Se realizaron las pruebas para obtener los datos de motores de inducción con cojinetes, en perfecto estado y con diferentes tipos de falla, para establecer los patrones de datos.
- Se entrenaron las ANN y de las SVM, con los datos que se obtuvieron para motores en perfecto estado y con diferentes tipos de fallas en los cojinetes.
- Se verificó si las ANN y las SVM discriminan en forma correcta entre cojinetes sanos o con falla a partir de los datos que no hayan sido utilizados en el entrenamiento.
- Se construyeron las funciones discriminantes usando un paquete de estadística.
- Se discriminaron si las funciones discriminantes diferencian correctamente el estado de los cojinetes de un motor de inducción, a través de los datos del espectro de corriente estatórica.

Fase III Comparar los métodos emergentes (ANN, SVM y análisis discriminante) para determinar cuál discrimina mejor las condiciones de los cojinetes de motores de inducción jaula de ardilla.

- Establecer las técnicas de comparación para las técnicas de datos utilizadas.

- Identificar la técnica de análisis de datos que indica una mejor proporción del estado de falla en los cojinetes de un motor de inducción a través del espectro de corriente estatórica.

3.4. Técnicas e instrumentos de recolección de datos

El procedimiento utilizado para diagnosticar fallas en los rodamientos del motor de inducción consta de cuatro pasos a saber: adquisición de los datos, extracción de características, selección de características y clasificación, esto se muestra en la figura 3.3.

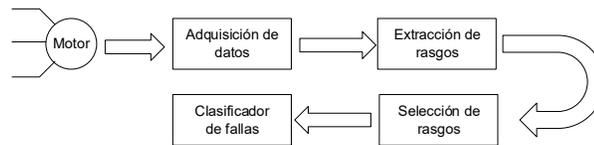


Figura 3.3: Arquitectura de diagnóstico de falla en cojinetes

Entre las técnicas que se utilizaron para la recolección de los datos están la medición de la corriente estatórica, y de la velocidad del rotor de los motores de inducción, que se encuentran en el laboratorio de mantenimiento predictivo de la Universidad de Carabobo y que conforman la muestra seleccionada para el estudio.

Las señales fueron obtenidas por medio de una tarjeta de adquisición de datos modelo USB-1208FS, desarrollado por la empresa Americana National Instruments, expandida en el 2006 por la adquisición de Measurement Computing; compañía que desarrolla tarjetas de adquisición de datos de bajo costo para interfaces ISA, PCI y USB para computadoras personales. La figura 3.4 muestra el esquema de montaje realizado.

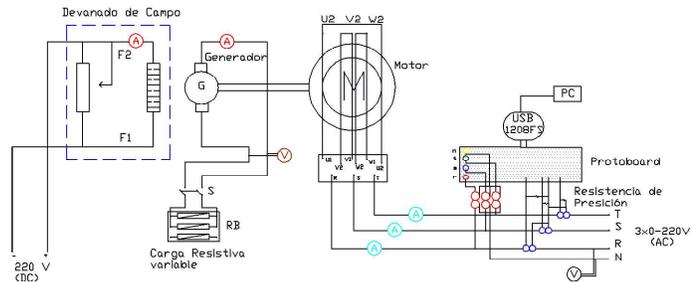


Figura 3.4: Esquema del montaje

3.5. Técnicas de procesamiento y análisis de datos

La información recolectada anteriormente fue almacenada en un computador en archivos de texto, todo esto por medio del software LABVIEW el cual viene incluido con la tarjeta de adquisición de datos. Se almacenaron tanto las variables de corriente en el dominio del tiempo como en el espectro en frecuencia. Luego de extraer ciertas características, se seleccionaron las mejores y finalmente se utilizan los clasificadores de fallas ver figura 3.3. El procesamiento de datos fue realizado por medio del software libre de programación PYTHON (Python Software Foundation, <https://www.python.org/>), en este se encuentra la librería scikit-learn (Pedregosa y otros, 2011), la cual se utilizó para entrenar las ANN y SVM, y a su vez el análisis lineal discriminante.

3.5.1. Extracción de rasgos de la corriente de armadura

Las variables extraídas, ya sea en el dominio del tiempo o de la frecuencia, puede resultar en una carga muy pesada para los clasificadores. Estos indudablemente realizarán de forma adecuada su tarea, pero los tiempos de computación aumentarán en la medida de que los datos aumenten; además se puede estar suministrando información que sea irrelevante en la clasificación.

Con el propósito de reducir la cantidad de información suministrada a los

clasificadores de falla se extrajeron métricas estadísticas o rasgos de la corriente de armadura, las cuales reflejan el contenido físico de la señal en el tiempo. Las siguientes características se extrajeron en el proceso de monitoreo (Jiménez y Jiménez, 2010; Bin Hasan, 2012):

- Media (μ): conocida como media aritmética, es la suma de los valores divididos entre el número de éstos tal como se muestra en la ecuación (3.1):

$$\mu = \frac{\sum X}{n} \quad (3.1)$$

- Raíz media cuadrática (RMS): es la raíz de la media de los cuadrados de los valores, ver ecuación (3.2):

$$RMS = \sqrt{\frac{\sum X^2}{n}} \quad (3.2)$$

- Media cuadrática (MC): es la media de la raíz cuadrada de la suma de los datos, ver ecuación (3.3):

$$MC = \frac{\sqrt{\sum X^2}}{n} \quad (3.3)$$

- Mediana (MD): es el valor central de un conjunto de datos obtenido una vez que estos se han ordenado de forma ascendente o descendente.
- Varianza (σ^2): es una medida de los cuadrados de la distancia de cada valor con respecto a su media tal como se muestra en la ecuación (3.4):

$$var = \sigma^2 = \frac{\sum (x - \mu)^2}{n - 1} \quad (3.4)$$

- Desviación estándar (σ): es una medida de la energía del contenido de la señal y se determina como la raíz cuadrada de la varianza ver ecuación (3.5):

$$sd = \sigma = \sqrt{\frac{\sum (x - \mu)^2}{n - 1}} \quad (3.5)$$

- Valor pico (X_{peak}): es la media de dos picos (máximo y mínimo) de todos los ciclos medidos y se determina según ecuación (3.6):

$$X_{peak} = \sum_{i=1}^N \left(\frac{|max_i| + |min_i|}{2N} \right) \quad (3.6)$$

donde: N es el número de ciclos

- Error estándar de la media (sdm): es la desviación estándar de la distribución de errores o fluctuaciones aleatorias y viene dado por la ecuación (3.7)

$$sdm = \frac{sd}{\sqrt{n}} \quad (3.7)$$

- Curtosis: mide cuan puntiaguda es una distribución con respecto a un estándar, se determina por la ecuación (3.8):

$$curtosis = \frac{\frac{1}{n} \sum (x - \mu)^4}{\left(\frac{1}{n} \sum (x - \mu)^2 \right)^2} \quad (3.8)$$

- Asimetría (SK): mide el grado de asimetría de la distribución respecto a su media, se determina según la ecuación (3.9):

$$SK = \frac{n}{n-1} \sum \frac{(x - \mu)^3}{sd} \quad (3.9)$$

- Factor de cresta (FC): es la razón entre el valor pico y el valor rms, se determina según la ecuación (3.10):

$$FC = \frac{X_{peak}}{RMS} \quad (3.10)$$

- Mínimo valor: es el mínimo valor de un conjunto de datos.
- Máximo valor: es el valor máximo valor de un conjunto de datos.
- Suma: es la suma de todos los valores de una señal.
- Rango: es la diferencia entre el máximo valor y el mínimo valor.

3.5.2. Normalización de la data

Normalizar la data se refiere a llevar todos los atributos a valores comprendidos entre -1 y +1 o entre 0 y 1. Esto debe realizarse porque los clasificadores son sensibles a la diferencia de orden de magnitud entre los atributos; por ejemplo supóngase que se tienen dos atributos, el valor rms que varía entre 4 y 7 amperios y la desviación estándar que se encuentra entre 0 y 1, debido a que estos últimos son mucho más pequeños en proporción a los primeros, el clasificador podría ser incapaz de aprender correctamente ante variaciones de este atributo. Adicionalmente las SVM requieren que los datos estén centrados alrededor de cero.

En este trabajo se utilizó la normalización por desviación estándar, en ésta a cada atributo se le resta la media y el resultado se divide entre la desviación estándar según se muestra en la ecuación (3.11).

$$v_{norm}[i] = \frac{v[i] - \mu}{\sigma} \quad (3.11)$$

donde: $v[i]$ es la característica i , $v_{norm}[i]$ es su valor normalizado, μ es su valor medio y σ es su desviación estándar

Al restar la media se garantiza que los datos están centrados alrededor de cero y, al dividir entre la desviación estándar se asegura que todos tendrán una varianza parecida.

3.5.3. Reducción de atributos

Al tener datos con un número alto de rasgos es probable que se tenga exceso de información o inclusive redundancia, para estos casos existen métodos para reducir la dimensión de los datos con la mínima pérdida de información. En el caso de datos numéricos, como ocurre en este trabajo, entre las técnicas más usadas se tienen: Análisis de Componentes Principales, Análisis Factorial y Transformada de Wavelet (Pérez y Santín, 2007).

Suponga que los datos consisten en un conjunto de vectores con n atributos, el Análisis de Componentes Principales (PCA) (de sus siglas en inglés, Principal Components Analysis) busca vectores ortogonales de k dimensiones (con $k \leq n$) que puedan ser usados para representarlos. Los datos originales quedan así proyectados en un espacio reducido resultando en una reducción de su dimensión. El procedimiento básico es el siguiente (Pérez y Santín, 2007):

1. Se normalizan los datos. Esto garantiza que los atributos de mayor valor no prevalezcan sobre los de menor valor.
2. Se determinan los vectores ortogonales. Estos se conocen como componentes principales (PC_1, PC_2, \dots, PC_n). Los datos de entrada se expresan como combinación lineal de estos vectores. En un principio, la dimensión de las componentes principales es igual a la de los datos originales.

$$\begin{aligned} PC_1 &= a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n \\ &\vdots \\ PC_n &= a_{n1}X_1 + a_{n2}X_2 + \dots + a_{nn}X_n \end{aligned}$$

donde: PC_i son las componentes principales, X_i los rasgos originales y a_{ij} los pesos de las variables

3. Las componentes principales se ordenan en orden decreciente. La primera componente será la de mayor varianza, la segunda será la siguiente con mayor varianza y así sucesivamente.
4. El tamaño de los datos se puede reducir eliminando las componentes de menor varianza.

La primera componente (PC_1) se obtiene de forma que su varianza sea máxima sujeta a la restricción de que la suma de los pesos a_{1j} al cuadrado sea igual a la unidad:

$$\max V(PC_1) \quad \text{sujeto a} \quad \sum_{j=1}^n a_{1j}^2 = 1$$

Se puede demostrar que para maximizar $V(PC_1)$ se toma el mayor valor propio (λ_1) de la matriz de covarianzas \mathbf{S} ver ecuaciones (2.11) y (2.12), para maximizar $V(PC_2)$ se toma el segundo mayor valor propio (λ_2) y así sucesivamente.

Cuando las variables originales están muy correlacionadas entre sí, la mayor parte de su variabilidad se puede expresar con muy pocas componentes. Si las variables originales estuvieran completamente no correlacionadas entre sí, entonces el PCA no tendría ningún sentido, ya que las componentes principales coincidirían con las originales.

Para escoger las componentes a retener se seleccionan aquellas que tengan las varianzas mas altas, de manera de que el intervalo de confianza sea mayor al 80 %. Cabe destacar que las varianzas se obtienen de la diagonal de \mathbf{S} .

3.5.4. Selección de características

Antes de que un grupo de características alimenten un clasificador, se deben seleccionar las que provean información relevante respecto a la falla, y las características irrelevantes o redundantes se deben descartar, esto se hace con el propósito de reducir las dimensiones y mejorar el funcionamiento del clasificador (Lei, He, y Zi, 2008).

Existen diferentes técnicas que realizan esta tarea entre algunas se tienen: algoritmo genético, entropía condicional y evaluación de distancias . En este trabajo se utilizará la técnica Eliminación Recursiva de Características (RFE), la cual calificará las características extraídas. En este método, la mejor característica tomará el valor de 1 como valor de RFE, el cual varía entre 0 y 1.

3.5.5. Clasificación del estado del cojinete

Como se explicó anteriormente se utilizó la librería `scikit-learn` de PYTHON para el análisis de datos. Esta librería incluye los siguientes paquetes:

feature_selection: con esta se extrajeron los rasgos.

decomposition: para realizar el análisis de componentes principales PCA.

svm: necesario para realizar el entrenamiento y uso de las SVM.

neural-network: usado en el estudio de las redes neuronales.

discriminant-analysis: usado en el estudio con el ALD.

metrics: con ésta se determinan las métricas que sirven para evaluar a los clasificadores.

Capítulo 4

RESULTADOS

En el presente capítulo se muestran los resultados obtenidos en esta investigación según el siguiente orden: data adquirida, preparación de la data y clasificación.

4.1. Data adquirida

Se hicieron mediciones de la corriente de armadura al motor con distintos niveles de carga. Cada muestreo de la corriente de armadura consiste en la adquisición de su magnitud en el dominio del tiempo, específicamente en 1 segundo y para un tiempo de muestreo igual a $500 \mu\text{s}$ se obtuvieron 2000 puntos. La figura 4.1 muestra la forma de onda correspondiente a uno de esos muestreos, para obtenerla se tomaron los primeros 33 puntos.

Por cada muestreo de corriente se generó un archivo de texto, el cual tiene dos columnas. La primera de ellas es el número de la muestra y se enumera a partir de 0, la segunda es el valor de la corriente que se está midiendo. Para ilustrar esto se tiene la tabla 4.1, en esta se muestran los valores de la corriente correspondientes a los primeros 10 puntos de la curva 4.1.

El proceso de adquisición de datos es el siguiente: inicialmente con el motor funcionando en vacío (corriente de armadura de 3 A) y con el cojinete en perfecto estado se hicieron 30 muestreos de la corriente de armadura, obteniéndose 30

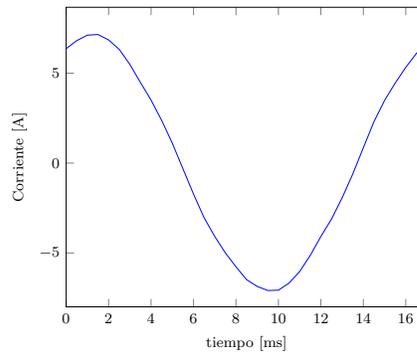


Figura 4.1: Muestreo de corriente en el dominio del tiempo

Tabla 4.1: Primeros 10 puntos de la figura 4.1

Muestra	Corriente (A)
0,000000	6,367314
1,000000	6,816880
2,000000	7,119648
3,000000	7,156348
4,000000	6,853579
5,000000	6,330615
6,000000	5,495708
7,000000	4,486479
8,000000	3,504775
9,000000	2,376274

archivos como los descritos anteriormente. Seguidamente se aumentó la carga hasta alcanzar 5 A y se hicieron 30 muestreos más. Finalmente se aumentó la carga hasta 7 A para las últimas mediciones. En total se tienen 90 muestreos de a corriente de armadura para a condición de cojinete en buenas condiciones, a los que se etiquetaron como ‘Sano’.

El proceso anterior se repite para cojinetes con distintos tipos de fallas: en

bola, pista interna y externa. Las etiquetas asignadas fueron: falla en bola ‘FB’, falla en pista interna ‘FPI’ y falla en pista externa ‘FPE’. La figura 4.2 muestra como se clasificaron los datos.

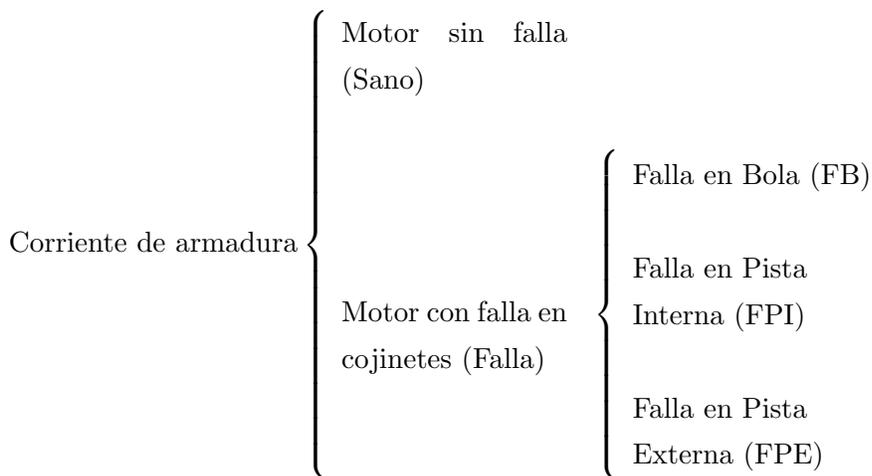


Figura 4.2: Clasificación de los datos

4.1.1. Extracción de rasgos en el dominio del tiempo

Para probar el comportamiento de los clasificadores ante las fallas se realizaron dos pruebas de clasificación: binaria y múltiple. La primera consiste en verificar que tan acertados son para distinguir si hay fallas o no, mientras que la segunda su capacidad para identificar el tipo de falla.

Para realizar la clasificación binaria del universo de datos (360) se seleccionaron 180 muestreos de la corriente del estator divididos de la manera siguiente: 90 con el cojinete en perfecto estado y 90 con las distintas fallas. Los datos se dividen en dos grupos uno de 75% datos para entrenar a los clasificadores y otro de 25% para probarlas, el primer grupo a su vez se subdivide en dos grupos, de 56% y 19% para entrenamiento y validación respectivamente, y con ello usar la técnica de validación cruzada. Esto se describe en la tabla 4.2. Desde el punto de vista matricial se tiene que para el motor ‘Sano’ hay 90 casos con 2000 atributos

(90×2000) y con falla 90×200 .

Tabla 4.2: *Descripción de los datos. Clasificación binaria*

Condición	Entrenamiento	Validación	Prueba	Total
Sano	50	22	18	90
Falla	50	22	18	90
Total	100	44	36	180

En la clasificación múltiple la data completa se divide en cuatro clases todas de 90×2000 , igualmente se dividen en 56 % para entrenar, 19 % validación y 25 % prueba ver tabla 4.3.

Tabla 4.3: *Descripción de los datos. Clasificación múltiple*

Condición	Entrenamiento	Validación	Prueba	Total
Sano	50	22	18	90
FB	50	22	18	90
FPE	50	22	18	90
FPI	50	22	18	90
Total	200	88	72	360

Las matrices de datos en cualquier caso son de 2000 atributos . Esta cantidad de datos en la actualidad, con la velocidad de los procesadores de las computadoras, no se considera una carga muy pesada para los clasificadores, pero si se quisiera aumentar el tiempo de muestreo, la cantidad de atributos aumentaría y llegará un momento en que será necesaria su reducción. Además está el inconveniente de que no todos los valores son relevantes para los clasificadores, es más se corre el riesgo de confundirlos. Con el fin de reducir el número de atributos, y por lo tanto la cantidad de datos, para cada caso o instancia se extraen 11 métricas estadísticas, descritas en el capítulo 3 (sección 3.5.1) : la mediana, valor medio,

valor rms, varianza, desviación estándar, valor pico, curtosis, asimetría, factor de cresta, máximo valor y rango.

La tabla 4.4 se muestran las características extraídas, en donde se plasman solo cuatro casos de la data. Para obtener por ejemplo las once características correspondientes a la condición ‘Sano’ y con carga de 5 A, se toma de la matriz de datos una fila cualquiera entre la 31 y la 60, obteniéndose un vector fila de 1×2000 . Finalmente a este vector se le aplican las ecuaciones de la sección 3.5.1.

Tabla 4.4: *Características extraídas en cada condición de operación*

Nº	Características	Sano		Falla en Bola(FB)	
		5 A	7 A	5 A	7 A
1	Media	0,0306	0,0164	0,0246	0,0192
2	FC	1,4214	1,4149	1,4164	1,4098
3	Curtosis	1,5033	1,4999	1,5030	1,4992
4	Mediana	0,0275	0,0917	0,0184	0,0000
5	RMS	4,9966	7,0261	5,0485	7,0549
6	Asimetría	-0,0002	0,0032	0,0019	0,0011
7	Desv Est	4,9978	7,0278	5,0497	7,0566
8	Varianza	24,978	49,390	25,499	49,7955
9	Max Val	7,1930	10,074	7,2114	10,0556
10	Rango	14,322	20,064	14,395	20,0469
11	Val Pico	7,1020	9,9412	7,1508	9,9461

De las características seleccionadas se pueden descartar las que no contribuyan a la clasificación correcta, para ello se hará uso de la matriz de covarianza que se muestra en la tabla 4.5, la numeración viene dada según la tabla 4.4. Con esta matriz se puede observar los datos que están muy correlacionados entre sí.

Se observa en la tabla 4.5 que la media no está relacionada con ninguna otra variable, pero por ejemplo las últimas 5 características están muy correlacionadas

Tabla 4.5: *Matriz de covarianza*

Nº	1	2	3	4	5	6	7	8	9	10	11
1 ^a	1,00	-0,03	-0,06	0,38	0,02	-0,21	0,02	0,02	0,02	0,02	0,02
2	-0,03	1,00	0,89	-0,09	-0,80	0,23	-0,80	-0,76	-0,79	-0,79	-0,79
3	-0,06	0,89	1,00	0,03	-0,79	0,02	-0,79	-0,76	-0,78	-0,78	-0,78
4	0,38	-0,09	0,03	1,00	-0,03	-0,64	-0,03	-0,03	-0,03	-0,03	-0,03
5	0,02	-0,80	-0,79	-0,03	1,00	-0,02	1,00	1,00	1,00	1,00	1,00
6	-0,21	0,23	0,02	-0,64	-0,02	1,00	-0,02	-0,02	-0,01	-0,02	-0,02
7	0,02	-0,80	-0,79	-0,03	1,00	-0,02	1,00	1,00	1,00	1,00	1,00
8	0,02	-0,76	-0,76	-0,03	1,00	-0,02	1,00	1,00	1,00	1,00	1,00
9	0,02	-0,79	-0,78	-0,03	1,00	-0,01	1,00	1,00	1,00	1,00	1,00
10	0,02	-0,79	-0,78	-0,03	1,00	-0,02	1,00	1,00	1,00	1,00	1,00
11	0,02	-0,79	-0,78	-0,03	1,00	-0,02	1,00	1,00	1,00	1,00	1,00

^a Numeración del los rasgos según la tabla 4.4

entre sí por lo que se puede pensar en descartar algunas. Para seleccionar los rasgos que no tengan dependencias entre sí se usó el Análisis de la Varianza (ANOVA) cuyo resultado se muestra en la tabla 4.6.

Del resultado de la ANOVA se seleccionaron 5 rasgos: Media, FC, Curtosis, Asimetría y Desv Est. Finalmente después de la extracción de rasgos en el dominio del tiempo y su posterior reducción, se tiene una matriz de datos para la clasificación binaria de 180×5 y para la múltiple 360×5 .

Al usar el ANOVA no se consideró el comportamiento de los clasificadores sino únicamente a los datos. Puede ocurrir que los clasificadores sean sensibles a algunas de las variables descartadas sin importar que sean redundantes o no. Para determinar el peso que cada variable pueda tener sobre la clasificación se usó una selección interactiva de características, en especial la RFE por sus siglas en inglés. Con la RFE se construye un modelo y partiendo de todos los rasgos se descartan los menos influyentes; se repite el proceso con otro modelo partiendo

Tabla 4.6: *Selección de rasgos por dos métodos distintos*

Nº	Rasgo	ANOVA	RFE
1	Media	Mantener	Descartar
2	FC	Mantener	Descartar
3	Curtosis	Mantener	Mantener
4	Mediana	Descartar	Descartar
5	RMS	Descartar	Mantener
6	Asimetría	Mantener	Descartar
7	Desv Est	Mantener	Mantener
8	Varianza	Descartar	Mantener
9	Max Val	Descartar	Descartar
10	Rango	Descartar	Descartar
11	Val Pico	Descartar	Mantener

de todas las características, al llegar a criterio de fiscalización se seleccionan las características. Los rasgos seleccionados por este método se muestran en la cuarta columna de la tabla 4.6, se observa que algunos rasgos que con el ANOVA se deben mantener, como por ejemplo el factor de cresta, ahora se descartan por ser irrelevantes para la clasificación; mientras que otros rasgos que se descartaban por estar correlacionados con otros ahora se deben mantener.

4.1.2. Extracción de componentes en el dominio de la frecuencia

Con el propósito de comparar la respuesta de los clasificadores ante distintos tipos de entradas se obtuvo el espectro en frecuencia a la señal de corriente en el dominio del tiempo, esto por medio del uso de la transformada rápida de Fourier. Cuando se obtiene este espectro hay que asegurarse que incluya frecuencias que permitan detectar las distintas fallas en los cojinetes $f_{cojinete}$, estas dependen

según la ecuación (2.6) de la frecuencia de la red f_1 en este caso 60 Hz y de las frecuencias características de falla $f_{1,0}$ estudiadas en el capítulo 2 sección (2.2.5), para determinar éstas últimas se obtienen partiendo de los datos de los cojinetes usados en los experimentos y mostrados en la tabla 4.7.

Tabla 4.7: *Datos del cojinete*

Parte	Símbolo	Dimensión
Jaula (diámetro)	D_j	3,8 cm
Bola (diámetro)	D_b	0,5 cm
Bola (cantidad)	N_b	9

La tabla 4.8 muestra las frecuencias características, con la que una bola se mueve en el cojinete $f_{1,0}$, y la de falla en cojinete $f_{cojinete}$. Cuando el movimiento de interés es la pista interna $f_{1,0}$ se sustituye por f_{BPI} de la ecuación (2.7), cuando es la pista externa se sustituye por f_{BPE} de la ecuación (2.8) y si es la rotación de la bola se sustituye por f_B de la ecuación (2.9).

Tabla 4.8: *Frecuencias características de movimiento de la bola en el cojinete*

Tipo de falla	Nivel de carga			
	Plena carga ¹		Vacío ²	
	$f_{1,0}$ (rev/s)	$f_{cojinete}$ (Hz)	$f_{1,0}$ (rev/s)	$f_{cojinete}$ (Hz)
FPI	144,3	204,3	152,8	212,8
FPE	110,7	170,7	117,2	177,2
FB	250,6	310,6	265,3	325,3

¹ La frecuencia de giro del rotor es $f_r = 28,3$ rev/s

² La frecuencia de giro del rotor es $f_r = 30$ rev/s

Las frecuencias características $f_{1,0}$ dependen de la frecuencia de giro del rotor f_r , la cual a su vez del nivel de carga, por lo que se calcularon tomando como

frecuencia de giro 1800 rpm ($f_r = 30$ rev/s) en vacío y 1700 rpm (28,3 rev/s) a plena carga. En la tabla 4.8 también se muestra la frecuencia de falla $f_{cojinete}$ y se calcula con la ecuación (2.6) sustituyendo $m = 1$ y $f_1 = 60$ Hz.

La frecuencia mínima en la que se puede detectar una falla es de 170 Hz y corresponde a la falla en pista externa, segundas columna y fila de la tabla 4.8, mientras que la máxima es 325 Hz y corresponde a la falla en bola. Según lo anterior deben estar incluidas en el espectro de Fourier las amplitudes que estén entre 170 y 325 Hz, esto serían 155 amplitudes ($325 - 170$) por lo que la matriz de datos, asociados al espectro de frecuencia, se convierte a una de 360×155 . Para reducir la cantidad de características se usó el PCA y se obtuvo que con dos componentes el porcentaje explicado de la varianza es de 90 %.

4.2. Evaluación de los clasificadores

Una vez usados los clasificadores es posible que den dos resultados: que clasifiquen un dato correctamente, describiendo realmente su condición, o que clasifiquen incorrectamente, haciendo una predicción errónea. Es importante saber el porcentaje con que se clasifican los datos pero también con el que se falla en la clasificación. En este trabajo se utilizaron varias herramientas con el propósito de evaluar el desempeño: matriz de confusión, exactitud, precisión, recall y f1-score.

Una *Matriz de Confusión* es una del tipo cuadrado que tiene tantas filas y columnas como clases existan. En la clasificación binaria la matriz es de 2×2 , mientras que en la clasificación múltiple para este trabajo es de 4×4 . En las filas se colocan los verdaderos valores pertenecientes a una clase, por ejemplo en la primera van los de la clase 1 en la segunda los de la clase 2 y así; en las columnas se colocan las predicciones realizadas ver figura 4.3.

En la figura 4.3 la nomenclatura *a* significa *Verdaderos de la Clase 1*, se refiere a los valores clasificados en esa clase y que en verdad pertenecen a la misma; *b* significa *Falsos de la Clase 2*, son los datos clasificados como 2 y pertenecen a otra clase; *c* significa *Falsos Clase 1*, son los valores clasificados en 1 y pertenecen

	Predicción en C_1	Predicción en C_2
Clase C_1	a	b
Clase C_2	c	d

Figura 4.3: Matriz de confusión, caso clasificación binaria.

a otra clase; d significa *Verdaderos Clase 2*, y son los valores clasificados como 2 y pertenecen a esa clase.

Una vez obtenida la matriz de confusión se usan unas métricas que sirven para interpretar los resultados:

- **Exactitud:** es el número de predicciones correctas o verdaderas divididas entre su total, en el caso de n clases se calcula según la ecuación 4.2:

$$\text{Exactitud} = \frac{a + d}{a + b + c + d}$$

- **Precisión:** es una métrica que mide que porcentaje de los que son clasificados en una clase en verdad pertenece a la misma. La precisión con la que se clasifican los datos en la clase C_1 , se calcula según la ecuación 4.2:

$$\text{Precisión}_i = \frac{a}{a + c} \times 100$$

- **Recall:** es una métrica que mide el porcentaje de los datos de una clase son clasificados correctamente. El recall con el que se clasifican los datos en C_1 se calcula según la ecuación 4.2:

$$\text{Recall} = \frac{a}{a + b} \times 100$$

- f1-score: es un promedio ponderado entre la precisión y el recall

$$\text{f1-score} = 2 \times \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}}$$

4.3. Clasificación usando Redes Neuronales Artificiales

Con los datos procesados se crearon las redes neuronales usando el modelo perceptron multicapa que se entrena con el algoritmo Backpropagation el cuál se optimiza usando el algoritmo *lbfgs*. Se utilizó una red de dos capas en donde el número de neuronas de la capa oculta se varió desde 1 hasta 20 para el caso de la clasificación binaria. El número máximo de iteraciones se fijó en 300. El parámetro α se varió en los siguientes valores: 1×10^{-5} , 1×10^{-4} , 1×10^{-3} , 1×10^{-2} , 1×10^{-1} , 1, 2, 3, 4, 5.

En primer lugar se quiere evaluar la sensibilidad de la red ante variaciones de las características o atributos de la data. En la tabla 4.9 se muestran los resultados de la respuesta de la red al entrenarse en cuatro condiciones distintas las cuales se describen a continuación:

Caso N° 1 Se entrena la red con 11 rasgos obteniéndose un porcentaje de entrenamiento de 93,00 % y una exactitud de 93,3 %.

Caso N° 2 Se entrena la red con los siguientes rasgos: Media,FC,Curtosis, Asimetría y Desv Est. Éstos fueron seleccionadas con el ANOVA. Se obtuvo un porcentaje de entrenamiento de 98,15 % y una exactitud de 88,9 %.

Caso N° 3 Se entrena la red con: Kurtosis, valor eficaz, desviación estándar, varianza y Xpeak. Se obtuvo un porcentaje de entrenamiento de 93,0 % y una exactitud de 93,3 %.

Tabla 4.9: *Efecto de variar rasgos en la ANN*

Nº	Rasgos	Resultado		Capa oculta	Tiempo entrenamiento
		Entrenamiento	Exactitud		
1	Todos	93,00 %	93,30 %	5	64,93 ms
2 ^a	1,2,3,6,7	98,00 %	88,90 %	3	52,76 ms
3 ^b	3,5,7,8,11	93,00 %	93,30 %	5	44,74 ms
4 ^c	3,7	96,00 %	95,60 %	5	33,53 ms

^a Rasgos seleccionados con el ANOVA.

^b Rasgos seleccionados con el RFE.

^c Rasgos comunes entre el ANOVA y RFE

Caso N° 4 Se entrena la red con: Kurtosis y Desviación estándar. Se obtuvo un porcentaje de entrenamiento de 96,0% y una exactitud de 95,6%.

De los casos anteriores se destaca el N° 4, en donde se usaron dos rasgos, comunes de la selección con el ANOVA y RFE, los porcentajes de aciertos son altos > 90% con la ventaja en la reducción de los tiempos de cálculo.

A continuación se va evaluar el comportamiento de la red ante variación de neuronas de la capa oculta. La tabla 4.10 muestra la cantidad de aciertos y el número de iteraciones obtenidas al variar la cantidad de neuronas de la capa oculta de 1 a 20.

En la tabla 4.10 se observa que los mejores resultados se obtuvieron cuando el número de neuronas de la capa oculta es menor a 5, a medida de que éste aumenta el porcentaje de aciertos disminuye, y a su vez empeoran los tiempos de cálculo.

Finalmente en la tabla 4.11 muestra todos resultados obtenidos al usar de la red de 4 neuronas en la capa oculta.

Para distinguir la habilidad de las ANN de distinguir entre las clases se usan la matriz de confusión mostrada en la tabla 4.12. Las suma de los valores de cada filas de la matriz son los valores verdaderos, en este caso hay 21 datos de cojinetes

Tabla 4.10: *Efecto de variar el número de neuronas ocultas en las ANN*

Rango	Cantidad	Cantidad de aciertos		Tiempo
	neuronas	Entrenamiento	Exactitud	
1-5	4	96,0 %	93,3 %	57,01 ms
6-10	6	84,0 %	91,1 %	84,26 ms
11-15	12	93,0 %	88,9 %	93,16 ms
15-20	17	82,0 %	86,7 %	103,41 ms

^a *Nota:* se usaron todos los rasgos de la señal de corriente.

Tabla 4.11: *Resultados de la clasificación en la Red Neuronal de 4 capas*

Descripción	Resultado
Mejor α	1×10^{-5}
Número de neuronas en la capa oculta	4
Máximo número de iteraciones	300
solver	lbfgs
Mejor resultado de la validación cruzada	87 %
Exactitud	93,6 %

sin falla y 24 con fallas. La suma de los valores de cada columna es el resultado de la predicción del clasificador, la primera columna indica que en 22 de las veces se predijo que no había falla, mientras que la segunda indica que en 23 veces se predijo que sí hubo.

Tabla 4.12: *Matriz de confusión. Clasificación con Redes Neuronales*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	20	1
	Falla	2	22

De la matriz de confusión se puede inferir lo siguiente: en 20 de 21 casos se clasificó correctamente a los datos como sin falla (95,2%), y en 22 de 24 casos (91,7%) con falla. La exactitud del clasificador mostrada en la tabla 4.11 se determina sumando los datos clasificados correctamente de cada clase entre los datos totales:

$$\text{Exactitud} = \frac{20 + 22}{20 + 1 + 2 + 22} \times 100 = 93,6\%$$

4.4. Clasificación múltiple usando Redes Neuronales

A continuación se probó la capacidad de las ANN de identificar correctamente el tipo de falla de los cojinetes. Al igual que en la clasificación simple se estudiaron 4 casos en los que se variaron los atributos de la data: todos los rasgos (11), los seleccionados con el ANOVA, los seleccionados con el RFE y con la desviación estándar y la curtosis. En cuanto a la arquitectura, como los resultados obtenidos con una sola capa oculta fueron muy bajos, se escogió una red de dos capas ocultas.

La entonación de la red se hizo por ensayo y error, el número de neuronas de las capas ocultas se varió en un rango de 10–35; la tabla 4.13 muestra los mejores resultados obtenidos.

Tabla 4.13: *Selección de arquitectura de la red neuronal ANN, caso 1*

Arquitectura	Resultado		Tiempo entrenamiento
	Validación cruzada	Exactitud	
11,14,10,4	74 %	72,2 %	236,8 ms
11,15,15,4	75 %	76,3 %	317,2 ms
11,17,15,4	75 %	78,9 %	321,2 ms
11,20,20,4	76 %	81,1 %	346,7 ms
11,23,20,4	79 %	84,3 %	340,8 ms
11,24,22,4	76 %	80,0 %	342,1 ms
11,26,25,4	76 %	78,9 %	381,3 ms
11,34,30,4	77 %	73,3 %	409,5 ms

De los casos mostrados en la tabla 4.13 destacan los que tienen un número de neuronas en la capa oculta cercano a 20. La mejor arquitectura es la que tiene 23 y 20 neuronas en la primera y segunda capa oculta respectivamente; con el mejor resultado de la validación cruzada 79 % y una exactitud de 84,3 %, el parámetro $\alpha = 1$. La tabla 4.14 muestra la matriz de confusión de la arquitectura seleccionada.

De la tabla 4.14 se tiene que se clasificaron correctamente 22 de 24 casos (91,7 %) sin falla, 21 de 24 (87,1 %) con falla en bola, 20 de 25 (80 %) con falla en pista externa y 12 de 16 (75 %) con falla en pista interna. La exactitud es entonces

$$\text{Exactitud} = \frac{22 + 21 + 20 + 12}{24 + 24 + 25 + 16} \times 100 = 84,3 \%$$

El siguiente caso de estudio es usar como atributos de la data las siguientes métricas estadísticas: media, factor de cresta, curtosis, asimetría y desviación

Tabla 4.14: *Matriz de confusión. Clasificación con Redes Neuronales*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	22	1	0	1
	FB	1	21	2	0
	FPE	2	2	20	1
	FPI	2	1	2	12

estándar; obtenidas del ANOVA. La tabla 4.15 muestra el comportamiento de la red al variar el número de neuronas de las capas ocultas.

Tabla 4.15: *Selección de arquitectura de la red neuronal ANN, caso 2*

Arquitectura	Resultado		Tiempo entrenamiento
	Validación cruzada	Exactitud	
5,9,5,4	69,3 %	78,9 %	187,2 ms
5,13,10,4	77,0 %	77,8 %	221,8 ms
5,15,15,4	77,0 %	83,3 %	297,8 ms
5,19,15,4	77,8 %	83,3 %	302,5 ms
5,15,13,4	78,5 %	78,9 %	298,6 ms
5,24,20,4	78,1 %	81,1 %	327,3 ms
5,28,25,4	79,3 %	83,3 %	355,7 ms
5,32,30,4	78,9 %	83,3 %	391,7 ms

^a Rasgos seleccionados con el ANOVA: media, factor de cresta, curtosis, asimetría y desviación estándar.

De los casos mostrados en la tabla 4.15 se selecciona la arquitectura de 5,19,15,4 neuronas con un mejor resultado en la validación cruzada de 77,8 % y una exactitud de 83,3 %. El número de neuronas de las capas ocultas es menor que cuando

se usaron los datos con todos los atributos. La tabla 4.14 muestra la matriz de confusión de la arquitectura seleccionada.

Tabla 4.16: *Matriz de confusión. Clasificación con Redes Neuronales*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	22	2	0	0
	FB	1	22	1	0
	FPE	0	3	19	3
	FPI	2	1	1	13

De la tabla 4.16 se tiene que se clasificaron correctamente 22 de 24 casos (91,7%) sin falla, 22 de 24 (91,7%) con falla en bola, 18 de 26 (69,2%) con falla en pista externa y 13 de 16 (81,3%) con falla en pista interna. La exactitud es entonces:

$$\text{Exactitud} = \frac{22 + 22 + 18 + 13}{24 + 24 + 26 + 16} \times 100 = 83,3\%$$

Las métricas que evalúan cómo las ANN discriminan el tipo de falla se muestran en la tabla 4.17. El peor recall fue para las fallas en pista 0,76 y el mejor caso cuando no hay falla 0,92. La predicción de la falla en bola es la de menor precisión 0,78. Cuando se predice que no hay falla en pista externa es 0,90.

A continuación se presentan los resultados de las ANN al usar como atributos de la data las siguientes métricas estadísticas: curtosis, valor rms, desviación estándar, varianza y el valor pico; obtenidas del RFE. La tabla 4.18 muestra el comportamiento de la red al variar el número de neuronas de las capas ocultas.

De los casos mostrados en la tabla 4.18 se selecciona la arquitectura de 5,25,25,4 neuronas con un mejor resultado en la validación cruzada de 77,8% y una exactitud de 73,3%. La tabla 4.19 muestra la matriz de confusión de la arquitectura seleccionada.

Tabla 4.17: *Evaluación de la clasificación múltiple con las ANN. Caso 2.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,85	0,92	0,88	24
FB	0,78	0,88	0,82	24
FPE	0,90	0,76	0,83	25
FPI	0,81	0,76	0,79	17
Promedio/Total	0,84	0,83	0,83	90

Tabla 4.18: *Selección de arquitectura de la red neuronal ANN, caso 3*

Arquitectura	Resultado		Tiempo entrenamiento
	Validación cruzada	Exactitud	
5 ^a ,5,4	54,8 %	42,2 %	153,1 ms
5,13,10,4	74,4 %	62,2 %	196,6 ms
5,15,15,4	77,8 %	68,9 %	270,9 ms
5,15,18,4	75,9 %	65,6 %	278,1 ms
5,20,20,4	77,0 %	66,7 %	294,7 ms
5,20,23,4	77,4 %	65,6 %	315,9 ms
5,25,25,4	77,8 %	73,3 %	333,8 ms
5,31,30,4	76,7 %	67,8 %	375,2 ms

^a Rasgos seleccionados con el RFE: curtosis, valor rms, desviación estándar, varianza y el valor pico.

De la tabla 4.19 se tiene que se clasificaron correctamente 15 de 24 casos sin falla dando una exactitud de 0,62, 23 de 24 (0,96) con falla en bola, 17 de 25 (0,68) con falla en pista externa y 11 de 17 (0,73) con falla en pista interna, estos valores se muestran en la tabla 4.20. La exactitud es entonces:

Tabla 4.19: *Matriz de confusión. Clasificación con Redes Neuronales*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	15	2	2	5
	FB	0	23	0	1
	FPE	2	3	17	3
	FPI	1	0	5	11

$$\text{Exactitud} = \frac{16 + 23 + 14 + 13}{24 + 24 + 25 + 17} \times 100 = 73,3\%$$

Tabla 4.20: *Evaluación de la clasificación múltiple con las ANN. Caso 3.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,83	0,62	0,71	24
FB	0,82	0,96	0,88	24
FPE	0,71	0,68	0,69	25
FPI	0,55	0,65	0,59	17
Promedio/Total	0,74	0,73	0,73	90

La precisión al predecir que hay falla en pista interna es muy pobre 0,55. En general el desempeño en este caso 3 es inferior a los anteriores.

A continuación se presentan el último caso en el que se usa como atributos de la data la curtosis y la desviación estándar. La tabla 4.21 muestra el comportamiento de la red al variar el número de neuronas de las capas ocultas.

De los casos mostrados en la tabla 4.21 se selecciona la arquitectura de 2,24,23,4 neuronas con un mejor resultado en la validación cruzada de 74,1% y una exactitud de 70,0%. La tabla 4.14 muestra la matriz de confusión de la arquitectura seleccionada.

Tabla 4.21: Selección de arquitectura de la red neuronal ANN, caso 4

Arquitectura	Resultado		Tiempo entrenamiento
	Validación cruzada	Exactitud	
2,6,5,4	55,6 %	34,4 %	155,7 ms
2,11,10,4	70,7 %	68,9 %	185,0 ms
2,18,15,4	72,2 %	70,0 %	262,0 ms
2,15,18,4	73,0 %	61,0 %	270,3 ms
2,22,10,4	73,7 %	66,7 %	285,0 ms
2,24,23,4	74,1 %	70,0 %	306,0 ms
2,29,25,4	75,9 %	66,7 %	331,5 ms
2,32,30,4	74,8 %	67,8 %	357,7 ms

^a Rasgos usados: curtosis y desviación estándar.

Tabla 4.22: Matriz de confusión. Clasificación con Redes Neuronales

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	13	2	6	3
	FB	1	23	0	0
	FPE	0	5	13	7
	FPI	1	0	2	14

De la tabla 4.22 se tiene que se clasificaron correctamente 13 de 24 casos (54,2%) sin falla, 23 de 24 (95,8%) con falla en bola, 13 de 25 (52,0%) con falla en pista externa y 14 de 17 (82,4%) con falla en pista interna. La exactitud es entonces:

$$\text{Exactitud} = \frac{13 + 23 + 13 + 14}{24 + 24 + 25 + 17} \times 100 = 70,0 \%$$

En la tabla 4.23 se muestra un resumen de los cuatro casos estudiados, en donde se comparan: tiempo promedio de entrenamiento, arquitectura, mejor resultado de la validación cruzada y la exactitud.

Tabla 4.23: *Resumen de la clasificación múltiple con ANN.*

Descripción	Resultados			
	Caso 1 ^a	Caso 2 ^b	Caso 3 ^c	Caso 4 ^d
Tiempo promedio de entrenamiento	340,8 ms	302,5 ms	333,8 ms	306,0 ms
Arquitectura	11,23,20,4	5,19,15,4	5,25,25,4	2,24,23,4
Resultado de la validación cruzada	79,0 %	77,8 %	77,8 %	74,1 %
Exactitud	84,3 %	83,3 %	73,3 %	70,0 %

^a Los datos tienen 11 atributos.

^b Rasgos seleccionados con el ANOVA: media, factor de cresta, curtosis, asimetría y desviación.

^c Rasgos seleccionados con el RFE: curtosis, valor rms, desviación estándar, varianza y el valor pico.

^d Rasgos seleccionados: curtosis y desviación estándar.

Se observa en la tabla 4.23 que las mejores exactitudes se obtuvieron cuando se entrenaron a las ANN con todos los rasgos (84,3 %) y cuando se usaron los 5 rasgos obtenidos del ANOVA (83,3 %); en todos los casos las arquitecturas fueron de 3 capas, y el número de neuronas de las capas ocultas estuvo cercano a 20, excepto en el caso dos. El tiempo promedio de entrenamiento estuvo entre 300 y 400 milisegundos.

4.5. Clasificación usando Máquinas de Vectores Soporte

En el estudio de las SVM se debe escoger primero la función kernel. En la revisión de la literatura se recomienda la tipo Gauseana para la clasificación de

fallas en los motores de inducción, por lo que en este trabajo se usó esa función. Seguidamente se debe identificar los parámetros de la función kernel que se deben entonar en el entrenamiento que, para este caso, son C y γ . Los valores con los que se ajustó C fueron: 10^{-3} , 10^{-2} , 10^{-1} , 1, 10, 100, 200 y 300; los valores con los que se ajustó γ fueron: 10^{-3} , 10^{-2} , 10^{-1} , 1, 10 y 100.

Al igual que con las ANN se estudiaron 4 casos, el primero mostrado en la tabla 4.24 es la respuesta de las SVM ante datos de entradas cuyos atributos son 11 métricas estadísticas.

Tabla 4.24: *Clasificación con SVM caso 1*

Descripción	Resultado
Tiempo promedio de entrenamiento	1,94ms
Mejor C	200
Mejor γ	0,1
Mejor resultado de la validación cruzada	88,9%
Exactitud	93,3%

La exactitud de las SVM fue 93,3% igual al mejor de los casos cuando se estudió la clasificación con ANN (ver tabla 4.10), sin embargo el tiempo de entrenamiento fue mucho menor. La capacidad del clasificador para discriminar en que condición se encuentra el motor, sin falla o con falla en cojinete, se entiende por medio de la tabla 4.25

Tabla 4.25: *Matriz de confusión. Clasificación con SVM*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	20	1
	Falla	2	22

De la matriz de confusión se puede inferir lo siguiente: en 20 de 21 muestras de la corriente del motor, con el cojinete sin falla, se clasificó correctamente por lo que el recall es 95 %, de los 22 casos clasificados como sin falla, 20 son verdaderos dando una precisión de 91 %, el f1-score es de 93 % esto se muestra en la tabla 4.26.

Tabla 4.26: *Evaluación de la clasificación con SVM. Caso 1.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,91	0,95	0,93	21
Falla	0,96	0,92	0,94	24
Promedio/Total	0,93	0,93	0,93	45

En cuanto a la capacidad de la SVM de identificar que hay falla se tiene que de 24 muestras con el cojinete con daños se acertó en la predicción 22 veces con lo que el recall es de 92 %; mientras que en 23 ocasiones se indicó que el había falla dando una precisión 96 %, el f1-score es de 94 %. La exactitud del clasificador mostrada en la tabla 4.24 se determina sabiendo que se clasificaron correctamente 42 muestras de 45 con lo que su valor es 93,3 %.

Los otros casos de estudios son: 5 rasgos obtenidos del ANOVA, 5 rasgos obtenidos del RFE y 2 rasgos comunes de los dos casos anteriores.

Si se comparan los resultados de la tabla 4.27 con los del estudio de la clasificación con ANN que se resumen en la tabla 4.9, se observa que los resultados en cuanto a exactitud para cada caso son similares; la pequeña diferencia estaría en el caso 2 donde las ANN ganan 88,9 % a 86,7 %, y en el caso 4 que gana 95,6 % a 93,3 %.

En cuanto a tiempo de entrenamiento las SVM fueron muy superiores con valores menores a 2 ms, los de las ANN son mayores a 30 ms.

A continuación en la tabla 4.28 se muestra la matriz de confusión asociada al caso de estudio 2.

Hay 21 muestras con el cojinete en buen estado de las cuales 16 son clasificadas correctamente con lo que el recall es 76 %, esto indica que no los está clasificando

Tabla 4.27: *Clasificación con SVM*

Descripción	Resultados			
	Caso 1 ^a	Caso 2 ^b	Caso 3 ^c	Caso 4 ^d
Tiempo promedio de entrenamiento	1,94 ms	1,69 ms	1,58 ms	1,94 ms
Mejor C	200	10	10	300
Mejor γ	0,1	0,1	1	1
Resultado de la validación cruzada	88,9 %	85,2 %	93,3 %	91,9 %
Exactitud	93,3 %	86,7 %	93,3 %	93,3 %

^a Los datos tienen 11 atributos.

^b Rasgos seleccionados con el ANOVA: media, factor de cresta, curtosis, asimetría y desviación.

^c Rasgos seleccionados con el RFE: curtosis, valor rms, desviación estándar, varianza y el valor pico.

^d Rasgos seleccionados: curtosis y desviación estándar.

Tabla 4.28: *Matriz de confusión. Clasificación con SVM*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	16	5
	Falla	1	23

de forma aceptable ver tabla 4.29. De las 17 ocasiones en las cuales se indicó que no había falla 16 eran correctas por lo que la precisión es 94 %, el f1-score es 0,84 debido a la bajo recall.

Se tienen 24 muestras de corriente con el cojinete fallado y en 23 ocasiones se aciertan con lo que el recall es 96 %. De las 28 veces en que se indica que hay falla 23 son verdad con lo que la precisión es 82 %, el f1-score es 0,88. La exactitud de

Tabla 4.29: *Evaluación de la clasificación con SVM. Caso 2.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,94	0,76	0,84	21
Falla	0,82	0,96	0,88	24
Promedio/Total	0,88	0,87	0,86	45

la red es 86,7% sabiendo que se acertaron en 40 de 45 ocasiones. En resumen la red se confunde cuando los cojinetes no tienen falla.

En la tabla 4.30 se muestra la matriz de confusión del caso 3, métricas estadísticas obtenidas del RFE. Se observa que hay 45 muestras en total, 21 son con el cojinete en buen estado y 24 con falla. En 20 ocasiones se indicó que no había falla y en 25 que sí.

Tabla 4.30: *Matriz de confusión. Clasificación con SVM*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	19	2
	Falla	1	23

De las 21 muestras con el cojinete en buen estado 19 fueron clasificadas correctamente con lo que el recall es de 90%, de las 20 veces en la que se indicó que no había falla 19 son correctas con lo que la precisión es 95% con un f1-score de 0,93. Ver tabla 4.31

De las 24 muestras con el cojinete con falla 23 fueron clasificadas correctamente por lo que el recall es 96%, de las 25 veces en las que se indicó que había falla 23 son clasificadas correctamente por lo que la precisión es 92%, el f1-score es 0,94. El promedio de cada métrica en este caso es de 0,93 dando un comportamiento superior al del caso 2.

Tabla 4.31: *Evaluación de la clasificación con SVM. Caso 3.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,95	0,90	0,93	21
Falla	0,92	0,96	0,94	24
Promedio/Total	0,93	0,93	0,93	45

En la tabla 4.32 se muestra la matriz de confusión del caso 4, en la que las métricas estadísticas con la curtosis y la desviación estándar. Se observa que hay 45 muestras en total, 21 son con el cojinete en buen estado y 24 con falla. En 20 ocasiones se indicó que no había falla y en 22 que sí. Los resultados son exactamente iguales a los del caso 3, por lo que lo que las métricas son las mostradas en la tabla 4.31.

Tabla 4.32: *Matriz de confusión. Clasificación con SVM*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	19	2
	Falla	1	23

En resumen en la clasificación binaria con SVM la exactitud con la que se clasifican los datos es de 93%, exceptuando el caso en el que se usan los rasgos extraídos con el ANOVA en donde se alcanzó un 88%. En este caso en particular se obtuvo el peor desempeño, con un recall de 0,76 cuando se clasifican corrientes con cojinetes en buen estado. En el resto de los casos los valores de las métricas fueron mayores a 0,90.

4.6. Clasificación múltiple usando Máquinas de Vectores Soporte

Para la clasificación múltiple las SVM usan la estrategia del uno contra uno, el número de clasificadores se determina según la ecuación (2.10) con $k = 4$ porque son 4 clases:

$$\text{Clasificadores} = 4 \left(\frac{4 - 1}{2} \right) = 6$$

Por lo que son 6 clasificadores, en donde cada uno se entrena con un par de clases distintas. La clasificación de un dato desconocido se realiza por medio de votación. En cuanto a la cantidad de atributos de los datos se tienen cuatro casos: 11 rasgos, 5 rasgos obtenidos con el ANOVA, 5 rasgos obtenidos por el RFE y 2 rasgos. La tabla 4.33 muestra un resumen de los resultados obtenidos del entrenamiento de las SVM.

Los tiempos de entrenamiento son menores 6 ms. La mayor exactitud fue de 88,9% cuando se usan datos con las 5 métricas obtenidas del RFE (caso 3), mientras el peor porcentaje 77,8% cuando se usaron las 5 métricas obtenidas del ANOVA. A continuación se analiza cada caso por separado empezando por el 1, datos de 11 atributos, la tabla 4.34 muestra su matriz de confusión.

La clase *Sin Falla* consta de 24 datos, *FB* de 24, *FPE* de 25 y *FPI* de 17 para un total de 90 datos. De la diagonal de la matriz se tiene que las predicciones verdaderas son: 20 de que no hay falla, 21 de falla en bola, 19 de falla en pista externa y 13 de pista interna.

De los 24 datos de la clase *Sin Falla* 20 fueron clasificados correctamente con lo que el recall es de $(20/24 = 0,83)$ ver tabla 4.35, en 25 ocasiones hay una predicción de que no hay falla por lo que la precisión es $(20/25 = 0,8)$ y su f-score:

$$\text{f-score} = 2 \times \frac{0,83 \times 0,8}{0,83 + 0,8} = 0,82$$

Al estudiar las métricas de todas las clases por medio de la tabla 4.35, se

Tabla 4.33: *Clasificación múltiple con SVM.*

Descripción	Resultados			
	Caso 1 ^a	Caso 2 ^b	Caso 3 ^c	Caso 4 ^d
Tiempo promedio de entrenamiento	6,22 ms	5,51 ms	5,30 ms	4,89 ms
Mejor C	200	10	300	100
Mejor γ	0,1	1	10	10
Resultado de la validación cruzada	74,8 %	77,8 %	88,5 %	81,9 %
Exactitud	81,1 %	77,8 %	88,9 %	80,0 %

^a Los datos tienen 11 atributos.

^b Rasgos seleccionados con el ANOVA: media, factor de cresta, curtosis, asimetría y desviación.

^c Rasgos seleccionados con el RFE: curtosis, valor rms, desviación estándar, varianza y el valor pico.

^d Rasgos seleccionados: curtosis y desviación estándar.

Tabla 4.34: *Matriz de confusión. Clasificación múltiple con SVM caso 1*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	20	4	0	0
	FB	0	21	2	1
	FPE	2	3	19	1
	FPI	3	1	0	13

observan que el peor recall es para las fallas en pista externa e interna con 0,76 mientras que el mejor es para la falla en bola con 0,88. Mientras que la peor precisión es la de la falla en bola con 0,72 y la mejor es cuando se predice que hay falla en pista externa con 0,90. Todos los f-score estuvieron entre 0,79 y 0,83. En

Tabla 4.35: *Evaluación de la clasificación múltiple con SVM. Caso 1.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,80	0,83	0,82	24
FB	0,72	0,88	0,79	24
FPE	0,90	0,76	0,83	25
FPI	0,87	0,76	0,81	17
Promedio/Total	0,82	0,81	0,81	90

resumen se confunden las fallas en las pistas, y en la predicción es la falla en bola la que presenta peor porcentaje.

A continuación se estudia el caso 2 que consiste en data de 5 atributos, los cuales son las métricas estadísticas de la corriente en el dominio del tiempo seleccionadas después de aplicar el ANOVA. La tabla 4.36 muestra su matriz de confusión de la que se desglosa la siguiente información: 24 instancias de cojinetes en buen estado, 24 con fallas en bola, 25 con fallas en pista externa y 17 en pista externa para un total de 90 datos.

En cuanto a los verdaderos positivos se tiene que 18 casos han sido clasificados correctamente como *Sin Falla*, 21 como *FB*, 19 como *FPE* y 12 *FPI*. En 22 ocasiones se predijo que no había fallas, en 26 que había falla en bola, en 25 en pista externa y 17 en pista interna.

Cuando se clasifican los 24 datos de cojinetes sin falla se predicen correctamente 18 por lo que el recall es ($18/25 = 0,75$) ver tabla 4.37, como en 22 ocasiones del total de la data se predijo que no había falla se tiene que la precisión es de ($18/22 = 0,82$) con lo que el f-score es:

$$\text{f-score} = 2 \times \frac{0,75 \times 0,82}{0,75 + 0,82} = 0,78$$

En la clasificación de cojinetes sin falla el desempeño de este clasificador fue pobre (f-score < 0,8). La tabla 4.37 muestra las métricas estadística que evalúan

Tabla 4.36: *Matriz de confusión. Clasificación múltiple con SVM caso 2*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	18	2	2	2
	FB	1	21	2	0
	FPE	1	2	19	3
	FPI	2	1	2	12

Tabla 4.37: *Evaluación de la clasificación múltiple con SVM. Caso 2.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,82	0,75	0,78	24
FB	0,81	0,88	0,84	24
FPE	0,76	0,76	0,76	25
FPI	0,71	0,71	0,71	17
Promedio/Total	0,78	0,78	0,78	90

el desempeño de las SVM en cada clase, el peor caso es la falla en pista interna con valores de recall y precisión de 0,71, mientras que el mejor fue la clasificación de falla en bola. En general el desempeño de las SVM para este caso fue mediocre, con el promedio de cada métricas igual a 0,78.

A continuación se estudia el comportamiento de las SVM cuando se entrenan con datos de 5 atributos, los cuales son las siguientes métricas estadísticas: curtosis, valor rms, desviación estándar, varianza y el valor pico. El tiempo de entrenamiento fue de 5,17 ms, con un mejor porcentaje en la técnica de validación cruzada de 88,5 %.

Primeramente se muestra en la tabla 4.38 la matriz de confusión obtenida. Se tienen 24 datos de prueba con el cojinete sin falla, 24 con falla en bola, 25 en pista externa y 17 en pista interna.

Tabla 4.38: *Matriz de confusión. Clasificación múltiple con SVM caso 3*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	21	0	3	0
	FB	0	23	0	1
	FPE	2	0	22	1
	FPI	1	0	2	14

Hay 24 predicciones de que no hay falla de las cuales 21 son verdaderas, 23 de que hay falla en bola con 23 verdaderas, 27 en pista externa con 22 verdaderas y 16 en pista interna con 14 verdaderas.

Como de las 24 instancias de cojinetes en buen estado 21 son clasificadas correctamente se tiene que la exactitud es ($21/24 = 0,88$), por lo que la discriminación de esta condición es buena. En 24 ocasiones se predice que no hay fallas dando una predicción de ($21/24 = 0,88$) ver tabla 4.39, el f-score es entonces:

$$\text{f-score} = 2 \times \text{frac},88 \times 0,880,88 + 0,88 = 0,88$$

Tabla 4.39: *Evaluación de la clasificación múltiple con SVM. Caso 3.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,88	0,88	0,88	24
FB	1,00	0,96	0,98	24
FPE	0,81	0,88	0,85	25
FPI	0,88	0,82	0,85	17
Promedio/Total	0,89	0,89	0,89	90

En general las métricas que evalúan el desempeño del clasificador son buenas según se muestra en la tabla 4.39, sin embargo es importante resaltar que en el

caso de la discriminación de falla en bola la exactitud es 0,96 mientras que la precisión es perfecta.

Finalmente se presenta el último caso en donde se evalúa el desempeño de las SVM cuando se tienen datos con dos atributos que son la curtosis y la desviación estándar. La tabla 4.40 muestra la matriz de confusión obtenida. Se tienen 24 instancias de cojinetes sin falla, 24 con falla en bola, 25 de pista externa y 17 de pista interna.

Tabla 4.40: *Matriz de confusión. Clasificación múltiple con SVM caso 4*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	18	0	6	0
	FB	0	23	0	1
	FPE	1	4	18	2
	FPI	0	1	3	13

En cuanto a la predicción se tiene que hay 19 sin falla con 18 verdaderas, 28 con falla en bola con 23 verdaderas, 27 en pista externa con 18 verdaderas y 16 en pista interna con 13 verdaderas.

De las 24 instancias sin falla 18 son clasificadas correctamente por lo que el recall es de $(18/24 = 0,75)$ ver tabla 4.41, lo que indica que no se discrimina bien esa condición. Como de 19 predicciones 18 son verdaderas la precisión es $(18/19 = 0,95)$ lo que indica que el número de falsos positivos muy bajo. El f-score es:

$$\text{f-score} = 2 \times \frac{0,75 \times 0,95}{0,75 + 0,95} = 0,84$$

A pesar de que 0,84 es un buen valor de f-score en este caso se observa que es debido a la precisión. Observando la tabla 4.41 se tiene que el promedio de las métricas que evalúan el clasificador es un aceptable 0,80, sin embargo cuando

Tabla 4.41: *Evaluación de la clasificación múltiple con SVM. Caso 4.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,95	0,75	0,84	24
FB	0,82	0,96	0,88	24
FPE	0,67	0,72	0,69	25
FPI	0,81	0,76	0,79	17
Promedio/Total	0,81	0,80	0,80	90

se trata de las fallas en pista externa el recall es muy bajo (0,67) y la precisión (0,72).

4.7. Clasificación partiendo del Análisis Discriminante

Para usar el análisis discriminante se dividen los datos en dos grupos Y_1 cuando el cojinete está en buenas condiciones y Y_2 cuando tiene alguna falla. La matriz de datos del grupo Y_1 es de 11 características por 30 observaciones (11×30) y la de Y_2 de (11×60). Luego para grupo se determina la media de cada característica:

Se determinó la covarianza de cada grupo \mathbf{S}_1 y \mathbf{S}_2 y la covarianza combinada \mathbf{S}_{pl} según la ecuación (2.13) la cual se muestra a continuación para mayor comprensión.

$$\mathbf{S}_{pl} = \frac{\sum_{i=1}^k (n_i - 1) \mathbf{S}_i}{\sum_{i=1}^k (n_i - 1)} \quad (4.1)$$

donde: S_{pl} es la matrix de covarianza combinada, S_i la matrix de covarianza de la población i , n_i es la cantidad de mediciones realizadas en la población i .

En este caso el número de grupos es dos ($k = 2$), los valores de n_1 y n_2 son 30 y 60 respectivamente; sustituyendo esto en la ecuación (4.1) se tiene:

Tabla 4.42: *Valores medios de cada características*

Característica	Grupos	
	\bar{y}_1	\bar{y}_2
Media	0,1935	-0,0967
FC	0,4637	-0,2318
Curtosis	-0,3468	0,1734
Mediana	0,3142	-0,1571
RMS	-0,0219	0,0110
Asimetría	-0,4702	0,2351
Desv Est	-0,0219	0,0110
Varianza	-0,0208	0,0104
Max Val	-0,0148	0,0074
Rango	-0,0146	0,0073
Val Pico	-0,0153	0,0076

$$\mathbf{S}_{pl} = \frac{(30 - 1)\mathbf{S}_1 + (60 - 1)\mathbf{S}_2}{(30 - 1) + (60 - 1)}$$

La matriz de covarianza combinada tiene dimensiones 11×11 . Cada observación consiste en un vector de 11 componentes, este se convierte en un escalar al multiplicarlo por el vector \mathbf{a} determinado por la ecuación (2.15):

$$\mathbf{a}' = \begin{pmatrix} -13,0 \\ 70,0 \\ -16,9 \\ 1,3 \\ 4,57 \cdot 10^5 \\ -0,929 \\ -4,52 \cdot 10^5 \\ -1811,0 \\ 60,1 \\ -578,0 \\ -2822,0 \end{pmatrix}$$

El criterio de clasificación es el siguiente:

$$\text{Cojinete en buen estado si } \mathbf{a}'\mathbf{y} \geq \frac{1}{2}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{S}_{pl}^{-1}(\bar{\mathbf{y}}_1 + \bar{\mathbf{y}}_2) = 3,6748$$

$$\text{Falla en cojinete si } \mathbf{a}'\mathbf{y} < 3,6748$$

Al igual que con las ANN y SVM se estudiaron 4 casos en los que se variaron la cantidad de atributos de la data. Es importante destacar que se realizó el PCA para que ninguna de los atributos esté correlacionadas entre sí.

El primer caso consiste en 11 métricas estadísticas obtenidas del análisis de la corriente en el dominio del tiempo. El tiempo de cálculo fue 3,80 ms, el mejor resultado de la validación cruzada 93,3% y la exactitud 95,6%. La tabla 4.43 muestra la matriz de confusión de este caso.

Se tiene 21 datos con cojinetes en buen estado y 24 con fallas. Los verdaderos positivos son 20 sin falla y 23 con falla. En el caso de las 21 instancias sin falla 20 se clasificaron correctamente con lo que el recall es $(20/21 = 0,95)$; igualmente se predijeron 21 veces que no había falla dando una predicción de $20/21 = 0,95$ esto se observa en la tabla 4.44. El f-score es:

Tabla 4.43: *Matriz de confusión. Clasificación con Redes Neuronales*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	20	1
	Falla	1	23

$$\text{f-score} = 2 \times \frac{0,95 \times 0,95}{0,95 + 0,95} = 0,95$$

Tabla 4.44: *Evaluación de la clasificación con ALD. Caso 1.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,95	0,95	0,95	21
Falla	0,96	0,96	0,96	24
Promedio/Total	0,96	0,96	0,96	45

El comportamiento cuando hay falla en cojinetes o se predice la misma es similar exactitud (0,96) y precisión (0,96). Estos valores indican que el desempeño del ALD para detectar si hay falla o no en cojinetes es igual a las ANN y SVM.

Cuando se utilizan datos de 5 atributos, estos son rasgos obtenidos con el ANOVA el tiempo de cálculo para crear la función discriminante fue de 3,60 ms con un mejor resultado de validación cruzada de 81,5%. En la tabla 4.45 se muestra la matriz de confusión de la que se observa que hay 21 muestras sin fallas y 24 con fallas para un total de 45, de éstas se acertó en la predicción en 32 (14 + 18) veces con lo que la exactitud es 71,1% o 22% por debajo del caso 1.

De las 21 instancias sin falla 14 son clasificadas de forma correcta por lo que el recall es 0,67 (14/21). De la misma manera con los datos que contienen falla el recall es 0,75. Se hacen 20 predicciones de que no hay fallas con lo que la precisión es 0,70 (14/20), estos resultados se muestran en la primera fila de la tabla 4.46.

Tabla 4.45: *Matriz de confusión. Clasificación con Redes Neuronales*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	14	7
	Falla	6	18

Tabla 4.46: *Evaluación de la clasificación con ALD. Caso 2.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,70	0,67	0,68	21
Falla	0,72	0,75	0,73	24
Promedio/Total	0,71	0,71	0,71	45

El f-score de la clase sin falla se determina con el recall y la precisión:

$$\text{f-score} = 2 \times \frac{0,70 \times 0,67}{0,70 + 0,67} = 0,68$$

Los resultados para la clase sin falla, mostrados en la tabla 4.46 son recall igual a 0,75, precisión 0,72 y f-score 0,73.

Cuando se utilizan datos de 5 atributos, aquellos seleccionados con el RFE se tiene que el tiempo empleado para crear la función discriminante fue de 2,80 ms, el mejor resultado en la validación cruzada 91,9%. Para probar el desempeño de se usan 45 datos, 21 con el cojinete en buen estado y 24 con fallas, obteniendo la matriz de confusión mostrada en la tabla 4.47.

Para la clase sin falla los verdaderos positivos son 19 y de la clase sin falla 22 con lo que la exactitud es 0,911 (41/45). Al acertar 19 veces de las 21 en la clase sin falla se obtiene un recall de 0,90, y al predecir 19 de 21 la precisión es 0,90 igualmente. Con lo que el f-score es 0,90. Éstas métricas son mostradas en la tabla 4.48.

Tabla 4.47: *Matriz de confusión. Clasificación con Redes Neuronales*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	19	2
	Falla	2	22

Tabla 4.48: *Evaluación de la clasificación con ALD. Caso 3.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,90	0,90	0,90	21
Falla	0,92	0,92	0,92	24
Promedio/Total	0,91	0,91	0,91	45

En cuanto a la clase sin falla las métricas son iguales a 0,92. En general el desempeño para este caso es muy superior al caso 2 es cercano a cuando se usan todos los rasgos.

Si se usan dos rasgos (caso 4) tiempo empleado para crea la función discriminante es 4,20 ms, con un mejor resultado de validación cruzada de 82,2%. La matriz de confusión es mostrada en la tabla 4.49.

Tabla 4.49: *Matriz de confusión. Clasificación con Redes Neuronales*

		Predicción	
		Sin Falla	Falla
Valores verdaderos	Sin Falla	14	7
	Falla	2	22

Son 45 instancias 21 con el cojinete en buen estado y 24 con falla de los que se acertó 36 veces dando una exactitud igual 0,800 (36/45). De las 21 instancias

sin falla se acertaron en 14 para un recall de 0,67 (14/21) indicando una baja discriminación, se predijo en 16 ocasiones que no había fallas dando una precisión de 0,88 (14/16) como se muestra en la tabla 4.50. El f-score para esta clase es:

$$\text{f-score} = 2 \times \frac{0,67 \times 0,88}{0,67 + 0,88} = 0,76$$

Tabla 4.50: *Evaluación de la clasificación con ALD. Caso 4.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,88	0,67	0,76	21
Falla	0,76	0,92	0,83	24
Promedio/Total	0,81	0,80	0,80	45

En cuanto a la clase 2 la precisión es 0,76 y el recall 0,92. El promedio de las métricas es 0,80.

En resumen con el ALD se logran altos valores de discriminación en la detección de si hay fallas en cojinetes.

4.8. Clasificación Múltiple partiendo del Análisis] Lineal Discriminante

Para lograr la clasificación múltiple con el ALD el procedimiento consiste en construir una función discriminante entre cada par de grupo, entonces un dato desconocido quedará clasificado en el grupo al que esté a la menor distancia.

Al usar datos cuyos atributos son las métricas estadísticas obtenidas de la señal de corriente en el dominio del tiempo el desempeño del ALD es muy inferior a las SVM y ANN. Al crear la función discriminante el mejor resultado de la técnica de validación cruzada fue 42,6 % con un tiempo de 5 ms; la matriz de confusión se muestra en la tabla 4.51.

Tabla 4.51: *Matriz de confusión. Clasificación múltiple con ALD con señal en el dominio del tiempo*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	14	2	3	5
	FB	2	13	3	6
	FPE	4	3	10	8
	FPI	1	1	2	13

De la matriz de confusión se calculan las métricas que evalúan al ALD y son mostradas en la tabla 4.52. El recall más bajo 0,40 es para la falla en pista externa y la precisión más baja es 0,41 al predecir la falla en pista interna. En promedio para todas las clases las métricas estuvieron por debajo de 0,6.

Tabla 4.52: *Evaluación de la clasificación múltiple con el ALD y la señal en el dominio del tiempo.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,67	0,58	0,62	24
FB	0,68	0,54	0,60	24
FPE	0,56	0,40	0,47	25
FPI	0,41	0,76	0,53	17
Promedio/Total	0,59	0,56	0,56	90

Es posible que en la extracción de rasgos de la corriente en el dominio del tiempo se haya perdido información necesaria para que el ALD pueda separar los grupos de una manera más eficiente. Con el propósito de mejorar el desempeño del ALD se aumentaron la cantidad de atributos de los datos, para ello se trabaja con la señal de corriente en el dominio de la frecuencia en un rango de 170 a

325 Hz. La figura 4.4 muestra con fines ilustrativos dos señales adquiridas, una sin falla y la otra con falla en bola.

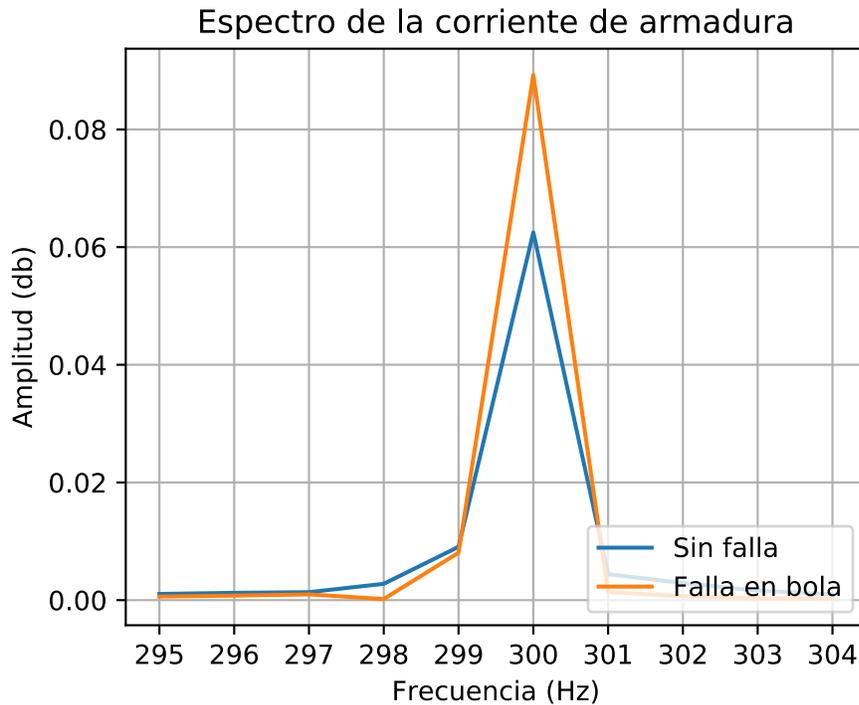


Figura 4.4: Comparación de dos espectros de la corriente armadura.

Luego de obtener la señal el dominio de la frecuencia se utiliza el PCA para que los datos no estén correlacionados entre sí, finalmente se crearon las funciones discriminantes. El tiempo de cálculo fue de 32,58 ms con un mejor porcentaje de validación cruzada de 49,6 %, la discriminación del tipo de falla en cojinete mejora con respecto al caso anterior como se muestra en la matriz de confusión dada por la tabla 4.53 en donde la exactitud es 82,2 %.

Se observa en la matriz de confusión que cuando hay una falla en pista interna no hay falsos negativos, por lo que el recall es de 1,00, mientras que para la falla en pista externa es de 0,60 según se muestra en la tabla 4.54. La mayor precisión se obtuvo al predecir la falla en bola 0,90, mientras que cuando se predice que no hay fallas en cojinetes la precisión disminuye hasta 0,73.

Tabla 4.53: *Matriz de confusión. Clasificación múltiple con ALD con señal en el dominio de la frecuencia*

		Predicción			
		Sin Falla	FB	FPE	FPI
Valores verdaderos	Sin Falla	19	0	4	1
	FB	0	23	1	0
	FPE	7	2	15	1
	FPI	0	0	0	17

Tabla 4.54: *Evaluación de la clasificación múltiple con el ALD y la señal en el dominio de la frecuencia.*

	Precisión	Recall	f1-score	Total
Sin Falla	0,73	0,79	0,76	24
FB	0,92	0,96	0,94	24
FPE	0,75	0,60	0,67	25
FPI	0,89	1,00	0,94	17
Promedio/Total	0,82	0,82	0,82	90

En comparación con las SVM el desempeño fue más pobre, ya que fueron superadas 82,2% vs 88% en exactitud, tiempo de cálculo 32,58 ms vs 5,32 ms; el mejor porcentaje en la técnica de validación cruzada 49,6% vs 88,5% y en las métricas que evalúan el desempeño 0,82 vs 0,89.

Al comparar con la red neuronal, que fue entrenada con 5 atributos (caso 2), las métricas en promedio son similares, sin embargo en la red son más uniformes.

CONCLUSIONES

En tiempo de vida del motor de inducción pueden ocurrir fallas en algunas de sus partes que, aun sea en estado incipiente, deben ser atendidas con prontitud para evitar que causen mayores daños o que se interrumpa el proceso. En este trabajo se estudió como las fallas incipientes en cojinete afectan a la corriente de armadura del motor, y se compararon algunos de los métodos emergentes que se usan para analizarla como son: el análisis lineal discriminante, las máquinas de vectores soporte y las redes neuronales.

Los daños en cojinetes se clasifican como fallas en bola o en pistas, externa e interna. Se comprobó que una falla en estado incipiente produce pequeñas alteraciones en la corriente ya sea en el dominio del tiempo o en el espectro en frecuencia. Estas alteraciones pudieron ser observadas por medio del análisis de la corriente, a la que se le extrajeron ciertas métricas estadísticas que sirvieron no solo para indicar si había falla, sino el tipo de falla en cojinete. Cuando se usó el análisis espectral se observó, tal y como se esperaba de la revisión bibliográfica, que se producen cambios en las bandas de frecuencia.

En el estudio de los clasificadores se observó que su eficacia depende de una correcta preparación de la data. Los mejores resultados se obtuvieron cuando se realizó un proceso de normalización. Adicionalmente se usó un análisis de componentes principales para evitar que los datos estén correlacionados entre sí.

Se realizaron dos estudios para comparar el desempeño de los clasificadores. El primero consistió en observar con que grado se detecta la condición del cojinete, es decir si hay o no falla. El segundo caso, se refiere al estudio de la capacidad de

los clasificadores en indicar el tipo de falla en cojinetes.

Se observó que los tres clasificadores detectan si hay fallas o no con una exactitud cercana al 93 %, siendo el ALD y las ANN los de mejor desempeño con 95 %.

En cuanto al estudio del desempeño de los clasificadores para discriminar el tipo de falla en cojinete son las SVM las de mejor resultado con una exactitud de hasta 89 %. Las ANN tuvieron una exactitud de 80 %, mientras que en el ALD 82 %.

Dos factores importantes que evalúan la discriminación del tipo de falla son la precisión y el recall. En las ANN la precisión al predecir una falla en bola fue de 0,80, mientras es recall al clasificar las fallas en pista fue de 0,78 siendo éstas sus peores evaluaciones.

La precisión con que el ALD predice que hay falla en pista externa es 0,75 mientras el recall con que se clasifica esa misma falla es 0,60. La peor precisión de las SVM fue de 0,81 para la falla en pista externa, y un recall de 0,82 para la falla en pista interna.

En general el mejor desempeño fue el de las máquinas de vectores soporte ya que, además de tener la mejor exactitud, clasifica mejor la condición del cojinete.

RECOMENDACIONES

Entre los trabajos que se pueden derivar de este tema para futuras investigaciones se tienen:

1. Detección de fallas en rodamiento en donde se indique el grado de severidad, con el propósito de indicar cuanto tiempo de operación le queda al motor sin que se produzca su parada.
2. Uso de aprendizaje no supervisado para detección de fallas incipientes en motores de inducción.
3. Estudio del método del análisis de la corriente de armadura, para detectar fallas incipientes, en motores de inducción que contengan variadores de velocidad.
4. Estudio del efecto que tienen las cargas de ciclo variable en el análisis de corriente para la detección de fallas en cojinete.

Bibliografía

- Alguindigue, I. E., Loskiewicz-Buczak, A., y Uhrig, R. E. (1993). Monitoring and diagnosis of rolling element bearings using artificial neural networks. *IEEE transactions on industrial electronics*, 40(2), 209–217.
- Bagheri, B., Ahmadi, H., y Labbafi, R. (2010). Application of data mining and feature extraction on intelligent fault diagnosis by artificial neural network and k-nearest neighbor. En *Electrical machines (icem), 2010 xix international conference on* (pp. 1–7).
- Benbouzid, M. E. H. (2000). A review of induction motors signature analysis as a medium for faults detection. *IEEE Transactions on Industrial Electronics*, 47(5), 984–993.
- Bin Hasan, M. (2012). *Current based condition monitoring of electromechanical systems: model-free drive system current monitoring: faults detection and diagnosis through statistical features extraction and support vector machines classification* (phdthesis). University of Bradford.
- Filippetti, F., Franceschini, G., Tassoni, C., y Vas, P. (2000). Recent developments of induction motor drives fault diagnosis using ai techniques. *IEEE transactions on industrial electronics*, 47(5), 994–1004.
- Fitzgerald, A. E., Kingsley, C., Umans, S. D., y James, B. (2003). *Electric machinery* (Vol. 5). McGraw-Hill New York.
- Giordanelli, F. (1997). *Mantenimiento predictivo de equipos rotativos basado en el análisis de vibraciones* (Tesis de Master). Universidad Simón Bolívar.
- Gonzales, R., y J, M. (2009). *Detección de fallas en los cojinetes del motor de*

inducción a través del análisis espectral de corriente aplicando máquinas de vectores soporte (Tesis de Master, Facultad de Ingeniería, Universidad de Carabobo). Universidad de Carabobo.

- Gunn, S. R., y cols. (1998). Support vector machines for classification and regression. *ISIS technical report*, 14.
- Hagan, M. T., Demuth, H. B., Beale, M. H., y cols. (1996). *Neural network design* (Vol. 20). Pws Pub. Boston.
- Hong, W.-P., Yoon, C.-S., y Kim, D.-H. (2005). Stator current processing-based technique for bearing damage detection in induction motor. *ICCAS2005, June2-5*.
- Hurtado, J. (2007). El proyecto de investigación. *Metodología de la investigación holística*, 183.
- Jiménez, C., y Jiménez, J. (2010). Clasificación de señales de voz utilizando transformada de wavelet y máquinas de vectores de soporte [Trabajo de Ascenso]. *Trabajo de ascenso. Facultad de Ingeniería. Universidad de Carabobo*.
- Jin, X., y Chow, T. W. (2013). Anomaly detection of cooling fan and fault classification of induction motor using mahalanobis-taguchi system. *Expert Systems with Applications*, 40(15), 5787–5795.
- Kowalski, C. T., y Orłowska-Kowalska, T. (2003). Neural networks application for induction motor faults diagnosis. *Mathematics and Computers in Simulation*, 63(3), 435–448.
- Lei, Y., He, Z., y Zi, Y. (2008). A new approach to intelligent fault diagnosis of rotating machinery. *Expert Systems with Applications*, 35(4), 1593–1600.
- Mago, M., y Monagas, E. (2004). *Fallas más comunes en los motores de inducción de empresas del sector industrial del edo carabobo*. Universidad de Carabobo.
- Neelam, M., y Ratna, D. (2007). An approach of condition monitoring of induction motor using MCSA. *International Journal of Systems Applications, Engineering & Development*, 1(1), 13–17.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O.,

- ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Peña, D. (2002). *Análisis de datos multivariantes* (Vol. 24). McGraw-Hill Madrid.
- Pérez, C., y Santín, D. (2007). Minería de datos: técnicas y herramientas. *Madrid, España: International Thompson Ediciones Paraninfo*.
- Rencher, A. C. (2003). *Methods of multivariate analysis* (Vol. 492). John Wiley & Sons.
- Schoen, R. R., Habetler, T. G., Kamran, F., y Bartfield, R. (1995). Motor bearing damage detection using stator current monitoring. *IEEE transactions on industry applications*, 31(6), 1274–1279.
- Sin, M., Soong, W., y Ertugrul, N. (2003). Induction machine on-line condition monitoring and fault diagnosis—a survey. En *Australasian universities power engineering conference* (Vol. 28, pp. 1–6).
- Trigeassou, J.-C. (s.f.). Fault diagnosis of the induction machine by neural networks. *Electrical Machines Diagnosis*, 227–270.
- Wen, X. (2011). *A hybrid intelligent technique for induction motor condition monitoring* (phdthesis). University of Portsmouth.
- Widodo, A., y Yang, B.-S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, 21(6), 2560–2574.