

Universidad Autónoma de San Luis Potosí  
Facultad de Ingeniería  
Centro de Investigación y Estudios de Posgrado

Título de la Tesis

Diagnóstico de Fallas en Motores de Inducción Utilizando Señales Eléctricas y Mecánicas

Tesis que para obtener maestría en  
Ingeniería Eléctrica  
Especialidad Control Automático

Presenta

Ing. José Daniel Martínez Morales

San Luis Potosí, S.L.P.

Ñoviembre 2009



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ  
FACULTAD DE INGENIERÍA  
CENTRO DE INVESTIGACIÓN Y ESTUDIOS DE POSGRADO

MAESTRÍA EN INGENIERÍA ELÉCTRICA  
ESPECIALIDAD EN CONTROL AUTOMÁTICO

"DIAGNÓSTICO DE FALLAS EN MOTORES DE INDUCCIÓN UTILIZANDO  
SEÑALES ELÉCTRICAS Y MECÁNICAS"

PRESENTA

José Daniel Martínez Morales

Dr. Daniel Ulises Campos Delgado  
(Director de tesis)

Dra. Elvia Ruth Palacios Hernández  
(Co-directora de tesis)

Dr. Ricardo Álvarez Salas  
(Sinodal titular 1)

Dr. José Salomé Murguía Ibarra  
(Sinodal titular 2)

Dra. Nancy Visairo Cruz  
(Sinodal suplente)



SAN LUIS POTOSÍ, S.L.P.

NOVIEMBRE 2009

---

## Agradecimientos

---

Agradezco a Dios y a mi familia por apoyarme en los momentos difíciles. Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo mediante el proyecto (CO1-AINAT-01-31.31) ref. 104912, y el apoyo económico otorgado a través de una beca para realizar los estudios de maestría. A mis asesores, el Dr. Daniel Ulises Campos Delgado y la Dra. Elvia Ruth Palacios Hernández por la asesoría proporcionada durante el desarrollo de este trabajo. También deseo agradecer a los profesores del Posgrado en Ingeniería Eléctrica que contribuyeron a mi formación académica y profesional.

---

## Resumen

---

En este trabajo se desarrolló un algoritmo que aborda el problema de diagnóstico de fallas de baleros, desbalance y desalineamiento en motores de inducción (MI) por medio de la medición de señales eléctricas y mecánicas. El algoritmo se diseñó a partir de datos experimentales obtenidos del kit de pruebas “Machine Fault Simulator” mediante la tarjeta de adquisición de datos NI-Compaq-DAQ-9172, para su posterior procesamiento en la plataforma de LabView. Se utilizaron el análisis espectral de las corrientes de línea de estator, el módulo del vector de Park, y las vibraciones mecánicas como medios de diagnóstico. A partir del conocimiento de las frecuencias de falla reportadas en la literatura, se extraen únicamente los componentes espectrales que contienen información referente a las fallas de interés. Las máquinas de soporte vectorial (máquinas de aprendizaje), se utilizan como clasificadores para llevar a cabo el diagnóstico, el objetivo de éstas máquinas consiste en clasificar la información característica de las fallas mediante hiperplanos óptimos en un mapa de características, el diseño de las máquinas de soporte vectorial conlleva la solución de un problema dual de optimización cuadrática. Estas se diseñan sólo para algunas condiciones de operación del MI, y se selecciona aquella que genere el mayor número de predicciones correctas para patrones de entrada nuevos, esto se realiza mediante el principio de minimización de riesgo estructural desarrollado por Vladimir N. Vapnik (2000). Por último,

se evaluó el desempeño del algoritmo para condiciones de velocidad y par de carga variable del motor de inducción, obteniéndose un porcentaje de precisión del 99 % en el diagnóstico para los casos de validación.

---

## Índice general

---

<b>Índice de Figuras</b> . . . . .	VII
<b>Índice de Tablas</b> . . . . .	XI
<b>Nomenclatura</b> . . . . .	XIII
<b>INTRODUCCIÓN</b>	<b>1</b>
Diagnóstico de Fallas en Motores de Inducción . . . . .	1
Objetivos generales . . . . .	5
Estado del Arte . . . . .	6
Aportaciones del trabajo de tesis . . . . .	14
Descripción del Documento . . . . .	14
<b>1. MÁQUINAS DE APRENDIZAJE</b>	<b>16</b>
1.1. Problemas de aprendizaje . . . . .	20
1.2. Máquinas de aprendizaje para clasificación de patrones (RNA) . . . . .	21
1.3. Máquinas de soporte vectorial . . . . .	25
1.3.1. Características del problema general de optimización . . . . .	28
1.3.2. SVM: Para patrones linealmente separables . . . . .	30

1.3.3.	SVM: Para patrones no separables linealmente . . . . .	34
1.3.4.	SVM no lineales . . . . .	37
1.3.5.	Multclasificación SVM uno vs. resto . . . . .	43
1.3.6.	Multclasificación SVM uno vs. uno . . . . .	43
<b>2.</b>	<b>MEDIOS DE DIAGNÓSTICO</b>	<b>45</b>
2.1.	Monitoreo de las corrientes de estator . . . . .	45
2.2.	Monitoreo de la potencia instantánea de entrada del MI . . . . .	48
2.3.	Monitoreo del vector de corrientes de Park . . . . .	50
2.4.	Monitoreo de las vibraciones mecánicas del MI . . . . .	50
2.5.	Monitoreo de los medios de diagnóstico para 3 tipos de falla. . . . .	53
2.6.	Generación de características . . . . .	60
<b>3.</b>	<b>DISEÑO E IMPLEMENTACIÓN DE UNA MÁQUINA DE SOPORTE VECTO-</b>	
	<b>RIAL</b>	<b>79</b>
3.1.	Selección del modelo óptimo . . . . .	79
3.2.	Implementación del algoritmo de diagnóstico mediante LabView . . . . .	88
	<b>CONCLUSIONES</b>	<b>94</b>
	Trabajo Realizado . . . . .	94
	Conclusiones . . . . .	95
	Trabajo Futuro . . . . .	96
	<b>APÉNDICE</b>	<b>97</b>
	<b>A. Implementación en LabView del algoritmo de diagnóstico</b>	<b>97</b>
	<b>BIBLIOGRAFÍA</b>	<b>103</b>

---

## Índice de figuras

---

1.	<i>Tipos de falla en un sistema según su comportamiento en el tiempo.</i>	2
2.	<i>Distribución del porcentaje de fallas en los componentes del MI [2], [3].</i>	3
3.	<i>Sistema de detección de fallas en baleros del MI basado en redes neuronales [16].</i>	7
4.	<i>Dibujo del simulador de fallas (1: encoder; 2: MI trifásico; 3: sensor de temperatura; 4: acoplamiento del eje; 5: accionamiento de frecuencia variable del motor; 6: sensor de corriente; 7: sensor de vibración en el eje x; 8: balero delantero; 9: disco rotacional; 10: sensor de vibración del eje y; 11: eje del motor; 12: disco rotacional 2; 13: balero posterior).</i>	8
5.	<i>Diagrama a bloques del esquema de diagnóstico usando RNA y el vector de corrientes de Park [17].</i>	9
6.	<i>Diagrama a bloques del esquema de diagnóstico usando RNA [18].</i>	10
7.	<i>Diagrama a bloques del esquema de diagnóstico basado en lógica difusa [19].</i>	11
8.	<i>Esquema de diagnóstico difuso basado en el patrón de Concordia [20].</i>	12
1.1.	<i>Estructura anidada de las funciones de aproximación.</i>	19
1.2.	<i>Diagrama de una neurona artificial</i>	22
1.3.	<i>a) RNA monocapa, b) RNA multicapa.</i>	23



1.4. <i>Ilustración de la idea de un hiperplano para la clasificación de patrones en 2 dimensiones</i> . . . . .	26
1.5. <i>Definición de regiones de la superficie de decisión.</i> . . . . .	27
1.6. <i>Cálculo del margen entre dos vectores de soporte.</i> . . . . .	30
1.7. <i>Patrones no separables linealmente.</i> . . . . .	35
1.8. <i>Obtención de multiplicadores de Lagrange <math>\alpha</math>, usando el Kernel <math>K(x_i, x_j) = \exp\left(-\frac{\ x_i - x_j\ ^2}{2\sigma^2}\right)</math>.</i> . . . . .	41
1.9. <i>Función de decisión <math>d(x_0, \alpha^*)</math> para un ejemplo de SVM.</i> . . . . .	42
1.10. <i>Hiperplano óptimo en el espacio de entrada <math>\mathbb{R}^2</math> para un ejemplo de SVM.</i> . . . . .	42
1.11. <i>Región no clasificable para el esquema de multclasificación uno vs. resto</i> . . . . .	43
2.1. <i>Geometría de un balero.</i> . . . . .	48
2.2. <i>Circuito para la medición de las vibraciones mecánicas del motor de inducción.</i> . . . . .	52
2.3. <i>Zonas relacionadas a la severidad de los defectos en un rodamiento [42].</i> . . . . .	54
2.4. <i>PSD de la corriente de línea de los MI sin falla y desalineado (dB).</i> . . . . .	55
2.5. <i>PSD del módulo del vector de Park para un MI sin falla y desalineado (dB).</i> . . . . .	55
2.6. <i>PSD de la potencia instantánea total para un MI sin falla y desalineado (dB).</i> . . . . .	56
2.7. <i>PSD de la señal de vibración en dirección radial para un MI sin falla y desalineado (magnitud).</i> . . . . .	56
2.8. <i>PSD de la corriente de línea para un MI sin falla y desbalanceado (dB).</i> . . . . .	57
2.9. <i>PSD del módulo del vector de Park para un MI sin falla y desbalanceado (dB).</i> . . . . .	58
2.10. <i>PSD de la señal de vibración en dirección radial para un MI sin falla y desbalanceado (magnitud).</i> . . . . .	58
2.11. <i>PSD de la señal de vibraciones del MI sin falla y con falla en baleros (magnitud).</i> . . . . .	59
2.12. <i>Generación en Labview de las características <math>c(i_a)</math> y <math>c(\ i_{dq}\ )</math>.</i> . . . . .	63
2.13. <i>Extracción de la característica <math>c(P_{abc})</math>.</i> . . . . .	64
2.14. <i>PSD de la señal de vibraciones del MI sin falla y con falla en baleros (magnitud).</i> . . . . .	64
2.15. <i>Voltaje de línea fundamental <math>V_{ab}</math> para <math>f_s = 30</math> Hz.</i> . . . . .	65
2.16. <i>Voltaje de línea fundamental <math>V_{ab}</math> para <math>f_s = 60</math> Hz.</i> . . . . .	65

2.17. <i>Potencia instantánea total de un MI para diferentes frecuencias de alimentación.</i>	65
2.18. <i>Kit de pruebas “Machine Fault Simulator”.</i>	68
2.19. <i>Mapa característico <math>c_1(v)</math> vs. <math>c_2(v)</math>.</i>	68
2.20. <i>Mapa característico <math>c_1(v)</math> vs. <math>c_3(v)</math>.</i>	71
2.21. <i>Mapa característico <math>c_1(v)</math> vs. <math>c_4(v)</math>.</i>	72
2.22. <i>Mapa característico <math>c_1(v)</math> vs. <math>c(i_a)</math>.</i>	72
2.23. <i>Mapa característico <math>c_1(v)</math> vs. <math>c(\ i_{dq}\ )</math>.</i>	73
2.24. <i>Mapa característico <math>c_2(v)</math> vs. <math>c_3(v)</math>.</i>	73
2.25. <i>Mapa característico <math>c_2(v)</math> vs. <math>c_4(v)</math>.</i>	74
2.26. <i>Mapa característico <math>c_2(v)</math> vs. <math>c(i_a)</math>.</i>	75
2.27. <i>Mapa característico <math>c_2(v)</math> vs. <math>c(\ i_{dq}\ )</math>.</i>	75
2.28. <i>Mapa característico <math>c_3(v)</math> vs. <math>c_4(v)</math>.</i>	76
2.29. <i>Mapa característico <math>c_3(v)</math> vs. <math>c(i_a)</math>.</i>	76
2.30. <i>Mapa característico <math>c_3(v)</math> vs. <math>c(\ i_{dq}\ )</math>.</i>	77
2.31. <i>Mapa característico <math>c_4(v)</math> vs. <math>c(i_a)</math>.</i>	77
2.32. <i>Mapa característico <math>c_4(v)</math> vs. <math>c(\ i_{dq}\ )</math>.</i>	78
2.33. <i>Mapa característico <math>c(i_a)</math> vs. <math>c(\ i_{dq}\ )</math>.</i>	78
3.1. <i>Clasificación binaria (MI sin falla vs. MI desalineado) con kernel gaussiano y <math>\sigma = 2</math>.</i>	88
3.2. <i>Clasificación binaria (MI desalineado vs. MI desbalanceado) con kernel gaussiano y <math>\sigma = 4</math>.</i>	89
3.3. <i>Esquema general de diagnóstico de fallas mediante SVM.</i>	90
3.4. <i>Implementación de una SVM para clasificación binaria de un MI sin falla y con desalineamiento.</i>	91
3.5. <i>Panel de control que indica la presencia de una falla.</i>	92
A.1. <i>Medición de los medios de diagnóstico mediante LabView.</i>	98
A.2. <i>Extracción de las características de clasificación.</i>	99

A.3. Construcción del vector de características. . . . .	100
A.4. Datos de entrenamiento en mathscript para calcular la salida de las SVM. . . .	101
A.5. Esquema de votación para determinar el tipo de falla en el MI. . . . .	102

---

## Índice de Tablas

---

1.	Resumen de las causas y clases de fallas generadas en un MI [4], [5]. . . . .	4
1.1.	Datos de entrenamiento para un ejemplo de SVM . . . . .	39
2.1.	Resumen de técnicas para diagnosticar fallas en un MI. . . . .	46
2.2.	Relación entre los rangos de frecuencias de falla y el número de muestras en la PSD. . . . .	63
2.3.	Características basadas en la potencia instantánea total del MI sin falla y con desalineamiento. . . . .	66
2.4.	Especificaciones del equipo de adquisición. . . . .	67
2.5.	Características para un MI sin falla y desalineado. . . . .	69
2.6.	Características para un MI con desbalance y falla en baleros. . . . .	70
3.1.	Selección de SVM para clasificar un MI sin falla ( $y = 1$ ) y un MI con desalineamiento ( $y = -1$ ) . . . . .	82
3.2.	Selección de SVM para clasificar un MI sin falla ( $y = 1$ ) y un MI con desbalance ( $y = -1$ ) . . . . .	83

3.3. Selección de SVM para clasificar un MI sin falla ( $y = 1$ ) y un MI con falla en baleros ( $y = -1$ ) . . . . .	84
3.4. Selección de SVM para clasificar un MI con desalineamiento ( $y = 1$ ) y un MI con desbalance ( $y = -1$ ) . . . . .	85
3.5. Selección de SVM para clasificar un MI con desalineamiento ( $y = 1$ ) y un MI con falla en baleros ( $y = -1$ ) . . . . .	86
3.6. Selección de SVM para clasificar un MI con falla en baleros ( $y = 1$ ) y desbalance ( $y = -1$ ) . . . . .	87
3.7. Porcentajes de clasificación correcta de la fase de prueba del esquema de diagnóstico . . . . .	90
3.8. Características para un MI con desalineamiento. . . . .	93

---

## Nomenclatura

---

FFT	<i>Transformada Rápida de Fourier (por sus siglas en inglés)</i>
MI	<i>Motor de inducción</i>
PSD	<i>Densidad espectral de potencia (por sus siglas en inglés)</i>
RNA	<i>Red neuronal artificial</i>
RPM	<i>Revoluciones por minuto</i>
SVM	<i>Máquina de soporte vectorial (por sus siglas en inglés)</i>
VC	<i>Vapnik-Chervonenkis</i>
$\ \cdot\ $	<i>Norma Euclidiana</i>
$\langle\cdot,\cdot\rangle$	<i>Producto interno</i>
$\nabla(\cdot)$	<i>Vector gradiente</i>
$\alpha$	<i>Vector de multiplicadores de Lagrange</i>
$\beta$	<i>Ángulo de contacto entre los elementos rodantes del balero y la pista externa</i>
$bd$	<i>Diámetro de los elementos rodantes del balero</i>
$D(k)$	<i>Transformada discreta de Fourier</i>
$\eta$	<i>La cota para el riesgo de una clasificación errónea se mantiene con probabilidad <math>1 - \eta</math></i>

$f_{bnq}$	<i>Frecuencias de falla de baleros en el espectro de las corrientes de estator</i>
$f_m$	<i>Frecuencias de falla de desbalance y desalineamiento en el espectro de las corrientes de estator</i>
$f_r$	<i>Frecuencia de rotación del MI en Hz.</i>
$f_s$	<i>Frecuencia de alimentación del MI</i>
$h$	<i>Dimensión de Vapnik-Chervonenkis</i>
$H_{ij}$	<i>Matriz Hessiana</i>
$i_f$	<i>Regla de decisión</i>
$(i_d, i_q)$	<i>Componentes del vector de corrientes de Park</i>
$\ i_{dq}\ $	<i>Módulo del vector de Park</i>
$K(x, x_i)$	<i>Función kernel</i>
$l$	<i>Número de datos de entrenamiento</i>
$\mathcal{L}$	<i>Lagrangiano</i>
$\Gamma$	<i>Intervalo de confianza</i>
$p$	<i>Número de pares de polos del motor de inducción</i>
$\phi$	<i>Ángulo del factor de potencia</i>
$P_{abc}$	<i>Potencia instantánea total de entrada</i>
$(P_{ab}, P_{cb})$	<i>Potencias instantáneas parciales de entrada</i>
$pd$	<i>Diámetro del cojinete de los elementos rodantes del balero</i>
$R$	<i>Diámetro de la esfera más pequeña que incluye a todos los datos de entrenamiento en el espacio característico</i>
$\mathbb{R}$	<i>Números reales</i>
$R_{emp}$	<i>Riesgo empírico del error de clasificación para el conjunto de entrenamiento</i>
$\sigma$	<i>Ancho de la función de base radial</i>
$w$	<i>Vector de pesos de la máquina de aprendizaje</i>
$\omega_0$	<i>Frecuencia angular del voltaje de alimentación</i>

$\omega_f$	<i>Frecuencia de falla en las corrientes de fase del MI</i>
$x$	<i>Vector de datos a ser clasificados</i>
$c(i_a)$	<i>Característica relacionada a una falla de desalineamiento y desbalance en el espectro de la corriente de estator</i>
$c(\ i_{dq}\ )$	<i>Característica relacionada a una falla de desalineamiento y desbalance en el espectro del módulo del vector de Park</i>
$c(P_{abc})$	<i>Característica relacionada a una falla de desalineamiento y desbalance en el espectro de la potencia instantánea total</i>
$c_1(v)$	<i>Característica relacionada a una falla de desbalance en el espectro de la señal de vibración</i>
$c_2(v)$	<i>Característica relacionada a una falla de desalineamiento en el espectro de la señal de vibración</i>
$c_3(v)$	<i>Característica relacionada a una falla de desalineamiento en el espectro de la señal de vibración</i>
$c_4(v)$	<i>Característica relacionada a una falla de baleros en el espectro de la señal de vibración</i>
$y_i$	<i>Clase asignada al patrón de entrada <math>x_i</math>, <math>y_i \in \{+1, -1\}</math></i>



---

# INTRODUCCIÓN

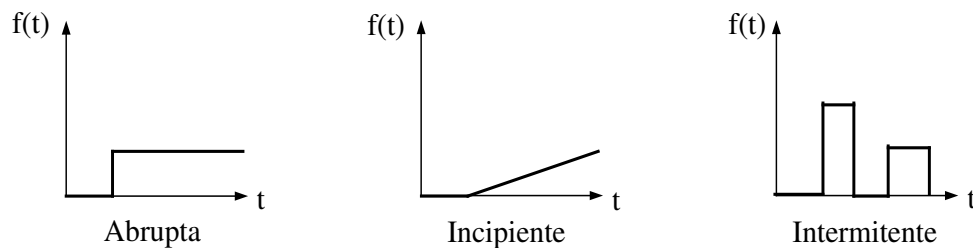
---

## **Diagnóstico de Fallas en Motores de Inducción**

El motor de inducción (MI), es la máquina eléctrica más utilizada en el entorno industrial debido a su bajo costo, alta confiabilidad, robustez y menor mantenimiento comparado con cualquier otra máquina rotativa, al carecer de escobillas, sin embargo, el mantenimiento de éstas máquinas ha cobrado una gran importancia en las últimas décadas debido a las exigencias de la calidad en las líneas de producción de las empresas. En este sentido, mediante el monitoreo de las condiciones de las máquinas de inducción, es posible detectar el desarrollo de los problemas relacionados con éstas y efectuar acciones correctivas solamente cuando así se requiera, evitando paros innecesarios de los MI y las pérdidas económicas asociadas. Debido a esto, el diagnóstico de fallas en MI ha cobrado gran interés en los últimos años en la comunidad científica. Las fallas son el resultado del desgaste normal, mal uso o un diseño erróneo de los componentes de un sistema y se definen como una desviación no permitida de por lo menos una característica o parámetro que indique el comportamiento normal del mismo.

En la Figura 1 se muestra una clasificación del tipo de fallas, entre ellas están las fallas abruptas que puede ser la interrupción de la fuente de energía del sistema; incipiente como por

ejemplo el desgaste en los componentes eléctricos o mecánicos del sistema; o intermitente como un falso contacto. En este trabajo se estudia el diagnóstico de fallas del tipo incipiente en MI.



**Figura 1:** Tipos de falla en un sistema según su comportamiento en el tiempo.

El diagnóstico de fallas involucra las siguientes acciones [1]:

- Detección de la Falla: determina la presencia de una falla en el sistema.
- Aislamiento e Identificación de las Fallas: encuentra en que componente del sistema ocurrió la falla y su severidad.

En un MI, las fallas se clasifican según sus componentes, es decir, fallas de rotor, de estator, de rodamientos, e incluso de los dispositivos externos como son los actuadores. Según estudios realizados sobre motores en compañías relacionadas con la industria petroquímica, terminales de gas y refinerías [2], [3], se encontró, que a pesar de las discrepancias debido al tipo de motores seleccionados en ambos estudios, en general, se mantienen las siguientes tendencias:

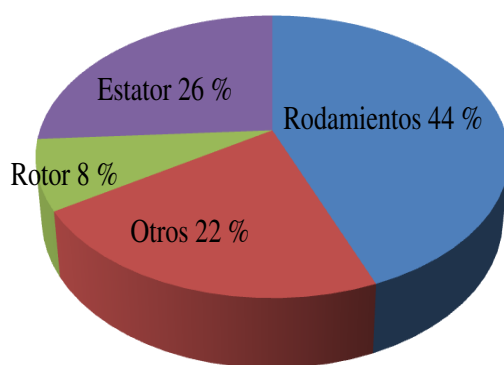
- Si el motor opera de manera continua por un tiempo prolongado tiene un porcentaje más alto de probabilidad de falla.
- Para un mayor número de horas diarias promedio de operación del motor, se muestra un incremento en el porcentaje de fallas.
- La tasa de fallas<sup>1</sup> es más alta para motores que se encuentran a la intemperie.

<sup>1</sup>La tasa de fallas se define como la frecuencia con la cual un motor falla en alguno de sus componentes y puede ser expresado en fallas/horas

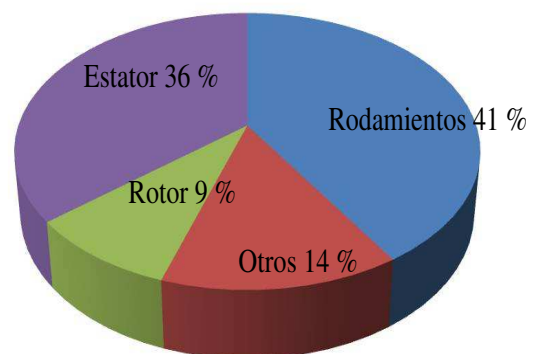
- La tasa de fallas es más alta para los MI de dos polos.
- Con respecto a la antigüedad, los motores que tienen una menor o mayor antigüedad tienen una tasa de fallas alta, tal como lo indica la curva de forma de bañera mostrada en [1].
- Las causas que contribuyen a la presencia de la mayoría de las fallas son: el deterioro normal por envejecimiento, así como lubricación inapropiada y vibración alta.

En la Figura 2 se muestran las distribuciones de fallas presentadas en [2], [3], en estas se aprecia que las fallas de baleros ocupan cerca del 40 % del total de las fallas detectadas, mientras que las fallas en los devanados ocupan entre el 26 % y 36 %, siguiendo las fallas debidas a los dispositivos externos con un 14 %, sólo estos tres tipos de fallas representan alrededor del 90 % del total. Mientras que, entre el 8 % y 9 % de las fallas se ocasionan en el rotor. Por lo anterior, el desarrollo de nuevas estrategias de monitoreo con los equipos en funcionamiento, es vital en el sector industrial debido a que es posible evitar el paro de las líneas de producción o incurrir en gastos innecesarios en mantenimiento. En la Tabla 1 se muestran las causas relacionadas con los tipos de fallas del MI reportadas en la literatura [4], [5]. Para llevar a cabo el diagnóstico de fallas se han propuesto diversos métodos en la literatura [1]:

Estudio EPRI (Electric Power Research Institute) [2]



Estudio IEEE [3]



**Figura 2:** Distribución del porcentaje de fallas en los componentes del MI [2], [3].

**Tabla 1:** Resumen de las causas y clases de fallas generadas en un MI [4], [5].

<b>Tipos de falla</b>	<b>Causas</b>
Baleros	Corrosión, contaminación, lubricación o instalación inapropiada, sobrecalentamiento, vibración alta, esfuerzos mecánicos radiales y axiales debido a una deflexión del eje.
Fallas en los devanados del estator	Mal funcionamiento del sistema eléctrico, deterioro normal del aislamiento por envejecimiento, humedad excesiva, sobrecalentamiento, mantenimiento inadecuado, vibración alta, descargas eléctricas y protecciones eléctricas inadecuadas.
Desalineamiento	Torque en forma de pulsos, vibraciones, variaciones de velocidad.
Fallas de excentricidad	Mala instalación del eje del rotor, ovalidad del núcleo del estator, deflexión del eje, envejecimiento normal de los baleros o desalineamiento.

### **Procesamiento de señales**

Los métodos basados en señales no incorporan modelo alguno del sistema a tratar, entre estos métodos se encuentran el análisis de firmas de corrientes del motor (MCSA, por sus siglas en inglés), biespectros, análisis wavelet, entre otros, [6], [7], [8]. Para llevar a cabo estos métodos es necesaria la disponibilidad de sistemas de adquisición de gran capacidad.

### **Basados en modelos**

Estos métodos de diagnóstico de fallas hacen uso de redundancia analítica, es decir, modelos matemáticos del sistema monitoreado, la idea básica es comparar las mediciones con los valores obtenidos computacionalmente a través del modelo, para generar señales llamadas residuos que indiquen la presencia de una falla en el sistema [9], [10], [11]. Estos residuos sólo deben ser sensibles ante fallas y no a perturbaciones, ruido en las mediciones, e incertidumbres en el modelo.

### **Basados en conocimiento**

Estos métodos utilizan la información de los expertos en diagnóstico, la información utilizada proviene de los medios de diagnóstico para un MI con y sin falla para realizar el diseño de máquinas de aprendizaje, las cuales posteriormente podrán reconocer y clasificar los patrones o características de acuerdo a la condición de falla. Entre estos métodos se encuentran las redes neuronales artificiales (RNA), lógica difusa, y las máquinas de soporte vectorial [12], [13], [14].

## **Objetivos generales**

En esta tesis se presenta el desarrollo de un algoritmo de diagnóstico basado en máquinas de soporte vectorial para aislar las fallas de rodamientos, desbalance y desalineamiento en motores de inducción. Los objetivos generales de esta tesis son

- Estudiar las propiedades de los residuos basados en la potencia total de alimentación y el radio del vector de Park para diagnosticar fallas en el rotor.

- 
- Generar residuos mecánicos utilizando mediciones de vibración en el MI y evaluar su robustez.
  - Integrar la información proporcionada por los medios de diagnóstico eléctricos y mecánicos mediante las máquinas de soporte vectorial para mejorar las propiedades de clasificación del algoritmo global de diagnóstico de fallas.
  - Desarrollar una plataforma experimental utilizando una tarjeta de adquisición de datos NI Compaq-DAQ para implementar el algoritmo de diagnóstico.

## Estado del Arte

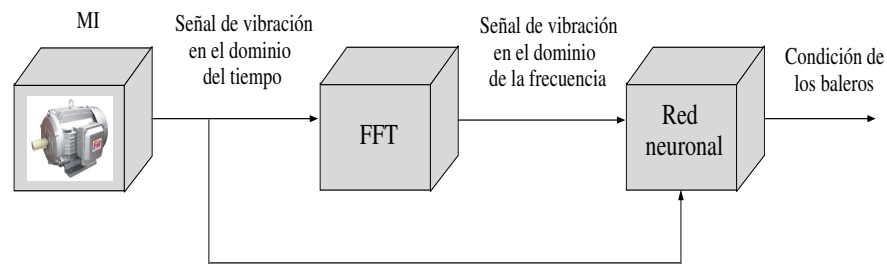
Los métodos de diagnóstico de fallas también pueden dividirse en dos grandes categorías: métodos de clasificación o reconocimiento de patrones, y métodos de inferencia. Los métodos de clasificación se utilizan cuando no existe conocimiento disponible acerca de la relación que existe entre los síntomas y las fallas del sistema. Mientras que el uso de los métodos de inferencia involucra relaciones básicas entre las fallas y los síntomas; una forma de realizar este tipo de esquemas de diagnóstico es expresando el conocimiento cualitativo en forma de reglas *si - entonces*. La clasificación o inferencia de las fallas se ha llevado a cabo durante mucho tiempo mediante redes neuronales, sistemas difusos o una fusión de estos, sin embargo también se ha reportado en la literatura el uso de máquinas de soporte vectorial (SVM, por sus siglas en inglés) como un método eficiente para la clasificación de fallas [15].

A continuación se presenta una revisión sobre algunos esquemas que se han reportado en la literatura sobre diagnóstico de fallas en motores de inducción usando técnicas de soft computing. El término *soft computing* hace referencia a los métodos que emplean algoritmos de inteligencia computacional como por ejemplo: lógica difusa, redes neuronales, esquemas neuro-difusos, algoritmos genéticos, maquinas de soporte vectorial, entre otros [12], [14].

En específico, para el problema de diagnóstico en el MI, usualmente un experto examinaría la señal en el dominio del tiempo y el espectro en frecuencia de las vibraciones mecánicas pro-

ducidas en el MI para determinar si existe algún defecto dentro de un balero. Sin embargo, la experiencia se ha ganado sobre un largo periodo de tiempo y sería un caro e ineficiente uso de recursos utilizar a los expertos para determinar una condición de falla. Por lo tanto, si una red neuronal se puede entrenar para emular el conocimiento de los expertos en señales de vibración, el diagnóstico de fallas en baleros puede realizarse de manera más eficiente y a un costo reducido.

Debido a la capacidad de aprendizaje, las redes neuronales artificiales (RNA) han recibido una gran atención en el campo del diagnóstico de fallas en motores de inducción. En [16], Bo Li *et al.* presentan un esquema para diagnosticar fallas en los baleros de un MI que utiliza redes neuronales junto con el análisis de las señales de vibración de los baleros en el dominio tiempo/frecuencia. El esquema propuesto se muestra en la Figura 3.



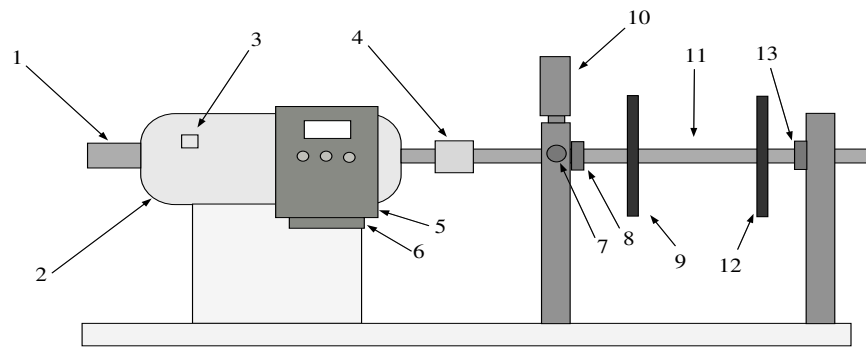
**Figura 3:** Sistema de detección de fallas en baleros del MI basado en redes neuronales [16].

Resulta conveniente monitorear los niveles de vibración de los baleros debido a que son movimientos periódicos cuyas frecuencias características ya han sido reportadas. En el proceso de diagnóstico, los sensores adquieren las señales de vibración en el dominio del tiempo, para después convertirlas al dominio de la frecuencia empleando la técnica de la transformada rápida de Fourier (FFT, por sus siglas en inglés). Las características de vibración de los baleros en el dominio de la frecuencia y del tiempo son aplicadas a la red neuronal con el propósito de construir una máquina automática que realice el diagnóstico de fallas. Después de que las redes neuronales son entrenadas, estas contienen el conocimiento experto y son capaces de identificar correctamente las diferentes causas de la vibración de los baleros.

Para verificar si el sistema de diagnóstico diseñado es capaz de clasificar correctamente las

diferentes condiciones de falla en los baleros, se utilizó el software MotorSim desarrollado en la plataforma Matlab/Simulink, este software se puede utilizar para generar señales de vibración en el dominio del tiempo para diferentes condiciones de operación y carga en una manera rentable (de bajo costo) y eficiente, para después comparar los resultados con las señales de vibración del MI obtenidas experimentalmente.

Los datos reales de la vibración de los baleros fueron generados con un simulador de fallas en máquinas como el que se muestra en la Figura 4, construido por SpectraQuest. Este simulador permite la medición de las vibraciones del motor operando bajo una gran variedad de condiciones de falla.

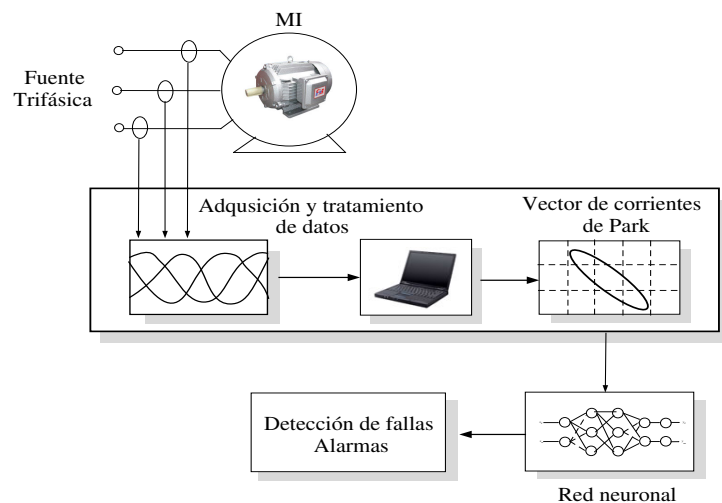


**Figura 4:** Dibujo del simulador de fallas (1: encoder; 2: MI trifásico; 3: sensor de temperatura; 4: acoplamiento del eje; 5: accionamiento de frecuencia variable del motor; 6: sensor de corriente; 7: sensor de vibración en el eje x; 8: balero delantero; 9: disco rotacional; 10: sensor de vibración del eje y; 11: eje del motor; 12: disco rotacional 2; 13: balero posterior).

En [17], se presenta una metodología basada en el vector de corrientes de Park, mediante la cual es posible diagnosticar fallas eléctricas en motores de inducción. Una RNA aprende los patrones del vector de Park de las corrientes del estator para poder discernir entre un motor de inducción con o sin fallas. En este trabajo se consideró una RNA unidireccional (FFNN, por sus siglas en inglés), donde el número de neuronas en la capa de entrada y salida depende de la aplicación en cuestión, mientras que el número de neuronas en la capa oculta se debe ajustar durante el periodo de aprendizaje de tal manera que la red se pueda entrenar eficientemente. En



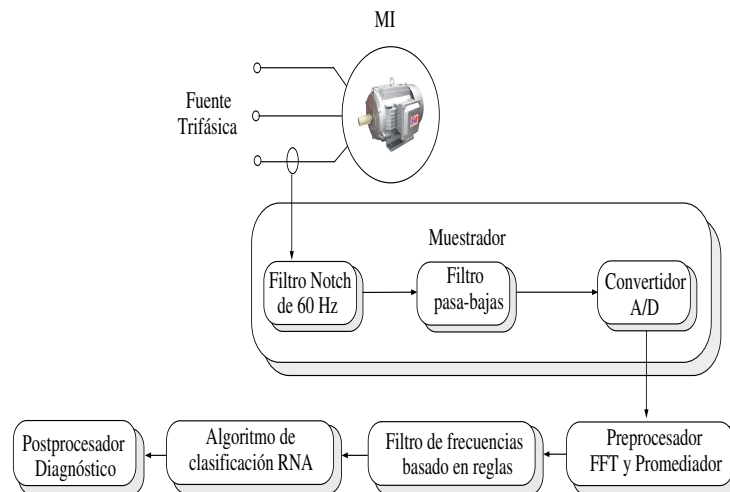
este estudio se utilizó un algoritmo de propagación hacia atrás (backpropagation) para entrenar a la red neuronal. En el esquema de diagnóstico mostrado en la Figura 5, se analizan los patrones del vector de corrientes de Park mediante un sistema de clasificación basado en una red neuronal, para esto es necesario atribuir a cada patrón una etiqueta que describa el estado del MI al instante de la captura del patrón. Entonces la entrada a la red es un patrón y la salida es una etiqueta (0 ó 1) que indica la clase de falla ocurrida. En este trabajo se considera la clasificación de las fallas como: circuito abierto en una fase y desbalance en el voltaje del estator. Los experimentos fueron llevados a cabo bajo condiciones de torque de carga variable.



**Figura 5:** Diagrama a bloques del esquema de diagnóstico usando RNA y el vector de corrientes de Park [17].

En [18], se presenta un método de detección de fallas en línea, utilizando RNA para aprender las características espectrales de un motor libre de fallas. Este espectro aprendido puede contener algunos armónicos debido a la carga, el cual corresponde a condiciones normales de operación. Una de las características más notables del esquema es la utilización de un filtro selectivo de frecuencia con el objeto de reducir el número de armónicos que son monitoreados continuamente. Este filtro sólo deja pasar aquellos componentes armónicos que se conoce son representativos de una falla hacia un algoritmo de clasificación basado en una red neuronal. En este esquema como en los anteriores, se busca automatizar el análisis de las señales medidas ya sea en el dominio del

tiempo o en el dominio de la frecuencia incorporando un sistema experto o una red neuronal en el sistema de monitoreo en línea, para determinar si una falla se está desarrollando en el MI, sin la necesidad de un experto en mantenimiento. El sistema de monitoreo en línea de las corrientes contiene las 5 secciones de procesamiento ilustradas en la Figura 6.

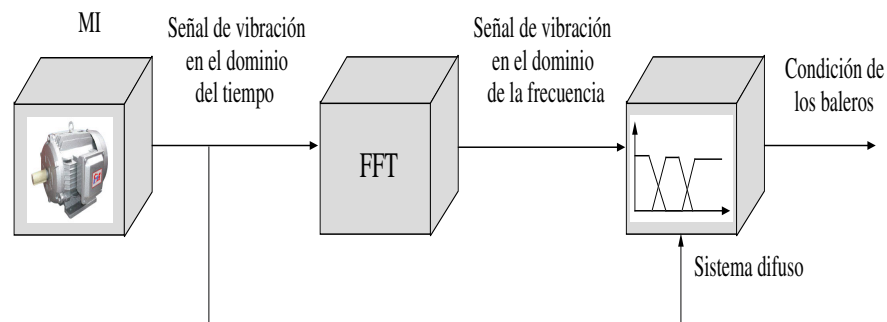


**Figura 6:** Diagrama a bloques del esquema de diagnóstico usando RNA [18].

El muestreador tiene como objetivo eliminar la componente de excitación a 60 Hz a través de un filtro de muesca (notch filter), la señal analógica se amplifica y pasa a través de un filtro pasa-bajas para reducir todos los componentes indeseables de alta frecuencia, la amplificación maximiza el uso del rango de entrada del convertidor A/D (analógico - digital). El preprocesador convierte la señal muestreada al dominio de la frecuencia usando un algoritmo de la FFT, de esta manera es posible eliminar la señal de ruido presente en el espectro promediando un número determinado de muestras del espectro. El filtro de frecuencia basado en reglas elimina una gran cantidad de información del espectro que no está relacionada con una condición de falla. Una red neuronal se entrena para todas las condiciones posibles de operación de la máquina, esta red se utiliza para clasificar los datos obtenidos del filtro. Para prevenir un diagnóstico erróneo, el postprocesador produce una alarma solamente cuando se observan perfiles de falla persistentes. Este esquema se ha utilizado exitosamente para diagnosticar fallas en los baleros, ruptura de

barras del rotor y fallas de excentricidad. Sin embargo, las redes neuronales por sí mismas no pueden proporcionar conocimiento heurístico del motor o la detección de falla del proceso.

Para tomar ventaja del conocimiento heurístico sobre las condiciones de falla en el MI, es posible incorporar este conocimiento al esquema de diagnóstico mediante el uso de lógica difusa. En [19], se analiza el espectro de frecuencia de las señales de vibración de los baleros usando un esquema de diagnóstico basado en lógica difusa como el que se muestra en la Figura 7.

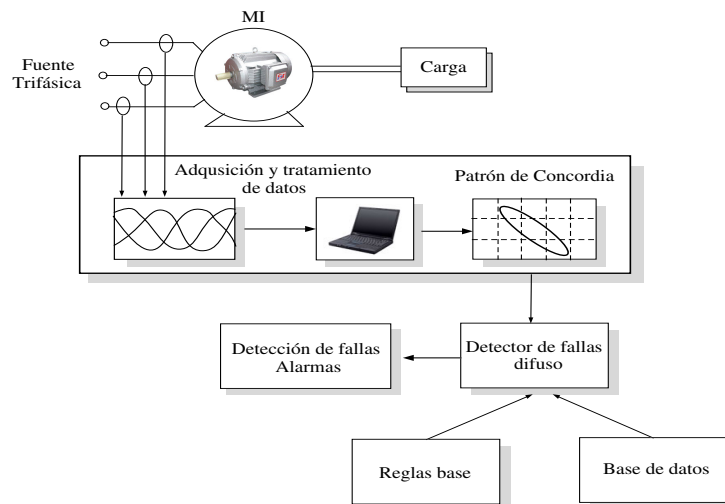


**Figura 7:** Diagrama a bloques del esquema de diagnóstico basado en lógica difusa [19].

El esquema de diagnóstico hace uso de las 5 frecuencias características reportadas en la literatura para identificar posibles defectos en los baleros, sin embargo, se discuten sólo tres tipos de falla: holgura del balero, defecto en la pista interior del balero y defectos en los elementos rodantes del balero. Para determinar la presencia de una falla se monitorean estas frecuencias características en el espectro de frecuencias, y adicionalmente se monitorea la amplitud de la señal de vibración en el dominio del tiempo para determinar la severidad de la falla. Las reglas difusas y las funciones de membresía se construyen a partir del conocimiento heurístico, por ejemplo si la amplitud del componente espectral relacionado con la holgura del balero es alta entonces existe tal falla.

En [20], se propone una estrategia de diagnóstico de fallas en el estator de un MI mediante lógica difusa, utilizando como medio de diagnóstico el patrón de Concordia de las corrientes del estator. El esquema de diagnóstico ilustrado en la Figura 8 se emplea en un MI que opera con diferentes niveles de par de carga, este par se produce mediante un generador de cd de excitación

separada conectado a un resistor variable. El MI se acciona inicialmente sin fallas para determinar el patrón de Concordia de referencia. Se consideran fallas de voltaje desbalanceado provocadas al conectar una resistencia en una de las fases, así como una falla de circuito abierto en una de las fases. Estas fallas provocan una deformación en el patrón de Concordia que es casi circular en una condición libre de fallas.



**Figura 8:** Esquema de diagnóstico difuso basado en el patrón de Concordia [20].

En [21], se presenta un estudio comparativo del diagnóstico de fallas en baleros basado en el vector de corrientes de Park usando redes neuronales y el diagnóstico basado en el vector de Concordia de las corrientes del estator usando un detector de fallas difuso. Los conjuntos difusos se han usado para el diagnóstico de fallas. Sin embargo la mayoría de estos esquemas son estáticos, es decir, se forma o construye un sistema de inferencia difuso general y no se permiten cambios a través de los experimentos. La idea detrás de la fusión de estas dos tecnologías es usar la capacidad de aprendizaje de las redes neuronales para implementar y automatizar los sistemas difusos, los cuales aprovechan la capacidad de razonamiento.

En [22], se presenta un estudio comparativo de los esquemas ANFIS (Adaptive Network Based Fuzzy Inference System) y FALCON (Fuzzy Adaptive Learning Control/Decision Network), evaluando la precisión del diagnóstico de fallas y el costo computacional que implican. La idea

---

principal es monitorear las corrientes del estator, la velocidad y el par de carga para diagnosticar una falla incipiente de fricción del MI. Aunque ambas estructuras proporcionan resultados aceptables bajo una condición de torque de carga variable, se encontró que el detector de fallas ANFIS tiene un tiempo de convergencia menor que el esquema FALCON.

También se han reportado esquemas de diagnóstico mediante redes neuronales que son optimizadas en sus parámetros, como el número de capas ocultas y las características de la función de transferencia de la neurona, mediante algoritmos genéticos [23], [24], en lugar de utilizar únicamente el algoritmo de aprendizaje de propagación hacia atrás. Sin embargo, el costo computacional es muy elevado. En [23], Samanta *et al.* utilizan las señales de vibración del MI como medio de diagnóstico para detectar fallas en los baleros, mientras que en [24], se utilizan como medio de diagnóstico las corrientes del estator para detectar una gran variedad de fallas como desalineamiento, barras rotas del rotor y fallas en baleros, además se encuentra que las diferentes condiciones de carga del MI afectan la construcción de la red y por lo tanto los resultados.

En [25], se presentan consideraciones sobre el diseño de una RNA unidireccional (feedforward) como desempeño, implementación, tamaño y entrenamiento. La red se entrena mediante un algoritmo de propagación hacia atrás para realizar la detección de fallas en motores de inducción. Con el fin de seleccionar la mejor configuración de la red neuronal para una aplicación específica, se agrega al sistema de diagnóstico un sistema difuso basado en razonamiento heurístico y variables lingüísticas para configurar la estructura de la red, basándose en las siguientes preguntas:

- 1.-¿El error de entrenamiento es suficientemente pequeño?
- 2.-¿El tiempo de entrenamiento es demasiado grande?
- 3.-¿Hay muchas o muy pocas capas y nodos ocultos empleándose?

Estas preguntas tienen respuestas “difusas” en el sentido de que tienen un cierto grado de imprecisión. La lógica difusa permite evaluar cuantitativamente las respuestas a estas preguntas de una manera sistemática.

Por otro lado, se ha demostrado que las máquinas de soporte vectorial (SVM, por sus siglas en inglés) son un método efectivo cuando se aplica a problemas de reconocimiento de patrones.

En [26] y [27], se han presentado dos esquemas de diagnóstico basados en máquinas de soporte vectorial. El esquema presentado en [26] es capaz de diagnosticar una ruptura en las barras del rotor del MI analizando el espectro de frecuencias de las corrientes del estator como medio de diagnóstico. Mientras que con el esquema presentado en [27], es posible clasificar los tipos de falla como ruptura en las barras del rotor y corto circuito en los devanados del estator usando como medio de diagnóstico el patrón del vector de corrientes de Park.

## **Aportaciones del trabajo de tesis**

Tomando como base el trabajo hasta ahora realizado en el área de diagnóstico de fallas aplicado al MI, en este proyecto de tesis se propone un algoritmo de diagnóstico, con base en la información proporcionada por medios de evaluación eléctricos, como la corriente de línea de estator, el módulo del vector de Park; y medios de evaluación mecánicos como las señales de vibración, las cuales se integran a través de máquinas de soporte vectorial. Las aportaciones del trabajo se resumen a continuación: i) Análisis y generación de características basadas en la corriente de línea, el módulo del vector de Park, y vibraciones mecánicas del MI para las fallas de desalineamiento, desbalance y rodamientos, bajo un régimen de operación variable. ii) Diseño de máquinas de soporte vectorial basado en la teoría de Vapnik, para crear un algoritmo de diagnóstico eficiente. iii) Implementación del algoritmo de diagnóstico a través de la tarjeta de adquisición de datos NI Compaq-DAQ, utilizando el lenguaje de programación de LabView, el cual es utilizado ampliamente en el sector industrial.

## **Descripción del Documento**

Este trabajo está organizado como se muestra a continuación:

En el Capítulo 1, se describe el principio de funcionamiento de las máquinas de aprendizaje, especialmente se describe el funcionamiento de las RNA y las SVM, tanto para clasificar patrones linealmente separables como SVM no lineales, así como el principio de minimización del riesgo

estructural del que se hace uso para el diseño de las SVM, y dos esquemas de multclasificación.

En el Capítulo 2, se describen los medios de diagnóstico usados en este trabajo para detectar las fallas en el MI, los cuales incluyen la corriente de línea, el módulo del vector de Park, y una señal de vibración mecánica. También se muestra la extracción de las características que son usadas para el diseño de la SVM multiclase mediante LabView.

En el Capítulo 3, se muestra el diseño e implementación de la SVM que clasifica las fallas de desalineamiento, desbalance y baleros en un MI, mediante LabView, usando la tarjeta de adquisición de datos NI-Compaq-DAQ-9172. También se muestran los resultados obtenidos al evaluar el esquema de diagnóstico para diferentes condiciones de velocidad y par de carga del MI.

Finalmente, se dan las conclusiones que se obtuvieron al desarrollar este trabajo de tesis y se sugiere trabajo a futuro. En el apéndice se muestra con más detalle la implementación del esquema de diagnóstico en LabView. Por último, se muestran las referencias asociadas con el trabajo de tesis.

# CAPÍTULO 1

---

## MÁQUINAS DE APRENDIZAJE

---

Con la tecnología actual se tiene la capacidad para almacenar y procesar grandes cantidades de datos, además la mayoría de los dispositivos de adquisición de datos son digitales y registran los datos de manera fiable. Considere, por ejemplo, una empresa del sector industrial que tiene decenas de motores en operación para llevar a cabo sus procesos productivos, y que periódicamente se monitorean y registran las condiciones de cada motor (potencia, vibración mecánica, par de carga etc.) y sus características (caballaje, no. de polos, etc.). Estos datos almacenados serán útiles solamente si son analizados y convertidos en información que se pueda usar, por ejemplo, para hacer predicciones o pronóstico. No se sabe exactamente cuando un motor requiere de mantenimiento, si se supiera no sería necesario analizar los datos, pero como esto no es posible, una opción es coleccionar la información relevante y extraer la respuesta de estos. Aunque no se conozcan los detalles del proceso a estudiar del cual se obtienen los datos (por ejemplo, la operación de un motor en un ambiente hostil), se asume que no es aleatorio, es decir, existen ciertos patrones fijos en los datos. Esta es la base de las máquinas de aprendizaje, estos patrones pueden ser útiles para entender el proceso o hacer predicciones. Asumiendo que los datos posteriores no



difieren mucho de los pasados, las predicciones consecuentes pueden ser correctas.

El enfoque de las maquinas de aprendizaje es aprender automáticamente a reconocer patrones y tomar decisiones inteligentes en base a datos, estas decisiones involucran el asignar a un cierto patrón una etiqueta o salida deseada, comúnmente llamada clase. A los datos que se emplean para el aprendizaje de la máquina junto con su salida deseada, se le conoce como conjunto de entrenamiento. El aprendizaje consiste en encontrar una dependencia entre los datos de entrada y salida de un sistema usando un número limitado de observaciones. En seguida se presentan los 3 componentes del modelo general del proceso de aprendizaje:

i) Un generador de datos  $G$ , que produce vectores aleatorios  $x \in X$ , independientes e idénticamente distribuidos (i.i.d.) de acuerdo a una función fija pero desconocida  $P(x)$  de distribución de probabilidad, es decir, que no varía durante el proceso de aprendizaje.

ii) Un supervisor  $S$  que genera un valor de salida  $y \in Y$  para cada vector de entrada  $x$ , de acuerdo a la función de distribución condicional  $P(y|x)$ , también fija pero desconocida.

iii) Una máquina de aprendizaje capaz de implementar un conjunto de funciones  $f(x, \alpha)$ ,  $\alpha \in \Omega$  donde  $\Omega$  representa un conjunto de parámetros.

Es decir, en el proceso de aprendizaje se adquiere una secuencia

$$D := \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}, \quad (1.1)$$

de pares entrada/salida (conjunto de entrenamiento) que son usados para construir una función  $f : X \rightarrow Y$ , tal que  $f(x, \alpha)$  sea una buena aproximación de las  $y$  posibles respuestas, a una muestra arbitraria  $x$ .

Al considerar que los datos  $x$  son generados de acuerdo a una función de probabilidad fija pero desconocida, se asume que es posible observar los datos con ruido, así mismo al considerar que

la probabilidad condicional  $P(y|x)$  es desconocida, se asume que no se tiene alguna descripción de la relación entre los valores de entrada y salida, es decir, la información que otorga  $x$  puede ser insuficiente para determinar la respuesta correcta  $y$ .

Algunas razones por las cuales una misma entrada  $x$  puede producir diferentes valores de salida son:

- i) El proceso a estudiar es determinístico, pero existe ruido en las observaciones.
- ii) El proceso a estudiar es estocástico.

Con el objeto de seleccionar la mejor aproximación a la respuesta del supervisor, se define una función de costo  $L(y, f(x, \alpha))$  que mide la pérdida o discrepancia entre la respuesta  $y$  del supervisor y la respuesta  $f(x, \alpha)$  proporcionada por la máquina de aprendizaje. El valor esperado de la función de pérdida se define mediante la función de riesgo

$$R(f) = \int L(y, f(x, \alpha)) dP(x, y). \quad (1.2)$$

Ahora el objetivo es encontrar una función  $f(x, \alpha)$  que minimice la función de riesgo  $R(f)$  aún cuando la función de distribución conjunta  $P(x, y)$  es desconocida y la única información disponible es el conjunto de entrenamiento. Para sobrellevar esta situación ( $P(x, y)$  desconocida), se utiliza el *principio inductivo de minimización del riesgo empírico*, el cual establece que es posible reemplazar la función de riesgo  $R(f)$  por la función de riesgo empírico  $R_{emp}(f)$  dada por

$$R_{emp}(f) = \frac{1}{l} \sum_{i=1}^l L(y_i, f(x_i, \alpha)), \quad (1.3)$$

donde  $l$  es el número de datos en el conjunto de entrenamiento (1.1). Se espera que el riesgo empírico  $R_{emp}(f)$  converga al riesgo esperado  $R(f)$  conforme el número de datos de entrenamiento aumenta  $l \rightarrow \infty$ . Sin embargo, en algunas situaciones prácticas el número de datos de entrenamiento  $l$  es limitado. Para abordar esta situación se desarrolló el *principio inductivo de minimización de riesgo estructural* [28], [29], el cual afirma que la siguiente cota sobre el riesgo

$R(f)$  se mantiene con una probabilidad de  $1 - \eta$

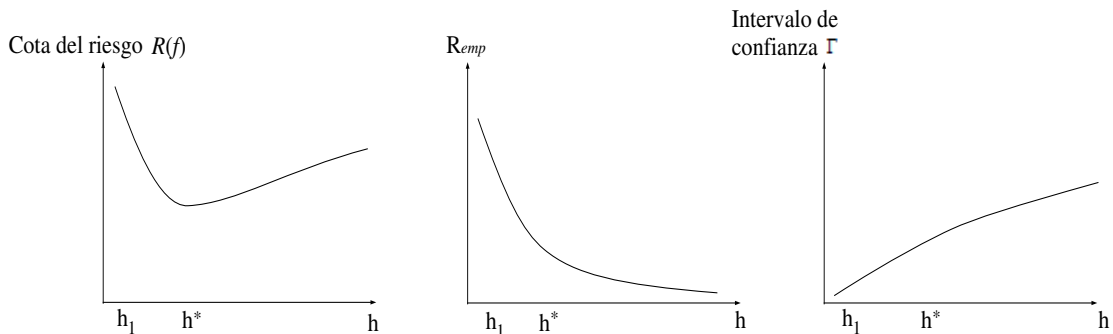
$$R(f) \leq R_{emp}(f) + \Gamma \left( \frac{h}{l}, \frac{\ln \eta}{l} \right), \quad (1.4)$$

donde  $\Gamma \left( \frac{h}{l}, \frac{\ln \eta}{l} \right)$  se le conoce como intervalo de confianza y está dado por

$$\Gamma \left( \frac{h}{l}, \frac{\ln \eta}{l} \right) = \sqrt{\frac{h[(\ln(2l/h) + 1) - \ln(\eta/4)]}{l}}, \quad (1.5)$$

donde  $h$  es la dimensión de VC (Vapnik-Chervonenkis) de la función de aproximación  $f(x, \alpha)$  y se define como el número máximo de datos de entrenamiento que pueden ser separados por  $f(x, \alpha)$  [28]. Una función  $f(x, \alpha)$  puede separar un conjunto de patrones  $x_1, x_2, \dots, x_l$ , si y sólo si, para cada conjunto de entrenamiento de la forma  $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$  existe un valor de  $\alpha$  tal que el error entre la salida real  $y$  y la salida aproximada por  $f(x, \alpha)$  sea cero.

De (1.5) se desprende que cuando  $h$  se incrementa también lo hace el intervalo de confianza  $\Gamma$ , y dado que  $l \geq h$ , el riesgo empírico (1.3) disminuye. En la Figura 1.1 se muestra el comportamiento del riesgo estructural  $R(f)$  dado por (1.4) como la suma del intervalo de confianza (1.5) y el riesgo empírico (1.3). El resultado del proceso de aprendizaje es una función de aproxima-



**Figura 1.1:** Estructura anidada de las funciones de aproximación.

ción  $f(x, \alpha)$ , la cual pertenece a un conjunto de funciones ( $f \in H$ ). Asumir que el conjunto de funciones de aproximación  $H$  se restringe a un conjunto más pequeño, y así sucesivamente hasta formar un conjunto anidado de funciones en el sentido de que se cumple lo siguiente:

$$H_1 \subset H_2 \subset H_3 \subset \dots \subset H_{n-1} \subset H_n \subset \dots \subset H,$$

donde cada conjunto siempre contiene al anterior, el cual contiene funciones de aproximación con una estructura menos compleja, por ejemplo,  $H_n$  puede ser un conjunto de polinomios de grado  $n$ . Asimismo las redes neuronales con  $n$  neuronas son consideradas un subconjunto de las redes neuronales con  $n + 1$  neuronas. La minimización del riesgo estructural se basa en la selección de una estructura apropiada de la máquina de aprendizaje con el fin de minimizar (1.5).

## 1.1. Problemas de aprendizaje

La formulación del proceso de aprendizaje como la minimización de una función de riesgo comprende los siguientes problemas:

### Clasificación de patrones

Asumir que la salida del supervisor toma únicamente 2 valores, es decir,  $y = \{+1, -1\}$  y sea  $f(x, \alpha)$ , un conjunto de *funciones indicador*<sup>1</sup>. Ahora considere la siguiente función de pérdida

$$L(y, f(x, \alpha)) = \begin{cases} +1 & \text{si } y = f(x, \alpha), \\ -1 & \text{si } y \neq f(x, \alpha). \end{cases} \quad (1.6)$$

Para esta función de pérdida, (1.2) determina la probabilidad de los errores en la clasificación. Por lo tanto, el problema es encontrar una función que minimize (1.2) con la función de pérdida dada por (1.6) cuando la función de distribución conjunta  $P(x, y)$  es desconocida, y la única información disponible es el conjunto de entrenamiento (1.1).

### Regresión

Asumir que  $y \in \mathbb{R}$  es la respuesta del supervisor, y sea  $f(x)$  la función de regresión definida como

$$f(x) = \int y dP(y|x).$$

---

<sup>1</sup>Funciones que sólo pueden tomar los valores  $+1$  y  $-1$ , estas funciones indican si un patrón de entrada pertenece a una u otra clase.

Ahora considere la siguiente función de pérdida

$$L(y, f(x, \alpha)) = (y - f(x, \alpha))^2. \quad (1.7)$$

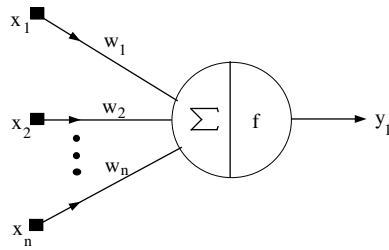
donde  $f(x, \alpha)$  es una aproximación de  $f(x)$ . El problema de regresión consiste en minimizar la función de riesgo (1.2) con la función de pérdida (1.7), cuando la probabilidad  $P(x, y)$  es desconocida pero se tiene acceso al conjunto de datos (1.1).

## 1.2. Máquinas de aprendizaje para clasificación de patrones (RNA)

En la literatura sobre máquinas de aprendizaje [30], [31], [32], se encuentra una variedad de enfoques clásicos para llevar a cabo la clasificación de patrones. Entre estos está el método de discriminantes lineales propuesto por Fisher, que conlleva el uso de combinaciones lineales de los patrones de entrada, el procedimiento se basa en aproximar los límites entre las clases a través de hiperplanos bajo la suposición de que los patrones son linealmente separables. La teoría de decisión de Bayes, se utiliza para construir reglas de clasificación, donde se asume el conocimiento de una densidad de probabilidad para el patrón de entrada, además de una distribución de probabilidad para la salida dada por el supervisor. Se han propuesto métodos como las redes neuronales artificiales y las máquinas de soporte vectorial, que son útiles cuando no existe información suficiente para suponer distribuciones de probabilidad del vector de patrones de entrada para diferentes clases.

Las redes neuronales artificiales RNA emulan el comportamiento de las neuronas del cerebro humano. Una RNA es una colección de unidades computacionales idénticas llamadas neuronas, en la Figura 1.2 se muestra un diagrama de una neurona.

Las entradas y salidas deseadas de la neurona son los correspondientes valores del conjunto de entrenamiento. Al conjunto de valores asociados a las conexiones entre las entradas y la neurona se le conoce como vector de pesos  $w$ . A la función que determina el nivel de activación de la neurona se le llama función de activación o de transferencia. En distintos modelos de RNA se



**Figura 1.2:** Diagrama de una neurona artificial

han propuesto diferentes funciones de activación [31], [32], como por ejemplo

i) Función sigmoide logarítmica

$$f(n) = \frac{1}{1 + e^{-n}}. \quad (1.8)$$

ii) Función lineal

$$f(n) = n. \quad (1.9)$$

iii) Función signo

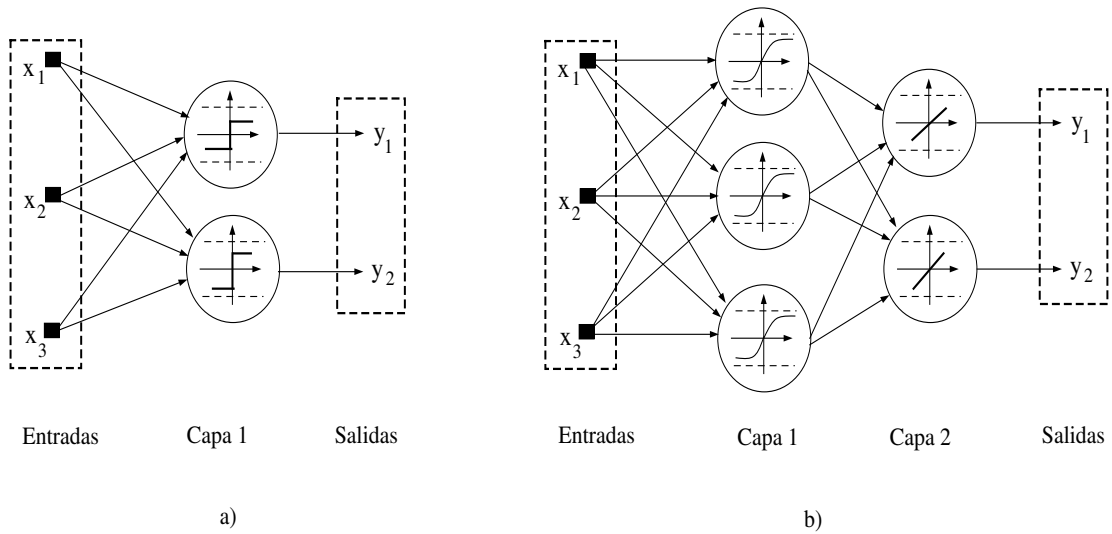
$$f(n) = \begin{cases} 1 & \text{si } n \geq 0, \\ -1 & \text{si } n < 0. \end{cases} \quad (1.10)$$

iv) Función Gaussiana de base radial

$$f(n) = e^{-\frac{|n-c|^2}{\rho}}, \quad (1.11)$$

donde  $c$  y  $\rho$  son el centro y el ancho de la función gaussiana respectivamente. Los tipos de RNA se pueden clasificar según su estructura como RNA monocapa o RNA multicapa, como se ilustra en la Figura 1.3.

Las RNA monocapa utilizan a (1.10) como función de activación, mientras que las RNA multicapa utilizan a (1.8) como función de activación para la capa 1, conocida como capa oculta,



**Figura 1.3:** a) RNA monocapa, b) RNA multicapa.

y a (1.9) como función de activación para la capa 2 (capa de salida). Las RNA que utilizan (1.11) como función de activación se conocen como RNA de base radial. Las RNA monocapa sólo pueden ser aplicadas a problemas de clasificación donde los patrones son linealmente separables, mientras que las RNA multicapa y de base radial pueden abordar el problema de clasificación de patrones no linealmente separables [32].

La salida de una RNA monocapa (Figura 2.3a) con sólo una neurona es por lo tanto

$$y = f(\langle w, x_i \rangle + b) = f(w_1x_1 + w_2x_2 + \dots + w_nx_n + b),$$

donde  $b$  es un escalar conocido como umbral,  $x_i$  es el  $i$ -ésimo patrón de entrada,  $\langle \cdot, \cdot \rangle$  denota el producto interno entre 2 vectores, y  $x_i, w \in \mathbb{R}^n$ . El algoritmo de aprendizaje para las RNA monocapa se basa en una regla de corrección del error que cambia los pesos proporcionalmente al error entre la salida actual y la salida proporcionada por el supervisor. En seguida se describe el procedimiento general utilizado para entrenar una RNA monocapa con sólo una neurona [31], [32]. Para esto, considerar el error para el  $i$ -ésimo patrón de entrada  $e_i = d_i - y_i$  y la función de

costo siguiente

$$E(w) = \frac{1}{l} \sum_{i=1}^l (d_i - y_i)^2 = \frac{1}{l} \sum_{i=1}^l (d_i - f(\langle w, x_i \rangle + b))^2, \quad (1.12)$$

donde  $l$  es el número de datos de entrenamiento,  $d_i$  es la  $i$ -ésima salida real dada por el supervisor, y  $y_i$  es la  $i$ -ésima predicción de la red. Ahora, considerar una función de costo  $E(w)$  que es continuamente diferenciable y que depende de un vector de pesos desconocidos  $w$ , se quiere encontrar una solución óptima  $w^*$  tal que se satisface

$$E(w^*) \leq E(w).$$

La condición de optimalidad necesaria para minimizar  $E(w)$  con respecto al vector  $w$  es

$$\nabla E(w^*) = \left[ \frac{\partial E}{\partial w_1^*}, \frac{\partial E}{\partial w_2^*}, \dots, \frac{\partial E}{\partial w_n^*} \right] = 0,$$

donde  $\nabla E(w)$  es el vector gradiente de la función de costo. Un método de optimización ampliamente usado en el campo de *soft computing* como *regla de aprendizaje*<sup>2</sup> es el método de paso descendente (para problemas de minimización). Los cambios en el vector de pesos se realizan de acuerdo al siguiente algoritmo:

$$w_{p+1} = w_p - \eta_1 \nabla E(w),$$

donde  $p$  denota el número de iteración y  $\eta_1 > 0$  denota la tasa de aprendizaje, si  $\eta_1$  es muy pequeña el entrenamiento de la red es muy lento. La existencia de mínimos locales en la función de costo dificulta el entrenamiento, pues una vez alcanzado un mínimo el entrenamiento se detiene, cuando esto sucede sin satisfacer el porcentaje de error permitido por el diseñador es posible considerar el cambio de la estructura de la red (número de neuronas), comenzar el entrenamiento con un vector de pesos iniciales diferentes, o cambiar la tasa de aprendizaje  $\eta_1$ .

En [31], [32] se describe la regla de aprendizaje de propagación hacia atrás del error basada en el algoritmo de paso descendente del gradiente, comúnmente usado para realizar el entrenamiento

---

<sup>2</sup>Las reglas de aprendizaje son algoritmos usados para modificar el vector de pesos  $w$  iterativamente hasta que la máquina de aprendizaje produzca una salida deseada.



de las RNA multicapa, el cual involucra la inicialización de una matriz de pesos. Note que por la definición de la función de costo (1.12), para el entrenamiento de la RNA se está implementando el principio de minimización de riesgo empírico.

Las máquinas de soporte vectorial proveen una solución al problema de clasificación usando el principio de minimización de riesgo estructural y además con un menor número de parámetros por definir, al contrario de las RNA (tasa de aprendizaje, número de neuronas, número de capas ocultas, y vector o matriz de pesos iniciales). En la siguiente sección se describe brevemente el funcionamiento de las máquinas de soporte vectorial.

### 1.3. Máquinas de soporte vectorial

Considere el problema de separar el conjunto de vectores de entrenamiento pertenecientes a dos clases separables mediante un hiperplano también conocido como superficie de decisión, donde el conjunto de entrenamiento está dado por:

$$D = \{(x_1, y_1), \dots, (x_l, y_l)\},$$

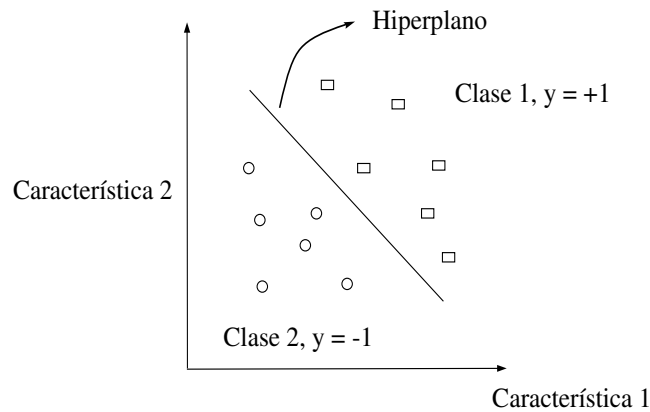
donde  $x_i \in \mathbb{R}^n$ , para  $i = 1, \dots, l$ , es el patrón de entrada a la máquina de aprendizaje,  $y_i \in \{-1, +1\}$  es la correspondiente clase de cada patrón (salida), y  $l$  es el número de patrones en el conjunto de entrenamiento. Esta idea se ilustra en la Figura 1.4. En seguida se presentan algunas definiciones con el objeto de formalizar el problema de clasificación de patrones.

**Definición 1.1:** Un hiperplano  $H$  se define como el conjunto  $H = \{x \in \mathbb{R}^n : w^T x = b\}$ , con  $w \in \mathbb{R}^n$ ,  $w \neq 0$ , y  $b \in \mathbb{R}$ .

**Definición 1.2:** Dado un hiperplano  $H$ , expresado por su ecuación  $w_1 x_1 + w_2 x_2 + \dots + w_n x_n = b$ , este define dos semiespacios cerrados dados por:

$$H_+ = \{x \in \mathbb{R}^n : w^T x \geq b\},$$

$$H_- = \{x \in \mathbb{R}^n : w^T x \leq b\}.$$



**Figura 1.4:** Ilustración de la idea de un hiperplano para la clasificación de patrones en 2 dimensiones

**Definición 1.3:** Un hiperplano es una variedad lineal de dimensión  $(n - 1)$ .

Es decir, en un espacio de una única dimensión (como una recta), un hiperplano es un punto, mientras que en un espacio bidimensional (como el plano  $x_1 - x_2$ ), un hiperplano es una recta. Asumir que los patrones de entrada  $x_i$  representados por  $y_i = 1$ , son linealmente separables de los patrones representados por  $y_i = -1$ . Entonces, la ecuación de una superficie de decisión en la forma de hiperplano que realiza la separación es

$$w^T x + b = 0,$$

donde  $x \in \mathbb{R}^n$ ,  $w$  es un vector de pesos, y  $b \in \mathbb{R}$  es un escalar. Este hiperplano define dos regiones:

$$w^T x_i + b \geq 0, \quad \text{si el índice } i \in [1, l] \text{ satisface } y_i = +1, \quad (1.13)$$

$$w^T x_i + b < 0, \quad \text{si el índice } i \in [1, l] \text{ satisface } y_i = -1. \quad (1.14)$$

Sin embargo, ningún vector de entrada satisface  $w^T x + b = 0$ , si esto ocurre el patrón  $x$  se encontraría sobre el hiperplano, y no pertenecería a ninguna clase. Por lo tanto, en vez de (1.13)

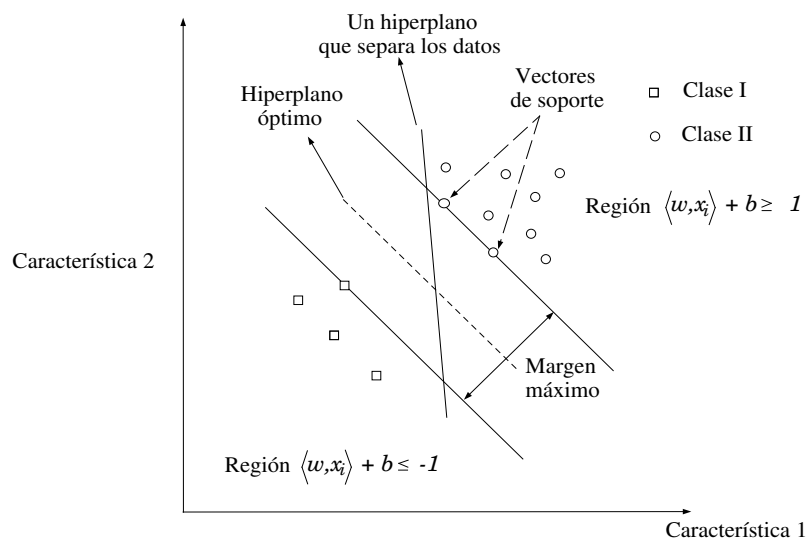
y (1.14) se consideran las siguientes desigualdades

$$\begin{aligned} w^T x_i + b &\geq 1, & \text{para } y_i = +1, \\ w^T x_i + b &\leq -1, & \text{para } y_i = -1. \end{aligned}$$

lo cual puede ser reescrito como

$$y_i(\langle w, x_i \rangle + b) \geq 1,$$

donde  $\langle \cdot, \cdot \rangle$  denota el producto interno entre dos vectores, esto se ilustra en la Figura 1.5. La distancia entre la superficie de decisión y los datos de entrenamiento más cercanos a esta se le conoce como margen. El hiperplano con el que se obtenga un margen mayor se le conoce como hiperplano óptimo. Los patrones de entrenamiento  $(x_i, y_i)$  que satisfacen  $w^T x_i + b = 1$  para  $y_i = +1$ , y  $w^T x_i + b = -1$  para  $y_i = -1$ , son llamados vectores de soporte, de ahí el nombre de máquina de soporte vectorial. Estos vectores juegan un papel importante en la operación de esta clase de máquinas de aprendizaje, ya que son los más cercanos a la superficie de decisión y por lo tanto tienen una influencia directa en la localización óptima del hiperplano.



**Figura 1.5:** Definición de regiones de la superficie de decisión.

A continuación, se presenta una breve introducción al problema general de optimización, para posteriormente encontrar el hiperplano óptimo de un problema de clasificación.

### 1.3.1. Características del problema general de optimización

En un problema general de optimización se busca encontrar el máximo o mínimo de una función, comúnmente sujeta a algunas restricciones. Es posible establecer el problema general de optimización como sigue

$$\begin{aligned} \min_w \quad & f(w), \quad w \in \Omega \subseteq \mathbb{R}^n, \\ \text{t. q.} \quad & g_i(w) \geq 0 \quad \text{para } i = 1, 2, \dots, k, \\ & h_i(w) = 0 \quad \text{para } i = 1, 2, \dots, m, \end{aligned}$$

donde  $f(w)$  es llamada función objetivo,  $g_i(w)$  y  $h_i(w)$  son llamadas, respectivamente, como restricciones de desigualdad e igualdad, y  $\Omega$  es la región donde está definida la función objetivo. A la región donde se satisfacen todas las restricciones se le conoce como región factible y se denota por

$$R = \{w \in \Omega : g_i(w) \geq 0, h_i(w) = 0, \forall i\}.$$

Al valor óptimo de la función objetivo se le llama valor del problema de optimización. Una solución al problema general de optimización es un punto  $w^* \in \mathbb{R}$  tal que no exista otro punto  $w \in \mathbb{R}$  para el cual  $f(w) < f(w^*)$ , a este punto se le conoce como mínimo global. Si las funciones  $f(w)$ ,  $g_i(w)$  y  $h_i(w)$  son lineales, el problema es conocido como optimización lineal, por otra parte, si  $f(x)$  es una función cuadrática, mientras que,  $g_i(w)$  y  $h_i(w)$  son lineales, el problema es conocido como optimización cuadrática, la optimización cuadrática es un tipo de optimización no lineal con restricciones.

Las condiciones que caracterizan un mínimo de una función continuamente diferenciable  $f(w)$ , con  $w \in \mathbb{R}^n$  se expresan en términos del vector de primeras derivadas parciales  $\nabla f$ , y la matriz  $H_{ij}$  ( $n \times n$ ) de segundas derivadas parciales, cuyos  $(i, j)$ -ésimos elementos están dados por

$$H_{ij} = \frac{\partial^2 f}{\partial w_i \partial w_j}. \quad (1.15)$$

**Definición 1.4:** A la matriz  $H_{ij}$  en (1.15) se le llama matriz Hessiana y se denota por  $\nabla^2 f$ .

La siguiente proposición da las condiciones de optimalidad, conocidas como condiciones de Karush-Kuhn-Tucker (KKT), para un problema con restricciones

$$\begin{aligned} & \underset{w}{\text{mín}} \quad f(w), \\ & \text{t. q.} \quad g_i(w) \geq 0, \quad \text{para } i = 1, \dots, k. \end{aligned}$$

**Proposición 1.1 [33]:** La solución óptima  $w^*$ , existe si y sólo si, se satisfacen las siguientes condiciones:  $i$ ) existen escalares  $\alpha^* = [\alpha_1^*, \dots, \alpha_k^*] \geq 0$  tales que para  $w^*$  se cumple

$$\alpha_i^* g_i(w^*) = 0, \quad \text{para } i = 1, 2, \dots, k, \quad (1.16)$$

y

$$\nabla \mathcal{L}(w^*, \alpha^*) = \nabla f(w^*) - \sum_{i=1}^k \alpha_i^* \nabla g_i(w^*) = 0,$$

donde  $\mathcal{L}(w, \alpha)$  se le conoce como función de Lagrange y está dada por

$$\mathcal{L}(w, \alpha) = f(w) - \sum_{i=1}^k \alpha_i g_i(w),$$

y a los escalares  $\alpha_1^*, \dots, \alpha_k^*$  se les denomina como multiplicadores de Lagrange.

Para el problema general de optimización, la función de Lagrange y las condiciones de KKT se pueden usar para definir un problema de optimización dual, donde las variables en el problema dual son los multiplicadores de Lagrange del problema original. En la literatura de optimización se recurre ampliamente al concepto de dualidad, el propósito es proveer una formulación alternativa a un problema de optimización, el cual será más conveniente computacionalmente. El problema original es llamado *primario*, y el transformado se le conoce como *dual*. El siguiente teorema introduce el problema dual utilizando la función de Lagrange.

**Teorema 1.1 [34]:** Sea  $w^*$  el mínimo de un problema primario de optimización:

$$\begin{aligned} & \underset{w}{\text{mín}} \quad f(w), \\ & \text{t. q.} \quad g_i(w) \geq 0, \quad \text{para } i = 1, \dots, k, \end{aligned}$$

entonces  $w^*$  y  $\alpha^*$  resuelven el problema dual

$$\begin{aligned} & \underset{w, \alpha}{\text{máx}} \quad \mathcal{L}(w, \alpha), \\ & \text{t. q.} \quad \nabla \mathcal{L}(w^*, \alpha^*) = 0, \quad \alpha^* \geq 0. \end{aligned}$$

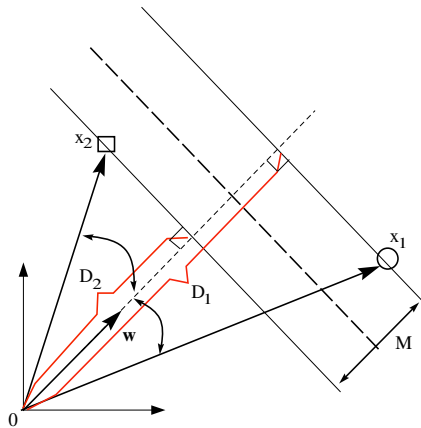
Además, el valor del problema de optimización primario y dual son iguales, es decir,

$$f(w^*) = \mathcal{L}(w^*, \alpha^*).$$

En la siguiente sección se detalla brevemente el problema de clasificación de patrones linealmente separables mediante una SVM.

### 1.3.2. SVM: Para patrones linealmente separables

Para encontrar el hiperplano óptimo que separe un determinado patrón de entrada, es necesario encontrar una expresión para el margen entre dos vectores de soporte y entonces plantear el problema de optimización. Para esto consideramos el problema bidimensional de clasificación, es decir,  $x \in \mathbb{R}^2$ , como el que se muestra en la Figura 1.6, en la que los símbolos  $\circ$  y  $\square$  representan dos tipos de clase.



**Figura 1.6:** Cálculo del margen entre dos vectores de soporte.

Es posible calcular el margen  $M$  que separa el hiperplano de los vectores de soporte  $x_1$  y  $x_2$  como  $M = D_1 - D_2$ , con

$$\begin{aligned} D_1 &= \|x_1\| \cos(\alpha), \\ D_2 &= \|x_2\| \cos(\beta), \end{aligned} \quad (1.17)$$

donde  $\alpha$  y  $\beta$  son los ángulos entre  $w$  y  $x_1$ , y  $w$  y  $x_2$  respectivamente y se obtienen a partir de

$$\begin{aligned} \cos(\alpha) &= \frac{\langle w, x_1 \rangle}{\|x_1\| \|w\|}, \\ \cos(\beta) &= \frac{\langle w, x_2 \rangle}{\|x_2\| \|w\|}, \quad \text{para } 0 \leq \alpha, \beta \leq \pi. \end{aligned} \quad (1.18)$$

Entonces sustituyendo (1.17) en (1.18), el margen resulta en

$$\begin{aligned} M &= D_1 - D_2 \\ &= \|x_1\| \cos(\alpha) - \|x_2\| \cos(\beta) \\ &= \|x_1\| \frac{w^T x_1}{\|x_1\| \|w\|} - \|x_2\| \frac{w^T x_2}{\|x_2\| \|w\|} \\ &= \frac{w^T x_1 - w^T x_2}{\|w\|} \\ &= \frac{w^T x_1 + b - (w^T x_2 + b)}{\|w\|} \\ &= \frac{2}{\|w\|}, \end{aligned}$$

donde se consideró el hecho de que  $x_1$  y  $x_2$  son vectores de soporte que satisfacen  $y_j(w^T x_j + b) = 1$ , con  $j = 1, 2$ , es decir,  $w^T x_1 + b = 1$  y  $w^T x_2 + b = -1$ . Para conseguir la maximización del margen  $M$ , se debe minimizar  $\|w\| = \sqrt{\langle w, w \rangle} = \sqrt{w_1^2 + w_2^2 + \dots + w_n^2}$ . Debido a que  $\sqrt{\langle w, w \rangle}$  es una función monótonica o no decreciente, la minimización de  $\|w\|$  es equivalente a la minimización de  $\|w\|^2$ . Entonces dado el conjunto de entrenamiento  $D = \{(x_i, y_i)\}_{i=1}^l$ , se tiene el siguiente problema de optimización con restricciones

$$\begin{aligned} \min_w \quad & f(w) = \min_w \frac{1}{2} w^T w, \\ \text{t. q.} \quad & y_i(\langle w, x_i \rangle + b) - 1 \geq 0, \quad \text{para } i = 1, \dots, l. \end{aligned}$$

Es posible resolver este problema usando el teorema (1.1), para esto se contruye la función de Lagrange como

$$\mathcal{L}(w, \alpha, b) = \frac{1}{2}w^T w - \sum_{i=1}^l \alpha_i [y_i(\langle w, x_i \rangle + b) - 1],$$

donde los escalares  $\alpha_i \geq 0$  son los multiplicadores de Lagrange. Posteriormente calculamos las condiciones de optimalidad dadas en la proposición (1.1), estas son

$$\frac{\partial \mathcal{L}(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^l \alpha_i y_i x_i = 0, \quad (1.19)$$

y

$$\frac{\partial \mathcal{L}(w, b, \alpha)}{\partial b} = \sum_{i=1}^l \alpha_i y_i = 0. \quad (1.20)$$

Al expandir la función de Lagrange  $\mathcal{L}(w, b, \alpha)$  y aplicar las condiciones de optimalidad (1.19) y (1.20) se tiene

$$\begin{aligned} \mathcal{L}(w, b, \alpha) &= \frac{1}{2}w^T w - \sum_{i=1}^l \alpha_i y_i \langle w, x_i \rangle - b \sum_{i=1}^l \alpha_i y_i + \sum_{i=1}^l \alpha_i \\ &= -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle + \sum_{i=1}^l \alpha_i, \end{aligned}$$

donde se consideró que:

$$w^T = \sum_{i=1}^l \alpha_i y_i x_i^T, \quad \text{y} \quad w^T w = \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle.$$

Entonces se establece el problema dual como: dado el conjunto de entrenamiento  $D = \{(x_i, y_i)\}_{i=1}^l$ , encuentre los multiplicadores de Lagrange  $\alpha = [\alpha_1, \dots, \alpha_l]$  tales que

$$\begin{aligned} \text{máx}_{\alpha} \mathcal{L}(\alpha) &= \text{máx}_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle, \\ \text{s. t.} \quad \sum_{i=1}^l \alpha_i y_i &= 0, \quad \alpha_i \geq 0 \text{ para } i = 1, \dots, l. \end{aligned}$$



Note que el problema dual depende sólo de los datos de entrenamiento, en especial del patrón de entrada en forma de un producto interno  $\langle x_i, x_j \rangle$ . Es posible entonces escribir este problema de optimización cuadrática en notación matricial como

$$\begin{aligned} \max_{\alpha} \mathcal{L}(\alpha) &= \max_{\alpha} -\frac{1}{2}\alpha^T H\alpha + f^T \alpha, \\ \text{t. q. } y^T \alpha &= 0, \quad \alpha \geq 0, \end{aligned}$$

donde  $\alpha = [\alpha_1, \dots, \alpha_l]^T$ ,  $y = [y_1, \dots, y_l]^T$ ,  $f$  es un vector con elementos 1 y dimensiones  $l \times 1$ , y  $H$  denota la matriz Hessiana que para este problema está dada por:

$$H_{ij} = y_i y_j \langle x_i, x_j \rangle.$$

De la condición de optimalidad (1.16), es decir,  $\alpha_i \{y_i (\langle w, x_i \rangle + b) - 1\} = 0$ , se observa que se debe satisfacer  $y_i (\langle w, x_i \rangle + b) = 1$  para  $\alpha_i \neq 0$ .

**Definición 1.5:** Los patrones de entrada  $x_i$  para los cuales  $\alpha_i \neq 0$  son llamados vectores de soporte.

Una vez que se determinaron los multiplicadores de Lagrange  $\alpha_i^*$ , es posible calcular el vector de pesos óptimo  $w^* = [w_1^*, \dots, w_l^*]$  usando la condición de optimalidad (1.19), es decir

$$w^* = \sum_{i=1}^l \alpha_i^* y_i x_i.$$

Además, para que se mantenga la condición de optimalidad  $\alpha_i \{y_i (\langle w, x_i \rangle + b) - 1\} = 0$ ,  $b^*$  se debe escoger como

$$b^* = y_i - w^T x_i.$$

Otra forma de calcular  $b^*$  es tomando el promedio entre los vectores de soporte como sigue

$$b^* = \frac{1}{l_{vs}} \sum_{i=1}^{l_{vs}} (y_i - w^T x_i),$$

donde  $l_{vs}$  es el número de vectores de soporte. Note que por la definición (1.5) tanto el vector de pesos óptimo  $w^*$  como  $b^*$  son calculados solamente usando los vectores de soporte. Ahora es posible definir la función de decisión como

$$\begin{aligned} d(x, w^*, b^*) &= (\langle w^*, x \rangle + b^*), \\ &= \sum_{i=1}^l \alpha_i^* y_i \langle x_i, x \rangle + b^*. \end{aligned}$$

Finalmente se establece la regla de decisión  $i_f$ , la cual indica si un patrón de entrada pertenece o no a una determinada clase

$$x_0 \in \begin{cases} \text{Clase I,} & \text{si } d(x_0, w^*, b^*) > 0, \\ \text{Clase II,} & \text{si } d(x_0, w^*, b^*) < 0. \end{cases}$$

Es decir

$$i_f = \text{sgn}(d(x_0, w^*, b^*)),$$

donde  $x_0$  es un nuevo patrón de entrada.

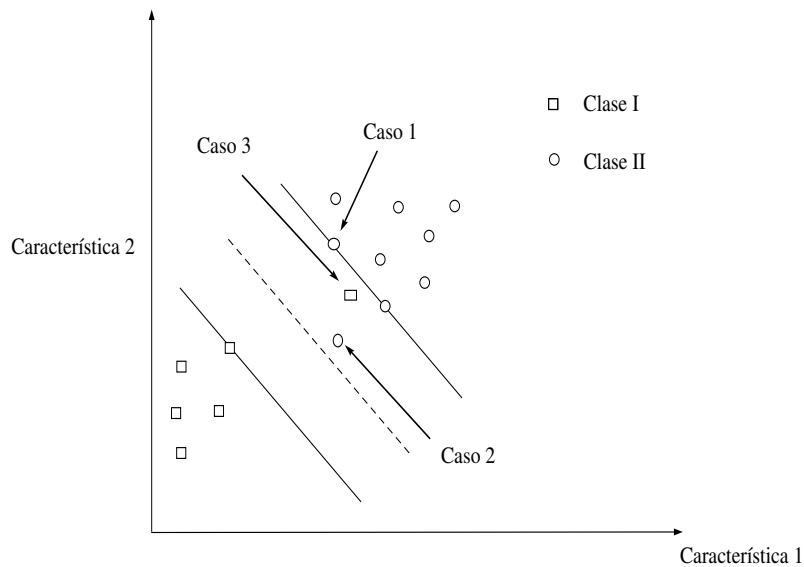
En [28], [29], Vapnik muestra que la dimensión VC,  $h$  de los hiperplanos con margen  $1/\|w\|$ , está acotada por

$$h \leq \text{mín}(R^2\|w\|^2, l) + 1,$$

donde  $R$  es el diámetro de la esfera más pequeña que incluye a todos los datos de entrenamiento. Por lo tanto, la minimización de  $\|w\|$  conduce a un valor pequeño de  $h$ , es decir, la minimización de  $\|w\|$  es una implementación del principio inductivo de minimización del riesgo estructural.

### 1.3.3. SVM: Para patrones no separables linealmente

En la Figura 1.7 se muestra un patrón de entrada que no es separable linealmente, aquí la idea es encontrar un hiperplano que incurra en el menor número de errores de clasificación posibles. Para esto se introducen las variables  $\xi_i \geq 0$ , las cuales compensan la desviación del patrón de entrada del margen. Existen tres posibles casos para la definición de  $\xi_i$  (véase Fig. 1.7), estas son:



**Figura 1.7:** Patrones no separables linealmente.

- i) El patrón se encuentra fuera del margen y por lo tanto es clasificado correctamente ( $\xi_i = 0$ ).
- ii) El patrón está dentro del margen pero está clasificado correctamente ( $0 < \xi_i < 1$ ).
- iii) El patrón se encuentra dentro del margen pero está mal clasificado ( $\xi_i > 1$ ).

Entonces las regiones que define el hiperplano están dadas por

$$y_i(w^T x_i + b) \geq 1 - \xi_i,$$

ó

$$y_i(w^T x_i + b) - 1 + \xi_i \geq 0, \quad \xi_i \geq 0. \quad (1.21)$$

Resumiendo, si  $\xi_i = 0$ , no hay error de clasificación. Si  $0 < \xi_i < 1$ , el patrón  $x_i$  está correctamente clasificado pero se encuentra en el margen, y si  $\xi_i > 1$ ,  $x_i$  está mal clasificado. Entonces, el número de falsas alarmas (clasificaciones erróneas) es el número de patrones de entrada para

los cuales  $\xi_i > 1$ , y el número de patrones no linealmente separables es aquel para el cual  $\xi_i > 0$ , por lo tanto es posible definir un error como

$$e = \sum_{i=1}^l \xi_i, \quad (1.22)$$

este se puede adherir al problema primario de optimización con el propósito de penalizar el error.

Ahora la función Lagrangiana del problema de optimización es

$$\mathcal{L}(w, b, \alpha, \beta, \xi) = \frac{1}{2}w^T w + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i (y_i [w^T x_i + b] - 1 + \xi_i) - \sum_{i=1}^l \beta_i \xi_i,$$

donde  $\alpha = [\alpha_1, \dots, \alpha_l]^T$ , y  $\beta = [\beta_1, \dots, \beta_l]^T$  son multiplicadores de Lagrange estrictamente positivos,  $C$  es una constante que penaliza el error (1.22), y además se consideraron las restricciones en (1.21). Ahora las condiciones de optimalidad dadas en la proposición (2.1) son

$$\begin{aligned} \frac{\partial \mathcal{L}(w, b, \alpha, \beta, \xi)}{\partial w} &= w - \sum_{i=1}^l \alpha_i y_i x_i = 0, \\ \frac{\partial \mathcal{L}(w, b, \alpha, \beta, \xi)}{\partial b} &= \sum_{i=1}^l \alpha_i y_i = 0, \\ \frac{\partial \mathcal{L}(w, b, \alpha, \beta, \xi)}{\partial \xi} &= \alpha_i + \beta_i = C, \end{aligned} \quad (1.23)$$

y

$$\begin{aligned} \alpha_i (y_i [w^T x_i + b] - 1 + \xi_i) &= 0, \\ \beta_i \xi_i &= 0, \\ \alpha_i \geq 0, \beta_i \geq 0, \xi_i \geq 0, & \quad \text{para } i = 1, \dots, l. \end{aligned}$$

Debido a que  $\beta_i = C - \alpha_i \geq 0$ , se debe cumplir  $0 \leq \alpha_i \leq C$ . Al sustituir las condiciones de optimalidad (1.23) en la función de Lagrange  $\mathcal{L}(w, b, \alpha, \beta, \xi)$  se genera el problema dual siguiente

$$\begin{aligned} \max_{\alpha} \mathcal{L}(\alpha) &= \max_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle, \\ \text{t. q. } \sum_{i=1}^l \alpha_i y_i &= 0, \quad 0 \leq \alpha_i \leq C \text{ para } i = 1, \dots, l. \end{aligned}$$

Observe que la función objetivo es la misma que para el caso anterior.

### 1.3.4. SVM no lineales

Las SVM no lineales son una alternativa para el caso en el que los patrones no son linealmente separables, y se necesite el mínimo número de errores posibles o incluso un error nulo. Las SVM no lineales crean una superficie no lineal que es capaz de clasificar los patrones. La idea básica de la SVM no lineal es mapear los datos de entrada  $x_i \in \mathbb{R}^n$  en vectores  $\phi(x_i) \in \mathbb{R}^f$ , con  $f > n$ . A  $\mathbb{R}^f$  se le conoce como espacio característico y el mapeo es denotado por

$$\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^f.$$

La superficie de decisión está dada ahora por

$$d(x, w, b) = \langle w, \phi(x) \rangle + b, \quad (1.24)$$

la cual es lineal en el espacio característico  $\mathbb{R}^f$ , debido a esto, es posible extender el procedimiento que se usó para el caso de patrones linealmente separables, en este caso se tiene el siguiente problema de optimización dual

$$\begin{aligned} \max_{\alpha} \mathcal{L}(\alpha) &= \max_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \langle \phi(x_i), \phi(x_j) \rangle, \\ \text{t. q.} \quad \sum_{i=1}^l \alpha_i y_i &= 0, \quad 0 \leq \alpha_i \leq C, \quad \text{para } i = 1, \dots, l. \end{aligned} \quad (1.25)$$

Al producto interno representado por  $\langle \phi(x_i), \phi(x_j) \rangle$  se le conoce como *kernel* y se denota por  $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ . El problema anterior se puede escribir en notación matricial como

$$\begin{aligned} \max_{\alpha} \mathcal{L}(\alpha) &= \max_{\alpha} -\frac{1}{2} \alpha^T H \alpha + f^T \alpha, \\ \text{t. q.} \quad y^T \alpha &= 0, \quad \alpha_i \in [0, C], \quad \text{para } i = 1, \dots, l. \end{aligned} \quad (1.26)$$

donde  $\alpha = [\alpha_1, \dots, \alpha_l]^T$ ,  $y = [y_1, \dots, y_l]^T$ ,  $H$  denota la matriz Hessiana  $H_{ij} = y_i y_j K(x_i, x_j)$ , y  $f$  es un vector de sólo componentes 1's y dimensiones  $l \times 1$ . Al sustituir la condición de optimalidad

$$w^* = \sum_{i=1}^l \alpha_i^* y_i \phi(x_i), \quad (1.27)$$

en (1.24) nos conduce al hiperplano óptimo dado por

$$d(x, \alpha^*, b^*) = \sum_{i=1}^l \alpha_i^* y_i K(x_i, x) + b^* = 0.$$

La regla de decisión de la SVM no lineal para la clasificación de los datos de entrada es

$$\begin{aligned} i_f(x, \alpha^*, b^*) &= \text{sgn}[d(x, \alpha^*, b^*)], \\ &= \text{sgn} \left[ \sum_{i=1}^l \alpha_i^* y_i K(x_i, x) + b^* \right]. \end{aligned}$$

Es decir,

$$x_0 \in \begin{cases} \text{Clase I,} & \text{si } i_f(x, \alpha^*, b^*) = +1, \\ \text{Clase II,} & \text{si } i_f(x, \alpha^*, b^*) = -1. \end{cases}$$

Note que para obtener el hiperplano óptimo no es necesario calcular explícitamente a  $w^*$  dada en (1.27). En seguida se presentan algunas funciones kernel citadas en la literatura de las máquinas de soporte vectorial [35], [36]:

*Polinomial*

$$K(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^2.$$

*Función Gaussiana de base radial*

$$K(x_i, x_j) = \exp \left( -\frac{\|x_i - x_j\|^2}{2\sigma^2} \right).$$

*Spline*

$$K(x_i, x_j) = \Pi(1 + \langle x_i, x_j \rangle) + \frac{1}{2} \langle x_i, x_j \rangle \text{mín}(x_i, x_j) - \frac{1}{6} \text{mín}(x_i, x_j)^3).$$

donde  $\text{mín}(x_i, x_j)$  denota el vector formado por los elementos mínimos del arreglo  $(x_i, x_j)$ ,  $\Pi$  denota el producto de cada elemento de un vector, y  $\text{mín}(\cdot, \cdot)^3$  denota el cubo de cada componente del vector  $\text{mín}(\cdot, \cdot)$ .

### Ejemplo ilustrativo

Considere el siguiente ejemplo en el que se desea construir una SVM que clasifique un patrón de entradas no separable linealmente y bidimensional, es decir,  $x \in \mathbb{R}^2$ , dado en la Tabla 1.1 junto con la clase deseada:

**Tabla 1.1:** Datos de entrenamiento para un ejemplo de SVM

Característica 1	Característica 2	Respuesta deseada, $y$
-1	-1	-1
-1	1	1
1	-1	1
1	1	-1

Al evaluar la función de Lagrange con  $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$ , y  $\sigma = 2$ , dada en (1.25) resulta

$$\begin{aligned}
 \mathcal{L}(\alpha) &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\
 &= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2} \alpha_1^2 e^{\left(-\frac{\|(-1 \ -1) - (-1 \ -1)\|^2}{2(2)^2}\right)} + \frac{1}{2} \alpha_1 \alpha_2 e^{\left(-\frac{\|(-1 \ -1) - (-1 \ 1)\|^2}{2(2)^2}\right)} \\
 &\quad + \frac{1}{2} \alpha_1 \alpha_3 e^{\left(-\frac{\|(-1 \ -1) - (1 \ -1)\|^2}{2(2)^2}\right)} - \frac{1}{2} \alpha_1 \alpha_4 e^{\left(-\frac{\|(-1 \ -1) - (1 \ 1)\|^2}{2(2)^2}\right)} + \dots \\
 &= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2} (\alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2) + \alpha_1 \alpha_2 e^{-0.5} + \alpha_1 \alpha_3 e^{-0.5} - \alpha_1 \alpha_4 e^{-1} \\
 &\quad - \alpha_2 \alpha_3 e^{-1} + \alpha_2 \alpha_4 e^{-0.5} + \alpha_3 \alpha_4 e^{-0.5},
 \end{aligned}$$

ó bien en notación matricial (1.26) se tiene

$$\mathcal{L}(\alpha) = -\frac{1}{2} \alpha^T H \alpha + f^T \alpha,$$

donde la matriz Hessiana  $H_{ij} = y_i y_j K(x_i, x_j)$  es

$$H_{ij} = \begin{pmatrix} 1 & -e^{-0.5} & -e^{-0.5} & e^{-1} \\ -e^{-0.5} & 1 & e^{-1} & -e^{-0.5} \\ -e^{-0.5} & e^{-1} & 1 & -e^{-0.5} \\ e^{-1} & -e^{-0.5} & -e^{-0.5} & 1 \end{pmatrix}.$$

Ahora, usando el hecho siguiente

$$\max_{\alpha} \mathcal{L}(\alpha) = \min_{\alpha} -\mathcal{L}(\alpha),$$

es posible resolver el problema de optimización (1.26) con ayuda del algoritmo computacional de Matlab *quadprog*, o empleando el bloque de programación cuadrática de *LabView*. En la Figura 1.8 se muestra la implementación en Labview de una solución al problema de optimización para el diseño de una SVM, de la cual se obtuvo

$$\alpha_1^* = \alpha_2^* = \alpha_3^* = \alpha_4^* = 6.459.$$

En la Figura 1.9 se presenta la función de decisión  $d(x, \alpha^*)$  en  $\mathbb{R}^3$ , dada por

$$d(x, \alpha^*) = \sum_{i=1}^l \alpha_i^* y_i K(x_i, x).$$

En la Figura 1.10 se muestra el hiperplano óptimo en  $\mathbb{R}^2$  dado por

$$\sum_{i=1}^l \alpha_i^* y_i K(x_i, x) = 0.$$

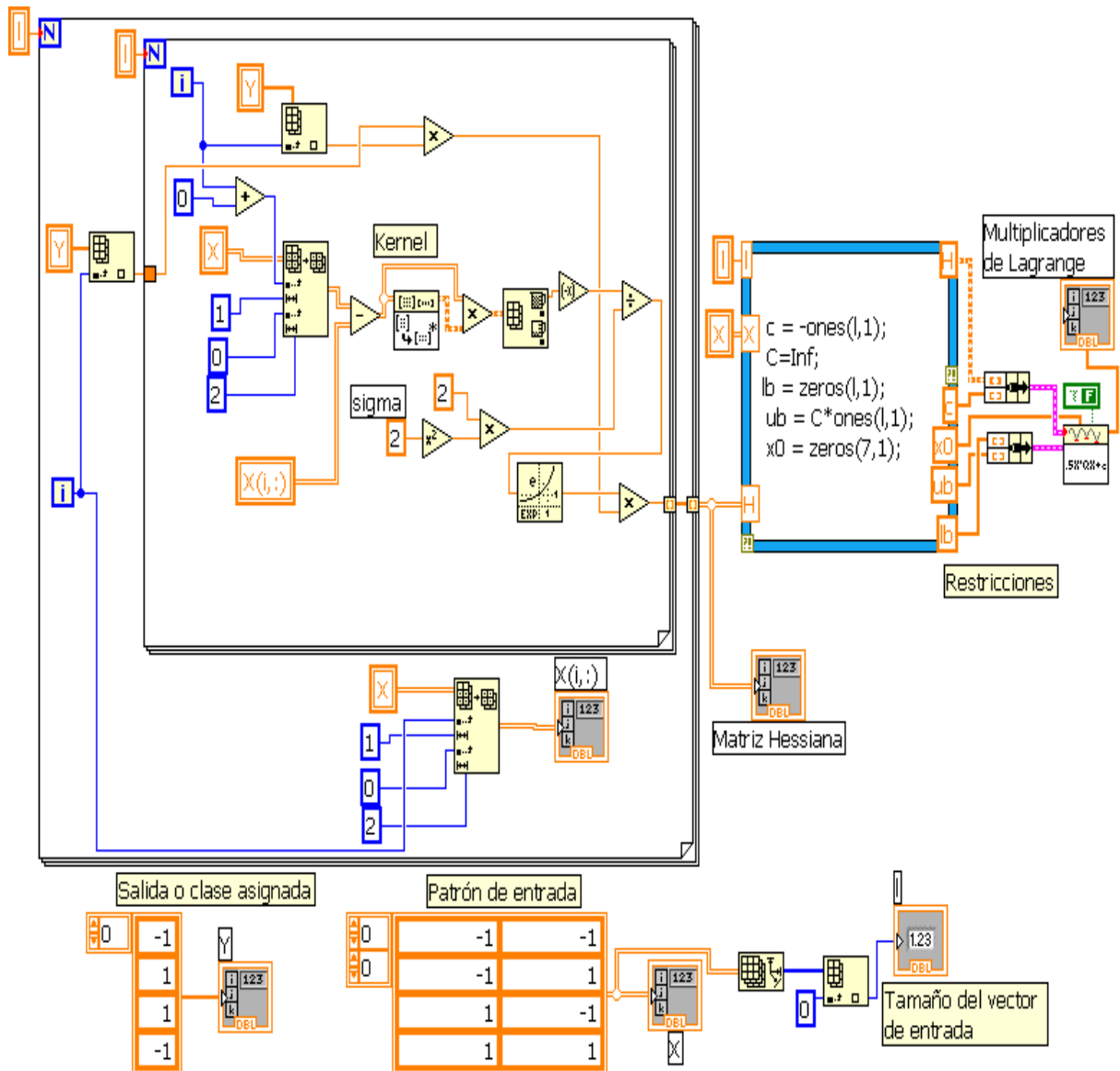
Observe que el hiperplano se forma de la intersección de la función de decisión con el plano de características. ▲

Es posible extender el problema de clasificación binaria a un problema de multclasificación. En las siguientes secciones se describen brevemente 2 esquemas de multclasificación basados en máquinas de soporte vectorial:

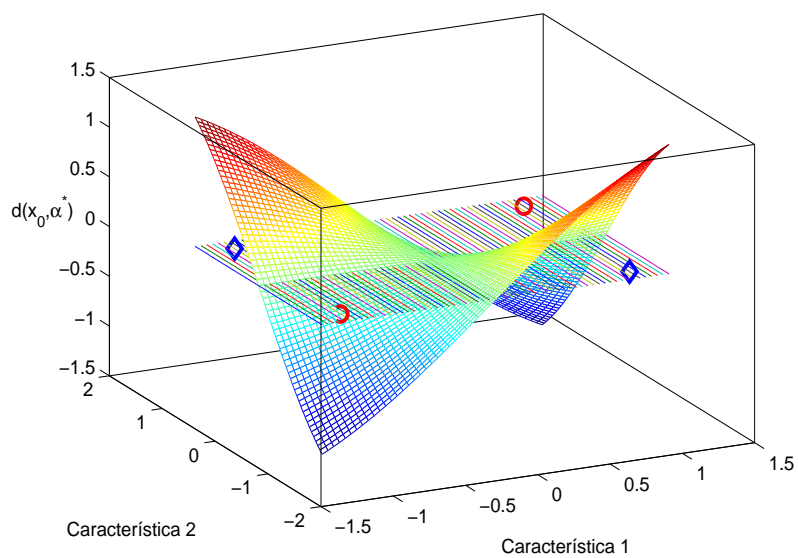
1.- SVM uno vs. resto. (One-against-all).

2.- SVM uno vs. uno. (One-against-one).

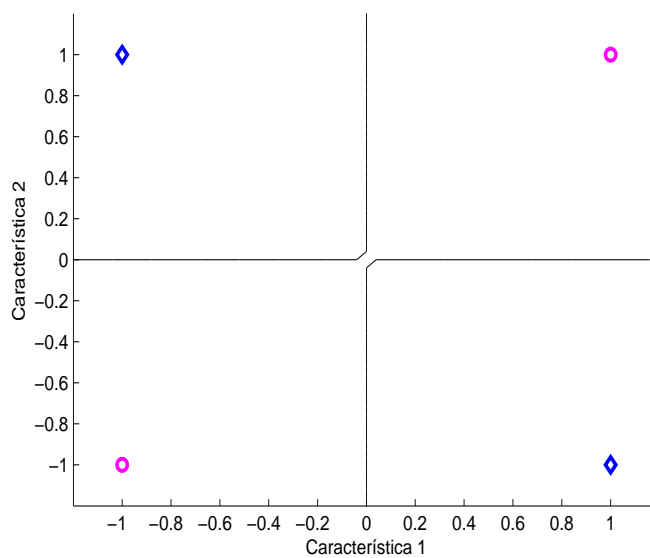




**Figura 1.8:** Obtención de multiplicadores de Lagrange  $\alpha$ , usando el Kernel  $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$ .



**Figura 1.9:** Función de decisión  $d(x_0, \alpha^*)$  para un ejemplo de SVM.



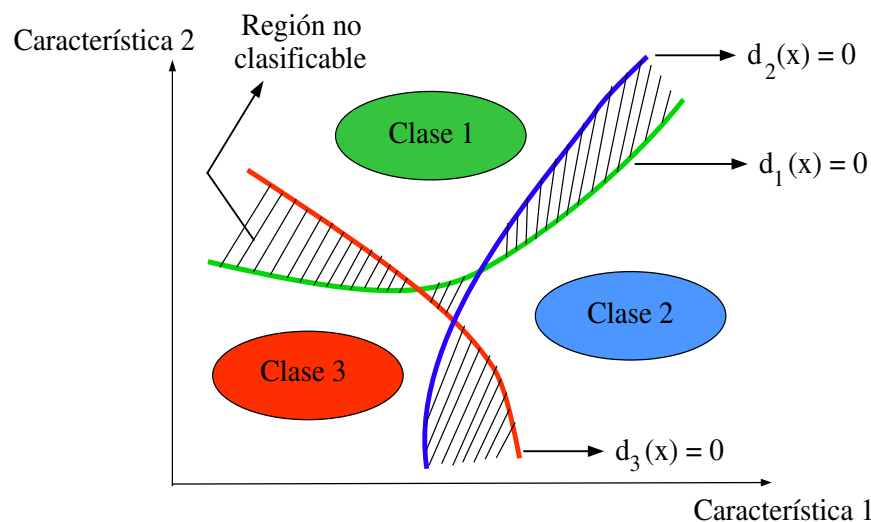
**Figura 1.10:** Hiperplano óptimo en el espacio de entrada  $\mathbb{R}^2$  para un ejemplo de SVM.

### 1.3.5. Multclasificación SVM uno vs. resto

En este tipo de esquema, el problema de clasificar los datos en  $n$  diferentes clases, se convierte en  $n$  problemas de clasificación binaria, es decir, se determinan  $n$  funciones de decisión que separan una clase de las restantes. Asumir que  $d_i(x) = \langle w_i, \phi(x) \rangle + b_i$  es la  $i$ -ésima función de decisión que separa la  $i$ -ésima clase del resto de las clases, entonces si para un nuevo patrón de entrada  $x$  se tiene

$$\text{sgn}(d_i(x)) = +1, \quad (1.28)$$

para un sólo valor de  $i$ , entonces  $x$  se clasifica como perteneciente a la clase  $i$ . Sin embargo, si no se satisface (1.28) para algún valor de  $i$  o se satisface para varios valores de  $i$  no es posible clasificar a  $x$ . Esta idea se ilustra en la Figura 1.11.



**Figura 1.11:** Región no clasificable para el esquema de multclasificación uno vs. resto

### 1.3.6. Multclasificación SVM uno vs. uno

En este esquema se determina una función de decisión para todos los pares de clases posibles, por lo tanto para este esquema se deben construir  $n(n - 1)/2$  funciones de decisión, donde  $n$  es

el número de clases. Sea

$$d_{ij}(x) = \langle w_{ij}, \phi(x) \rangle + b_{ij}, \quad (1.29)$$

la función de decisión que separa la  $i$ -ésima clase de la  $j$ -ésima clase. Si el nuevo patrón de entrada  $x$  pertenece a la región

$$R_i = \{x | \text{sgn}(d_{ij}) = +1, j = 1, \dots, n; j \neq i\}, \quad (1.30)$$

entonces se dice que  $x$  pertenece a la clase  $i$ . Si  $x$  no pertenece a  $R_i$  para  $i = 1, \dots, n$ , es posible llevar a cabo la clasificación mediante un esquema de votación. Para esto se calcula

$$d_i(x) = \sum_{j=1, j \neq i}^n \text{sgn}(d_{ij}(x)), \quad (1.31)$$

que entrega el número de veces que  $x$  fue clasificado como perteneciente a la clase  $i$ . Entonces se tiene la siguiente función de decisión

$$d(x) = \text{arg máx}_{i=1, \dots, n} d_i(x). \quad (1.32)$$

Es decir, si  $d(x)$  es máximo para un determinado valor de  $i$ , se dice que  $x$  pertenece a la  $i$ -ésima clase. El mayor número de votos que puede presentar  $x$  es  $n-1$ , entonces si  $x \in R_i$ ,  $d_i(x) = n-1$  y  $d_k(x) < n-1$  para  $k \neq i$ , el patrón  $x$  pertenece a la clase  $i$ .

---

### MEDIOS DE DIAGNÓSTICO

---

En la Tabla 2.1 se resumen los medios de diagnóstico que se han utilizado para evaluar diferentes tipos de fallas en un MI. En las siguientes secciones se describe el uso de los medios de diagnóstico de las corriente de línea de estator, el módulo del vector de Park, la potencia instantánea total y las vibraciones mecánicas del MI para estudiar fallas de desalineamiento, desbalance y falla en rodamientos.

#### **2.1. Monitoreo de las corrientes de estator**

##### **Falla de desalineamiento y desbalance**

Tanto el desalineamiento y el desbalance del rotor pueden producir oscilaciones en la longitud del entrehierro que causan variaciones en su densidad de flujo, produciendo corrientes armónicas

**Tabla 2.1:** Resumen de técnicas para diagnosticar fallas en un MI.

Monitoreo	Detección	Observación
Espectro de la corriente de estator en una fase [6], [37], [38]	Fallas en baleros, ruptura en la jaula del rotor, desalineamiento, desbalance	Dificultad de diagnóstico con torque de carga variable.
Espectro de la potencia instantánea del motor [39]	Fallas en la jaula del rotor	Alta sensibilidad.
Espectro del modulo del vector de corrientes de Park [40]	Fallas en devanados del estator	Niveles diferentes de par de carga dificultan la detección.
Pulsos de choque [5]	Fallas en baleros	La intensidad del choque varía con los cambios de carga.
Vibraciones [41]	Desbalances en la carga del motor, asimetría en la fuente de alimentación, fallas en el rotor, fallas en baleros	La detección de fallas en baleros es afectada por la velocidad de la máquina.
Emisiones acústicas [5]	Fallas en baleros	–
Campo magnético [5]	Ruptura en barras de rotor, corto circuito en el estator, pérdida de una fase, excentricidad	–
Temperatura [5]	Problemas de fricción, corto circuito en los devanados del estator	–

en el estator a las siguientes frecuencias [5], [6], [37]:

$$\begin{aligned} f_m &= f_s \left[ 1 \pm k \left( \frac{1-s}{p} \right) \right], \\ &= |f_s \pm k f_r|, \end{aligned}$$

donde  $f_s$  es la frecuencia de alimentación eléctrica,  $k = 1, 2, 3, \dots$ ,  $s$  es el deslizamiento,  $p$  es el número de pares de polos, y  $f_r$  es la velocidad mecánica del rotor en Hz.

### Falla de baleros

Este tipo de fallas generan corrientes de estator a las frecuencias dadas por [38], [41]:

$$f_{bnq} = |f_s \pm m f_{i,o}|,$$

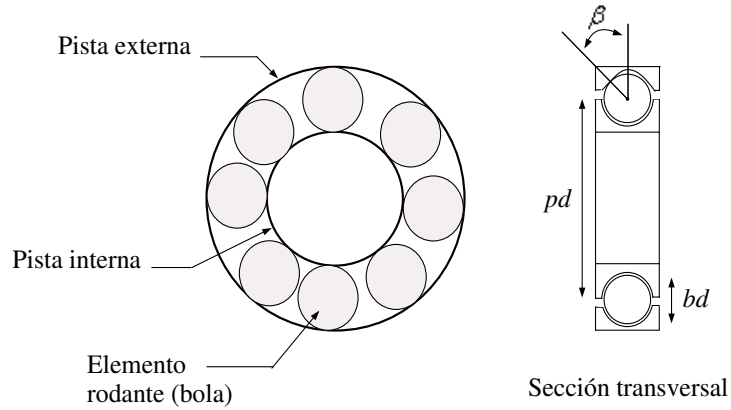
donde  $f_s$  es la frecuencia de alimentación eléctrica,  $m = 1, 2, 3, \dots$  es el número de armónico de las frecuencias características  $f_{i,o}$ , las cuales estarán determinadas por una de las siguientes ecuaciones dependiendo del tipo de defecto en los baleros. Para un defecto en el anillo interno o externo del balero:

$$f_{i,o} = \frac{n}{2} f_r \left[ 1 \pm \frac{bd}{pd} \cos \beta \right], \quad (2.1)$$

donde  $n$  es el número de bolas del balero,  $f_r$  es la velocidad mecánica del rotor en Hz,  $bd$  es el diámetro de la bola,  $pd$  es el diámetro del cojinete de la bola, y  $\beta$  es el ángulo de contacto de la bola con el anillo.

En la Figura 2.1 se indican las características de un balero que están relacionadas con las frecuencias (2.1) en el espectro de la corriente de línea del MI. Sin embargo, como se verá más adelante dependiendo de la severidad de la falla, estos componentes pueden o no estar presentes.

Aunque otro medios de diagnóstico como la potencia instantánea total y el módulo del vector de Park proveen información relacionada con las frecuencias de falla anteriores, estos además contienen información adicional ya que involucran más de una corriente de línea del MI.



**Figura 2.1:** Geometría de un balero.

## 2.2. Monitoreo de la potencia instantánea de entrada del MI

Bajo la suposición de que las fallas en el MI producen bandas laterales de la frecuencia de alimentación en las corrientes de estator, los voltajes de línea a línea, y corrientes de fase de un MI se pueden expresar en estado estacionario como sigue:

$$\begin{aligned}
 v_{ab}(t) &= \sqrt{2}V_{LL} \cos(\omega_0 t + \pi/6), \\
 v_{cb}(t) &= \sqrt{2}V_{LL} \cos(\omega_0 t + \pi/2), \\
 i_a(t) &= \sqrt{2}I_a \cos(\omega_0 t - \phi) + \sqrt{2}I_a^i \cos([\omega_0 - \omega_f]t - \phi) + \sqrt{2}I_a^d \cos([\omega_0 + \omega_f]t - \phi), \\
 i_c(t) &= \sqrt{2}I_c \cos(\omega_0 t - \phi + 2\pi/3) + \sqrt{2}I_c^i \cos([\omega_0 - \omega_f]t - \phi + 2\pi/3) \\
 &\quad + \sqrt{2}I_c^d \cos([\omega_0 + \omega_f]t - \phi + 2\pi/3),
 \end{aligned}$$

donde  $V_{LL} = \sqrt{3}V_p$ , es el valor rms del voltaje de línea a línea,  $\omega_0$  es la frecuencia angular de alimentación,  $\phi$  es el ángulo del factor de potencia, e  $I_a, I_c$  son los valores rms de las corrientes de línea, y se consideran bandas laterales en las corrientes en  $\omega_0 - \omega_f$  y  $\omega_0 + \omega_f$ , donde  $\omega_f$  indica la frecuencia inducida por una falla. La potencia instantánea total  $P_{abc}$  está dada por:

$$\begin{aligned}
 P_{abc} &= P_{ab} + P_{cb}, \\
 &= v_{ab}i_a + v_{cb}i_c.
 \end{aligned} \tag{2.2}$$



Al evaluar la ecuación (2.2) se obtiene:

$$\begin{aligned}
P_{abc} &= P_{ab} + P_{cb}, \tag{2.3} \\
&= V_{LL}I_a \cos(\phi + \pi/6) + V_{LL}I_c \cos(\phi - \pi/6) + (I_a V_{LL} - V_{LL}I_c) \cos(2\omega_0 t - \phi + \pi/6) + \\
&+ V_{LL}I_a^i \cos(\omega_f t + \pi/6 + \phi) + V_{LL}I_c^i \cos(\omega_f t + \phi - \pi/6) + V_{LL}I_a^d \cos(\omega_f t - \phi - \pi/6) + \\
&+ V_{LL}I_c^d \cos(\omega_f t - \phi + \pi/6) + (V_{LL}I_a^d - V_{LL}I_c^d) \cos([2\omega_0 + \omega_f]t - \phi + \pi/6) + \\
&+ (V_{LL}I_a^i - V_{LL}I_c^i) \cos([2\omega_0 - \omega_f]t - \phi + \pi/6).
\end{aligned}$$

De la expresión anterior se desprende que el espectro de  $P_{abc}$  tendrá componentes armónicos a las frecuencias  $2\omega_0$ ,  $\omega_f$  y  $2\omega_0 \pm \omega_f$ . Si se asume que  $I_L = I_a = I_c$ , es decir,

$$\begin{aligned}
i_{a,0}(t) &= \sqrt{2}I_L \cos(\omega_0 t - \phi) + \sqrt{2}I_c^i \cos([\omega_0 - \omega_f]t - \phi) + \sqrt{2}I_c^d \cos([\omega_0 + \omega_f]t - \phi), \\
i_{c,0}(t) &= \sqrt{2}I_L \cos(\omega_0 t - \phi + 2\pi/3) + \sqrt{2}I_c^i \cos([\omega_0 - \omega_f]t - \phi + 2\pi/3) + \\
&+ \sqrt{2}I_c^d \cos([\omega_0 + \omega_f]t - \phi + 2\pi/3).
\end{aligned}$$

Entonces la potencia instantánea total  $P_{abc,0}$  es:

$$\begin{aligned}
P_{abc,0} &= P_{ab,0} + P_{cb,0}, \\
&= 2V_{LL}I_L (\cos(\pi/6) \cos(\phi)) + 2V_{LL}I_c^i (\cos(\pi/6) \cos(\omega_f t + \phi)) + \\
&+ 2V_{LL}I_c^d (\cos(\pi/6) \cos(-\omega_f t + \phi)), \\
&= \sqrt{3}V_{LL}I_L \cos(\phi) + \sqrt{3}V_{LL}(I_c^i \cos(\omega_f t + \phi) + I_c^d \cos(\omega_f t - \phi)).
\end{aligned}$$

De la expresión anterior se desprende que el espectro de  $P_{abc,0}$  tendrá componentes armónicos sólo a la frecuencia  $\omega_f$ . Mientras que para el caso ideal,  $I_c^i = I_c^d = 0$  el espectro de  $P_{abc,0}$  sólo tiene el componente de cd  $\sqrt{3}V_{LL}I_L \cos(\phi)$ . En conclusión, se esperan componentes en el espectro de  $P_{abc}$  a las frecuencias  $2\omega_0$ ,  $\omega_f$  y  $2\omega_0 \pm \omega_f$ .

## 2.3. Monitoreo del vector de corrientes de Park

Para ver el efecto de una condición de falla sobre la magnitud del vector de Park considere lo siguiente. Sean las corrientes de línea de un MI bajo una condición de falla expresadas por:

$$\begin{aligned} i_a(t) &= I_a \cos(\omega_0 t + \phi) + I_a^i \cos([\omega_0 - \omega_f]t + \phi) + I_a^d \cos([\omega_0 + \omega_f]t + \phi), \\ i_b(t) &= I_b \cos(\omega_0 t + \phi - 2\pi/3) + I_b^i \cos([\omega_0 - \omega_f]t + \phi - 2\pi/3) + I_b^d \cos([\omega_0 + \omega_f]t + \phi - 2\pi/3), \\ i_c(t) &= I_c \cos(\omega_0 t + \phi + 2\pi/3) + I_c^i \cos([\omega_0 - \omega_f]t + \phi + 2\pi/3) + I_c^d \cos([\omega_0 + \omega_f]t + \phi + 2\pi/3), \end{aligned} \quad (2.4)$$

donde cada una de las corrientes se representa por una componente fundamental a la frecuencia  $\omega_0$ , y además contienen bandas laterales a las frecuencias  $\omega_0 - \omega_f$ , y  $\omega_0 + \omega_f$ , donde  $\omega_f$  es la frecuencia asociada a una falla. La transformación de Park se usa para obtener una representación bidimensional de las corrientes de estator del MI. Los componentes del vector de corrientes de Park son los siguientes:

$$\begin{cases} i_d = \sqrt{\frac{2}{3}}i_a - \frac{1}{\sqrt{6}}i_b - \frac{1}{\sqrt{6}}i_c, \\ i_q = \frac{1}{\sqrt{2}}i_b - \frac{1}{\sqrt{2}}i_c. \end{cases} \quad (2.5)$$

La magnitud al cuadrado del vector de Park es:

$$||i_{dq}||^2 = i_d^2 + i_q^2. \quad (2.6)$$

Al sustituir (2.4) en (2.5) y evaluar (2.6), se encuentra que el módulo del vector de Park tendrá componentes armónicos a las frecuencias  $2\omega_0$ ,  $2\omega_0 \pm \omega_f$ ,  $2\omega_0 \pm 2\omega_f$ ,  $\omega_f$  y  $2\omega_f$ , además de contar con una componente de cd [39].

## 2.4. Monitoreo de las vibraciones mecánicas del MI

Todas las máquinas giratorias producen vibración en función de la dinámica de la máquina, como el alineamiento y balance del rotor. Comúnmente el análisis de vibraciones se usa para determinar la operación y condición mecánica del equipo. Una gran ventaja es que el análisis de las vibraciones es capaz de identificar problemas desarrollándose antes de que estos lleguen a ser

serios y causen tiempo de inactividad no programado de un MI. Esto se puede lograr llevando a cabo un monitoreo regular de las vibraciones de la máquina, ya sea continuamente o a intervalos programados.

Es posible usar el monitoreo de los niveles de vibración para diagnosticar desalineamiento y desbalance antes de que estas condiciones resulten en un deterioro de los baleros del MI. Por último, el análisis de vibraciones se puede usar como parte de un programa general para aumentar la fiabilidad del equipo [42].

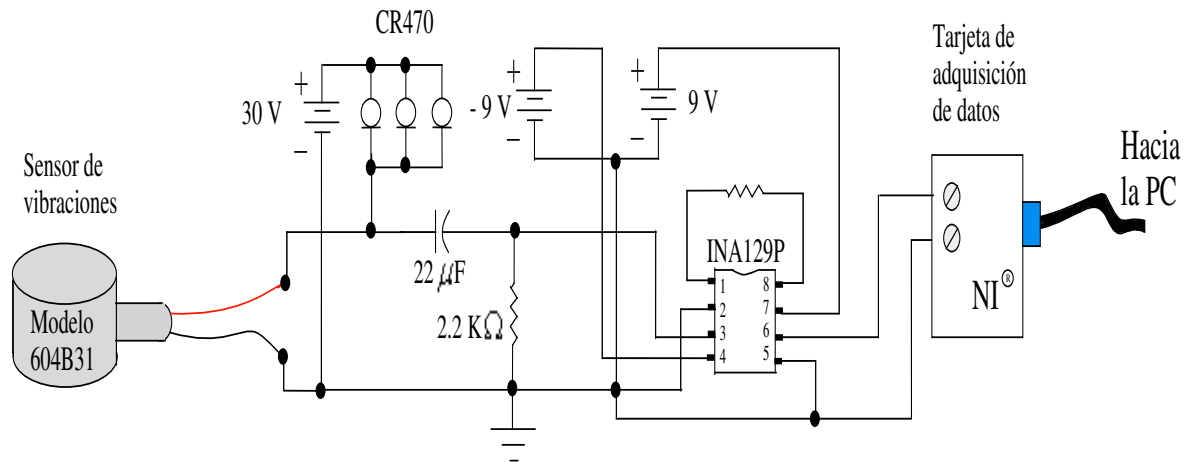
La medición de las vibraciones se lleva a cabo mediante acelerómetros, que son dispositivos de medida inercial que convierten el movimiento mecánico en una señal de voltaje. Estos dispositivos están basados en el principio del efecto piezoeléctrico para producir una señal que es proporcional a la aceleración. Los acelerómetros consisten de un cristal piezoeléctrico y una pequeña masa normalmente encerrada en una protección de metal [43].

En la Figura 2.2 se muestra el circuito usado para la medición de las vibraciones mecánicas del MI, en el cual se utiliza un amplificador de instrumentación INA129P [44] para amplificar la señal de baja amplitud producida por el sensor. Sin embargo, antes de la amplificación la señal pasa por un filtro pasa altas para eliminar la componente de CD dada por el acelerómetro modelo 604B31 [45], se utilizaron 3 diodos reguladores de corriente CR470 [46] en paralelo para incrementar la corriente suministrada al sensor, la cual debe ser cercana a 20mA cuando la extensión del cable del sensor es mayor a 5 metros [45].

En la literatura relacionada con el diagnóstico de fallas en MI [41], [42], se han reportado las frecuencias a las que se presenta vibración mecánica para fallas de desalineamiento, desbalance y deterioro en los rodamientos.

### **Falla de desalineamiento**

Aunque existen dos tipos de desalineamiento (paralelo y angular), es común encontrar una combinación de ambos, lo cual produce componentes en el espectro de la señal de vibración a las frecuencias  $f_r$ ,  $2f_r$  y  $3f_r$ , de los cuales el componente predominante puede ocurrir a  $2f_r$ , donde  $f_r$  es la velocidad de rotación, y además esto ocurre en dirección radial. Sin embargo, cuando



**Figura 2.2:** Circuito para la medición de las vibraciones mecánicas del motor de inducción.

el desalineamiento es severo se pueden encontrar otros armónicos desde  $3f_r$  hasta  $8f_r$ , o incluso toda una serie de armónicos de alta frecuencia.

### Falla de desbalance

Una distribución desigual de masa alrededor del eje giratorio puede generar componentes a la velocidad de rotación  $f_r$ , y además la amplitud de este componente varía directamente proporcional a la velocidad. Al igual que en un caso de desalineamiento, un desbalance en el rotor produce una mayor vibración en dirección radial.

### Falla de baleros

El espectro de frecuencias de una señal de vibración relacionada con una falla en rodamientos puede ser dividido en 4 zonas, que son observadas dependiendo de la severidad del defecto en el balero [42], estas son

- Zona A: Frecuencia de rotación del MI y sus armónicos.
- Zona B: Frecuencias de defectos en los baleros (2.1), es decir,  $f_i$  y  $f_o$  (5-30 krpm).
- Zona C: Frecuencias naturales de los componentes del balero (30-120 krpm).

- Zona D: Altas frecuencias (superior a 120 krpm).

En una primera etapa de una falla de baleros en la que todavía no existen defectos visibles en las pistas y elementos rodantes del balero, aparecen componentes en el espectro de frecuencias de la señal en el rango definido por la zona D.

La segunda etapa de desgaste de un balero relacionada con la zona C, se presenta debido a la oxidación y diminutas picaduras del rodamiento debidas al desgaste normal o a la fatiga del material de las pistas. Los elementos rodantes pasan sobre estas picaduras y se comienzan a generar las frecuencias naturales del balero que ocurren entre 30 y 120 krpm.

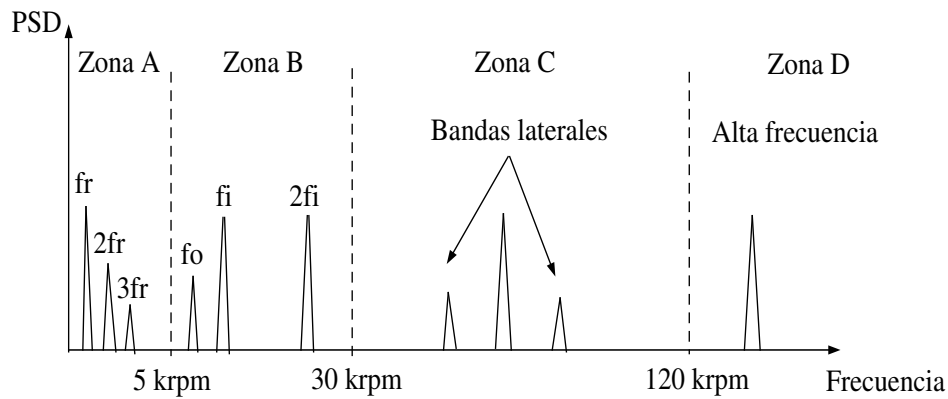
En la tercera etapa se tienen picaduras mucho más grandes que en la segunda etapa y es cuando aparecen componentes en el espectro en el rango definido por la zona B, relacionados a las frecuencias de defecto en pista interna y externa  $f_i$  y  $f_o$ , respectivamente. Según algunos estudios, cuando se presenta esta etapa, la vida útil restante del balero puede ser desde 1 hora a 1 % de su vida promedio.

En la etapa final, las picaduras se unen con otras y a su vez crean pistas muy ásperas, produciendo un crecimiento en los componentes del espectro de frecuencias a la velocidad del eje del MI y sus armónicos, es decir, en la zona A.

En la Figura 2.3 se muestra un bosquejo de la aparición de los componentes debido a una falla de baleros cuando se presentan las 4 etapas.

## **2.5. Monitoreo de los medios de diagnóstico para 3 tipos de falla.**

En esta sección se muestra el efecto que se produce en los medios de diagnóstico para cada uno de los casos de falla estudiados, los cuales son rotor desalineado, rotor desbalanceado y falla en baleros. Para los datos mostrados en esta sección se consideró una frecuencia de operación de 60 Hz y un par de carga de 5 lb-in, cabe mencionar que el factor para convertir las unidades lb-in a N · m es de 0.16.



**Figura 2.3:** Zonas relacionadas a la severidad de los defectos en un rodamiento [42].

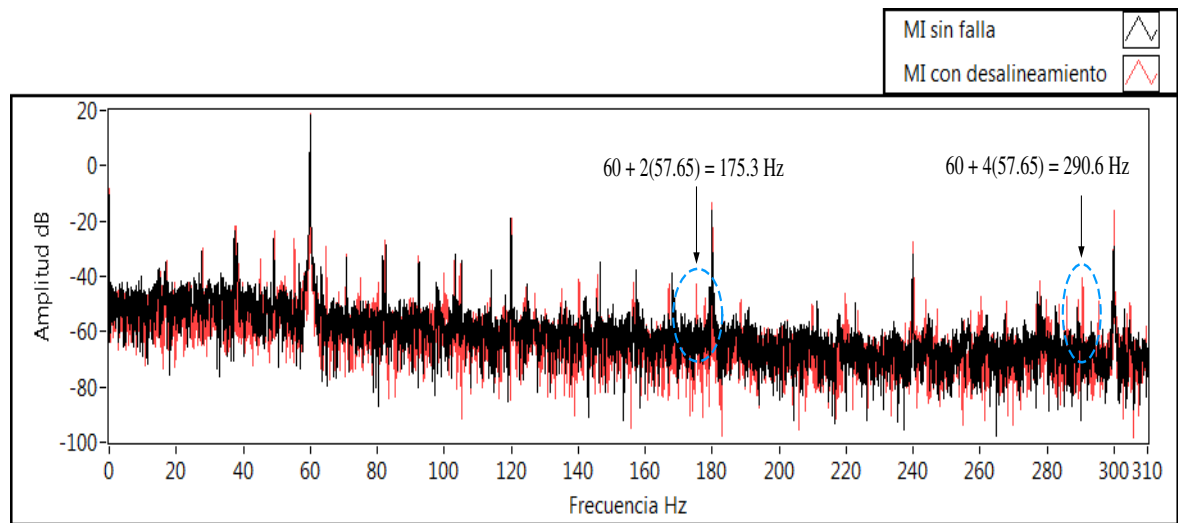
### MI con falla de desalineamiento

Debido a que una condición de desalineamiento produce variaciones en la longitud del entrehierro, es de esperarse que se produzcan componentes de falla en el espectro a las frecuencias  $f_m = |f_s \pm k f_r|$ .

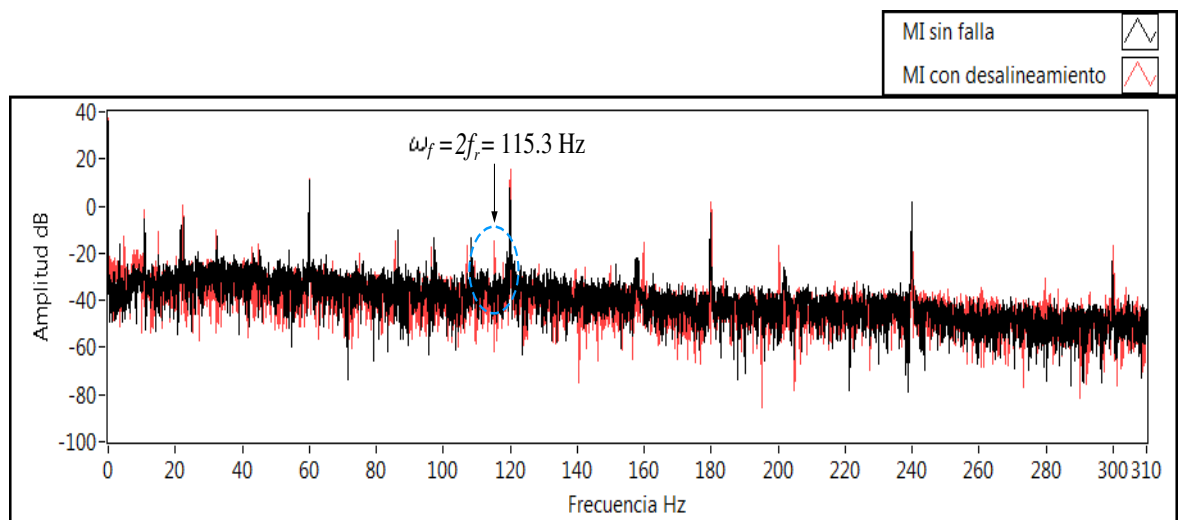
En la Figura 2.4 se muestra el espectro de una corriente de línea para un MI sin falla y un MI con desalineamiento, en este se observa que existe un componente a las frecuencias de falla  $f_{m1} = f_s + 2f_r = 60 + 2(57.65) = 175.3$  Hz, y  $f_{m2} = f_s + 4f_r = 60 + 4(57.65) = 290.6$  Hz. En la Figura 2.5 se muestra la PSD del módulo del vector de Park, para el cual se considera la frecuencia de falla  $f_{m1} - f_s = 2f_r = 115.3$  Hz. En la Figura 2.6 se muestra el efecto producido en la potencia instantánea total, que según la ec. (2.3) se deben generar componentes a las frecuencia  $\omega_f = 2f_r = 115.3$  Hz. Por último, en la PSD de la señal de vibraciones es de esperarse componentes en  $f_r$ ,  $2f_r$  y algunos armónicos de orden mayor de la velocidad del rotor  $f_r$ , en este caso aparecen  $10f_r$  y  $12f_r$  como se observa en la Figura 2.7.

### MI con rotor desbalanceado

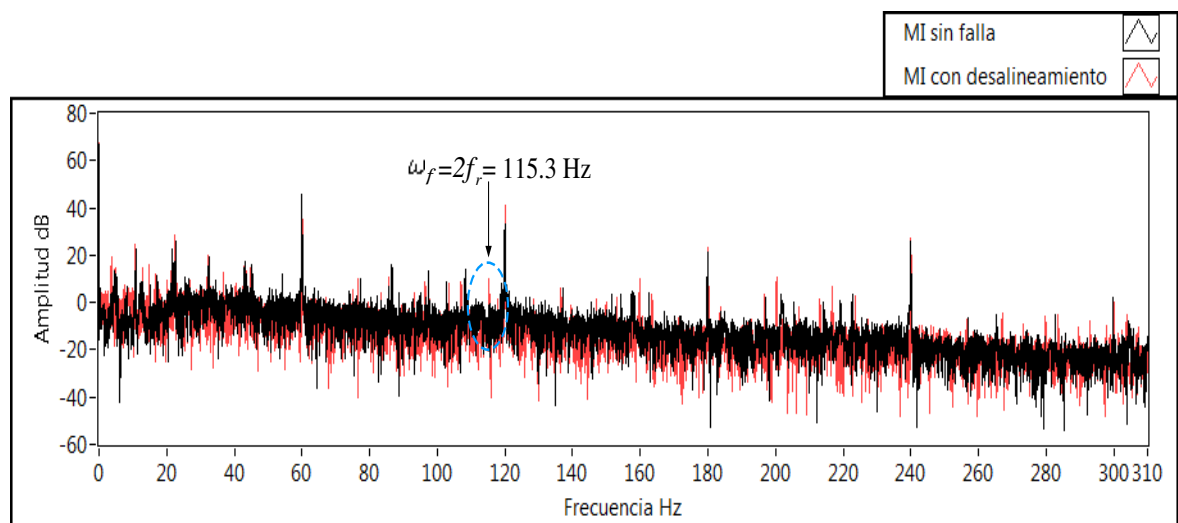
Al igual que una falla de desalineamiento, otra causa de una variación en la longitud del entrehierro lo produce un desbalance en el rotor, por lo que es de esperarse que se encuentren



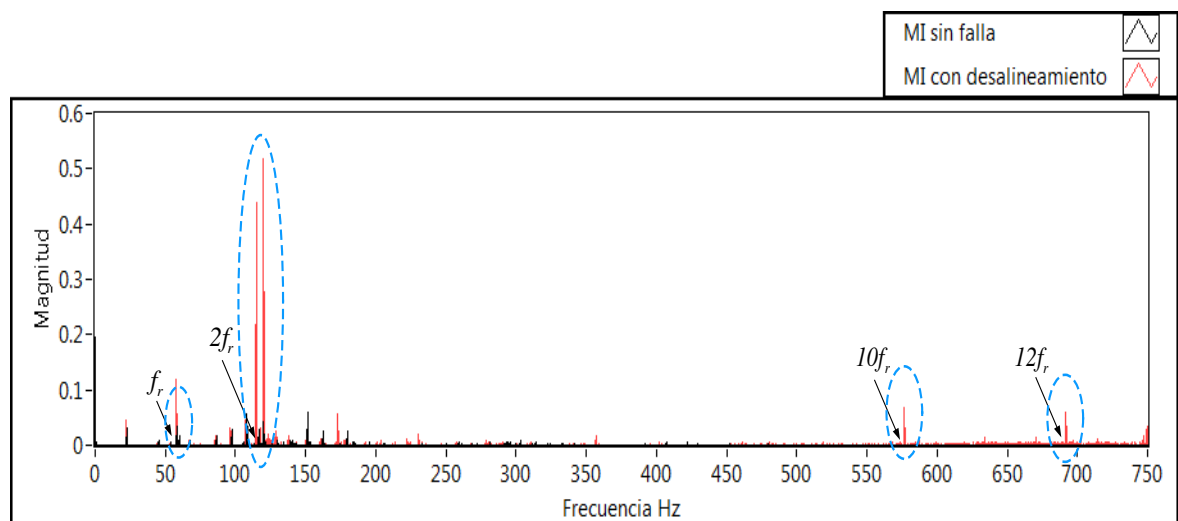
**Figura 2.4:** PSD de la corriente de línea de los MI sin falla y desalineado (dB).



**Figura 2.5:** PSD del módulo del vector de Park para un MI sin falla y desalineado (dB).



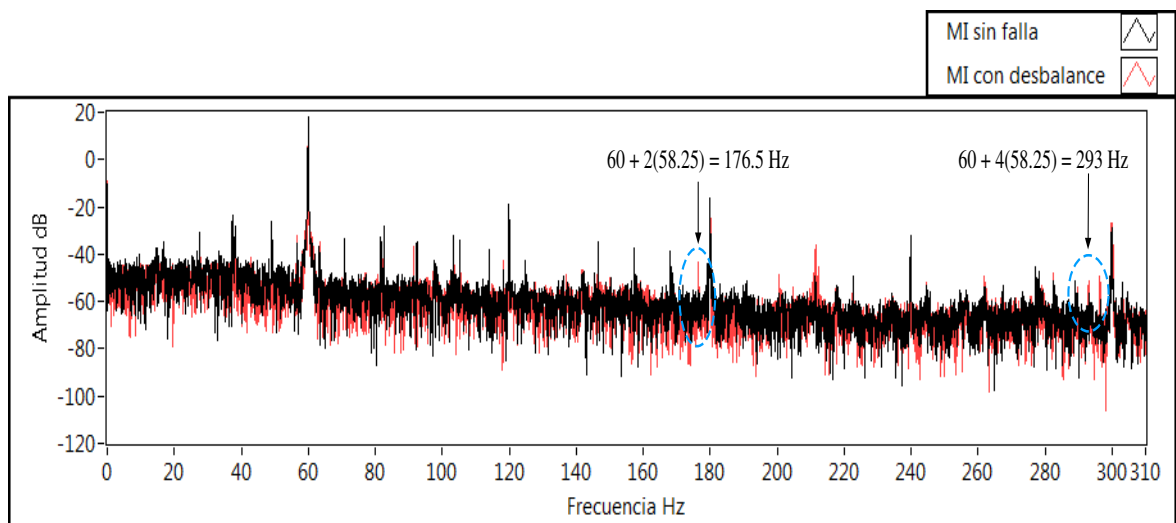
**Figura 2.6:** PSD de la potencia instantánea total para un MI sin falla y desalineado (dB).



**Figura 2.7:** PSD de la señal de vibración en dirección radial para un MI sin falla y desalineado (magnitud).



componentes a las frecuencias de falla  $f_m = |f_s \pm k f_r|$ . En la Figura 2.8 se observa la PSD de una corriente de línea para un MI sin falla y un MI desbalanceado, en esta se encuentran componentes en  $f_{m1} = f_s + 2f_r = 60 + 2(58.25) = 176.5 \text{ Hz}$ , y  $f_{m2} = f_s + 4f_r = 60 + 4(58.25) = 293 \text{ Hz}$ . A pesar de que el par de carga se ajustó en  $5 \text{ lb} - \text{in}$ , existe una diferencia entre las frecuencias de rotación para los casos de desalineamiento y desbalance, esto se debe a la tensión que se le da a las bandas que acoplan el eje del MI con el torque mecánico. Asimismo en el módulo del

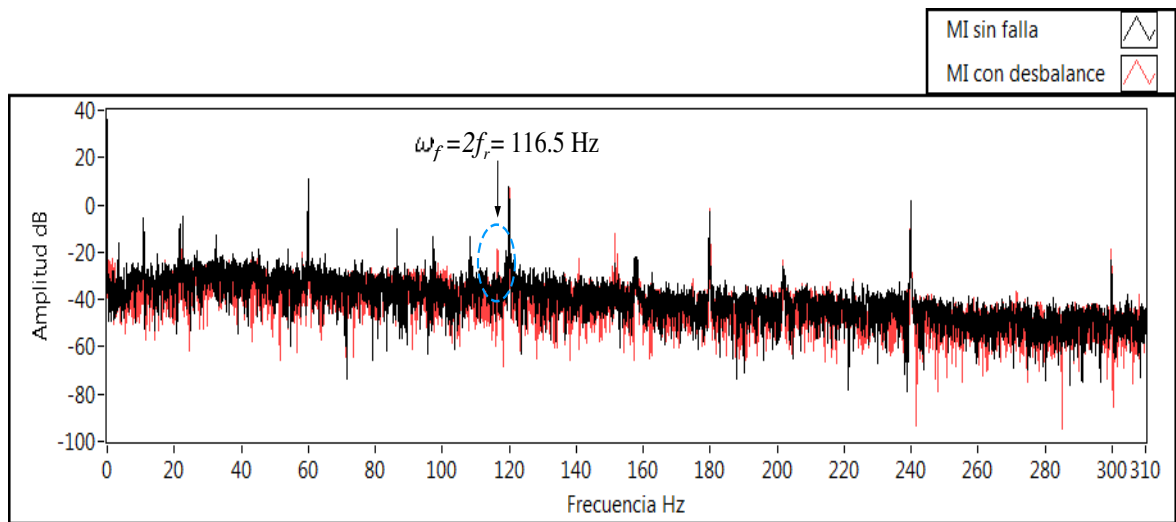


**Figura 2.8:** PSD de la corriente de línea para un MI sin falla y desbalanceado (dB).

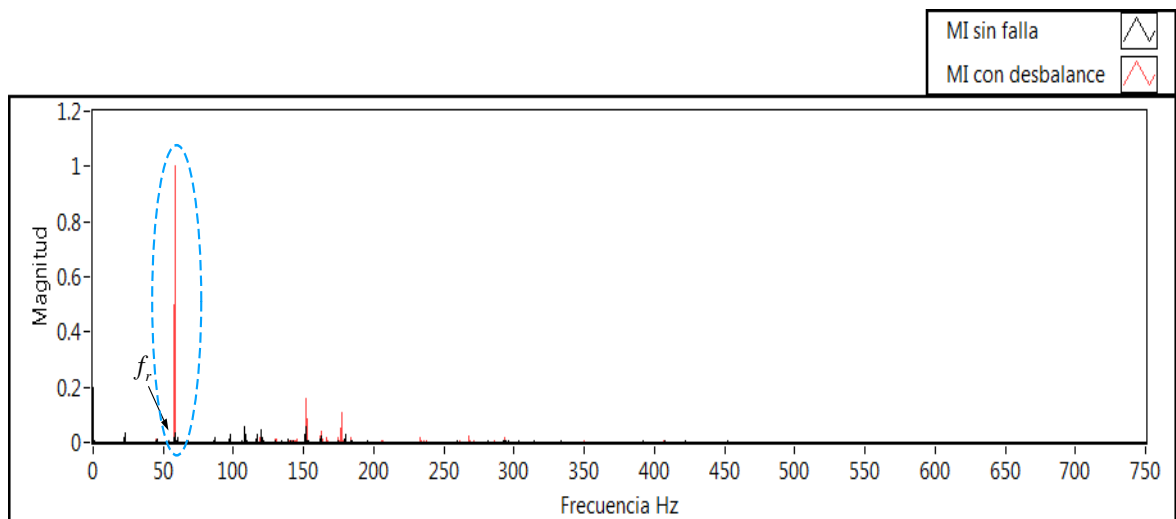
vector de Park se encuentran la frecuencia de falla  $\omega_f = 2f_r = 116.5 \text{ Hz}$ , esto se muestra en la Figura 2.9. También se encuentra el componente de falla a  $\omega_f = 2f_r \text{ Hz}$  en el PSD de la potencia instantánea total del MI. Sin embargo, a pesar que hay similitudes en los medios de diagnósticos eléctricos, para una falla de desalineamiento y desbalance, la PSD de la señal de vibraciones proporciona información adicional, como se muestra en la Figura 2.10, para un caso de desbalance se encuentra un componente predominante a la frecuencia de rotación del eje  $f_r$ .

### MI con falla en baleros

Una falla de baleros es más difícil de diagnosticar ya que las frecuencias características de la falla no dependen directamente de la velocidad de rotación del eje del MI, y además producen

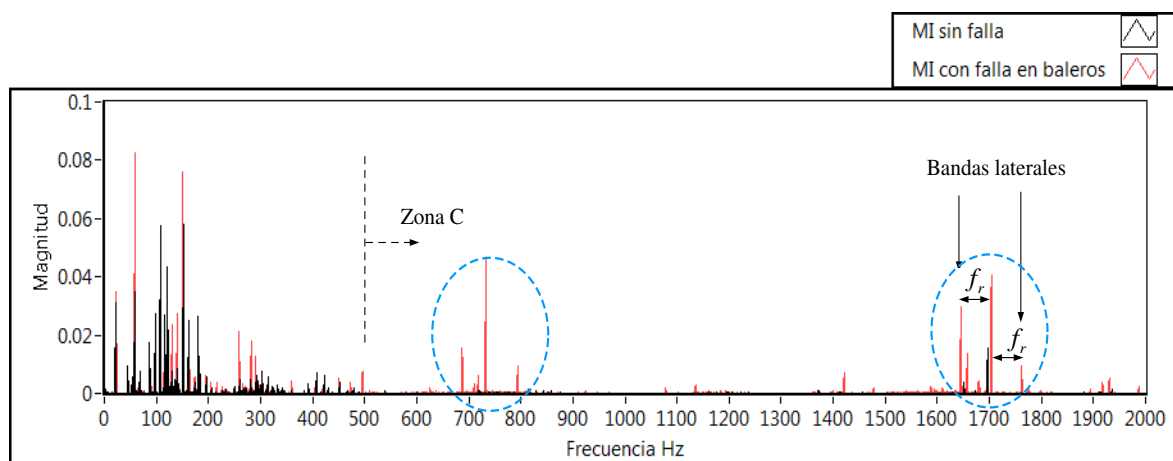


**Figura 2.9:** PSD del módulo del vector de Park para un MI sin falla y desbalanceado (dB).



**Figura 2.10:** PSD de la señal de vibración en dirección radial para un MI sin falla y desbalanceado (magnitud).

señales de vibración de menor amplitud que para una falla de desalineamiento [42]. Sin embargo según el deterioro del balero es posible encontrar algunas características en la PSD de las vibraciones. En la Figura 2.11 se muestra el espectro de la señal de vibración para un MI sin falla y un MI con falla incipiente en baleros, en esta se muestra que los baleros del MI se encuentran en una segunda etapa de desgaste pues sólo se encuentran componentes relacionados con la falla en la zona C descrita anteriormente. Observe además, que en comparación con las demás fallas los niveles de magnitud de estos componentes son mucho menores. En este caso, no se encuentran componentes a las frecuencias conocidas como  $f_i$  y  $f_o$  pertenecientes a la tercera etapa de desgaste, que según las características de los baleros del MI considerado [47], [48], se estiman en  $f_i = 4.932f_r$  y  $f_o = 3.066f_r$ , considerando un ángulo de contacto  $\beta$  nulo, es decir, cerca del tercer y quinto armónico de la velocidad de rotación  $f_r$ . Cuando estos componentes se encuentran en el espectro de vibraciones es de esperar componentes de falla en  $f_{bmg} = |f_s \pm m f_{i,o}|$ , para  $m = 1, 2, 3, \dots$ , en el espectro de la corriente de línea, sin embargo, en este caso los baleros se encuentran en un nivel de degradación muy incipiente y por ello no existen. Los componentes encontrados en la zona C se encuentran con bandas laterales a una distancia de la velocidad de rotación  $f_r$ , como se muestra en la Figura 2.11.



**Figura 2.11:** PSD de la señal de vibraciones del MI sin falla y con falla en baleros (magnitud).

## 2.6. Generación de características

El razonamiento básico detrás de las características basadas en transformación es que un mapeo escogido apropiadamente puede explotar y remover redundancia en la información, que usualmente existe en el conjunto de datos obtenidos por los dispositivos de medición. Si la transformación se escoge adecuadamente, las características pueden exhibir propiedades de alta información comparadas con los datos de entrada originales. Enseguida se presenta la transformación usada para la extracción de características. Se define una función en el dominio de la frecuencia que nos da una medida de las contribuciones relativas de energía en función de la frecuencia para una señal dada, esta función se le conoce como densidad espectral de potencia (PSD, por sus siglas en inglés), y es descrita por la siguiente ecuación:

$$P_k = |D(k)|^2, \quad (2.7)$$

donde  $D(k)$  es la transformada discreta de Fourier dada por

$$D(k) = \sum_{n=0}^{N-1} d(n)e^{-j2\pi kn/N},$$

donde  $N$  es el número de muestras que tiene la señal en tiempo discreto  $d(n)$ ,  $0 \leq n \leq N - 1$ , y  $k = 0, 1, \dots, N - 1$ . En este caso la corriente de línea  $i_a$  y la magnitud del vector de Park  $\|i_{dq}\|$ , son señales muestreadas a una frecuencia de  $f = 40$  KHz, y con  $N = 800,000$ , para tener un tiempo de adquisición de 20 seg, con esto se logra una resolución en la PSD de  $\Delta f = f/N = 0.05$  Hz; cabe mencionar que la selección de la frecuencia de muestreo se realizó en base a la frecuencia de conmutación del inversor, la cual es de 8 KHz, y a la resolución deseada en el espectro.

La idea es crear características o patrones basados en los medios de diagnóstico para el diseño de una SVM-multiclase, para esto se definen las regiones en los medios de diagnóstico que indican la presencia de las fallas. Con base en lo observado en la PSD de la corriente de línea, el módulo del vector de Park, la potencia instantánea total y las señales de vibración, para diferentes niveles de par de carga, 1,3 y 5 lb-in y valores de la frecuencia de alimentación 30, 40, 50 y 60 Hz para los tres tipos de falla, se definen las siguientes características.

### **MI con desalineamiento**

En la PSD de la corriente de línea se toma la suma de las amplitudes a las frecuencias  $f_m = f_s + kf_r$ , con  $k = 1, 2, 3, 4$ , para sobrellevar la imprecisión de esta fórmula se toma en cuenta  $f_m \pm 0.5$  Hz. Tanto en la PSD del módulo del vector de Park como en la potencia instantánea total se toma la suma de las amplitudes a las frecuencias de  $f_r \pm 0.5$  Hz,  $2f_r \pm 0.5$  Hz,  $3f_r \pm 0.5$  Hz,  $4f_r \pm 0.5$  Hz. Por último en la PSD de la señal de vibraciones  $v(n)$  se toma la suma de las amplitudes a las frecuencias  $2f_r \pm 0.5$  Hz y  $3f_r \pm 0.5$  Hz, y como otra característica se tiene la suma de las amplitudes de los armónicos existentes entre el quinto hasta una frecuencia de 900 Hz. Por ejemplo, para  $f_r = 58$  Hz se tomaría en cuenta hasta el armónico 15, aunque las amplitudes de los armónicos de mayor orden son pequeños, la suma supera a la producida en una falla de desbalance o de baleros ya que en estas no se generan dichos armónicos.

### **MI con desbalance**

En este caso se consideran las mismas características que en el caso de desalineamiento para los medios de diagnósticos eléctricos, por generar componentes a la misma frecuencia de falla  $f_m$ . Sin embargo, en la PSD de las vibraciones se considera la amplitud a la frecuencia de rotación del MI  $f_r$ .

### **MI con falla en baleros**

La característica encontrada en este tipo de fallas fue en la PSD de la señal de vibración  $v(n)$ , se tomó en cuenta la suma de las amplitudes en el rango de 1500 a 1800 Hz, perteneciente a la zona característica de la etapa inicial de un defecto en baleros.

Entonces las 7 características a extraer de las mediciones eléctricas y mecánicas están dadas por:

$$c(i_a) = \sum_{k=k_s+k_r-10}^{k_s+k_r+10} P_k(i_a) + \sum_{k=k_s+2k_r-10}^{k_s+2k_r+10} P_k(i_a) + \sum_{k=k_s+3k_r-10}^{k_s+3k_r+10} P_k(i_a) + \sum_{k=k_s+4k_r-10}^{k_s+4k_r+10} P_k(i_a), \quad (2.8)$$

$$c(\|i_{dq}\|) = \sum_{k=k_r-10}^{k_r+10} P_k(\|i_{dq}\|) + \sum_{k=2k_r-10}^{2k_r+10} P_k(\|i_{dq}\|) + \sum_{k=3k_r-10}^{3k_r+10} P_k(\|i_{dq}\|) + \sum_{k=4k_r-10}^{4k_r+10} P_k(\|i_{dq}\|), \quad (2.9)$$

$$c(P_{abc}) = \sum_{k=k_r-10}^{k_r+10} P_k(P_{abc}) + \sum_{k=2k_r-10}^{2k_r+10} P_k(P_{abc}) + \sum_{k=3k_r-10}^{3k_r+10} P_k(P_{abc}) + \sum_{k=4k_r-10}^{4k_r+10} P_k(P_{abc}), \quad (2.10)$$

$$c_1(v) = \sum_{k=k_r-10}^{k_r+10} P_k(v), \quad (2.11)$$

$$c_2(v) = \sum_{k=2k_r-10}^{2k_r+10} P_k(v) + \sum_{k=3k_r-10}^{3k_r+10} P_k(v), \quad (2.12)$$

$$c_3(v) = \sum_{k=5k_r-10}^{5k_r+10} P_k(v) + \sum_{k=6k_r-10}^{6k_r+10} P_k(v) + \dots + \sum_{k=\lfloor 18000/k_r \rfloor k_r - 10}^{\lfloor 18000/k_r \rfloor k_r + 10} P_k(v), \quad (2.13)$$

$$c_4(v) = \sum_{k=30000}^{36000} P_k(v), \quad (2.14)$$

donde  $k_s = f_s/0.05$  es el número de muestra en la PSD correspondiente a la frecuencia de alimentación  $f_s$ , asimismo para  $k_r = f_r/0.05$  con respecto a la frecuencia de rotación.  $c(i_a)$  y  $c(\|i_{dq}\|)$  son características obtenidas a partir de la corriente de línea y el módulo del vector de Park respectivamente, relacionadas con las fallas de desalineamiento y desbalance,  $c_2(v)$  y  $c_3(v)$  son características obtenidas a partir de las señales de vibración relacionadas a una falla de desalineamiento. Por último  $c_1(v)$  y  $c_4(v)$  son características obtenidas a partir de las señales de vibración relacionadas con las fallas de desbalance y rodamientos respectivamente.

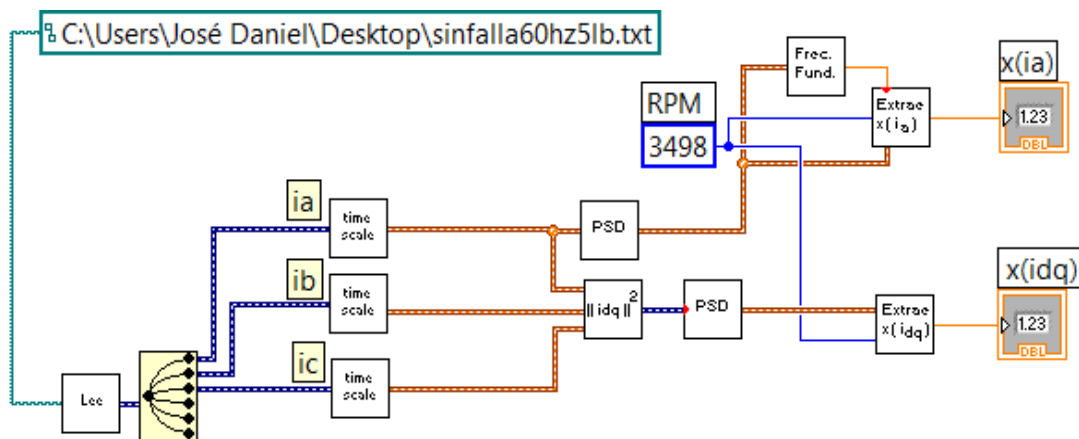
Los índices en las ecuaciones (2.8) a (2.14) son establecidos de acuerdo a la correspondencia

entre las frecuencias de falla y el rango de muestras en la PSD, esta relación se muestra en la Tabla 2.2.

**Tabla 2.2:** Relación entre los rangos de frecuencias de falla y el número de muestras en la PSD.

No. de muestras	$k_s + k_r \pm 10$	$k_r \pm 10$	$k = 30000$	$k = 36000$
Frecuencia	$f_s + f_r \pm 0.5 \text{ Hz}$	$f_r \pm 0.5 \text{ Hz}$	1500 Hz	1800 Hz

Para extraer las características de los medios de diagnóstico mediante LabView, primero se adquieren las señales de las corrientes de línea del MI utilizando la tarjeta de adquisición de datos NI-Compaq-DAQ-9172, estas se guardan en un archivo de texto *.txt* para posteriormente generar las características fuera de línea. En la Figura 2.12 se muestra el programa en LabView para generar las características relacionadas con la corriente de línea y el módulo del vector de Park. En



**Figura 2.12:** Generación en Labview de las características  $c(i_a)$  y  $c(\|i_{dq}\|)$ .

la Figura 2.13 se muestra el programa utilizado para la característica de la potencia instantánea total, y por último en la Figura 2.14 se muestra el programa para extraer las características relacionadas con la señal de vibración. En estos programas se utilizaron los bloques de Labview conocidos como *FFT power spectral density*, *array subset*, *for loop*, *round to - infinity*, entre otros, como operaciones básicas [49].

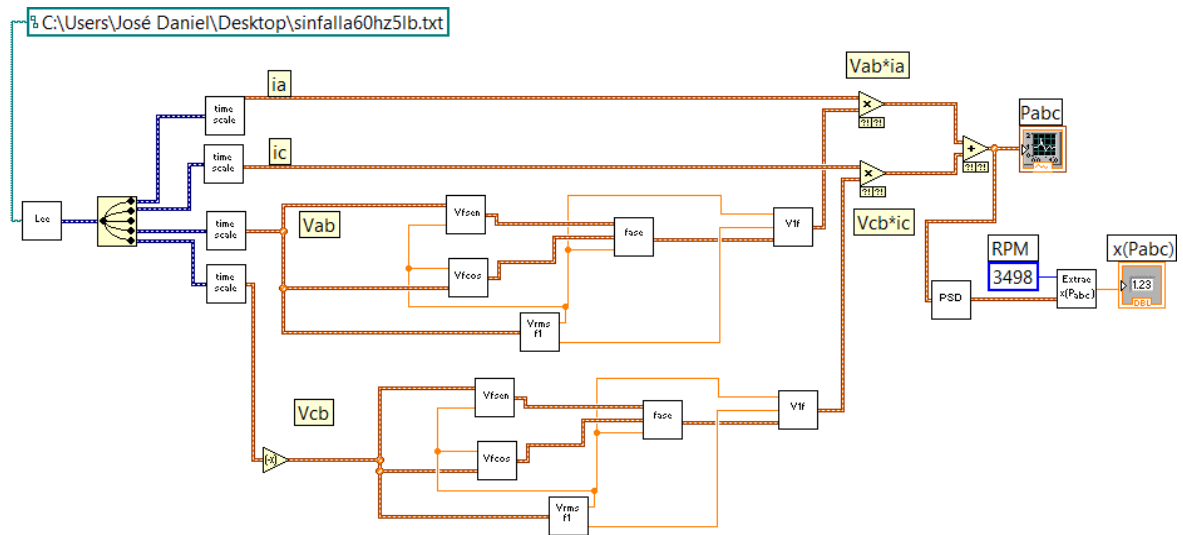


Figura 2.13: Extracción de la característica  $c(P_{abc})$ .

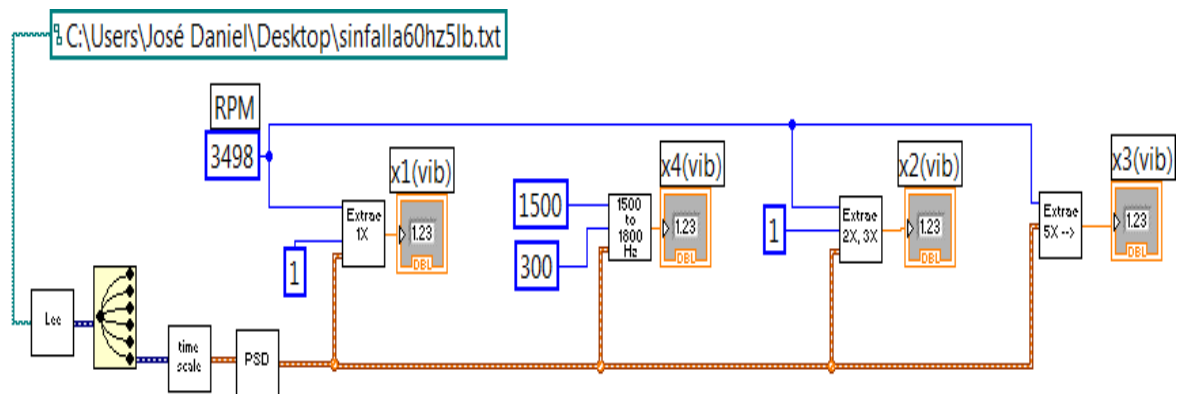
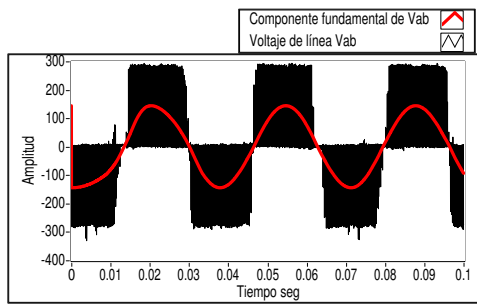
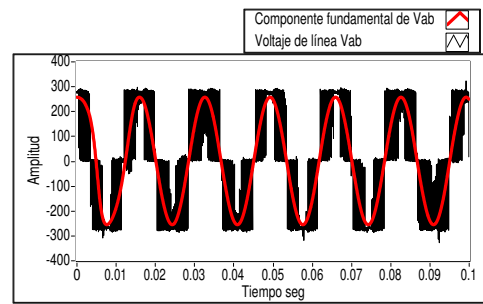


Figura 2.14: PSD de la señal de vibraciones del MI sin falla y con falla en baleros (magnitud).



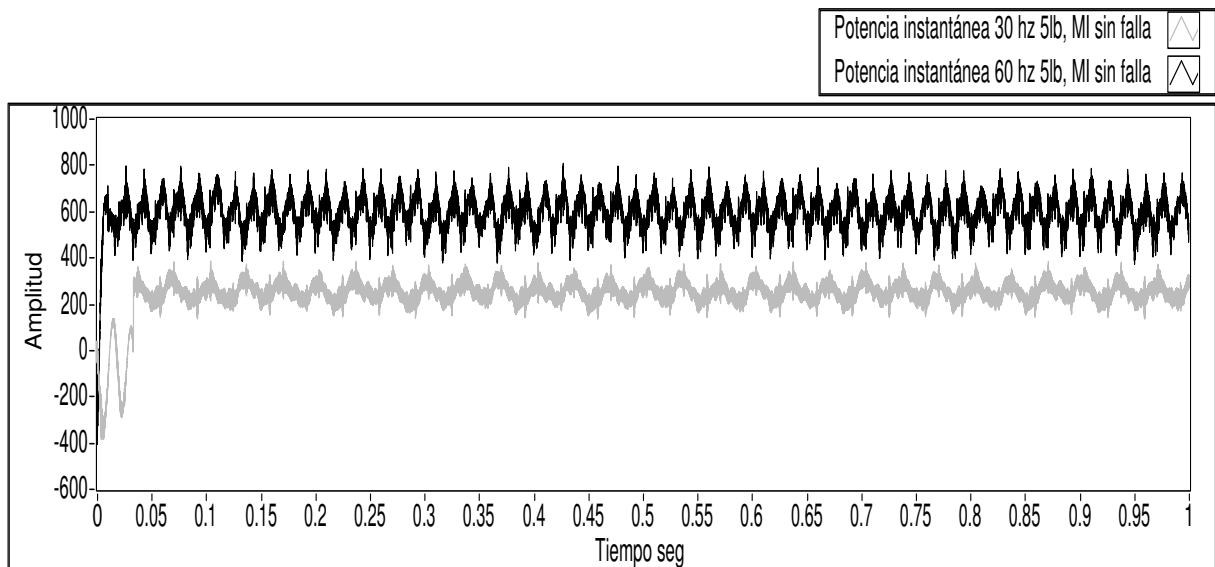


**Figura 2.15:** Voltaje de línea fundamental  $V_{ab}$  para  $f_s = 30$  Hz.



**Figura 2.16:** Voltaje de línea fundamental  $V_{ab}$  para  $f_s = 60$  Hz.

En el tipo de control voltz/Hertz constante siempre que la frecuencia de alimentación se modifique para variar la velocidad del motor, los voltajes de estator también deben cambiar (ver Figuras 2.15 y 2.16) en la misma proporción para mantener un flujo magnético constante. Este aumento en el voltaje fundamental de línea ocasiona que la amplitud de la potencia instantánea total se incremente considerablemente (ver Figura 2.17), produciendo que los valores de la característica  $c(P_{abc})$  de un MI sin falla con  $f_s = 60$  Hz estén por encima de los valores relacionados con una condición de falla con  $f_s = 30$  Hz, ya que no sólo se toman en cuenta las frecuencias de falla si no también las amplitudes. Este fenómeno se muestra en la Tabla 2.3 para un MI sin



**Figura 2.17:** Potencia instantánea total de un MI para diferentes frecuencias de alimentación.

falla y un MI con desalineamiento, se observa que  $c(P_{abc})$  para un MI sin falla con  $f_s = 60$  Hz no es mayor que para un MI desalineado con  $f_s = 60$  Hz, pero si lo es a los datos obtenidos para un MI desalineado con  $f_s = 30$  y  $40$  Hz. Sin embargo, este fenómeno no ocurre para los demás medios de diagnóstico. Por este motivo, y debido a que la potencia instantánea genera información similar en frecuencia a la magnitud del vector de Park, se optó por no utilizarla en la implementación final de diagnóstico.

**Tabla 2.3:** Características basadas en la potencia instantánea total del MI sin falla y con desalineamiento.

$c(P_{abc})$		
MI sin falla	MI desalineado	Operación
1.34213	5.20112	$f_s = 30\text{Hz}$ , 1 lb - in
1.47678	5.37076	$f_s = 30\text{Hz}$ , 3 lb - in
2.57366	4.02936	$f_s = 30\text{Hz}$ , 5 lb - in
3.86881	12.2292	$f_s = 40\text{Hz}$ , 1 lb - in
2.50633	9.51911	$f_s = 40\text{Hz}$ , 3 lb - in
3.6782	11.3149	$f_s = 40\text{Hz}$ , 5 lb - in
3.49401	11.6463	$f_s = 50\text{Hz}$ , 1 lb - in
6.44817	17.0455	$f_s = 50\text{Hz}$ , 3 lb - in
5.3894	17.1955	$f_s = 50\text{Hz}$ , 5 lb - in
11.0028	34.7342	$f_s = 60\text{Hz}$ , 1 lb - in
9.69274	40.0474	$f_s = 60\text{Hz}$ , 3 lb - in
20.969	24.7664	$f_s = 60\text{Hz}$ , 5 lb - in

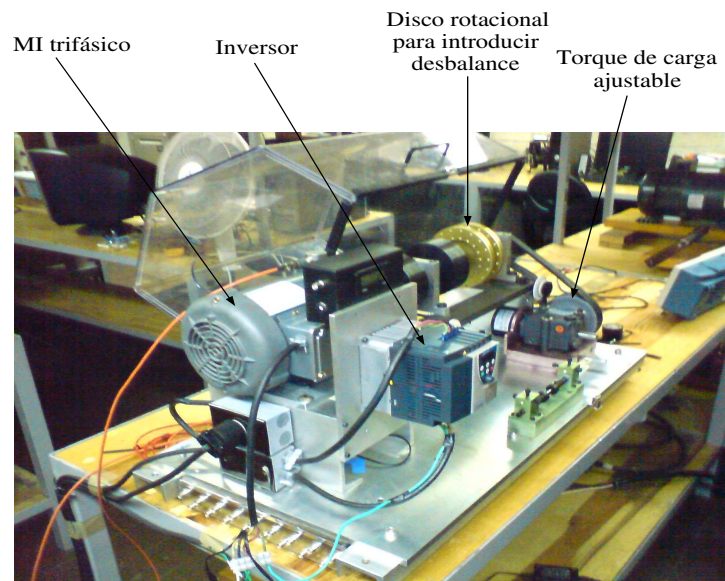
Los datos se obtuvieron usando el kit “Machine Fault Simulator” mostrado en la Figura 2.18, mediante la tarjeta de adquisición de datos NI Compaq-DAQ-9172. Los MI usados tienen las siguientes especificaciones: 60 Hz, 1/2 HP, 3450 rpm, 2 polos, 230/469 V, y 2. 2/1. 1 A. Mientras que las especificaciones de la tarjeta de adquisición de datos y el equipo de cómputo se muestran en la Tabla 2.4.

**Tabla 2.4:** Especificaciones del equipo de adquisición.

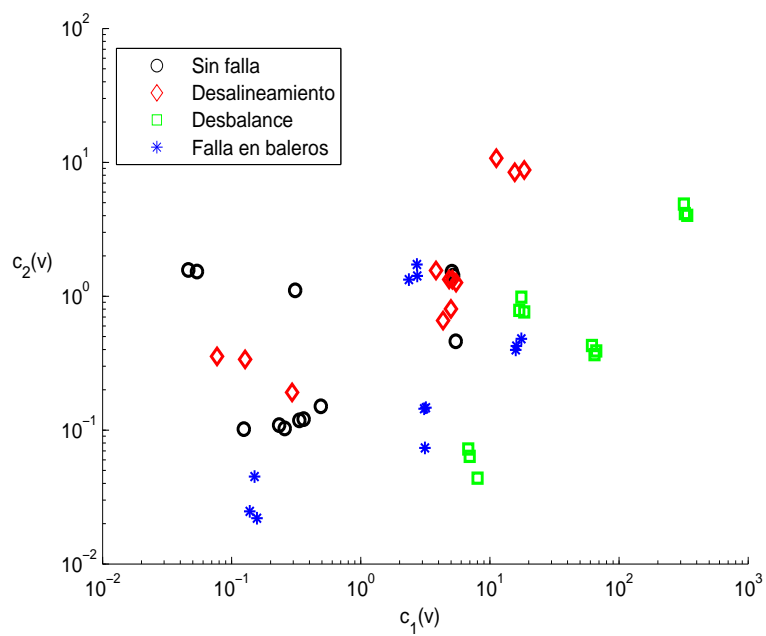
Tarjeta NI Compaq-DAQ-9172
32 entradas analógicas de una sola terminal o 16 diferenciales Resolución de 16 bits, rango de muestreo de 250 kS/s Rangos de entrada programables de $\pm 200$ mV a $\pm 10$ V
PC
Procesador Pentium Dual Core, 2. 5 GHz 4 GB de memoria RAM

En las Tablas 2.5, y 2.6 se muestran los valores de las características para cada uno de los escenarios de falla a estudiar, los datos presentados no han sido escalados, sin embargo, para evitar posteriormente problemas de precisión numérica se escalarán de la siguiente manera  $c(i_a) * 10000$ ,  $c(|i_{dq}|) * 100$ ,  $c_1(v) * 100$ ,  $c_2(v) * 10$ ,  $c_3(v) * 10$  y  $c_4(v) * 10$ .

A partir de la Figura 2.19 se muestran los mapas de características en escala logarítmica, que nos indican que los datos son no linealmente separables, y por lo tanto será necesaria la construcción de SVM no lineales. Como era de esperarse para la característica  $c_1(v)$ , los datos relacionados con una falla de desbalance se encuentran más alejados de las otras clases, también se muestra que las amplitudes crecen conforme la velocidad aumenta, sin embargo no sucede lo mismo para los datos de  $c_2(v)$  (ver Figura 2.19) relacionados a una falla de desalineamiento excepto para cuando  $f_s = 60$  Hz. Sin embargo los datos proporcionados por  $c_3(v)$  si permiten diferenciar, al menos visualmente una falla de desalineamiento, esto se muestra en la Figura 2.20.



**Figura 2.18:** Kit de pruebas “Machine Fault Simulator”.



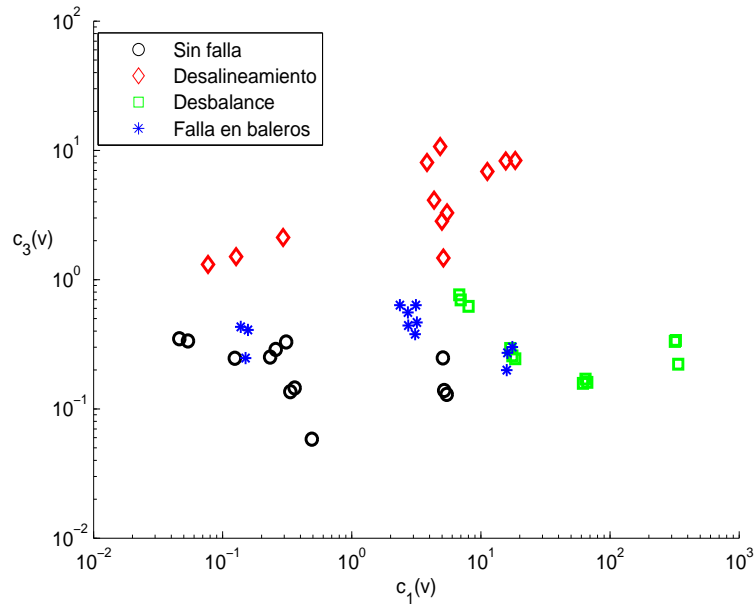
**Figura 2.19:** Mapa característico  $c_1(v)$  vs.  $c_2(v)$ .

**Tabla 2.5:** Características para un MI sin falla y desalineado.

$c(i_a)$	$c(\ i_{dq}\ )$	$c_1(v)$	$c_2(v)$	$c_3(v)$	$c_4(v)$	Régimen
<b>MI sin falla</b>						
$4.99 \times 10^{-5}$	$3.61 \times 10^{-3}$	$2.32 \times 10^{-3}$	$1.08 \times 10^{-2}$	$2.51 \times 10^{-2}$	$5.06 \times 10^{-2}$	30 Hz, 1lb-in
$7.19 \times 10^{-5}$	$5.35 \times 10^{-3}$	$1.24 \times 10^{-3}$	$1.01 \times 10^{-2}$	$2.46 \times 10^{-2}$	$6.32 \times 10^{-2}$	30 Hz, 3lb-in
$1.12 \times 10^{-4}$	$5.46 \times 10^{-3}$	$2.57 \times 10^{-3}$	$1.02 \times 10^{-2}$	$2.88 \times 10^{-2}$	$7.3 \times 10^{-2}$	30 Hz, 5lb-in
$4.21 \times 10^{-5}$	$7.01 \times 10^{-3}$	$4.9 \times 10^{-3}$	$1.5 \times 10^{-2}$	$5.82 \times 10^{-3}$	$7.94 \times 10^{-2}$	40 Hz, 1lb-in
$3.9 \times 10^{-5}$	$7.36 \times 10^{-3}$	$3.61 \times 10^{-3}$	$1.21 \times 10^{-2}$	$1.45 \times 10^{-2}$	$5.85 \times 10^{-2}$	40 Hz, 3lb-in
$8 \times 10^{-5}$	$8.16 \times 10^{-3}$	$3.33 \times 10^{-3}$	$1.18 \times 10^{-2}$	$1.36 \times 10^{-2}$	$4.9 \times 10^{-2}$	40 Hz, 5lb-in
$2.93 \times 10^{-5}$	$8.29 \times 10^{-3}$	$4.61 \times 10^{-4}$	$1.57 \times 10^{-1}$	$3.48 \times 10^{-2}$	$1.43 \times 10^{-1}$	50 Hz, 1lb-in
$6.14 \times 10^{-5}$	$1.06 \times 10^{-2}$	$5.38 \times 10^{-4}$	$1.53 \times 10^{-1}$	$3.35 \times 10^{-2}$	$1.68 \times 10^{-2}$	50 Hz, 3lb-in
$4.65 \times 10^{-5}$	$1.39 \times 10^{-2}$	$3.1 \times 10^{-3}$	$1.1 \times 10^{-1}$	$3.29 \times 10^{-2}$	$8.84 \times 10^{-2}$	50 Hz, 5lb-in
$9.13 \times 10^{-5}$	$1.33 \times 10^{-2}$	$5.19 \times 10^{-2}$	$1.42 \times 10^{-2}$	$1.38 \times 10^{-2}$	$9.16 \times 10^{-2}$	60 Hz, 1lb-in
$7.86 \times 10^{-5}$	$1.54 \times 10^{-2}$	$5.07 \times 10^{-2}$	$1.52 \times 10^{-2}$	$2.47 \times 10^{-2}$	$8.79 \times 10^{-2}$	60 Hz, 3lb-in
$7.75 \times 10^{-5}$	$3.39 \times 10^{-2}$	$5.43 \times 10^{-2}$	$4.61 \times 10^{-2}$	$1.29 \times 10^{-2}$	$1.81 \times 10^{-1}$	60 Hz, 5lb-in
<b>MI con desalineamiento</b>						
$2.17 \times 10^{-4}$	$1.41 \times 10^{-2}$	$2.93 \times 10^{-3}$	$1.91 \times 10^{-2}$	$2.12 \times 10^{-1}$	$2.88 \times 10^{-2}$	30 Hz, 1lb-in
$2.08 \times 10^{-4}$	$1.67 \times 10^{-2}$	$1.26 \times 10^{-3}$	$3.37 \times 10^{-2}$	$1.5 \times 10^{-1}$	$2.26 \times 10^{-2}$	30 Hz, 3lb-in
$2.72 \times 10^{-4}$	$2.02 \times 10^{-2}$	$7.7 \times 10^{-4}$	$3.55 \times 10^{-2}$	$1.31 \times 10^{-1}$	$2.32 \times 10^{-2}$	30 Hz, 5lb-in
$2.68 \times 10^{-4}$	$3.23 \times 10^{-2}$	$4.33 \times 10^{-2}$	$6.58 \times 10^{-2}$	$4.11 \times 10^{-1}$	$7.67 \times 10^{-1}$	40 Hz, 1lb-in
$2.94 \times 10^{-4}$	$3.06 \times 10^{-2}$	$4.98 \times 10^{-2}$	$8.02 \times 10^{-2}$	$2.83 \times 10^{-1}$	$7.77 \times 10^{-2}$	40 Hz, 3lb-in
$2.89 \times 10^{-4}$	$4.75e^{-2}$	$5.46 \times 10^{-2}$	$1.26 \times 10^{-1}$	$3.28 \times 10^{-1}$	$8.28 \times 10^{-2}$	40 Hz, 5lb-in
$2.54 \times 10^{-4}$	$2.55 \times 10^{-2}$	$3.82 \times 10^{-2}$	$1.55 \times 10^{-1}$	$8.05 \times 10^{-1}$	$1.85 \times 10^{-1}$	50 Hz, 1lb-in
$2.83 \times 10^{-4}$	$4.92 \times 10^{-2}$	$4.83 \times 10^{-2}$	$1.33 \times 10^{-1}$	1.07	$1.55 \times 10^{-1}$	50 Hz, 3lb-in
$3.02 \times 10^{-4}$	$5.49 \times 10^{-2}$	$5.12 \times 10^{-2}$	$1.33 \times 10^{-2}$	$1.56 \times 10^{-1}$	$1.66 \times 10^{-1}$	50 Hz, 5lb-in
$4.25 \times 10^{-4}$	$6.51 \times 10^{-2}$	$1.12 \times 10^{-1}$	1.07	$6.86 \times 10^{-1}$	$4.19 \times 10^{-1}$	60 Hz, 1lb-in
$4.3 \times 10^{-4}$	$9.81 \times 10^{-2}$	$1.55 \times 10^{-1}$	$8.44 \times 10^{-1}$	$8.27 \times 10^{-1}$	$5.05 \times 10^{-1}$	60 Hz, 3lb-in
$3.69 \times 10^{-4}$	$6.92 \times 10^{-2}$	$1.84 \times 10^{-1}$	$8.77 \times 10^{-1}$	$8.35 \times 10^{-1}$	$4.47 \times 10^{-1}$	60 Hz, 5lb-in

**Tabla 2.6:** Características para un MI con desbalance y falla en baleros.

$c(i_a)$	$c(\ i_{dq}\ )$	$c_1(v)$	$c_2(v)$	$c_3(v)$	$c_4(v)$	Régimen
<b>MI con desbalance</b>						
$6.36 \times 10^{-5}$	$2.23 \times 10^{-3}$	$6.8 \times 10^{-2}$	$7.23 \times 10^{-3}$	$7.63 \times 10^{-2}$	$4.11 \times 10^{-3}$	30 Hz, 1lb-in
$2.9 \times 10^{-5}$	$2.2 \times 10^{-3}$	$6.98 \times 10^{-2}$	$6.35 \times 10^{-3}$	$6.96 \times 10^{-2}$	$5.43e^{-3}$	30 Hz, 3lb-in
$2.87 \times 10^{-5}$	$2.21 \times 10^{-3}$	$8.02 \times 10^{-2}$	$4.37 \times 10^{-3}$	$6.22 \times 10^{-2}$	$5.27 \times 10^{-3}$	30 Hz, 5lb-in
$1 \times 10^{-4}$	$1.2 \times 10^{-2}$	$6.16 \times 10^{-1}$	$4.28 \times 10^{-2}$	$1.57 \times 10^{-2}$	$8.88 \times 10^{-3}$	40 Hz, 1lb-in
$8.3 \times 10^{-5}$	$1.49 \times 10^{-2}$	$6.46 \times 10^{-1}$	$3.66 \times 10^{-2}$	$1.7 \times 10^{-2}$	$1.05 \times 10^{-2}$	40 Hz, 3lb-in
$1.14 \times 10^{-4}$	$9.17 \times 10^{-3}$	$6.66 \times 10^{-1}$	$3.9 \times 10^{-2}$	$1.6 \times 10^{-2}$	$1.04 \times 10^{-2}$	40 Hz, 5lb-in
$3.37 \times 10^{-4}$	$2.18 \times 10^{-2}$	$1.75 \times 10^{-1}$	$9.82 \times 10^{-2}$	$2.57 \times 10^{-2}$	$3.32 \times 10^{-2}$	50 Hz, 1lb-in
$2.87 \times 10^{-4}$	$1.66 \times 10^{-2}$	$1.68 \times 10^{-1}$	$7.83 \times 10^{-2}$	$2.93 \times 10^{-2}$	$3.6e^{-2}$	50 Hz, 3lb-in
$3.51 \times 10^{-4}$	$2 \times 10^{-2}$	$1.84 \times 10^{-1}$	$7.62 \times 10^{-2}$	$2.44 \times 10^{-2}$	$3.46 \times 10^{-2}$	50 Hz, 5lb-in
$6.31 \times 10^{-4}$	$5.87 \times 10^{-3}$	3.37	$4.02 \times 10^{-1}$	$2.21 \times 10^{-2}$	$3.89 \times 10^{-2}$	60 Hz, 1lb-in
$1.57 \times 10^{-4}$	$1.47 \times 10^{-2}$	3.23	$4.14 \times 10^{-1}$	$3.38 \times 10^{-2}$	$3.32 \times 10^{-2}$	60 Hz, 3lb-in
$7.74 \times 10^{-5}$	$1.08 \times 10^{-2}$	3.18	$4.87 \times 10^{-1}$	$3.32 \times 10^{-2}$	$3.4 \times 10^{-2}$	60 Hz, 5lb-in
<b>MI con falla en baleros</b>						
$2.31 \times 10^{-4}$	$5.91 \times 10^{-3}$	$1.5 \times 10^{-3}$	$4.49 \times 10^{-3}$	$2.46 \times 10^{-2}$	$5.16 \times 10^{-1}$	30 Hz, 1lb-in
$9.63 \times 10^{-5}$	$7.81 \times 10^{-3}$	$1.56 \times 10^{-3}$	$2.19 \times 10^{-3}$	$4.08e^{-2}$	$5.58 \times 10^{-1}$	30 Hz, 3lb-in
$8.28 \times 10^{-5}$	$9.04 \times 10^{-3}$	$1.38 \times 10^{-3}$	$2.47 \times 10^{-3}$	$4.3 \times 10^{-2}$	$4.55 \times 10^{-1}$	30 Hz, 5lb-in
$9.18 \times 10^{-5}$	$9.25 \times 10^{-3}$	$3.09 \times 10^{-2}$	$1.44 \times 10^{-2}$	$3.78 \times 10^{-2}$	1.36	40 Hz, 1lb-in
$6.81 \times 10^{-5}$	$5.84 \times 10^{-3}$	$3.19 \times 10^{-2}$	$1.47 \times 10^{-2}$	$4.65 \times 10^{-2}$	1.35	40 Hz, 3lb-in
$9.69 \times 10^{-5}$	$1.02 \times 10^{-2}$	$3.13 \times 10^{-2}$	$7.36 \times 10^{-3}$	$6.34 \times 10^{-2}$	1.07	40 Hz, 5lb-in
$5.04 \times 10^{-5}$	$7.99 \times 10^{-3}$	$2.74 \times 10^{-3}$	$1.41 \times 10^{-1}$	$4.41 \times 10^{-2}$	1.21	50 Hz, 1lb-in
$5.72 \times 10^{-5}$	$1.97 \times 10^{-2}$	$2.71 \times 10^{-2}$	$1.72 \times 10^{-1}$	$5.56 \times 10^{-2}$	1.49	50 Hz, 3lb-in
$2.97 \times 10^{-5}$	$6.41 \times 10^{-3}$	$2.36 \times 10^{-2}$	$1.33 \times 10^{-1}$	$6.34 \times 10^{-2}$	1.65	50 Hz, 5lb-in
$4.19 \times 10^{-5}$	$1.21 \times 10^{-2}$	$1.58 \times 10^{-1}$	$3.97 \times 10^{-2}$	$1.99 \times 10^{-2}$	2.59	60 Hz, 1lb-in
$7.92 \times 10^{-5}$	$1.67 \times 10^{-2}$	$1.6 \times 10^{-1}$	$4.24 \times 10^{-2}$	$2.71 \times 10^{-2}$	2.6	60 Hz, 3lb-in
$6.58 \times 10^{-5}$	$1.9 \times 10^{-2}$	$1.75 \times 10^{-1}$	$4.8 \times 10^{-2}$	$3 \times 10^{-2}$	2.11	60 Hz, 5lb-in



**Figura 2.20:** Mapa característico  $c_1(v)$  vs.  $c_3(v)$ .

Así mismo, la característica propia de una falla de baleros  $c_4(v)$  permite distinguir este tipo de falla de las demás, esto se ilustra mediante el mapa de características  $c_1(v)$  vs.  $c_4(v)$  mostrado en la Figura 2.21.

En las Figuras 2.22 y 2.23 se muestran los mapas que involucran las características de las mediciones eléctricas  $c(i_a)$  y  $c(\|i_{dq}\|)$  con  $c_1(v)$  respectivamente, en estas se puede observar que se puede lograr una separación de las fallas de desalineamiento y desbalance de los otros casos, ya que las características  $c(i_a)$  y  $c(\|i_{dq}\|)$  son propias de estas fallas. Sin embargo, se observa una mayor separación en el mapa  $c_1(v)$  vs.  $c(i_a)$ , que en el que involucra el módulo del vector de Park.

En las figuras subsecuentes se muestran las combinaciones restantes de las características, cabe hacer notar que los datos que aparecen separables en los mapas corresponden a las características representativas de las fallas, por ejemplo, en el mapa  $c_2(v)$  vs.  $c_3(v)$  (Figura 2.24) se ve claramente que una condición de desalineamiento es separable de las restantes ya que este mapa involucra la característica propia de la falla de desalineamiento  $c_3(v)$ .

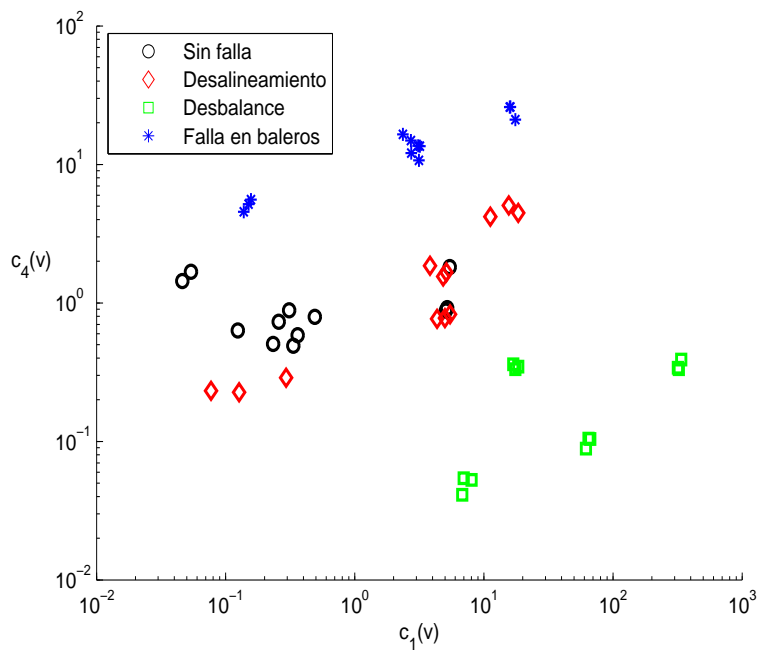


Figura 2.21: Mapa característico  $c_1(v)$  vs.  $c_4(v)$ .

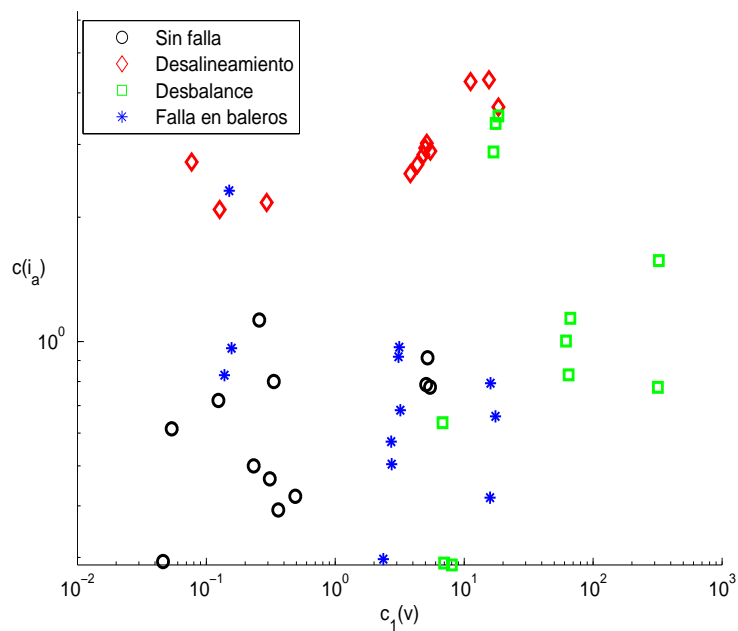
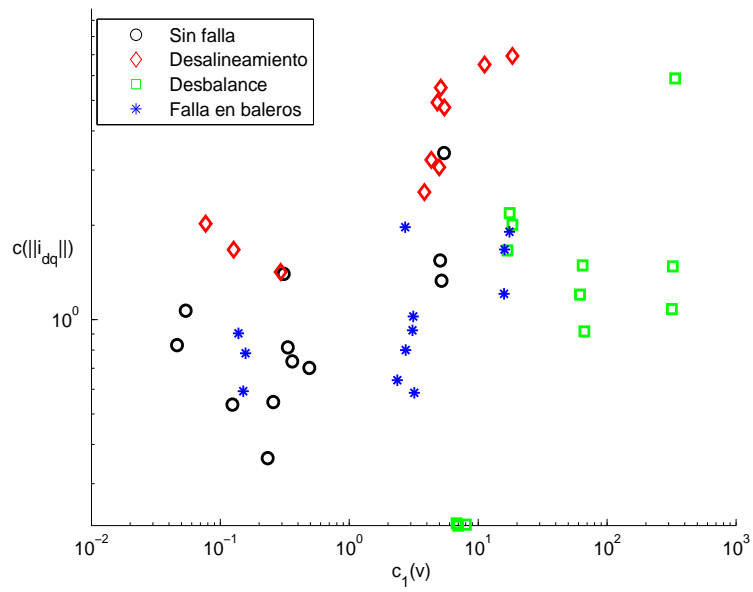
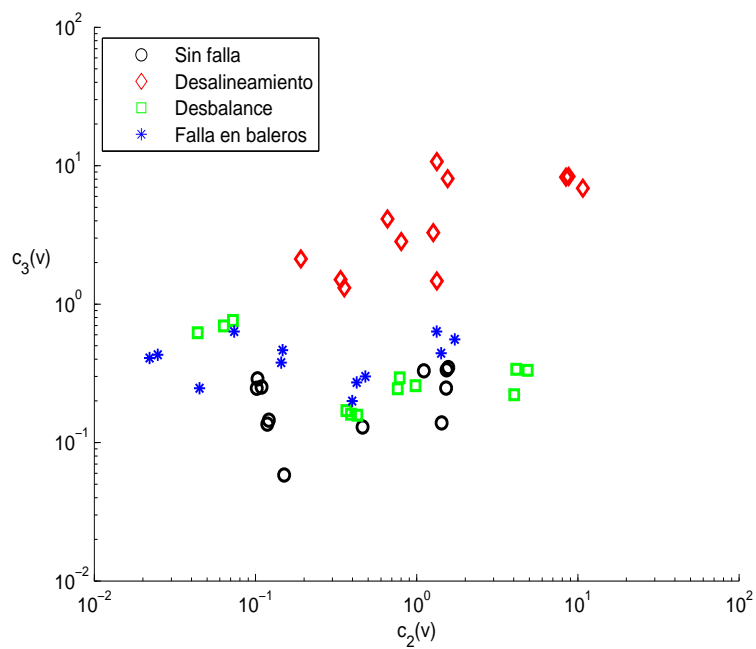


Figura 2.22: Mapa característico  $c_1(v)$  vs.  $c(i_a)$ .

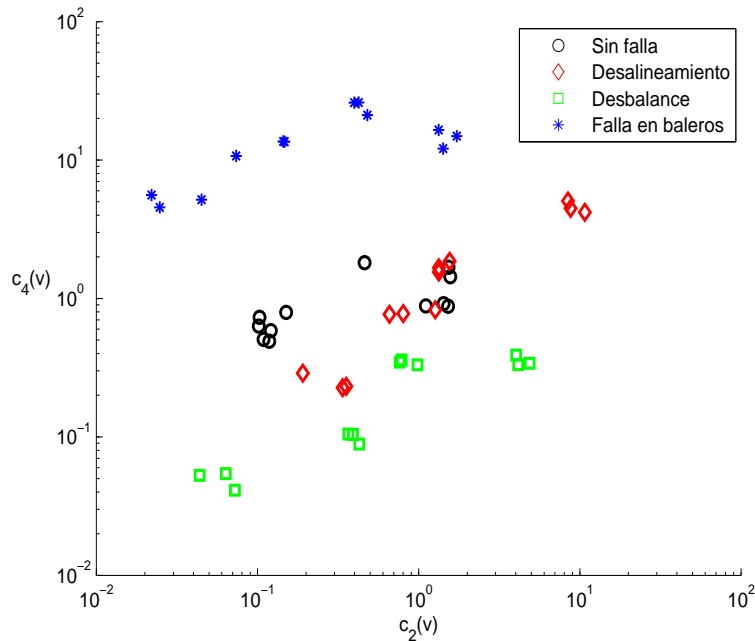




**Figura 2.23:** Mapa característico  $c_1(v)$  vs.  $c(\|i_{dq}\|)$ .



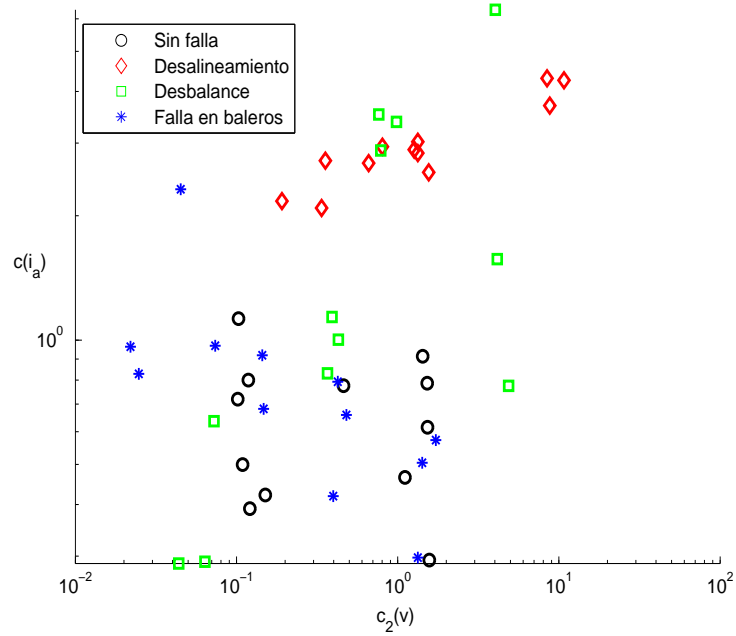
**Figura 2.24:** Mapa característico  $c_2(v)$  vs.  $c_3(v)$ .



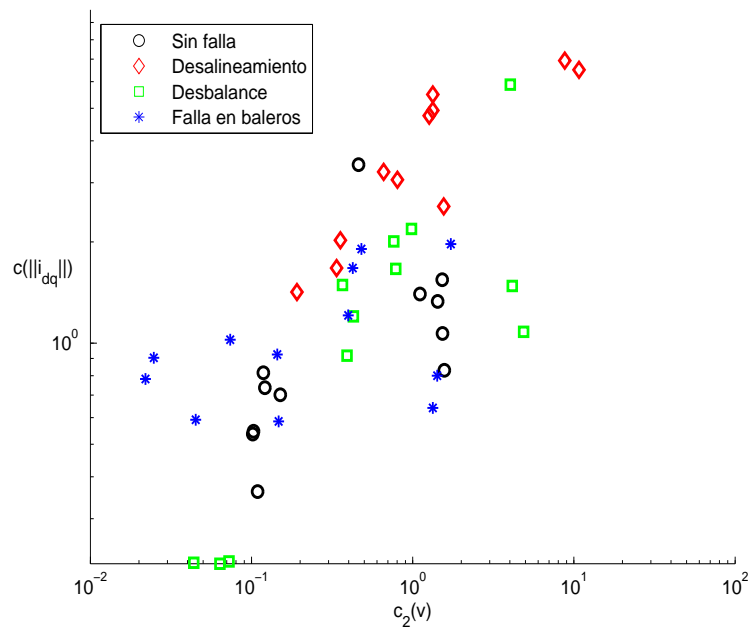
**Figura 2.25:** Mapa característico  $c_2(v)$  vs.  $c_4(v)$ .

De la misma manera se muestra la separabilidad de una falla de baleros en la Figura 2.25 a causa de la característica  $c_4(v)$ . Los mapas de características  $c_2(v)$  vs.  $c(i_a)$  y  $c_2(v)$  vs.  $c(\|i_{dq}\|)$  mostrados en las Figuras 2.26 y 2.27 pueden ser útiles para separar una condición de desalineamiento o desbalance de las restantes. Sin embargo, se hace necesaria la información proporcionada por las señales de vibración para poder separar entre sí a estas dos condiciones de falla, tal como se muestra en la gráfica  $c_3(v)$  vs.  $c_4(v)$  de la Figura 2.28.

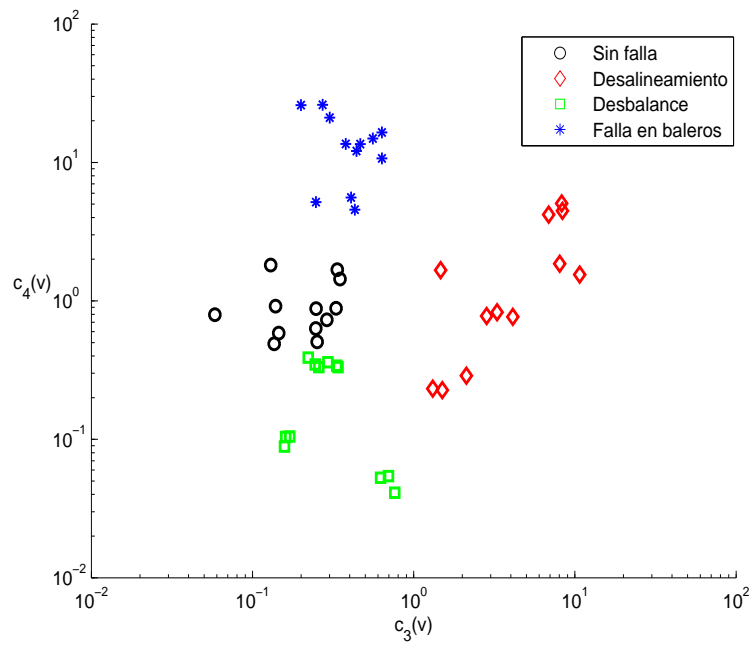
La conclusión anterior se puede verificar en los 5 mapas de características restantes, donde es posible lograr una separación de las fallas de desalineamiento y desbalance, sin embargo, puede ser complicado mediante los medios eléctricos de diagnóstico separar una de estas fallas de la otra pues producen efectos similares en las corrientes de línea del MI.



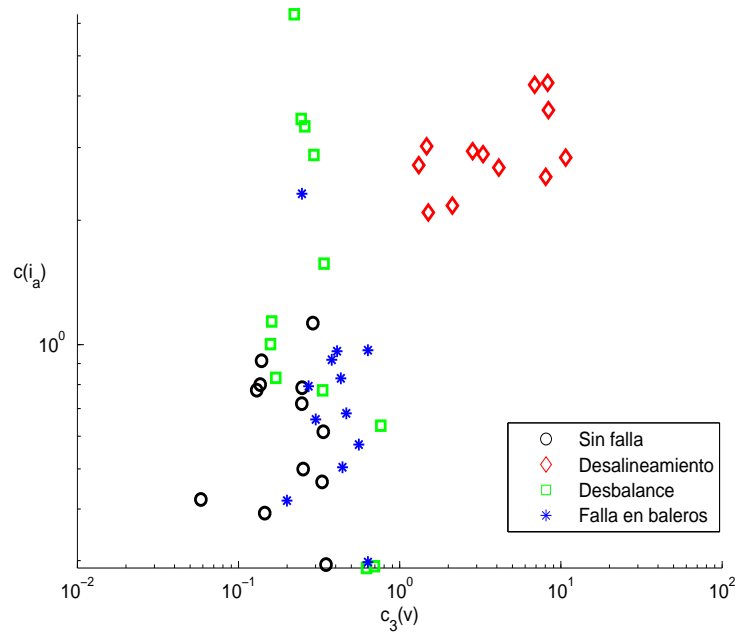
**Figura 2.26:** Mapa característico  $c_2(v)$  vs.  $c(i_a)$ .



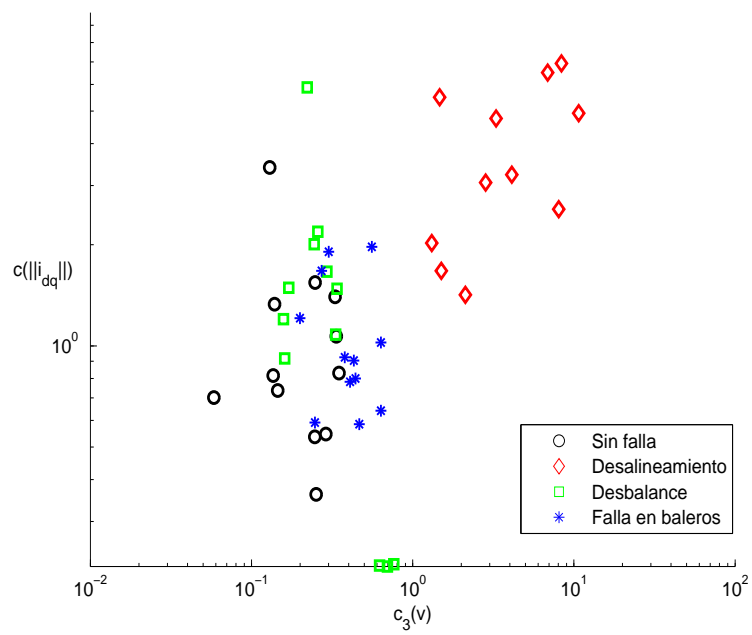
**Figura 2.27:** Mapa característico  $c_2(v)$  vs.  $c(\|i_{dq}\|)$ .



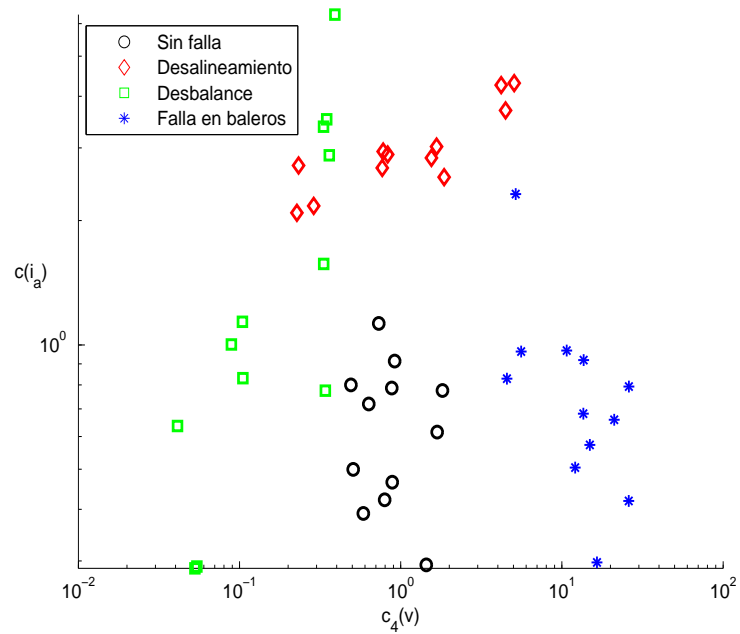
**Figura 2.28:** Mapa característico  $c_3(v)$  vs.  $c_4(v)$ .



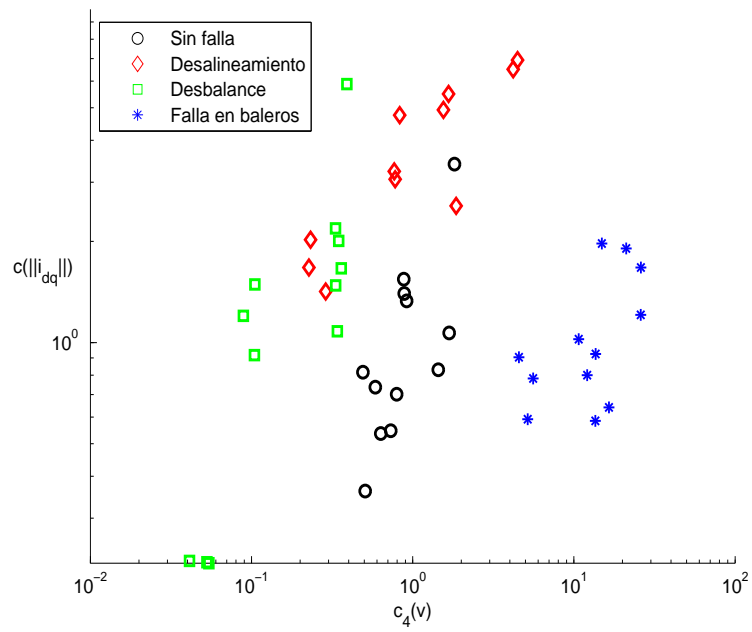
**Figura 2.29:** Mapa característico  $c_3(v)$  vs.  $c(i_a)$ .



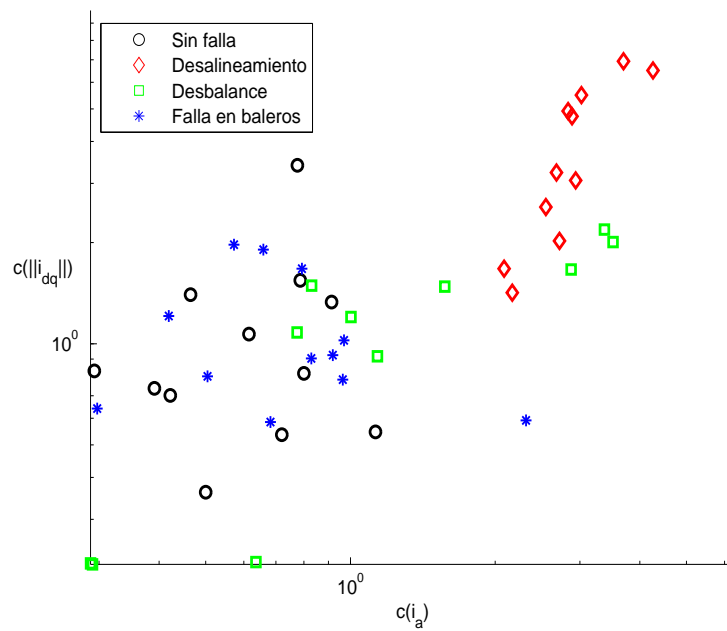
**Figura 2.30:** Mapa característico  $c_3(v)$  vs.  $c(\|i_{dq}\|)$ .



**Figura 2.31:** Mapa característico  $c_4(v)$  vs.  $c(i_a)$ .



**Figura 2.32:** Mapa característico  $c_4(v)$  vs.  $c(\|i_{dq}\|)$ .



**Figura 2.33:** Mapa característico  $c(i_a)$  vs.  $c(\|i_{dq}\|)$ .

---

# DISEÑO E IMPLEMENTACIÓN DE UNA MÁQUINA DE SOPORTE VECTORIAL

---

### 3.1. Selección del modelo óptimo

Al proceso de determinar el clasificador óptimo como aquel que tiene una capacidad de generalización mayor<sup>1</sup>, se le conoce como selección del modelo, en el cual para una SVM no lineal se busca un kernel y valores apropiados de sus parámetros. Para controlar la capacidad de generalización, se debe minimizar la dimensión VC,  $h$ , la cual está acotada por [28]:

$$h \leq \min(R^2 \|w\|^2, l) + 1,$$

donde  $R$  es el diámetro de la esfera más pequeña que incluye a todos los datos de entrenamiento en el espacio característico, y  $\|w\|$  es la norma euclidiana del vector de pesos, es decir, se debe construir el hiperplano que minimiza  $R^2 \|w\|^2$ . Para encontrar  $R^2$  se debe resolver el siguiente

---

<sup>1</sup>La capacidad de generalización consiste en generar predicciones correctas para patrones de entrada nuevos.

problema de optimización:

$$\begin{aligned} & \text{mín } R^2 \\ & \text{t. q. } \|\phi(x_i) - a\|^2 \leq R^2, \quad \text{para } i = 1, \dots, l, \end{aligned}$$

donde  $a$  es el centro de la esfera en el espacio característico. Para resolver este problema se plantea el problema de optimización dual, introduciendo el Lagrangiano

$$\mathcal{L}(R, a, \beta) = R^2 - \sum_{i=1}^l \beta_i (R^2 - (\phi(x_i) - a)^2),$$

de la condición de optimalidad dada en la proposición (2.1), se obtiene

$$\frac{\partial \mathcal{L}(R, a, \beta)}{\partial R} = 1 - \sum_{i=1}^l \beta_i = 0,$$

y

$$\frac{\partial \mathcal{L}(R, a, \beta)}{\partial a} = a - \sum_{i=1}^l \beta_i \phi(x_i) = 0. \quad (3.1)$$

De lo anterior resulta el problema dual siguiente

$$\begin{aligned} & \text{máx}_{\beta} \sum_{i=1}^l \beta_i K(x_i, x_i) - \sum_{i,j=1}^l \beta_i \beta_j K(x_i, x_j) \\ & \text{t. q. } \sum_{i=1}^l \beta_i = 1, \beta_i \geq 0 \quad \text{para } i = 1, \dots, l, \end{aligned}$$

donde  $\beta$  es el vector de multiplicadores de Lagrange. Para cada punto  $x$  en el espacio de entrada, la distancia entre  $\phi(x)$  y el centro de la esfera se obtiene a partir de

$$\begin{aligned} \text{máx } R^2(x) &= \text{máx } \|\phi(x) - a\|^2 \\ &= \text{máx } K(x, x) - 2 \sum_{j=1}^l \beta_j K(x_j, x) + \sum_{i,j=1}^l \beta_i \beta_j K(x_i, x_j) \end{aligned}$$

para cada patrón de entrada  $x$ , donde se usó (4.1) y la definición de kernel.

Por otra parte, es posible calcular  $\|w\|^2$  mediante

$$\|w\|^2 = \sum_{i,j=1}^l \alpha_i \alpha_j K(x_i, x_j) y_i y_j$$



De (1.5) se desprende que cuando  $h$  disminuye también lo hace el intervalo de confianza  $\Gamma$ . Por lo tanto, para una SVM que clasifica correctamente los datos de entrenamiento, es decir,  $R_{emp}(f) = 0$ , se deben escoger los parámetros que minimicen  $\Gamma$ . Para esto, se estima  $h$  mediante  $h_{est} = R^2 \|w\|^2 + 1$  o  $h_{est} = l + 1$ . Una vez establecido el criterio de selección de las SVM's, se procede a calcular  $h_{est}$  y el riesgo estructural  $R(f)$  que proporcionan una medida de la capacidad de generalización.

En las siguientes Tablas se muestran los valores obtenidos, para las siguientes funciones kernel, y valores en sus parámetros ( $\eta = 0.1$ ):

*Kernel Polinomial*

$$K(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^p.$$

para  $p = 2, 3, 4$ .

*Kernel de Función Gaussiana de base radial*

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right).$$

para  $\sigma = 1, 2, \dots, 10$ .

Cabe señalar que el algoritmo de multclasificación está basado en SVM uno vs. uno, para el cual se deben construir 6 SVM's, es decir, el problema de multclasificación se divide en 6 problemas de clasificación binaria.

En la Tabla 3.1, 3.2 y 3.3 se muestran los datos obtenidos para las SVM's que clasifican las fallas de desalineamiento, desbalance y baleros respectivamente de un MI sin falla, mientras que en las Tablas 3.4, 3.5 y 3.6 se muestran los valores obtenidos para clasificar los tres casos de falla estudiados entre sí.

De las tablas se observa que se alcanza una dimensión VC,  $h$  menor para un kernel RBF, por lo cual, resulta conveniente implementar una SVM-multiclase basada en este tipo de kernel para alcanzar una mayor capacidad de generalización. Para la implementación se escogió el valor de  $\sigma$  para el cual se obtuvo el valor más pequeño de  $h_{est}$  para cada una de las SVM.

**Tabla 3.1:** Selección de SVM para clasificar un MI sin falla ( $y = 1$ ) y un MI con desalineamiento ( $y = -1$ )

Kernel Polinomial	$\ w\ ^2$	$h_{est}$	$\Gamma$
$p = 2$	0.5593	25	2.358
$p = 3$	0.23189	25	2.358
$p = 4$	0.1154108	25	2.358
Kernel RBF	$\ w\ ^2$	$h_{est}$	$\Gamma$
$\sigma = 1$	13.2839	12.2638	1.7587
<b><math>\sigma = 2</math></b>	<b>13.4378</b>	<b>11.8853</b>	<b>1.7358</b>
$\sigma = 3$	17.095	14.2475	1.87
$\sigma = 4$	22.66017	17.7254	2.049
$\sigma = 5$	30.1542	22.103	2.2433
$\sigma = 6$	40.472	25	2.358
$\sigma = 7$	53.6182	25	2.358
$\sigma = 8$	68.9579	25	2.358
$\sigma = 9$	86.4286	25	2.358
$\sigma = 10$	105.912	25	2.358

**Tabla 3.2:** Selección de SVM para clasificar un MI sin falla ( $y = 1$ ) y un MI con desbalance ( $y = -1$ )

Kernel Polinomial	$\ w\ ^2$	$h_{est}$	$\Gamma$
$p = 2$	0.3284	25	2.358
$p = 3$	0.1818	25	2.358
$p = 4$	0.0948	25	2.358
Kernel RBF	$\ w\ ^2$	$h_{est}$	$\Gamma$
$\sigma = 1$	12.628	11.621	1.719
<b><math>\sigma = 2</math></b>	<b>10.925</b>	<b>9.752</b>	<b>1.597</b>
$\sigma = 3$	12.054	10.497	1.647
$\sigma = 4$	15.108	12.851	1.793
$\sigma = 5$	19.725	16.45	1.987
$\sigma = 6$	25.504	20.917	2.193
$\sigma = 7$	32.707	25	2.358
$\sigma = 8$	41.307	25	2.358
$\sigma = 9$	51.096	25	2.358
$\sigma = 10$	62.06	25	2.358

**Tabla 3.3:** Selección de SVM para clasificar un MI sin falla ( $y = 1$ ) y un MI con falla en baleros ( $y = -1$ )

Kernel Polinomial	$\ w\ ^2$	$h_{est}$	$\Gamma$
$p = 2$	0.4208	25	2.358
$p = 3$	0.2	25	2.358
$p = 4$	0.106819	25	2.358
Kernel RBF	$\ w\ ^2$	$h_{est}$	$\Gamma$
$\sigma = 1$	11.91	10.915	1.675
$\sigma = 2$	8.52	7.425	1.424
<b><math>\sigma = 3</math></b>	<b>8.571</b>	<b>7.074</b>	<b>1.395</b>
$\sigma = 4$	10.164	7.902	1.462
$\sigma = 5$	12.968	9.546	1.583
$\sigma = 6$	16.525	11.59	1.717
$\sigma = 7$	21.22	14.2	1.869
$\sigma = 8$	26.79	17.04	2.016
$\sigma = 9$	33.06	19.87	2.148
$\sigma = 10$	40.041	22.7	2.267

**Tabla 3.4:** Selección de SVM para clasificar un MI con desalineamiento ( $y = 1$ ) y un MI con desbalance ( $y = -1$ )

Kernel Polinomial	$\ w\ ^2$	$h_{est}$	$\Gamma$
$p = 2$	$9.516e^{-3}$	25	2.358
Kernel RBF	$\ w\ ^2$	$h_{est}$	$\Gamma$
$\sigma = 1$	18.117	17.117	2.0198
$\sigma = 2$	14.339	13.332	1.821
$\sigma = 3$	12.368	11.298	1.699
<b><math>\sigma = 4</math></b>	<b>11.729</b>	<b>10.512</b>	<b>1.648</b>
$\sigma = 5$	12.047	10.592	1.654
$\sigma = 6$	13.082	11.277	1.698
$\sigma = 7$	14.806	12.511	1.773
$\sigma = 8$	17.056	14.138	1.866
$\sigma = 9$	19.731	16.072	1.967
$\sigma = 10$	22.812	18.286	2.075

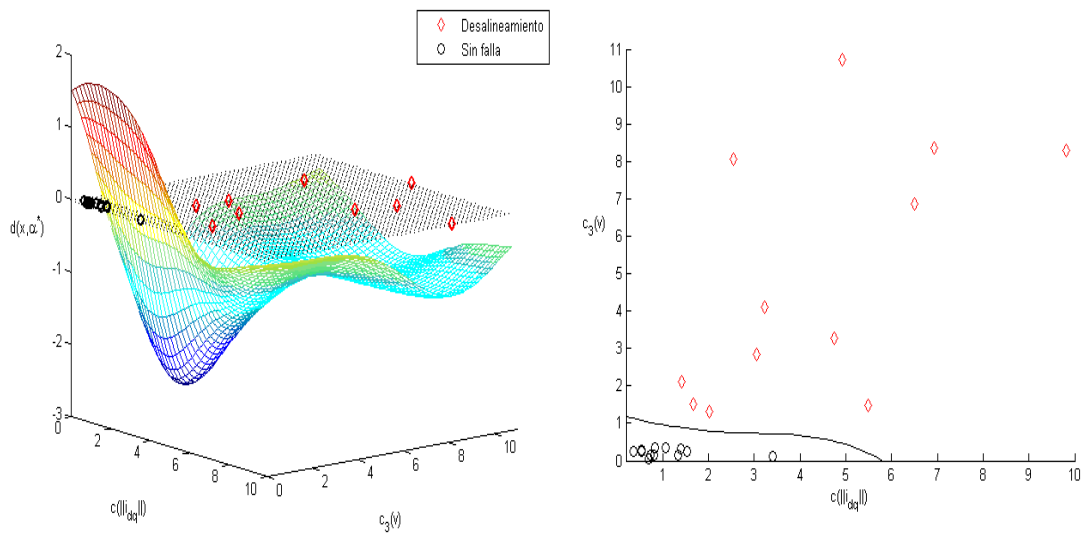
**Tabla 3.5:** Selección de SVM para clasificar un MI con desalineamiento ( $y = 1$ ) y un MI con falla en baleros ( $y = -1$ )

Kernel Polinomial	$\ w\ ^2$	$h_{est}$	$\Gamma$
$p = 2$	0.0121	25	2.358
$p = 3$	0.000946	25	2.358
$p = 4$	0.00008316	25	2.358
Kernel RBF	$\ w\ ^2$	$h_{est}$	$\Gamma$
$\sigma = 1$	17.44	16.44	1.986
$\sigma = 2$	12.62	11.6	1.718
$\sigma = 3$	10.48	9.36	1.5707
<b><math>\sigma = 4</math></b>	<b>9.787</b>	<b>8.45</b>	<b>1.5044</b>
$\sigma = 5$	10.153	8.476	1.5061
$\sigma = 6$	11.042	8.899	1.537
$\sigma = 7$	12.57	9.75	1.598
$\sigma = 8$	14.9	11.105	1.687
$\sigma = 9$	17.69	12.605	1.779
$\sigma = 10$	20.81	14.125	1.865

**Tabla 3.6:** Selección de SVM para clasificar un MI con falla en baleros ( $y = 1$ ) y desbalance ( $y = -1$ )

Kernel Polinomial	$\ w\ ^2$	$h_{est}$	$\Gamma$
$p = 2$	0.00226	25	2.358
$p = 3$	$8.629e^{-5}$	25	2.358
Kernel RBF	$\ w\ ^2$	$h_{est}$	$\Gamma$
$\sigma = 1$	16.897	15.897	1.959
$\sigma = 2$	12.606	11.606	1.718
$\sigma = 3$	10.776	9.77	1.598
$\sigma = 4$	9.71	8.665	1.5202
$\sigma = 5$	9.287	8.168	1.482
$\sigma = 6$	9.151	7.94	1.465
<b><math>\sigma = 7</math></b>	<b>9.176</b>	<b>7.854</b>	<b>1.458</b>
$\sigma = 8$	9.457	7.982	1.468
$\sigma = 9$	10.08	8.399	1.5
$\sigma = 10$	11.087	9.108	1.552

En la Figura 3.1 se muestra la clasificación binaria de los datos de un MI sin falla y un MI con desalineamiento a través de las características  $c_3(v)$  y  $c(\|i_{dq}\|)$ , donde, al aplicar la función signo a la superficie de decisión  $d(x, \alpha^*)$  es posible obtener un sólo valor para la falla de desalineamiento ( $y = -1$ ). En la Figura 3.2 se muestra la clasificación entre los datos de un MI desalineado y un MI con desbalance, a través de los medios de diagnóstico eléctricos  $c(i_a)$  y  $c(\|i_{dq}\|)$ , en esta se observa que aún cuando los datos son no linealmente separables, no se presentan errores de clasificación. Observe que la intersección de la función de decisión  $d(x, \alpha^*)$  con el plano de características forma el hiperplano óptimo.

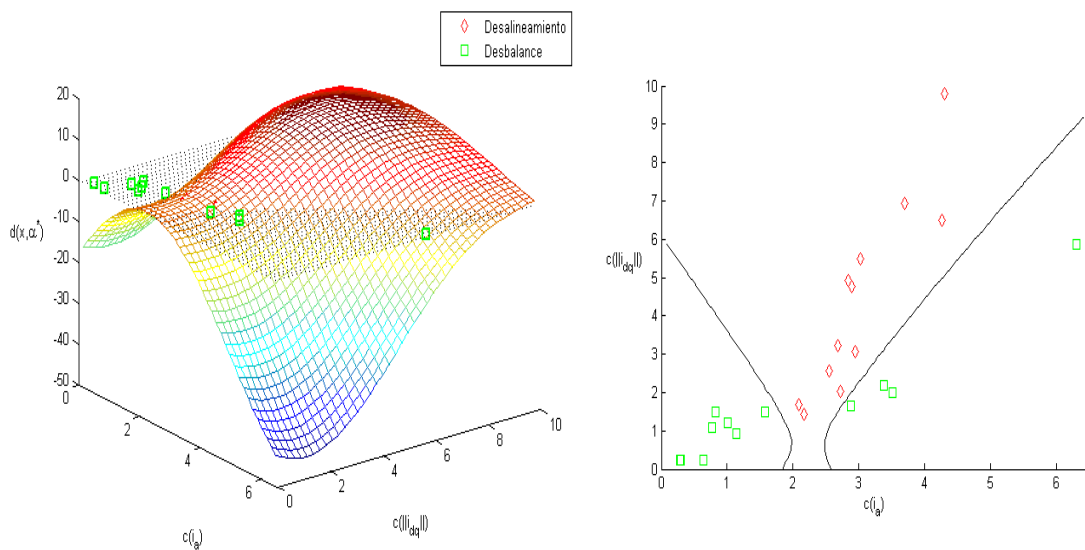


**Figura 3.1:** Clasificación binaria (MI sin falla vs. MI desalineado) con kernel gaussiano y  $\sigma = 2$ .

## 3.2. Implementación del algoritmo de diagnóstico mediante LabView

En la Figura 3.3 se muestra el esquema general de diagnóstico de fallas propuesto, en este se señala la medición de las corrientes de línea del MI, la vibración mecánica y la velocidad





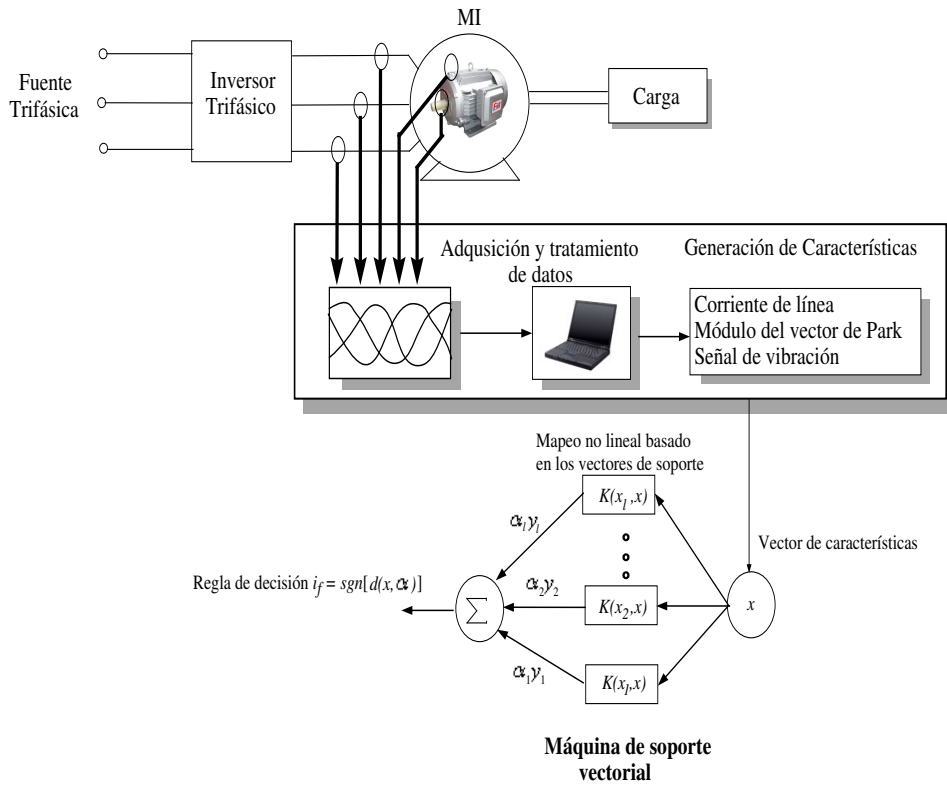
**Figura 3.2:** Clasificación binaria (MI desalineado vs. MI desbalanceado) con kernel gaussiano y  $\sigma = 4$ .

del rotor. También se muestra la estructura de una SVM para clasificación binaria, en la cual el patrón de entrada es mapeado a un espacio característico mediante la función kernel, el cual es multiplicado por las salidas deseadas y multiplicadores de Lagrange  $\alpha$  obtenidos de los datos de entrenamiento. Posteriormente se aplica la regla de decisión (función signo) que proporciona la predicción de una condición específica del MI.

En la Figura 3.4 se muestra el programa en LabView de una SVM que separa los datos de una falla de desalineamiento y un MI sin falla, esto se realiza para cada una de las 6 combinaciones de casos que se deben clasificar.

Por otra parte, en la Figura 3.5 se muestra el panel de control en LabView que nos indica una condición de desbalance del MI, así como la PSD de la señal de vibración.

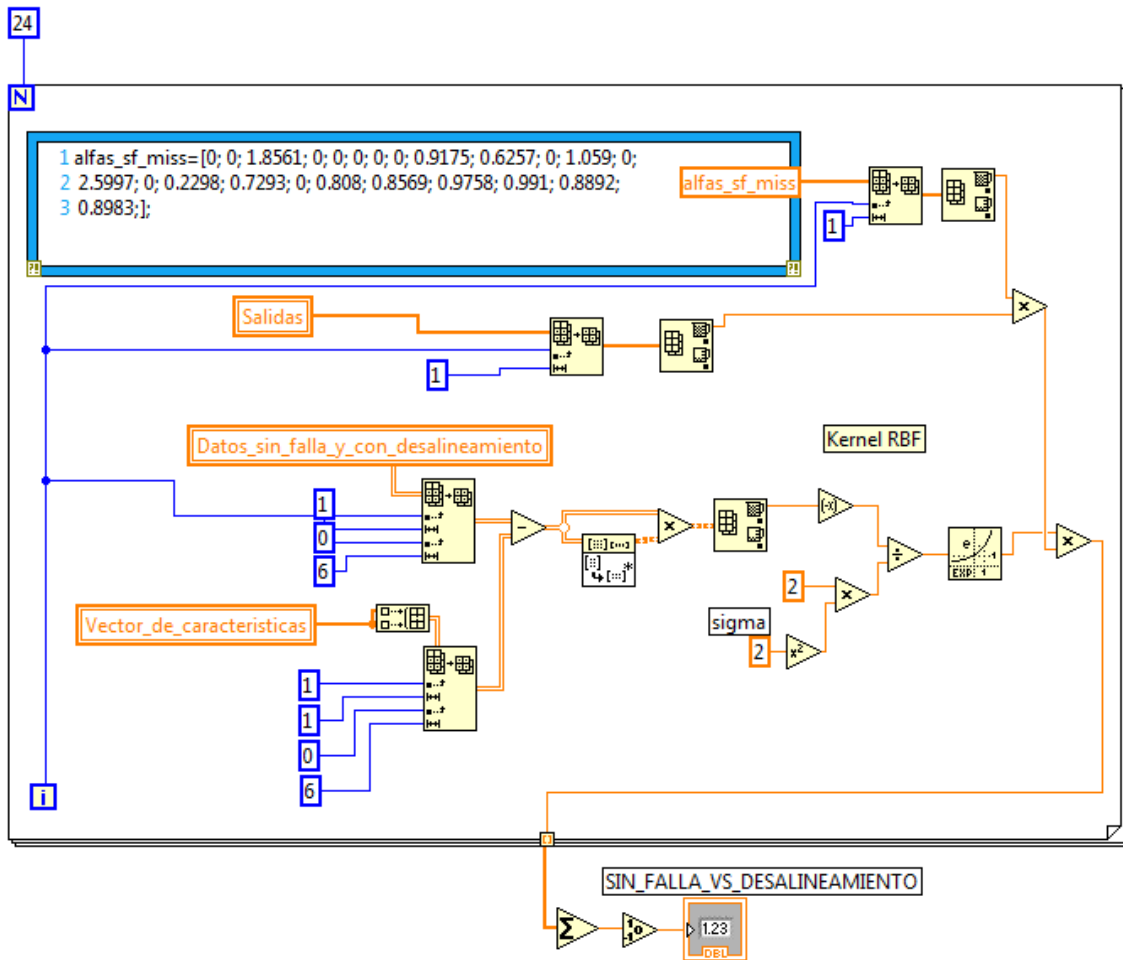
Los resultados de la etapa de prueba del esquema de diagnóstico propuesto se muestran en la Tabla 3.7, en esta se indica la condición de par de carga y los porcentajes de clasificación correcta. El algoritmo fue evaluado para las frecuencias de alimentación de 30, 35, 40, 45, 50, 55 y 60 Hz para cada nivel de par de carga presentado. Los resultados indican que las vibraciones mecánicas en conjunto con los medios de diagnóstico de la corriente de línea y el módulo del vector de Park



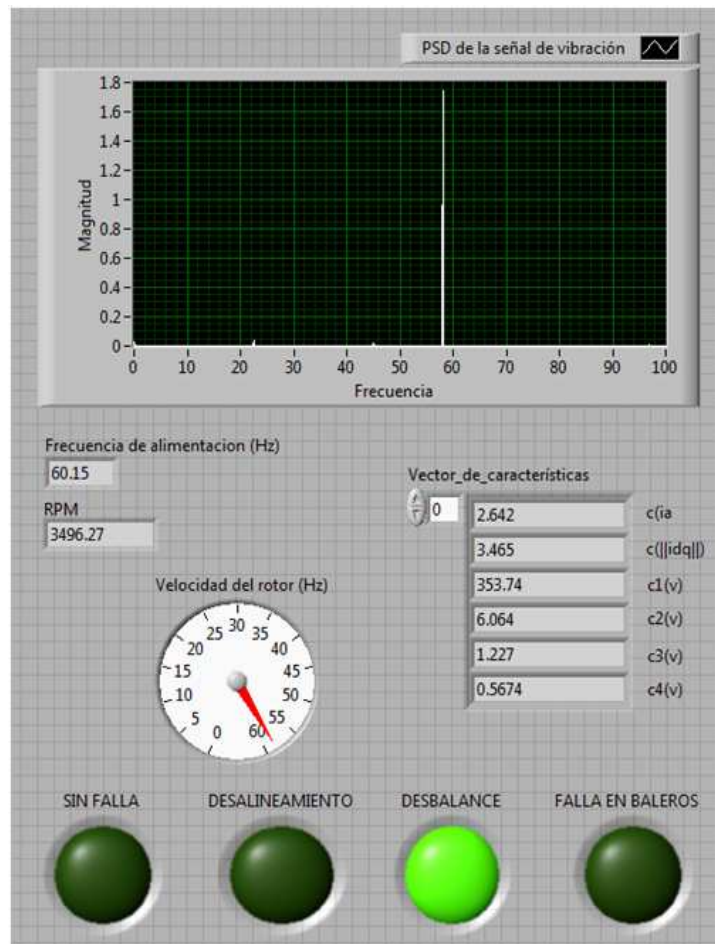
**Figura 3.3:** Esquema general de diagnóstico de fallas mediante SVM.

**Tabla 3.7:** Porcentajes de clasificación correcta de la fase de prueba del esquema de diagnóstico

% / Par de carga (lb-in)	1.5	2	3.5	4	5	Total
Sin Falla	100	100	100	100	100	100
Falla en baleros	100	100	100	100	100	100
Falla de desbalance	100	100	100	100	100	100
Falla de desalineamiento	100	85.7	100	100	100	97.14



**Figura 3.4:** Implementación de una SVM para clasificación binaria de un MI sin falla y con desalineamiento.



**Figura 3.5:** Panel de control que indica la presencia de una falla.

son buenos indicadores de la presencia de los casos de falla estudiados. Se observó un error de clasificación en la condición de operación de 2 lb-in a una frecuencia de alimentación de 45 Hz, se indicó un MI sin falla cuando en realidad estaba presente un MI con desalineamiento. Esto se debe a que las características extraídas para este caso (ver Tabla 3.8) son muy similares a las de una condición sin falla. Al no contar con un dispositivo para alinear correctamente al MI sin falla,

**Tabla 3.8:** Características para un MI con desalineamiento.

$c(i_a)$	$c(\ i_{dq}\ )$	$c_1(v)$	$c_2(v)$	$c_3(v)$	$c_4(v)$	Régimen
$1.7 \times 10^{-4}$	$1.15 \times 10^{-2}$	$2.95 \times 10^{-2}$	$1.69 \times 10^{-1}$	$1.54 \times 10^{-1}$	$9.27 \times 10^{-2}$	45 Hz, 2lb-in

es posible que algunos datos fueran extraídos con cierto grado de desalineamiento, en el ámbito industrial se utilizan dispositivos láser para alinear este tipo de máquinas. A pesar de que existe una diferencia entre los valores de las características  $c(i_a)$  y  $c_3(v)$  de la Tabla 3.8 y los valores de un MI sin falla, no ocurre lo mismo para el resto de las características. Sin embargo, en general el algoritmo de diagnóstico presenta un buen desempeño para diferentes regímenes de operación del MI. Cabe mencionar que ninguna de las condiciones para las que fue evaluado el algoritmo fue usada para realizar el entrenamiento de las SVM, es decir, a partir de un número determinado de condiciones de entrenamiento, fue posible generalizar para el resto de las condiciones.

---

## CONCLUSIONES

---

### **Trabajo Realizado**

En este trabajo se realiza el diagnóstico de fallas de desalineamiento, desbalance y baleros en un MI bajo condiciones de velocidad y par de carga variable. El trabajo realizado a lo largo de la tesis se resume en lo siguiente:

- a) Revisión bibliográfica, en la cual, se obtuvo el conocimiento necesario sobre las fallas de desalineamiento, desbalance y baleros en un MI, utilizando medios de diagnóstico eléctricos y mecánicos.
- b) Estudio del funcionamiento de las máquinas de aprendizaje, en especial de las máquinas de soporte vectorial como clasificadores.
- c) Implementación de una etapa de instrumentación para adquirir las señales de vibración mecánica del MI.
- d) Reproducción y análisis espectral de los datos para cada tipo de falla a estudiar del MI bajo diferentes regímenes de operación, mediante el kit de pruebas “Machine Fault Simulator” de Spectra Quest.

- e) Implementación en LabView del algoritmo de diagnóstico de fallas basado en máquinas de soporte vectorial.
- f) Evaluación del desempeño del algoritmo de diagnóstico para condiciones de velocidad y par de carga variable.

## Conclusiones

Los porcentajes de clasificación correcta para los casos de falla estudiados fueron altos, solamente ocurrió un error en la clasificación de un MI con problema de desalineamiento, sin embargo, no se registraron falsas alarmas. Es posible concluir que el análisis espectral de las corrientes de estator, módulo del vector de Park y vibraciones mecánicas constituyen medios de diagnóstico confiables para llevar a cabo el diagnóstico de fallas del tipo incipiente en MI.

Se observó que conforme se incrementa la velocidad mecánica del rotor, los datos que indican una condición de falla se encuentran más separados entre sí. Además, para una falla incipiente de baleros, no se encontraron componentes en el espectro de la corriente de estator, siendo necesaria la inclusión de mediciones mecánicas para diagnosticar este tipo de fallas.

Para la selección de las SVM se utilizó el principio de minimización de riesgo estructural, el cual nos proporciona una medida de la capacidad de generalización de estas máquinas de aprendizaje. Por último, el diseño del algoritmo de diagnóstico basado en SVM se llevó a cabo en la plataforma de LabView, este algoritmo tarda aproximadamente 25 segundos en generar las características a partir de los medios de diagnóstico eléctricos y mecánicos, y procesarlas a través de las máquinas de soporte vectorial para diagnosticar una falla en el MI.

Las herramientas de clasificación basadas en datos son especialmente apropiadas para diagnosticar MI que son fabricados en grandes cantidades.

Con este algoritmo de diagnóstico es posible predecir una falla de baleros, la cual ocupa cerca del 50 % de las fallas presentadas en un MI, además de las fallas mecánicas relacionadas con el rotor como el desalineamiento y desbalance que ocupan cerca del 10 % de las fallas en un MI. Por

lo tanto, es posible reducir los costos por mantenimiento del mismo, y a su vez puede ser usado por un experto como herramienta complementaria para realizar un diagnóstico más preciso.

La implementación del sistema de diagnóstico utilizando la tarjeta de *National Instruments*, NI-Compaq-DAQ-9172, es una buena opción ya que permite la medición de señales en un rango amplio que va desde  $\pm 200$  mV a  $\pm 10$  V, esto es ideal para el caso en que las señales de vibración presenten mayor amplitud. Por otra parte, el lenguaje de programación gráfico de Lab-View permite trabajar con programas de otra área de aplicación como Matlab, y además cuenta con numerosas herramientas de presentación como gráficas e indicadores que son complicados de realizar en plataformas como C++. Sin embargo, debido a que se requiere la medición de 5 señales durante un tiempo de 20 segundos, se debe disponer de una computadora con suficiente memoria RAM para el almacenamiento temporal de los datos y así obtener una ejecución eficiente del algoritmo.

## Trabajo Futuro

Aunque las fallas de rodamientos constituyen un gran porcentaje de los problemas presentes en un MI, las fallas ocasionadas en los devanados del estator también ocurren de manera significativa en estas máquinas, por lo tanto, es importante considerar el problema de diagnóstico de este tipo de fallas.

También resulta interesante realizar un estudio comparativo de la técnica presentada en este trabajo y técnicas de diagnóstico basadas en modelos, evaluando el desempeño de los esquemas en un escenario real de fallas del tipo incipiente en un MI, es decir, bajo diferentes regímenes de operación, con el fin de proporcionar la solución más confiable.

Además, es posible considerar el problema de diagnóstico de fallas en los actuadores y sensores de un MI, para el cual, uno de los principales retos es conseguir un tiempo de detección pequeño.

Por último, la implementación del esquema de diagnóstico propuesto para fallas en MI, en una plataforma digital (DSP) que no dependa de una computadora también representa un reto.



---

### Implementación en LabView del algoritmo de diagnóstico

---

En las siguientes figuras se muestra la implementación en Labview del esquema de diagnóstico basado en máquinas de soporte vectorial. En la Figura A.1 se muestra el programa para adquirir las señales mediante la tarjeta de adquisición de datos NI-Compaq-DAQ 9172.

En seguida se generan las características a partir de la PSD de las señales de corriente, módulo del vector de Park y vibraciones, mediante el programa mostrado en la Figura A.2. Se construye el vector de características mediante el programa de la Figura A.3, este vector, será procesado por las SVM's.

En la Figura A.4 se muestran los datos de entrenamiento que son utilizados para formar la función de decisión en el mapa de características, para esto se utilizó la herramienta *mathscript*, en la que se pueden ejecutar comandos de Matlab.

Una vez obtenido el resultado de las 6 SVM's se lleva a cabo un esquema de votación mostrado en la Figura A.5, y se determina por mayoría el tipo de falla presente en el MI.

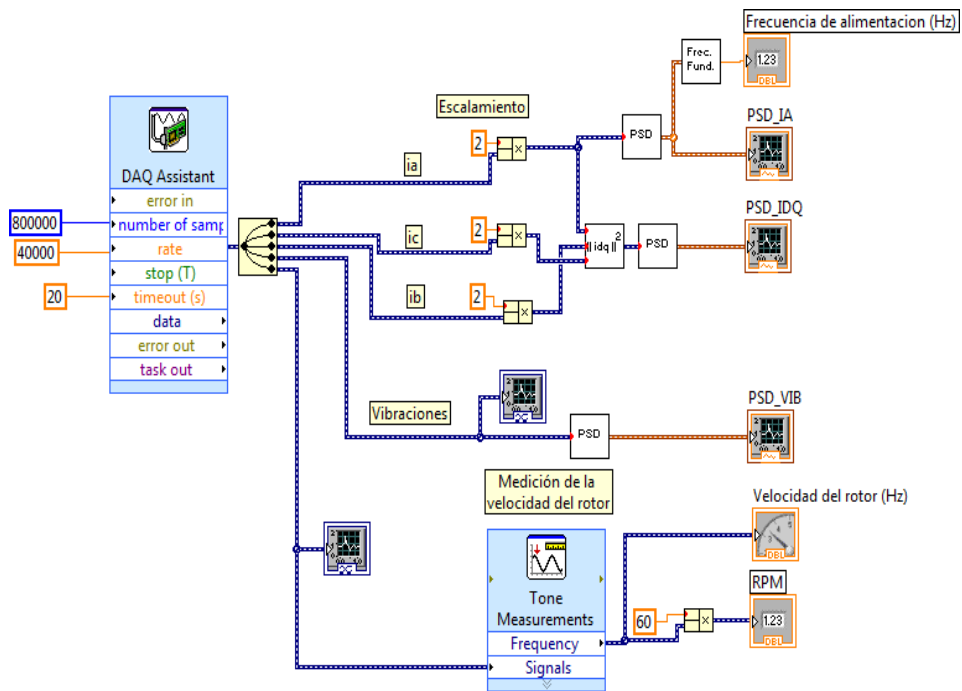
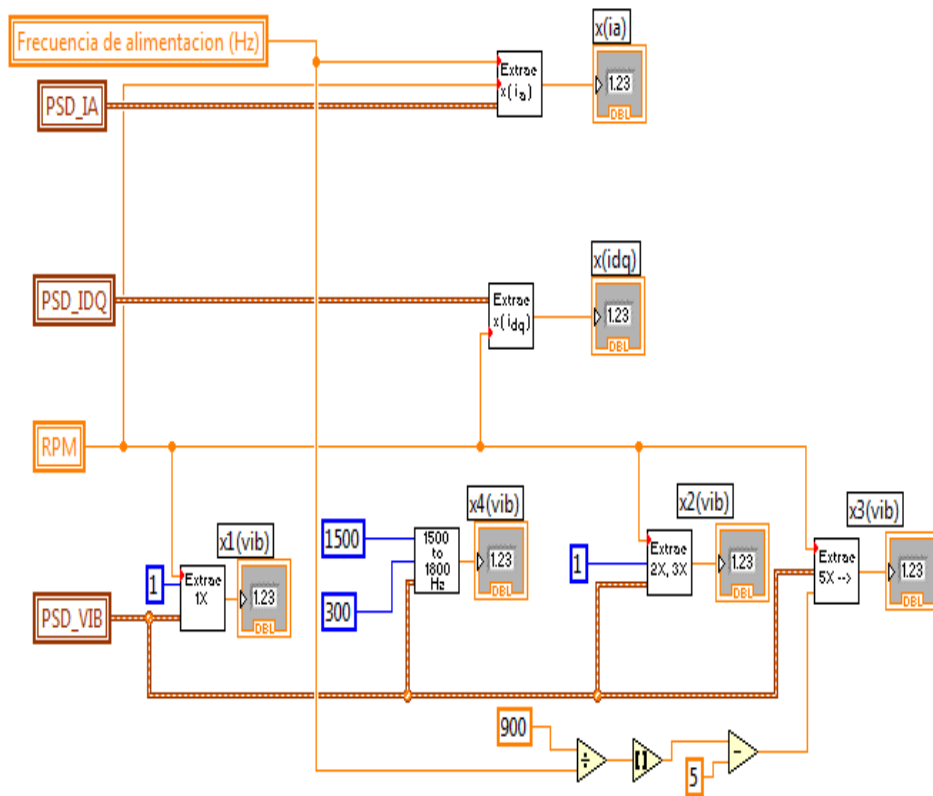
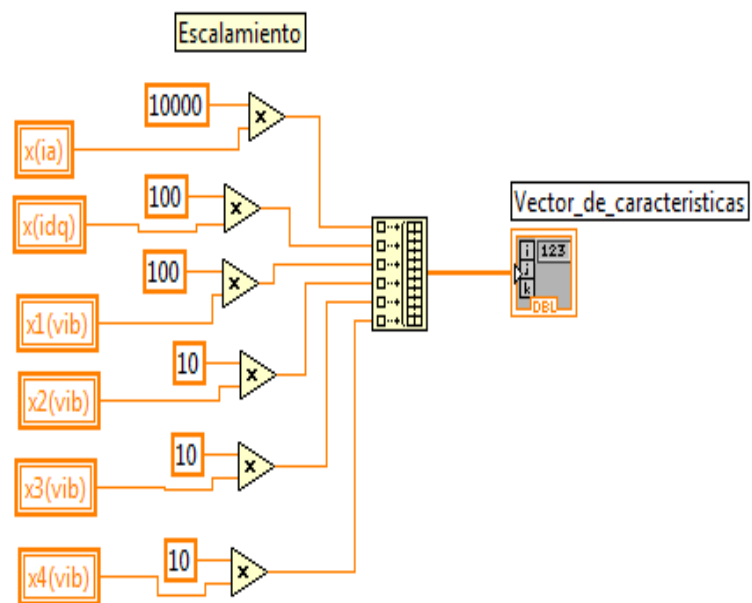


Figura A.1: Medición de los medios de diagnóstico mediante LabView.



**Figura A.2:** Extracción de las características de clasificación.



**Figura A.3:** Construcción del vector de características.

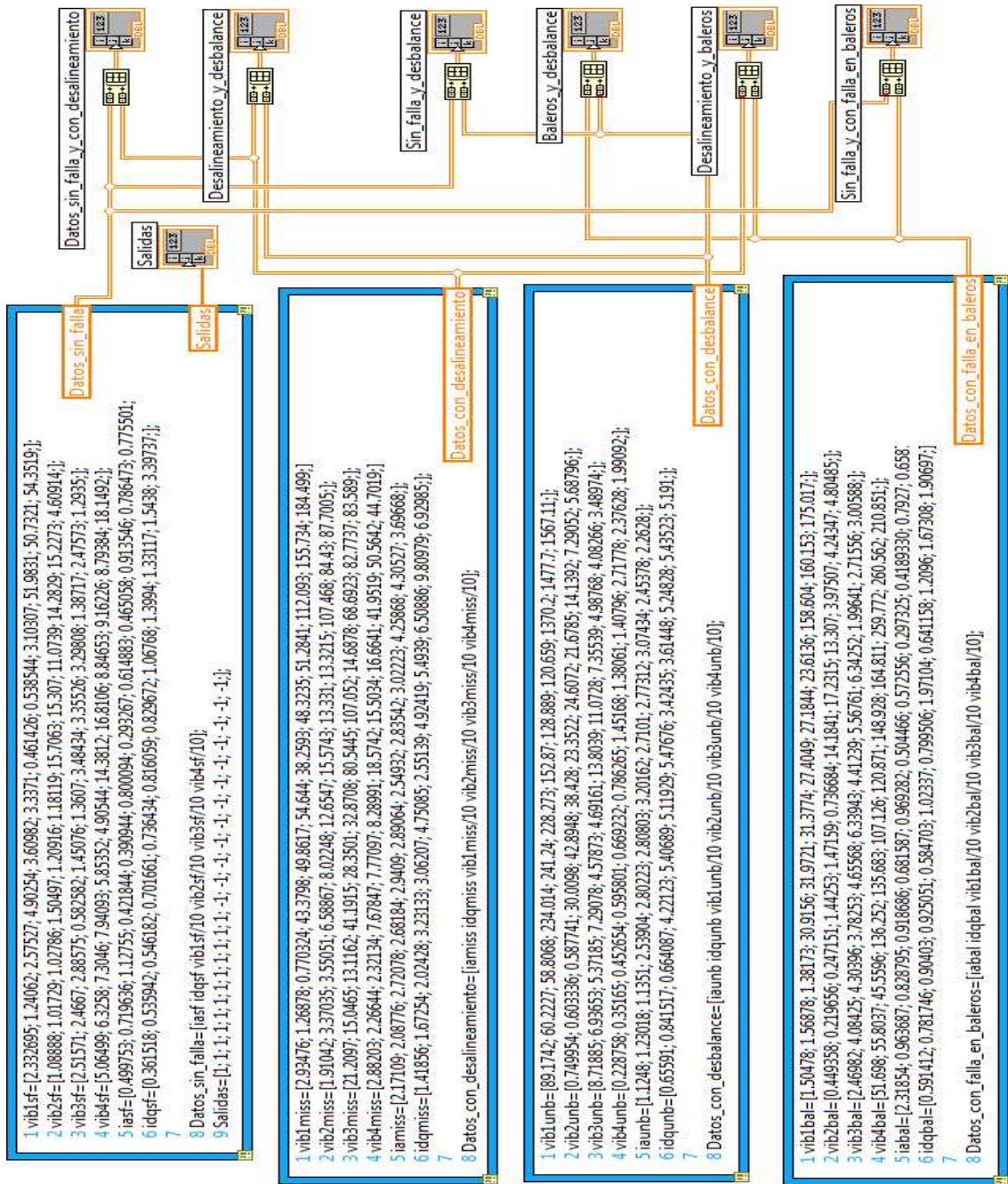
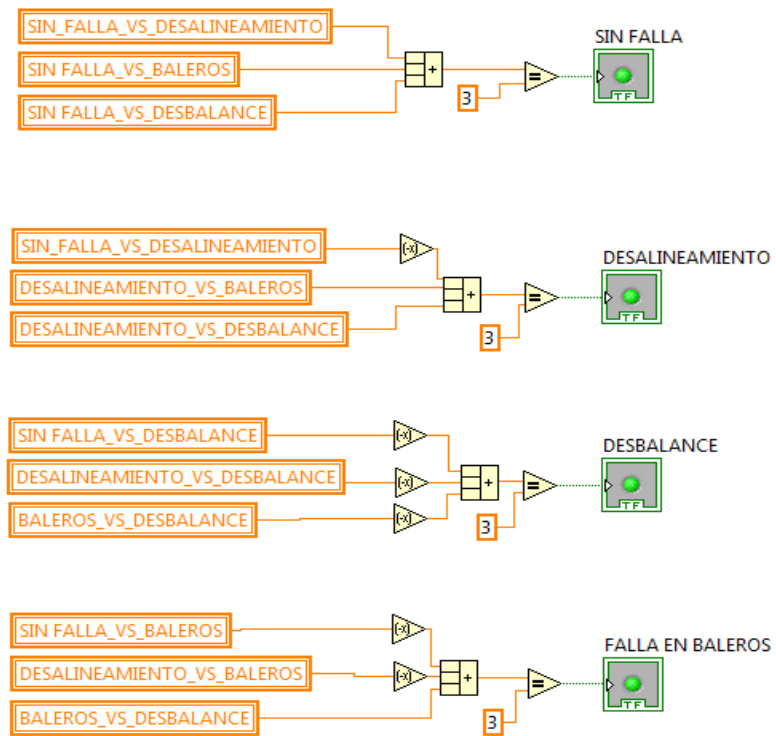


Figura A.4: Datos de entrenamiento en mathscript para calcular la salida de las SVM.



**Figura A.5:** Esquema de votación para determinar el tipo de falla en el MI.

---

## Bibliografía

---

- [1] R. Isermann. Fault Diagnosis Systems. An introduction from fault detection to fault tolerant. *Springer*, (2006).
- [2] IEEE Committe Report, Report of large motor reliability survey of industrial and commercial installations, Part I, *IEEE Trans. Ind. Applicat.*, **IA-21**: 853-872, (1985).
- [3] O. V. Thorsen, and M. Dalva, Failure Identification and Analysis for High-Voltage Induction Motors in the Petrochemical Industry, *IEEE Trans. Ind. Applicat.*, **35**: 810-818, (1999).
- [4] A. Siddique, G. S. Yadava, and B. Singh, A review of stator fault monitoring techniques of induction motors, *IEEE Trans. Energy Conversion*, **20**: 106-114, (2005).
- [5] S. Nandi, H. A. Toliyat, and X. Li, Condition monitoring and fault diagnosis of electrical motors- a review, *IEEE Trans. Energy Conversion*, **20**: 719-729, (2005).
- [6] M. Benbouzid, G. B. Kliman, What stator current processing-based technique to use for induction motor faults diagnosis, *IEEE Trans. Energy Conversion*, **18**: 238-244, (2003).

- 
- [7] J. Cusido, L. Romeral, Juan A. Ortega, J. A. Rosero, and Antonio García Espinosa. Fault detection in induction machines using power spectral density in wavelet decomposition. *IEEE Trans. Ind. Electron.*, **55**: 633-643, (2008).
- [8] D. U. Campos-Delgado, J. S. Murguía, and O. Ramírez-Rodríguez. Multivariable fault diagnosis in induction motor. En *Congreso Nacional de Control Automático (AMCA)*, Monterrey, México, (2007).
- [9] R. Isermann. Process fault detection on modeling and estimation methods - a survey, *Automatica*, **20**: 387-404, (1984).
- [10] R. Isermann, and P. Ball. Trends in the application of model-based fault detection and diagnosis of technical processes, *Control Engineering Practice*, **5**: 709-719, (1997).
- [11] P. M. Frank. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy. A survey and some new results. *Control Engineering Practice*, **26**: 459-474, (1990).
- [12] S. Quiang, X. Z. Gao, and X. Zhuang. State-of-the-art in soft computing-based motor fault diagnosis. *IEEE International Conference on Control Applications*, 1381-1386, (2003).
- [13] F. Filippetti, G. Franceschini and C. Tassoni. Recent developments of induction motor drives fault diagnosis using AI techniques. *IEEE Trans. Ind. Electron.*, **47**: 994-1004, (2000).
- [14] R. J. Patton, F. J. Uppal and C. J. Lopez-Toribio. Soft computing approaches to fault diagnosis for dynamic systems: a survey., *4th IFAC Symposium on Fault Detection Supervision and Safety for Technical Processes*, 298-311, (2000).
- [15] H. J. Shin, Dong-Hwan Eom, and Sung-Shick Kim. One-class support vector machines-an application in machine fault detection and classification, *Computers & Industrial Engineering*, **48**: 395-408, (2005).



- 
- [16] B. Li, Mo-Yuen Chow, Y. Tipsuwam, and J. C. Hung, Neural-network-based motor rolling bearing fault diagnosis, *IEEE Trans. Ind. Electron.*, **47**: 1060-1069, (2000).
- [17] H. Nejjari and M. Benbouzid, Monitoring and diagnosis of induction motors electrical faults using a current Park's vector pattern learning approach, *IEEE Trans. Ind. Applicat.*, **36**: 730-735, (2000).
- [18] R. Schoen, B. Lin, T. Habetler, J. Schlag, and S. Farag, An unsupervised, on-line system for induction motor fault detection using stator current monitoring, *IEEE Trans. Ind. Applicat.*, **31**: 1280-1286, (1995).
- [19] G. Goddu, B. Li, Mo-Yuen Chow and J. Hung, Motor bearing fault diagnosis by a fundamental frequency amplitude based fuzzy decision system, *IEEE Trans. Ind. Electron.*, **47**: 1961-1965, (1998).
- [20] F. Zidani, M. Benbouzid, D. Diallo and M. Said , Induction motor stator faults diagnosis by a current Concordia pattern-based fuzzy decision system, *IEEE Trans. Energy Conversion*, **18**: 469-475, (2003).
- [21] I. Onel and M. Benbouzid, Induction motor bearing failure detection and diagnosis: Park and Concordia transform approaches comparative study, *IEEE Trans. on Mechatronics*, **13**: 257-262, (2008).
- [22] S. Altug, M. Chow, and H. Trussell, Fuzzy inference systems implemented on neural architectures for motor fault detection and diagnosis, *IEEE Trans. Ind. Electron.*, **46**: 1069-1079, (1999).
- [23] B. Samanta, K. Al-Balushi, and S. Al-Araimi, Bearing fault detection using artificial neural networks and genetic algorithm, *Journal on Applied Signal Processing*, **3**: 366-377, (2004).

- 
- [24] T. Han, B. Yang, W. Choi, and Jae-Sik Kim, Fault diagnosis system of induction motors based on neural network and genetic algorithm using stator current signals, *International Journal of Rotating Machinery*, **2006**: 1-13, (2006).
- [25] M. Chow, R. Sharpe, and J. Hung, On the application and design of artificial neural networks for motor fault detection- Part II , *IEEE Trans. Ind. Electron.*, **40**: 189-196, (1993).
- [26] C. Zhitong et al., Support vector machine used to diagnose the fault of rotor broken bars of induction motors, *IEEE Electrical Machines and Systems*, **2**: 891-894, (2003).
- [27] I. Aydin, M. Karakose, and E. Akin, Artificial immune based support vector machine algorithm for fault diagnosis of induction motors. *International Aegean Conference on Electrical Machines and Power Electronics*, 217-221, (2007).
- [28] V. Vapnik. The nature of statistical learning theory. *Springer*, (2000).
- [29] V. Vapnik. Statistical learning theory. *John Wiley & Sons*, (1998).
- [30] E. Alpaydin. Introduction to machine learning. *The MIT Press*, (2004).
- [31] V. Kecman. Learning and soft computing. *The MIT Press*, (2001).
- [32] S. Haykin. Neural Networks: A comprehensive foundation. *Prentice Hall*, (1999).
- [33] M. Bartholomew. Nonlinear optimization with financial applications. *Kluwer Academic Publishers*, (2005).
- [34] W. Sun, Y. Yuan. Optimization theory and methods: Nonlinear programming. *Springer*, (2006).
- [35] S. Abe. Support vector machine for pattern classification. *Springer*, (2005).
- [36] L. Wang. Support vector machines: Theory and applications. *Springer*, (2006).

- 
- [37] M. Benbouzid, A review of induction motors signature analysis as a medium for faults detection, *IEEE Trans. Ind. Electron.*, **47**: 984-993, (2000).
- [38] R. Schoen, T. Habetler, F. Kamran, and R. Bartheld, Motor bearing damage detection using stator current monitoring, *IEEE Trans. Ind. Applicat.*, **31**: 1274-1279, (1995).
- [39] A. Trzynadlowski, and E. Ritchie, Comparative investigation of diagnostic media for induction motors: a case of rotor cage faults, *IEEE Trans. Ind. Electron.*, **47**: 1092-1099, (2000).
- [40] S. Cruz, and A. Marques, Stator winding fault diagnosis in three-phase synchronous and asynchronous motors, by the extended Park's vector approach, *IEEE Trans. Ind. Electron.*, **37**: 1227-1233, (2001).
- [41] N. Tandon and A. Parey. Condition Monitoring of Rotary Machines, Chapter 5, In Condition Monitoring and Control for Intelligent Manufacturing, Lihui Wang , Robert X. Gao. *Springer*, (2006).
- [42] P. Girdhar. Practical machinery vibration, analysis & predictive maintenance. *Newnes*, (2004).
- [43] C. de Silva. Vibration fundamentals and practice. *CRC Press*, (1999).
- [44] Datasheet. Precision, Low Power instrument amplifier INA129P. *Burr-Brown Corporation*, (1995). <http://www.datasheetarchive.com/search.php?q=ina129p>
- [45] Installation and Operating Manual. Low-cost Triaxial Industrial ICP Accelerometer Model 604B31. *IMI sensors*, (1998). [http://www.pcb.com/contentstore/docs/PCB\\_Corporate/IMI/products/Manuals/604B31.pdf](http://www.pcb.com/contentstore/docs/PCB_Corporate/IMI/products/Manuals/604B31.pdf)
- [46] Datasheet. Current regulator diodes CR470. *Vishay Siliconix*, (2001). <http://www.datasheetarchive.com/search.php?q=cr470>

- [47] User operating manual for machinery fault simulator<sup>TM</sup>, *SpectraQuest, Inc*, (2003).
- [48] Applied vibration analysis, Volumen 1 y 2, *SpectraQuest, Inc*, (2003).
- [49] R. Bishop. Learning with LabView 7 express. *Prentice Hall*, (2004).