

**CORPORACIÓN MEXICANA DE INVESTIGACIÓN EN MATERIALES S.A.
DE C.V.**

DIVISIÓN DE ESTUDIOS DE POSGRADO



POSGRADO INTERINSTITUCIONAL EN CIENCIA Y TECNOLOGÍA

**DIAGNÓSTICO Y DETECCIÓN DE FALLAS EMPLEANDO UN MÉTODO DE
SISTEMAS INTELIGENTES**

**POR
CÉSAR OCTAVIO HERNÁNDEZ MORALES**

**TESIS
MAestrÍA EN CIENCIA Y TECNOLOGÍA EN SISTEMAS DE
MANUFACTURA AVANZADA**

SALTILLO, COAHUILA, MEX.

20 DE FEBRERO DEL 2015

Diagnóstico y detección de fallas empleando un método de sistemas inteligentes

Por

César Octavio Hernández Morales

Tesis

Presentada al Programa Interinstitucional en Ciencia y Tecnología

Sede

Corporación Mexicana de Investigación en Materiales S.A. de C.V.

Como requisito parcial para obtener el grado académico de

Maestro en Ciencia y Tecnología
En Sistemas de Manufactura Avanzada

Programa Interinstitucional en Ciencia y Tecnología COMIMSA / CONACyT

Saltillo Coahuila, 20 de febrero del 2015.

**CORPORACIÓN MEXICANA DE INVESTIGACIÓN EN MATERIALES S.A. DE
C.V.**

**GERENCIA DE DESARROLLO HUMANO
DIVISIÓN DE ESTUDIOS DE POSGRADO**

Los miembros del Comité Tutorial recomendamos que la Tesis “**Diagnóstico y detección de fallas empleando un método de sistemas inteligentes**”, realizada por el alumno César Octavio Hernández Morales matrícula 1206MA3098 sea aceptada para su defensa como Maestro en Ciencia y Tecnología en Sistemas de Manufactura Avanzada.

El Comité Tutorial



Director de Tesis
Dr. Juan Pablo Nieto González



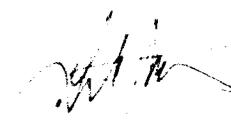
Tutor en Planta
Dra. Indira Gary
Escamilla Salazar



Asesor
Dr. Elías Gabriel
Carrum Siller



Asesor
Dr. Mario Cantú
Sifuentes



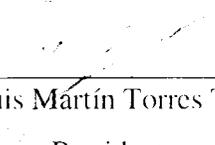
Vo. Bo.
Dr. Felipe Arturo Reyes Valdés
Coordinador de Posgrado

**CORPORACIÓN MEXICANA DE INVESTIGACIÓN EN MATERIALES S.A. DE
C.V.**

**GERENCIA DE DESARROLLO HUMANO
DIVISIÓN DE ESTUDIOS DE POSGRADO**

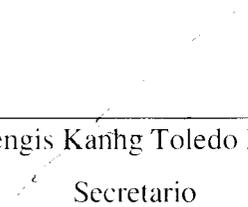
Los abajo firmantes, miembros del Jurado de Examen de Grado del alumno **CÉSAR OCTAVIO HERNÁNDEZ MORALES**, una vez leída y revisada la Tesis titulada **“DIAGNÓSTICO Y DETECCIÓN DE FALLAS EMPLEANDO UN MÉTODO DE SISTEMAS INTELIGENTES”**, aceptamos que la referida tesis revisada y corregida sea presentada por el alumno para aspirar al grado de Maestría en Ciencia y Tecnología en Sistemas de Manufactura Avanzada durante el Examen de Grado correspondiente.

Y para que así conste firmamos la presente a los 20 días del mes de Febrero del año 2015



Dr. Luis Martín Torres Treviño

Presidente



Dr. Gengis Kanhg Toledo Ramírez

Secretario



Dr. Juan Pablo Nieto González

Vocal

DEDICATORIA

A MI ABUELA Y A MI MADRE.

Domitila Morales Belmonte y Guadalupe Hernández Morales

AGRADECIMIENTOS

A Dios: Por darme la oportunidad de terminar esta etapa de mi vida, por darme una familia increíble y porque ha llenado mi vida de gente que quiero y admiro.

A mi Mamá: Gracias mami por ser un gran ejemplo. Quiero que alguien, algún día llegue a estar tan orgulloso de mí, como yo lo estoy de ti. Gracias por amarme.

A mi Abuelita: Mami muchas gracias, por cuidarme, educarme, consentirme y quererme.

A mis Tíos: Jesús y Rafael gracias por ser un ejemplo y un papá para mí.

A mis primos: Herman, Edson, Rafita, Paola Said y Adriana gracias por estar conmigo en todos los momentos de mi vida los quiero mucho.

A mi tutor: Dr. Juan Pablo gracias por ayudarme a alcázar esta meta, por tolerarme y por guiarme durante este proceso.

A mi Asesor: Dr. Elías Gracias por ser un apoyo y un guía en este periodo.

A los Doctores Luis e Indira: Por ayudarme a obtener mi caso de estudio gracias por su tiempo y disposición.

A Mis Amigos: Muchas gracias por estar conmigo ser las personas que me han hecho disfrutar de la vida y la han hecho más amena, Esme, Yessy, Karla, Olga, Mayote, Isra y Vanne. Muchas gracias, Carlita y Liz muchas gracias porque comenzamos este camino juntos. Liz muchas gracias por ser mi conciencia, por soportarme, por hacerme reír, enojar, te quiero mucho, gracias porque en este tiempo he tenido el honor de conocer y de vivir muchas cosas juntos.

A mis amigos que están lejos muchas gracias por que siempre han estado en mi vida aunque sea lejos seguimos disfrutando muchas cosas juntos, gracias a todos, muy en especial a Edgar, Mauricio, Hugo, Tony, Samuel, Cesar, Víctor y Jonathan y los que me faltan muchas gracias.

A CONACYT: Gracias por el financiamiento para realizar mi maestría.

A COMIMSA: Muchas gracias, muy en especial a Dr. Arturo, Dr. Pedro, Pily, Maru, Mary, Memo y Blanquita.

AUTOBIOGRAFÍA

César Octavio Hernández Morales nació en Orizaba, Veracruz el 25 de septiembre de 1986. Es ingeniero industrial graduado del Instituto Tecnológico de Orizaba donde estudió entre los años 2006 - 2009. Realizó sus residencias profesionales en la empresa International Paper Empaques Industriales de México S.A De C.V. Planta Ixtac. En el período Agosto 2009 – Abril 2010. En septiembre del 2012 inicia sus estudios de maestría en COMIMSA.

Ha realizado diferentes publicaciones enfocadas a la detección y diagnóstico de fallas entre las cuales se encuentran las siguientes: "Fault detection and diagnosis of electrical networks using a fuzzy system and euclidian distance" en el "12th Mexican International Conference on Artificial Intelligence, MICAI 2013". Ciudad de México, Noviembre 2013, "Detección y Diagnóstico de Fallas en SEP's Combinando Lógica difusa con distancias Euclidianas y de Mahalanobis" en el "X Congreso Internacional sobre Innovación y Desarrollo Tecnológico", Abril de 2014, Cuernavaca Morelos, México, "Detección y Diagnóstico de Fallas en sistemas eléctricos de potencia (SEP) Combinando Lógica difusa, métricas y una red neuronal probabilística" en el "6o Congreso Mexicano de Inteligencia Artificial COMIA 2014, Mayo 2014", Zumpango Edo. De México, México, "Detección de Fugas en un sistema de Distribución de Gas Metano Empleando Lógica Difusa una Red Neuronal y una Máquina de Soporte Vectorial" en el "XVII Seminario de Investigación del PICYT 2014", Querétaro, Qro. CIDESI. En proceso de publicación, por ultimo "Fault Detection and Diagnosis of Electrical Networks Using a Fuzzy System, Probabilistic Neural Network and Support Vector Machine".

Las áreas de investigación en las que se desarrolla se enfocan a la detección y diagnóstico de fallas con metodologías de inteligencia artificial así como el control de procesos.

RESUMEN

En la actualidad, mantener los procesos o sistemas en control es una técnica que día a día se va haciendo más común en la mayoría de las empresas, esto por la gran necesidad de hacer sus procesos cada vez más autónomos, y con ello, el deseo de mantener una calidad constante en los productos y servicios, la cual pueda satisfacer los requerimientos que ya han sido establecidos por los clientes, estos requerimientos, se describen, como las características con las cuales debe cumplir el producto o servicio, y son dependientes de las variables que intervienen directamente en el proceso o sistema.

Debido a esto, muchos sistemas autónomos, están volcando sus procesos a la tendencia de detectar y diagnosticar las fallas presentes, para posteriormente poder recuperar su sistema o proceso y con ello mantener constante la calidad que se desea del producto o servicio.

Para realizar la tarea de detección y diagnóstico existen distintos tipos de métodos, los más usuales son con frecuencia los basados en modelos matemáticos que determinan, donde se han presentado las fallas, pero también, existe otra tendencia que realiza el análisis de los datos históricos del proceso, donde, para realizar la tarea de detección y diagnóstico, son utilizados los datos obtenidos directamente del proceso, para con su análisis se tenga conocimiento de las condiciones en que está funcionando el sistema.

Considerando los métodos antes mencionados y teniendo la necesidad en la mayoría de las empresas de localizar las variaciones que tienen los distintos tipos de procesos, en el presente trabajo se desea realizar un sistema de monitoreo para poder detectar y diagnosticar las fallas presentes en un proceso o sistema, proponiendo una nueva metodología, para poder

realizar la tarea de identificar el tipo de falla y poder localizar aquellas variables que se encuentren fuera de las condiciones normales de operación, con el fin de hacer los procesos y sistemas autónomos, más simples y más seguros, para en pasos posteriores, sea recupere el sistema o proceso y con ello poder garantizar la calidad de los productos, por lo tanto se propone una nueva metodología que emplea técnicas de softcomputing, la cual está compuesta por dos etapas. La primera utiliza un sistema difuso para realizar la detección y la segunda emplea una red neuronal probabilística en conjunto con una máquina de soporte vectorial para facilitar la tarea del diagnóstico.

La metodología propuesta fue valida en dos casos de estudio en el primer caso de estudio, se presenta un análisis en un sistema eléctrico de potencia el cual tiene presencia de distintas fallas. El segundo caso de estudio fue realizado en un proceso de mecanizado en el cual se analizan distintas variables para obtener el acabado requerido por el cliente, este analiza los parámetros promedio provenientes de una máquina de torno y una fresadora.

En el presente trabajo se encontrará que la metodología propuesta tiene una buena eficiencia en los dos casos de estudio y al compararlo con el estado del arte, se tiene una metodología competitiva ya que dicha propuesta hace uso de cada una de las cualidades de los tres sistemas inteligentes empleados con lo cual se crea un sistema redundante para poder alcanzar una buena detección y diagnóstico de las fallas presentes en el proceso o sistema.

ABSTRACT

Currently, maintain the processes or systems in control is a technique that every day becomes more common in most companies, this is great need for their processes will be increasingly autonomous, and thus, the desire to maintain constant the products with quality and that the services can meet of requirements that have been established by clients, these requirements are described as the characteristics which must comply the product or service, and these are dependent on the variables directly involved in the process or system.

As result, many autonomous systems are turning their processes to the tendency to detect and diagnose faults present in the process or system, for later to recover your system or process and thereby maintain constant the desired quality in product or service.

To perform the task of detection an diagnosis are different types of methods, the most common are often based on mathematical models that determine where failures have occurred, but, there is another trend that performs the analysis of historical data process, where, for the task of detection and diagnosis are used the data obtained directly from the process, for his analysis and have knowledge of conditions in which the system is working.

Considering the above methods and having the need for most companies to find the variations in different types of processes, in this work we want to develop a monitoring system to detect and diagnose faults in a process or system, generating a new methodology to perform the task of identifying the type of fault and to locate those variables that are outside normal operating conditions, in order to make the process and simpler and safer autonomous systems, for in later steps, it recovers the system or

process and thus to guarantee the quality of the products, therefore, it proposes a new framework that use soft computing techniques, which consists of two steps. The first step uses a fuzzy system for detection and the second step employs a probabilistic neural network in conjunction with a support vector machine to facilitate the task of diagnosis.

The proposed methodology was validated in two case studies, the first case study; it is presented an analysis in a power system that is present different faults. In the second case study, was conducted in a machining process in which analyzed to different variables for obtain the finish required by the customer; it analyzes the average parameters from a machine turning and a machine milling.

In this was framework found that the proposed methodology has a good efficiency in the two cases of studies and to compared with the state of art analyzed, the framework proposed has a competitive methodology, since this methodology makes use of each of the qualities of the three intelligence system, thereby, is presented a redundant system to achieve good detection and diagnosis of faults in the system process or system.

ÍNDICE

I. INTRODUCCIÓN	1
II. ESTADO DEL ARTE	8
II.1. Reconocimiento de patrones en la detección y diagnóstico de fallas	8
III. PLANTEAMIENTO DEL PROBLEMA.....	14
III.1. Descripción del Problema	14
III.2. Preguntas de Investigación.....	15
III.3. Hipótesis.....	16
III.4. Objetivos	16
III.4.1. Objetivo general	16
III.4.2. Objetivos específicos	16
III.5. Justificación.....	17
III.6. Alcances y delimitaciones.....	17
III.6.1. Alcances	17
III.6.2. Delimitaciones.....	18
III.7. Resultados e impacto esperados.....	19
III.7.1. Resultados.	19
III.7.2. Impactos.	19
IV. MARCO TEÓRICO	21
IV.1. Conceptos básicos de detección y diagnóstico	21
IV.2. Métodos de monitoreo de procesos	26
IV.2.1. Métodos analíticos y basados en el conocimiento	27
IV.2.2. Pretratamiento de datos, Métodos de estadística univariante, multivariante para realizar la detección y diagnóstico de fallas	30
IV.2.2.1. <u>Pretratamiento de datos</u>	30
IV.2.2.2. <u>Estadística univariante</u>	31
IV.2.2.3. <u>T² de hotelling</u>	32

IV.2.2.4. <u>Análisis de componentes principales</u>	35
IV.2.2.5. <u>Discriminante de Fisher</u>	39
IV.2.2.6. <u>Análisis de correlación canónica</u>	42
IV.2.2.7. <u>Mínimos cuadrados parciales</u>	45
IV.2.2.8. <u>Distancia estadística</u>	48
IV.2.3. <u>Computación Suave Para Realizar la Detección y Diagnóstico</u>	49
IV.2.3.1. <u>Sistemas difusos</u>	50
IV.2.3.2. <u>Algoritmos evolutivos</u>	51
IV.2.3.3. <u>Redes neuronales</u>	52
IV.2.3.4. <u>Razonamiento probabilístico</u>	53
IV.3. <u>Sistemas inteligentes empleados en la metodología para realizar la detección y diagnóstico de fallas</u>	54
IV.3.1. <u>Lógica difusa</u>	54
IV.3.1.1. <u>Funciones de membresía</u>	56
IV.3.1.2. <u>Operadores difusos</u>	57
IV.3.1.3. <u>Implicaciones difusas</u>	58
IV.3.1.4. <u>Defussificador</u>	59
IV.3.2. <u>Red Neuronal Probabilística</u>	59
IV.3.3. <u>Máquina de soportes vectoriales</u>	62
IV.3.3.1. <u>La Máquina de soporte vectorial de dos clases</u>	63
IV.3.3.2. <u>Máquina de soporte vectorial multiclase</u>	67
IV.3.3.3. <u>Clasificación basada en las funciones de toma continua</u>	68
V. METODOLOGÍA	73
V.1. <u>Pasos de la metodología</u>	73
V.2. <u>Metodología de forma gráfica</u>	77
VI. EXPERIMENTACIÓN	78
VI.1. <u>Caso de estudio 1</u>	78
VI.1.1. <u>Etapas de detección para el sistema eléctrico</u>	81
VI.1.2. <u>Etapas de diagnóstico para el sistema eléctrico</u>	85
VI.2. <u>Caso de estudio 2</u>	97

VI.2.1. Etapa de detección para el proceso de mecanizado.....	102
VI.2.2. Etapa de diagnóstico para el proceso de mecanizado.....	104
VII. CONCLUSIONES	106
VIII. PRODUCTOS DERIVADOS DEL TRABAJO	110
IX. REFERENCIAS.....	115

ÍNDICE DE TABLAS

Tabla VI. 1 <i>Reglas difusas para la detección de fallas asimétricas.</i>	82
Tabla VI. 2 <i>Reglas difusas para la detección de fallas simétricas y asimétricas.</i>	82
Tabla VI. 3 <i>Comparación de métricas estadísticas.</i>	92
Tabla VI. 4 <i>Comparación de las técnicas empleadas en la detección y diagnóstico del sistema eléctrico.</i>	95
Tabla VI. 5 <i>Reglas difusas para la detección en el proceso de mecanizado.</i>	103
Tabla VI. 6 <i>Comparación de la metodología en el proceso de mecanizado.</i> .	105
Tabla VII. 1 <i>Comparación del estado del arte contra la metodología propuesta.</i>	108

ÍNDICE DE FIGURAS

Figura I. 1 <i>Representación esquemática general de un sistema o proceso (adaptado de [1]).</i>	1
Figura I. 2 <i>Esquema típico de un proceso de transformación (adaptado de [1]).</i>	2
Figura I. 3 <i>Proceso o sistema, entradas y salidas (adaptado de [2]).</i>	4
Figura I. 4 <i>Esquema de método de control (adaptado de [3]).</i>	6
Figura III. 1 <i>Alcances y delimitaciones de la metodología.</i>	18
Figura IV. 1 <i>Clasificación de métodos de diagnóstico (adaptado de [23])</i>	23
Figura IV. 2 <i>Métodos de computación suave para detectar y diagnosticar fallas.</i>	50
Figura IV. 3 <i>Función de pertenencia a la "alta", donde el eje horizontal representa la velocidad del vehículo y el eje vertical representa el valor de pertenencia a la "alta" (tomado de [34]).</i>	55
Figura IV. 4 <i>Proceso difuso. El cual describe la metodología a seguir para ocupar esta herramienta.(tomado de Toolbox de MatLab®).</i>	56
Figura IV. 5 <i>Red Neuronal Probabilística (tomada de [35]).</i>	61
Figura IV. 6 <i>La separación de los hiperplanos con márgenes pequeños y grandes en la clasificación binaria (tomada de [36]).</i>	65
Figura IV. 7 <i>SVM-dos clase con funciones de decisión discreta y continua (tomada de [36]).</i>	69
Figura IV. 8 <i>Diagrama esquemático de UVT-SVM con funciones de decisión continuas (tomada de [36]).</i>	70
Figura IV. 9 <i>Diagrama esquemático de UVU-SVM con funciones de decisión continuas (tomada de [36]).</i>	72
Figura V. 1 <i>Metodología Propuesta.</i>	77
Figura VI. 1 <i>Diagrama unifilar del sistema propuesto por el IEEE (adaptado de [40]).</i>	79
Figura VI. 2 <i>Modo de operación normal.</i>	80

Figura VI. 3 <i>Falla simétrica</i>	80
Figura VI. 4 <i>Falla asimétrica</i>	80
Figura VI. 5 <i>PCA en modo de operación normal</i>	83
Figura VI. 6 <i>PCA en modo de falla</i>	83
Figura VI. 7 <i>Metodología para el diagnóstico de fallas Simétricas</i>	87
Figura VI. 8 <i>Metodología para el diagnóstico de fallas Asimétricas</i>	89
Figura VI. 9 <i>Máquina de torno de dos ejes CNC Okuma LB15</i>	98
Figura VI. 10 <i>Cilindro a mecanizar compuesto de los tres materiales, aluminio 6062, acero 1018 y acero 4140</i>	99
Figura VI. 11 <i>Rugosímetro Mitutoyo SJ-301</i>	99
Figura VI. 12 <i>Uso del rugosímetro en cada cilindro mecanizado</i>	100
Figura VI. 13 <i>Centro de mecanizado EMCO PC MILL 125</i>	101
Figura VI. 14 <i>Placas por ser mecanizadas hechas de los tres materiales, aluminio 6062, acero 1018 y del acero 4140</i>	101
Figura VI. 15 <i>Utilización del rugosímetro en cada placa mecanizada</i>	102

I. INTRODUCCIÓN

José Francisco Vilar [1] describe un control de proceso o sistema como un método de mejora que pretende reducir sistemáticamente la variación de los parámetros dentro de un proceso y obtener las características requeridas para los productos. Las herramientas para realizar esta tarea son el seguimiento, el control y el monitoreo de los parámetros causantes de estas variaciones. La aplicación de un programa de control se basa en el estudio, análisis y monitoreo de las variables. Lo cual permitirá entender cuáles son las causas que producen variación en los resultados y como se pueden mejorar estos procesos. La Figura I.1 muestra el esquema general de un sistema o proceso que cuenta con un suministrador de materiales que serán transformados mediante un proceso en un producto o servicio. La calidad del producto se evalúa comparando sus características con los requerimientos del cliente y posteriormente esta información realimenta al proceso.

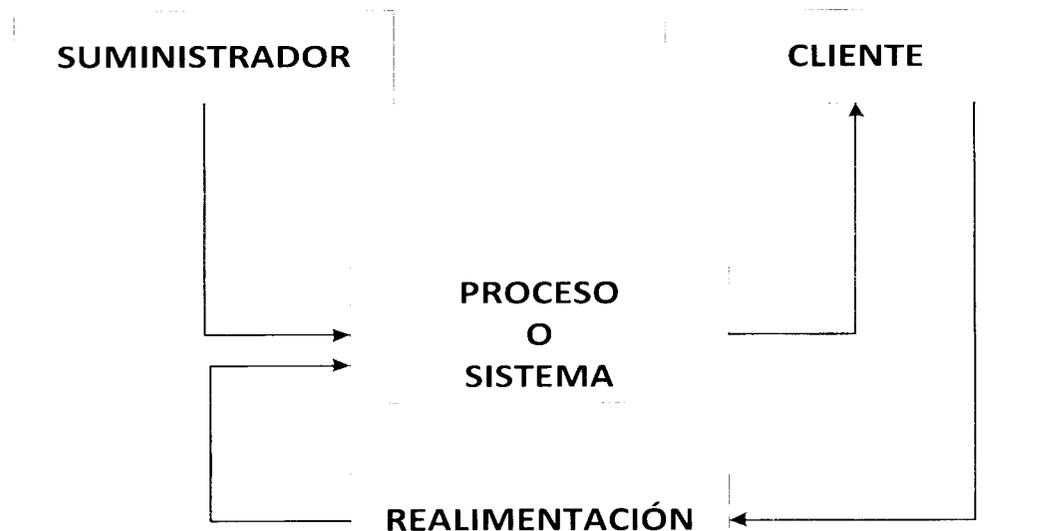


Figura I. 1 Representación esquemática general de un sistema o proceso (adaptado de [1]).

La Figura I.2 muestra el esquema general de un proceso de transformación.

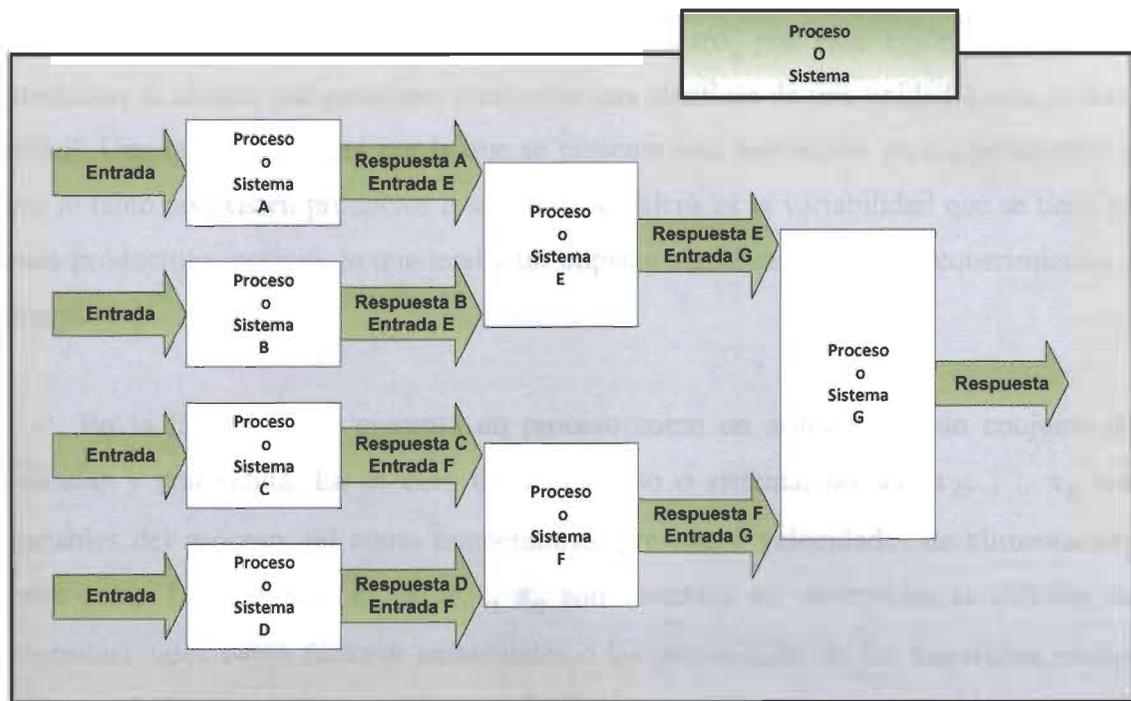


Figura I. 2 Esquema típico de un proceso de transformación (adaptado de [1]).

El cual recibe un número de entradas del entorno del cual está rodeado, transforma esas entradas en respuestas y entrega esta nueva entrada a un nuevo entorno. Se puede entender como entradas a todos aquellos aspectos que tengan una directa interacción con el proceso o sistema, para poder tener una respuesta la cual cumpla con las características que ya han sido determinadas por un cliente, y así integrar este producto o servicio a su nuevo entorno. El proceso como se muestra, puede estar conformado por subprocesos los cuales son las entradas de procesos secundarios y juntos integran un proceso o sistema de transformación.

Como ya se ha mencionado el control de un proceso o sistema se ha convertido en una importante estrategia empresarial para muchas organizaciones, fabricantes, distribuidores. Lo anterior debido a que la calidad es una ventaja competitiva. Debido a

que tener variación excesiva en un proceso a menudo genera residuos. Por ejemplo, considera el desperdicio de dinero, tiempo y esfuerzo, la reducción de la variación en los procesos pretende eliminar la variabilidad en los parámetros de entrada. Lo que para la mayoría de las organizaciones resulta difícil y caro, por esta razón proporcionar productos al cliente que presenten características idénticas de una unidad a otra es muy difícil. Una razón importante por la que se presenta esta desviación en los parámetros y por lo tanto no existen productos o servicios idénticos es la variabilidad que se tiene en cada producto o servicio, lo que tendrá un impacto significativo en los requerimientos a cumplir.

En la Figura I.3 se presenta un proceso como un sistema con un conjunto de entradas y una salida. En el caso de un proceso o sistema, las x_1, x_2, \dots, x_p son variables del proceso, tal como temperaturas, presiones, velocidades de alimentación, entre otras. Las entradas z_1, z_2, \dots, z_q son variables no observadas (o difíciles de controlar), tales como factores ambientales o las propiedades de los materiales crudos proporcionados por un proveedor externo. El proceso o sistema transforma la entrada de materias primas, partes, componentes y subconjuntos en un producto terminado, el cual al ser evaluado se determinaran las variables de salida, dicha salida tiene varias características que tiene que ser cumplidas; las cuales son llamadas características de calidad, es decir, una medida de la calidad del proceso, producto o servicio. Lo que se busca con la implementación de una nueva metodología es controlar las variables de entrada, para obtener continuamente la calidad requerida por el cliente, ya que al haber una variación en los parámetros de entrada, se reflejara en las características de calidad, caracterizándose como fallas [1] y [2].

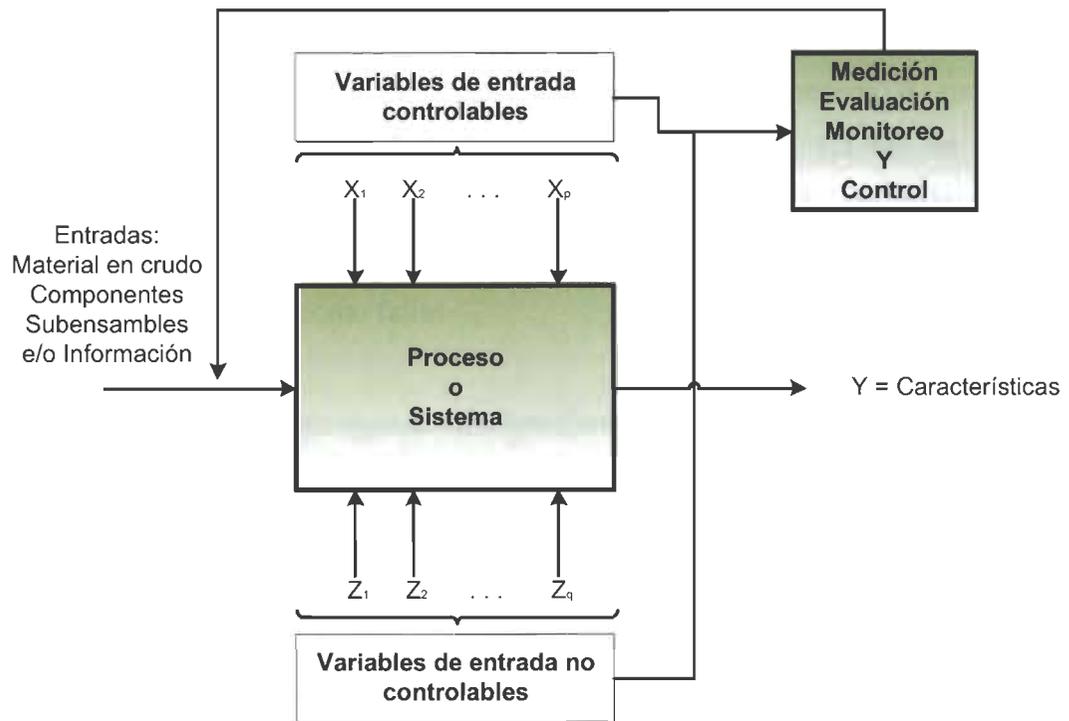


Figura I. 3 Proceso o sistema, entradas y salidas (adaptado de [2]).

Esto lleva a definir el siguiente trabajo como un proceso donde se tienen definidas las características de calidad, en un estado que se llamara modo de operación normal, y se desea evaluar a las variables de entrada del proceso que están impactando directamente en la obtención de dichas características de calidad. Si se presenta variación dentro de las variables controladas se les denominarán fallas del proceso o sistema.

Como se ha mencionado en los procesos de hoy en día y en las industrias manufactureras, se está teniendo un gran impulso por mejorar la calidad en los productos y servicios. Lo anterior con la finalidad de satisfacer los estándares y asegurar que las condiciones del proceso son aceptables para satisfacer la seguridad, las regulaciones ambientales y los requerimientos del cliente, ésta es una ardua tarea debido a que hoy en día los sistemas cuentan con un gran número de variables de operación para mantener el proceso bajo control, como se ha esquematizado en la Figura I.3.

Si se ejemplifica al área de control, se puede mencionar que los controles de proceso estándar, están diseñados para mantener a los procesos o sistemas en condiciones de operación satisfactorias, para compensar los efectos de cualquier perturbación y los cambios ocurridos durante el proceso, esto depende de si los parámetros se desvían de sus valores normales de operación, donde se consideran a estos cambios o perturbaciones como fallas.

Para que los procesos operen satisfactoriamente de acuerdo a las especificaciones es necesario detectar, diagnosticar y remover las fallas presentes en ellos. Estas tareas están asociadas con el proceso de monitoreo. Estos sistemas monitorean los procesos con el fin de que los productos o servicios cumplan con las características de calidad que requiere el cliente.

Los procedimientos para el monitoreo de procesos se esquematizan en la Figura I.4 y son los siguientes:

- Detección de la falla
- Identificación de la falla
- Diagnóstico de la falla
- Recuperación del proceso

José Francisco Vilar [1] y Ali Ajami et al. [3] definen una falla como una desviación no permitida de al menos una propiedad, característica o parámetro en condiciones aceptables habituales, de la operación normal de un sistema. La detección de fallas es determinar cuándo una falla está presente en el sistema. Aislamiento de la falla se utiliza para determinar el tipo y la localización de una falla. La identificación de fallas es para determinar el tamaño y la variación en el tiempo de una falla. Por último, el diagnóstico de las fallas es para determinar el tipo, tamaño, ubicación y el tiempo de detección. Y para finalizar el ciclo de control, el sistema se interviene para recuperar su

funcionamiento normal. De acuerdo con [3], la Figura I.4 muestra la metodología general empleada para la detección y diagnóstico.



Figura I. 4 Esquema de método de control (adaptado de [3]).

El monitoreo de procesos puede ser clasificado y asociado con el manejo de datos, el análisis de datos y basados en el conocimiento.

- **El manejo de datos:** hace mediciones basadas directamente en los datos del proceso. La fuerza de las técnicas de manejo de datos está en su capacidad para transformar la alta dimensión a una dimensión inferior, en el que se captura la información importante del proceso. Al calcular algunas estadísticas significativas para los operadores e ingenieros del proceso, un esquema de monitoreo de proceso para un sistema a gran escala se puede mejorar de manera significativa. El principal inconveniente del manejo de datos es que solo mide los datos de proceso.
- **El análisis de datos:** hace uso de los modelos matemáticos, al hacer un enfoque cualitativo para desarrollar las mediciones del monitoreo del proceso. El enfoque analítico es aplicable a sistemas ricos en información, donde los modelos son satisfactorios ya que disponen de suficientes sensores para la recolección de información. La mayoría de métodos analíticos se basan en la estimación paramétrica, diseños basados en observadores y/o relaciones de paridad. La mayoría de las aplicaciones del enfoque analítico han sido en los sistemas con un número relativamente pequeño de entradas, salidas debido a que es difícil aplicar el enfoque analítico de un sistemas a gran escala (es decir, si el sistema que

contienen un gran número de entradas, salidas y / o estados), requiere modelos detallados con el fin de ser eficaces. Modelos detallados de sistemas a gran escala son difíciles de obtener dados los cruces asociados con un sistema multivariante. La principal ventaja de este enfoque analítico es la capacidad de incorporar la comprensión física del proceso en el plan de seguimiento del proceso.

- **Los basados en el conocimiento:** Tratan de simular el razonamiento humano, utiliza modelos cualitativos para desarrollar medidas de control de proceso. El enfoque basado en el conocimiento es especialmente adecuado para sistemas en los que los modelos matemáticos detallados no están disponibles. La mayoría de las medidas basadas en el conocimiento se basan en el análisis causal, sistemas expertos, y en el reconocimiento de patrones. Al igual que el enfoque analítico, la mayoría de las aplicaciones del enfoque basado en el conocimiento han sido utilizados en sistemas con un número relativamente pequeño de entradas, salidas y estados. La construcción de los modelos de fallas para un sistema a gran escala puede requerir una gran cantidad de esfuerzo.

Estos métodos serán utilizados para realizar la tarea de monitoreo en sistemas complejos los que se definen como sistemas con alta correlación en sus variables y con ruido dentro de ellas. Con la implementación de estos métodos se intenta poder alcanzar el objetivo primordial de este trabajo el cual es la detección y el diagnóstico de las fallas presentes en el proceso [4].

II. ESTADO DEL ARTE

II.1.Reconocimiento de patrones en la detección y diagnóstico de fallas

Juan Andrés Cadena et al [5] presenta una herramienta de diagnóstico basada en metodologías de inteligencia artificial como las redes neuronales probabilísticas, para la detección de fallas en transformadores. Utilizando los resultados entregados por ensayos realizados sobre el aceite de un transformador, a través del análisis de los gases disueltos, se obtienen los conjuntos de datos para el entrenamiento y prueba de la red neuronal para posteriormente llevar a cabo el diagnóstico de las fallas.

Fernando et al [6] presenta un enfoque de la clasificación de la superficie y textura de los descriptores de color en el problema de la clasificación de azulejos cerámicos, los cuales deben ser agrupados de acuerdo a su apariencia. Para realizar esta tarea se realiza una clasificación utilizando los K-vecinos más cercanos (K-NN por su nombre en inglés K nearest neighbors), junto con un diseño de experimento y una regresión logística se genera la mejor clasificación y selección de las combinaciones de factores que proporcionen los mejores resultados para cada modelo de azulejo.

Juan Pablo Nieto González et al [7] propone un método de diagnóstico de fallas en sistemas eléctricos de potencia, que es capaz de localizar eventos de fallas múltiples utilizando datos históricos para construir un sistema de diagnóstico completo, para ayudar a la detección de fallas, selecciona un red neuronal probabilística como detector de fallas debido a su sencillez de aprendizaje y por la cantidad de datos que se manejan.

Ribhan Zafira Abdul Rahman et al [8] se centra en la aplicación de redes neuronales en la detección y diagnóstico de fallas. El objetivo de este trabajo es detectar

y diagnosticar las fallas de un reactor continuo de tanque agitado. La detección de fallas se realiza mediante el uso de las señales de error. Cuando la señal de error es cero o casi cero, el sistema está en condición normal de operación, y cuando se produce la falla, las señales de error distintivamente se alejan de cero. El diagnóstico de la falla se realiza al identificar la amplitud del error de la salida del reactor con una red neuronal back-propagation.

P. Subbaraj et al [9] presenta un enfoque basado en red neuronal para la detección de fallas en un actuador neumático en un sistema crítico como el de un enfriador de un sistema de aspersión de agua en la industria del cemento. Los datos necesarios para el desarrollo del modelo de la red neuronal se obtuvieron a través de los datos en tiempo real del funcionamiento del sistema. Los datos de entrenamiento generados se normalizan y se aplican a la red neuronal con una salida correspondiente, para poder conocer la relación de entrada y de salida. La alimentación del modelo de red neuronal se entrena hacia delante como una "feed-forward" pero se utiliza el método de propagación hacia atrás "back-propagation".

Karim Salahshoor et al [10] realiza un monitoreo de la condición de las máquinas para tener una reducción de los costos de mantenimiento, mejorando la productividad y la disponibilidad de la maquinaria. Este trabajo presenta una metodología para una turbina de vapor industrial utilizando una fusión de una máquina de soporte vectorial (SVM por sus siglas en inglés Support Vector Machine) con un clasificador ANFIS (neuro-fuzzy sistema de inferencia adaptativo), para mejorar la detección de errores y tareas de diagnóstico. Los estudios de simulación indican que la fusión resultante de un SVM individual supera en ciertas fallas a los sistemas ANFIS para detectar y diagnosticar fallas incipientes en turbinas de vapor y viceversa.

M. Arunachalam et al. [11] desarrolla agentes inteligentes basados en redes neuronales para predecir las fallas y garantizar la rápida dinámica de re-enrutamiento de las redes. Detecta las fallas en los nodos teniendo en cuenta las dependencias entre las fallas, en los canales de diferentes longitudes de onda. Se utiliza para esta tarea una red neuronal basada en la detección de fallas distribuido en un algoritmo de localización en la red de fibra óptica.

Karim Salahshoor et [12] se centra en el diagnóstico de fallas en una turbina de vapor industrial por un sistema de inferencia de adaptación neuro-difuso (ANFIS) utiliza la metodología como un poderoso sistema de manejo de datos. El principal motivo es debido al régimen de aprendizaje adaptativo inherente, que se incorpora en la estructura del ANFIS, lo que le permite realizar adecuadamente la tarea de diagnóstico. La técnica ANFIS se utiliza como un clasificador del patrón de la falla por discriminación. El sistema propuesto se evalúa a través de un conjunto de las diversas fallas posibles, introducidas en una simulación de la turbina de vapor que comprende las secciones de presión, extracciones de vapor, separadores de humedad y los actuadores relacionados.

S. Deng et al [13] presenta un análisis con una Máquina de soporte vectorial “SVM por sus siglas en inglés support vector machine” y una red neuronal de retro-propagación “back-propagation neural network BPNN” para el diagnóstico de fallas de un cañón de campaña de defensa aérea. Estos métodos inteligentes excluyen el error humano en el diagnóstico de fallas, y hacen que sea posible diagnosticar rápidamente y con precisión. Los resultados experimentales muestran que tanto la SVM y BPNN proporcionan una precisión excelente en el diagnóstico de fallas. Examina que cuando las muestras de entrenamiento no son suficientemente grandes, los modelos SVM multiclase tienen una mejor precisión que los modelos de diagnóstico de la BPNN, esto cuando el número de sesiones de entrenamiento son pequeñas. Cuando son grandes no hay una diferencia entre ellas.

Tony Boutros et al [14] detecta y diagnostica los fallos en un proceso de mecanizado y máquinas rotativas por medio de modelos discretos ocultos de Markov, este método es aplicado en fallas mecánicas para poder detectar el desgaste de la herramienta, fracturas y fallas en los rodamientos, en una primera etapa detecta el estado de la herramienta y en la etapa posterior induce la localización de la falla.

A. Zadeh et al [15] se enfoca en la detección y diagnóstico de averías ya que tienen un papel eficaz para el funcionamiento seguro y la larga vida útil de los sistemas. Propone un algoritmo flexible único para la clasificación de la condición de funcionamiento de una bomba centrífuga sobre la base de un SVM de hiper-parámetros de optimización y redes neuronales artificiales, la SVM con un algoritmo genético y la SVM con un optimizador de enjambre de partículas son algoritmos que se han considerado como un algoritmo flexible para realizar una clasificación precisa en el área de fabricación. Además, analiza la superioridad del algoritmo híbrido propuesto (SVM con GA y PSO) se muestra mediante la comparación de sus resultados con el rendimiento SVM y la red neuronal.

Ali Ajami et al [3] se centra en una técnica estadística para detectar y diagnosticar las fallas en una turbina de potencia "modelo V94.2" el diagnóstico se realiza con un análisis de componentes independientes (ICA) lo que hace este enfoque es analizar los principales factores que inducen una anomalía, este método realiza una división de una señal multivalente en subcomponentes en los cuales supone una independencia estadística. Se dice que este sistema es eficaz en sistemas ruidosos y muchas veces es mejor para detectar las fallas que el análisis de componentes principales (PCA) en fallas con pequeñas magnitudes.

Manjeevan Seera et al [16] presenta un modelo de computación suave híbrido el cual utiliza un modelo ANFIS y un árbol de regresión para la detección y el diagnóstico de las fallas de un motor, la detección la hace al clasificar las condiciones de falla fuera de línea en el motor de inducción. Los datos analizados son los armónicos de la señal de la densidad espectral de potencia, y las pruebas son realizadas bajo diferentes condiciones de funcionamiento.

P. Subbaraj et al [17] detecta y diagnóstica las fallas en la industria del cemento para la operación segura de la planta. Desarrolla un sistema Adaptivo de Inferencia Neuro-difuso "ANFIS", basa la detección de fallas y el diagnóstico en la válvula neumática utilizada en un sistema de enfriamiento en la pulverización de agua en la industria del cemento. El modelo ANFIS se utiliza para detectar y diagnosticar la ocurrencia de diversas fallas en la válvula neumática en el sistema de pulverización del agua más fría. A partir de los resultados de la simulación se observa que ANFIS se desempeña mejor que la ANN.

Zhimin Du et al [18] detecta fallas múltiples que se produjeron en los edificios en la calefacción, ventilación y sistemas de aire acondicionado, con lo cual se produce un mayor consumo de energía. Hace uso de una red neuronal junto con un análisis de Cluster para hacer el diagnóstico utilizó estas técnicas debido a que los sesgos de los sensores son difíciles de descubrir en los edificios reales.

Jeevanand Seshadrinath et al [19] realiza un análisis de vibraciones para el diagnóstico de fallas entre máquinas de inducción, utiliza una red neuronal probabilística. La extracción de características para modelar la PNN se hace significativa al utilizar un doble árbol de la transformada Wavelet, dando una verdadera representación del espacio de entrada.

Yousef Shatnawi et al [20] diagnostica las fallas en un motor de combustión interna para un tipo de movimiento especial alternativo y la rotación de la máquina, la cual, es un elemento esencial de todos los automóviles. Este trabajo propone un sistema eficaz y automatizado que analiza la señal de sonido emitida por un motor y esta se ocupa como el portador de información de las fallas, una transformada Wavelet es utilizada como la herramienta de extracción de características, y finalmente el análisis es extendido al utilizar una red neuronal de retropropagación.

Fafa Chen et al [21] propone un novedoso modelo de diagnóstico de fallas basado en una máquina de vectores de soporte multi-núcleo "MSVM" con un optimizador caótico de enjambre de partículas "CPSO" para el diagnóstico de fallas en rodamientos de rodillos, con pequeñas muestras, la optimización de enjambre se desarrolla en este estudio para determinar los parámetros óptimos para MSVM con una alta precisión y una gran capacidad de generalización. Por otra parte, los vectores de características para el diagnóstico de fallas se obtienen a partir de la señal de vibración que son pre-procesados por el dominio del tiempo, frecuencia y el modo de descomposición empírica.

III. PLANTEAMIENTO DEL PROBLEMA

III.1. Descripción del Problema

La problemática en la gran mayoría de las empresas es tener una mala calidad, fallas y descomposturas en sus productos o servicios, que puedan ser costosas y pongan en riesgo la vida de personas o equipos. El aseguramiento de la calidad se puede lograr al mantener las variables de operación constantemente en un modo de operación normal, con ello se garantiza el cumplir con las características que han sido requeridas por los clientes. Al presentar las variables monitoreadas mucha variación durante la operación o servicio, generan imperfecciones en el acabado final, estas variaciones hacen que no se cumplan con las características requeridas. Estas características no alcanzadas pueden ejemplificarse como: malos dimensionamientos, una mala terminación en el acabado, pérdida de resistencia entre otras características que afectan el producto o servicio final.

Debido a que las variables controlan la calidad del proceso o sistema, tienen un efecto considerable sobre la calidad final. Hoy en día se está tendiendo a monitorear no solo los atributos del producto o servicio, como comúnmente se hace en el control de proceso clásico. Ahora se está adoptando la tendencia por hacer un monitoreo de las variables de entrada del sistema, para poder identificar con mayor rapidez las condiciones en que se está trabajando y determinar el estado de operación de cada una de las variables involucradas, y con ello los productos o servicios puedan cumplir con las especificaciones requeridas.

III.2. Preguntas de Investigación

Observadas en el planteamiento del problema anterior se generan las siguientes preguntas de investigación

- ¿De acuerdo a sus parámetros fuera de rango cuál es el tipo de falla que se presenta en el proceso o sistema?
- ¿La falla es generada por uno o varios parámetros?
- ¿La falla está dada por la correlación que tienen estas variables entre sí?
- ¿En qué tiempo se presenta la falla?
- ¿El empleo de sistemas inteligentes tendrá un mejor desempeño que empleando los métodos tradicionales para la detección y diagnóstico de las fallas un sistema o proceso?
- ¿El método de sistemas inteligentes a utilizar (Lógica difusa, red neuronal probabilística y una máquina de soporte vectorial) podrá distinguir entre la presencia de ruido o la presencia de una falla?

III.3. Hipótesis

Mediante la implementación de una metodología que utilice técnicas inteligentes, se podrá detectar y diagnosticar las fallas presentes en los procesos o sistemas con variables altamente correlacionadas y con presencia de ruido, al determinar el tipo y tiempo donde se presentan las fallas con un alto porcentaje de exactitud.

III.4. Objetivos

III.4.1. Objetivo general

Implementar una metodología que utilice sistemas inteligentes basada en lógica difusa, Red neuronal probabilística y una Máquina de soporte vectorial, para facilitar la detección y diagnóstico de las fallas, garantizando el buen funcionamiento del mismo.

III.4.2. Objetivos específicos

- Localizar y clasificar los parámetros fuera de especificación cuando el producto final presente defectos.
- Crear una base de datos que contenga información del sistema en modo de operación normal y de las fallas que ya se han presentado en él.
- Trabajar con presencia de ruido dentro de los procesos.
- Implementar las herramientas de sistemas inteligentes seleccionadas para realizar la detección y diagnóstico de las fallas.

III.5. Justificación.

Las industrias proveedoras, productoras y de servicios tienen que garantizar a sus clientes la buena calidad de sus productos y servicios, es por ello que hoy en día, para estas industrias es de vital importancia el diagnosticar y detectar las fallas identificando los parámetros que están fuera de las condiciones normales de operación, que afectan el buen funcionamiento del proceso y con ello se vea perjudicado el acabado final del producto, el tiempo de entrega y los costos. Actualmente los avances en la ciencia computacional han permitido llevar los métodos de sistemas inteligentes a áreas que no se pensaban inicialmente, por esta razón se propone la implementación de una metodología que haga uso de las técnicas de sistemas inteligentes para la detección y diagnóstico de las fallas presentes en un proceso o sistema. De esta manera se facilitará la detección y diagnóstico de las fallas localizando los parámetros que están fuera de las especificaciones o del modo de operación normal. Adicionalmente ayudará a identificar el tipo de falla, lo que hará más fácil el posterior ajuste del proceso, y así, eliminar rápidamente las variaciones que se presentan al elaborar un producto.

III.6. Alcances y delimitaciones

III.6.1. Alcances

La metodología propuesta facilitará la detección e identificación de las condiciones en que se encuentra el proceso y localizará los parámetros que se encuentren fuera de las condiciones normales de operación.

III.6.2. Delimitaciones.

- La metodología propuesta está limitada a aquellas variables que son mensurables a través de sensores dispuestos en el sistema que se monitorea.
- La principal limitante es la cantidad y calidad de los datos.
- La metodología no puede trabajar con ausencia de uno o varios datos ya que ésta pérdida de información hará que la metodología no funcione.
- La metodología se limitará solo a realizar la detección y diagnóstico de las fallas y no se enfocará a recuperar el proceso como se muestra en la Figura III.1. en los cuadros resaltados Tal recuperación se podrá llevar a cabo ya sea manualmente o en un trabajo posterior para cerrar el ciclo de control.

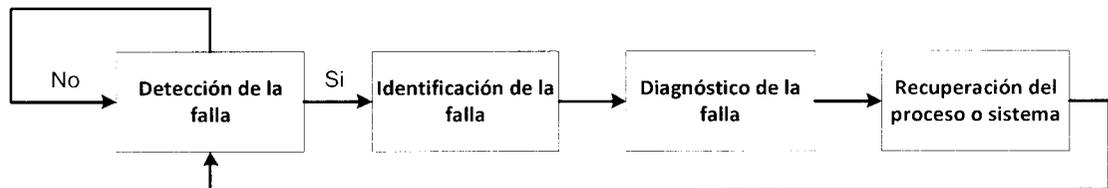


Figura III. 1 Alcances y delimitaciones de la metodología.

III.7. Resultados e impacto esperados

III.7.1. Resultados.

Implementar un sistema de monitoreo que realice la detección y diagnóstico de fallas en procesos y sistemas, donde en una primera etapa realice la detección de las condiciones de operación del proceso utilizando un sistema difuso y en una segunda etapa de diagnóstico, utilice una red neuronal probabilística junto con una máquina de vectores de soporte.

III.7.2. Impactos.

- Impacto científico:

La generación de una metodología capaz de detectar y diagnosticar las fallas presentes en los procesos o sistemas complejos al hacer uso de técnicas de sistemas inteligentes, "La cual sea expuesta en diferentes artículos y conferencias de divulgación científica".

- Impacto tecnológico:

La implementación en planta de la metodología propuesta para la detección y diagnóstico de fallas dentro de procesos o sistemas, la cual realice esta tarea de una forma más rápida y segura.

- Impacto social:

La metodología propuesta al facilitar la detección y el diagnóstico de las fallas, hará más seguros los procesos o sistemas y con ello se reducirán pérdidas humanas y económicas.

IV. MARCO TEÓRICO

IV.1. Conceptos básicos de detección y diagnóstico

A continuación se presentan los conceptos básicos necesarios del campo de la detección y diagnóstico de fallas tomados de [22]:

- **Falla:** Es una desviación no permitida de por lo menos una propiedad característica o parámetro de las condiciones aceptables / usuales / normales del sistema.
- **Fuera de servicio:** Una interrupción permanente de la habilidad de un sistema para realizar las funciones específicas para la cual es requerido.
- **Funcionamiento defectuoso:** Una irregularidad intermitente en la habilidad de un sistema para realizar las funciones específicas para la cual es requerido.
- **Error:** Una desviación entre un valor medido o calculado (de una variable de salida) y el valor verdadero, especificado o teóricamente correcto.
- **Perturbación:** Una entrada que actúa en un sistema produciendo una desviación temporal del estado actual, que no es considerado una falla.
- **Residuo:** Un indicador de falla, basado en una desviación, entre las medidas y cálculos basado en modelos de ecuaciones.
- **Síntoma:** Cambio de una cantidad respecto a un comportamiento normal.

- Detección de una falla: Determinar la presencia de una falla en un sistema y el tiempo de detección de la misma.
- Aislamiento de la falla: Después de la detección de una falla, determinar la clase y ubicación, pero sin identificar que falla es.
- Identificación de la falla: Después del aislamiento de la falla, determinar el tamaño y variación en tiempo del comportamiento de una falla. Se especifica claramente si es una falla conocida y cual es. En caso de ser desconocido se somete a análisis para determinar si es o no falla. En caso de ser una nueva falla se le identifica y clasifica.
- Diagnóstico de una falla: Después de la detección de la falla, se determina clase, tamaño, localización y su tiempo de detección. Incluye el aislamiento de la falla e identificación. Es la etapa donde se toma una decisión con respecto a la veracidad, gradualidad y gravedad de la falla.
- Monitoreo: Una tarea en tiempo-real continúa, que determina las condiciones de un sistema físico, almacenando información, reconociendo e indicando anomalías en el comportamiento.
- Supervisión: Monitoreo de un sistema físico, tomando acciones apropiadas para mantener el funcionamiento en el caso de las fallas.
- Protección: Medios por los que, si es posible, un comportamiento potencialmente peligroso del sistema se suprime, o medios por los que, las consecuencias de un comportamiento peligroso se evitan.
- Recuperación: Se eliminan las fallas presentes en el proceso para que esté pueda trabajar normalmente.

En los últimos años se han implementado varias alternativas para realizar la detección y diagnóstico en sistemas y/o procesos. Existen diferentes métodos estos pueden ser invasivos y no invasivos. Los métodos invasivos son aquellos que trabajan directamente con la planta como

ejemplo se tiene un modelo matemático y como ejemplo de métodos no invasivos son aquellos que utilizan distintas técnicas de análisis de datos de origen eléctrico, mecánico, químico, etc. [5]. Un ejemplo son las técnicas que utilizan los históricos del proceso, muchas de ellas son basadas en inteligencia artificial y los métodos estadísticos [10].

Los métodos de detección y diagnóstico de fallas se pueden clasificar en tres grandes grupos como lo muestra la Figura IV.1 de [23].



Figura IV. 1 Clasificación de métodos de diagnóstico (adaptado de [23])

Como se muestra en la Figura IV.1 los métodos de detección y diagnóstico son [7]:

- *Los basados en modelos cuantitativos:* Son métodos de detección de fallas que se basan en un modelo matemático del sistema. La ocurrencia de una falla es capturada por las discrepancias entre el comportamiento observado y aquel que está predicho por el modelo. Estos enfoques hacen uso de las estimaciones de estado, las técnicas de identificación de parámetros y las relaciones de paridad para generar residuos. La localización de fallas se basa en una interlínea de los grupos de componentes, que intervienen en cada una de las discrepancias detectadas. Sin embargo, a menudo es difícil y consume mucho tiempo para desarrollar modelos matemáticos precisos que caractericen a todos los fenómenos físicos que se producen en los procesos industriales.
- *Los basados en modelos cualitativos:* Son métodos de detección de fallas que utilizan razonamiento simbólico, el cual, generalmente combina diferentes tipos de conocimiento, como la teoría de grafos para analizar las relaciones entre las variables de un sistema. Una ventaja de estos métodos es que no es necesario un modelo explícito del sistema a diagnosticar. Los enfoques basados en el conocimiento, tales como los sistemas expertos pueden ser considerados como enfoques alternativos o complementarios de los modelos analíticos que no están disponibles.
- *Los modelos basados en el histórico de Proceso:* Son métodos de detección de fallas que sólo requieren una gran cantidad de datos históricos del proceso. Hay varias formas en las que estos datos pueden ser transformados y presentados como conocimiento previo de un sistema. Estas transformaciones se conocen como extracción de características, las que pueden ser cualitativas, como las utilizadas

por los sistemas expertos, y métodos de análisis de tendencias cualitativas o cuantitativas, como los que se utilizan en redes neurales, en un análisis de componentes principales o un análisis de mínimos cuadrados.

Desde otro punto de vista los modelos de detección y diagnóstico pueden ser clasificados como en línea y fuera de línea, en los que los sistemas en línea determina la presencia o ausencia de una falla mediante el uso de un algoritmo de diagnóstico y detección simultáneo con la planta y los sistemas fuera de línea utilizan un algoritmo que no requiere trabajar dentro de la planta y en tiempo real para realizar la detección y diagnóstico. Por otra parte los sistemas de diagnóstico también se pueden clasificar como sistemas de modelo libre o cerrado, en los cuales en los modelos libres no se utiliza un modelo matemático y los modelos cerrados utilizan un modelo matemático para diagnosticar la falla [3]. Los modelos cerrados o modelos matemáticos pueden ser muchas veces sustituidos por sistemas inteligentes como los son: los sistemas de lógica difusa, redes neuronales artificiales, algoritmos evolutivos y razonamiento probabilístico, estas técnicas muchas de las veces sustituyen o complementan los sistemas de control cuando los diseñadores de controladores, no pueden modelar matemáticamente el sistema o el modelo matemático del sistema es incompleto, debido a que es muy difícil de modelar. Los sistemas inteligentes son una nueva y eficaz manera de realizar esta tarea a partir del uso del histórico de los datos para el reconocimiento de patrones [24], estas técnicas hoy en día son parte de un conjunto de técnicas llamadas técnicas de computación suave.

A continuación se presenta una breve descripción de los métodos de monitoreo de procesos mostrados en la Figura IV.1.

IV.2. Métodos de monitoreo de procesos

Los métodos tradicionales de monitoreo de procesos consisten en límites de percepción y la detección de discrepancia. Los límites de detección generan alarmas cuando se cruza el umbral predefinido [4].

Existen ciertos métodos univariantes que no analizan la relación entre las variables y los métodos multivariantes que analizan la relación entre las variables. Uno de los métodos es el análisis de componentes principales (PCA). Este es empleado como una técnica óptima para reducir los términos evaluados por el sistema al obtener la varianza de los datos y la correlación entre las variables. Otra técnica es el análisis discriminante de Fisher (FDA) el cual al igual que el PCA es una técnica de reducción dimensional y se estudia dentro de la clasificación de parámetros está determina la porción en el espacio de observaciones que es más efectiva discriminar entre varias clases de datos, también están los mínimos cuadrados parciales (PLS) son métodos de descomposición de datos para maximizar la covarianza entre el bloque predictor (independiente) y el bloque (dependiente) para cada componente, en la misma línea se encuentra el análisis de correlación canónica su objetivo es buscar las relaciones entre dos grupos de variables y la validez de las mismas. Se diferencia del análisis de correlación múltiple en que éste sólo predice una variable dependiente a partir de múltiples independientes, mientras que la correlación canónica predice múltiples variables dependientes a partir de múltiples independientes. Las mediciones para estos métodos, pueden ser basadas totalmente en los datos históricos del proceso, al analizar los datos de entrada y salida. Los métodos de análisis son utilizados para detallar características del sistema o complementar a los modelos matemáticos [4].

Existen también otros métodos para el diagnóstico de fallas los más comúnmente utilizados para caracterizarlos son: la inclusión de residuos, la estimación de parámetros y estimación de estados. Los métodos más comúnmente utilizados para el análisis residual son denominados como métodos de redundancia analítica. En una situación preferente el residual será relativamente grande cuando la falla este presente, y pequeña con la presencia de perturbaciones, ruido y errores del modelo. Las tres principales formas de generar residuos son estimación de parámetros, observadores de estado y relaciones de paridad [4].

Otros métodos son los sistemas expertos basados en el razonamiento humano para diagnosticar las fallas, la experiencia puede ser modelada en términos de reglas. Las redes neuronales artificiales y los mapas auto organizados son considerados ejemplos de las técnicas de reconocimiento de patrones para el reconocimiento de datos y la clasificación de fallas.

IV.2.1. Métodos analíticos y basados en el conocimiento

Los métodos analíticos son aquellos métodos que son generados a partir de un modelo matemático para realizar la detección y el diagnóstico de las fallas, estos pueden ser clasificados como sigue [4]:

- *Estimación de parámetros.* Para la estimación de parámetros los residuos son diferencias entre los parámetros del modelo nominal y de los parámetros del modelo estimado, las desviaciones en los parámetros del modelo sirven para detectar y aislar las fallas.

- *Observadores de estado.* Son métodos de reconstrucción de la salida para el sistema, para las mediciones obtenidas y las estimadas son usadas como vectores de residuos.
- *Relaciones de paridad.* Estos métodos analizan la consistencia de los modelos matemáticos para el sistema con las mediciones. La relación de paridad está sujeta a una transformación dinámica lineal. Con la transformación de residuales se logra el diagnóstico y el aislamiento.

Estos métodos parten de las ecuaciones de estado del modelo y determinan que las fallas pueden ser encontradas en los actuadores en los sensores o en las plantas. A las ecuaciones de estado del modelo se le va agregando un vector de fallas, otro de ruido y perturbaciones para obtener una función de transferencia con la cual evaluar las condiciones del sistema. Como se muestra a continuación.

Partiendo de las ecuaciones de estado de un modelo para un sistema determinado [4]:

$$\mathbf{x}(t + 1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{1}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \tag{2}$$

Donde \mathbf{A} es una Matriz de Estado, \mathbf{B} es una Matriz de Entrada, \mathbf{C} es una Matriz de Salida, \mathbf{D} es una Matriz de Transmisión Directa, \mathbf{u} es un

vector que contiene cada una de las p entradas al sistema, y es un vector que contiene cada una de las q salidas del sistema, x es un vector que contiene cada una de las n variables de estado y t es un intervalo de tiempo.

Se le adicionan las fallas al sistema.

- Fallas en el actuador $\Delta u(t)$
- Fallas en el sensor $\Delta u(t)$
- Fallas en la planta $\Delta u_p(t)$

De la misma manera se les puede adicionar ruido o alguna perturbación.

- Ruido en el actuador $\delta u(t)$
- Ruido en el sensor $\delta u(t)$
- Ruido en la planta $\delta u_p(t)$

Al considerar esta adición de las fallas, el ruido y las perturbaciones las ecuaciones de estado quedan de la siguiente forma.

$$x(t+1) = (A + \Delta A)x(t) + (B + \Delta B)u(t) \quad (3)$$

$$y(t) = (C + \Delta C)x(t) + (D + \Delta D)u(t) \quad (4)$$

Lo anterior con el fin de poder analizar los residuos y poder hacer un comparativo para realizar una adecuada valoración de las condiciones de funcionamiento del sistema [4].

IV.2.2. Pretratamiento de datos, Métodos de estadística univariante, multivariante para realizar la detección y diagnóstico de fallas

La efectividad de las mediciones en el manejo de datos depende de la caracterización de las variaciones en los datos del proceso, hay dos tipos de variación, los casos comunes y los casos especiales en las que las variaciones son causadas por un ruido aleatorio y asociado a una falla atípica en el sistema [4].

IV.2.2.1. Pretratamiento de datos

Al extraer la información relevante de los datos de un eficiente monitoreo del proceso o sistema, es necesario obtener un conjunto de entrenamiento, que consta de evaluar con anterioridad los datos fuera de línea, para la implementación de un procesos en línea. Para realizar el monitoreo y desarrollo de las mediciones que se presentan dentro de los datos monitoreados es necesario dar un pretratamiento a los datos, el cual consiste en remover las variables (eliminar a aquellas variables menos significantes o irrelevantes dentro del proceso), autoescalar (estandarizar los valores analizados) y remover los valores atípicos (eliminar aquellos datos o estados que presente un comportamiento atípico del sistema y pueda influenciar el análisis del mismo).

El pretratamiento de los datos es relevante en los procesos de monitoreo, esto debido a que consiste en la remoción de variables, remoción de punto atípicos y la auto escalación, debido a que los datos de entrenamiento pueden contener variables que no tienen información relevante para el monitoreo del proceso, y estas variables pueden ser removidas para un futuro análisis, en los procesos, los datos a menudo necesitan ser escalados para evitar que algunas variables puedan dominar el monitoreo del proceso, especialmente los métodos basados en reducción dimensional como FDA y PCA. También es necesario en el manejo de datos remover los puntos atípicos ya que estos pueden ser errores en las mediciones y pueden tener una significativa relevancia en el análisis.

IV.2.2.2. Estadística univariante

La estadística univariante determina los límites que pueden ser usados para determinar el umbral para cada variable observada. Cuando es definido el umbral se convierte en un límite para las operaciones en control y si se presenta una violación en los datos en línea pueden indicar la presencia de una falla. Esta aproximación es vista típicamente usando las cartas de control y son referidas como límites de detección o valores de límites de control. Estos métodos ignoran la correlación entre las variables.

La efectividad del manejo de datos depende de las mediciones en la caracterización de los datos, la variación de los procesos. La efectividad se da al detectar las alteraciones en los procesos. Estas variaciones son fáciles de detectar con herramientas estadísticas al analizar el comportamiento del sistema.

IV.2.2.3. T^2 de hotelling

Es una de las herramientas más empleadas por la estadística multivariante para hacer el análisis de los procesos, ya que ésta, parte de los datos en el conjunto de entrenamiento, que consiste en m variables de observación y n observaciones para cada variable, pueden describirse en una matriz $X \in R^{n \times m}$ dada por:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (5)$$

Entonces la matriz de covarianza (es una matriz que contiene la covarianza entre los elementos de un vector, la cual indica el grado de variación conjunta de dos o más variables aleatorias) para el conjunto de entrenamiento es.

$$S = \frac{1}{n-1} X^T X \quad (6)$$

Se realiza una descomposición en eigenvalores (raíces características o también llamados valores propios de $X^T X$ de finidos por $\lambda_1, \lambda_2, \dots, \lambda_m$ que miden el grado de multicolinealidad en los datos) para la matriz S .

$$S = V \Lambda V^T \quad (7)$$

Donde Λ es una matriz diagonal de $m \times m$, cuyos elementos de la diagonal principal son los eigenvalores $\lambda_j (j = 1, 2, \dots, m)$ de S y V es una matriz ortogonal de $m \times m$, cuyas columnas son los eigenvectores de S . La proyección $y = V^T x$ da un vector de observación $x \in R^m$ desacoplado en el espacio de observación en un conjunto de variables no correlacionadas correspondientes a los elementos de y . La varianza de los elementos i^{th} de y es igual a las i^{th} de los eigenvalores en la matriz Λ . Asumiendo que S es invertible se tiene la definición

$$z = \Lambda^{-1/2} V^T x \quad (8)$$

El estadístico T^2 Hotelling's está dado por

$$T^2 = z^T z \quad (9)$$

El umbral para el estadístico T^2 de Hotelling's

Para obtener un apropiado umbral para el estadístico T^2 Hotelling basado en el nivel de significancia, α , este se debe determinar asumiendo las observaciones de una muestra al azar para una distribución normal multivariante. Se supone además un vector de medias y una matriz de covarianzas para caracterizar el modo normal de operación del proceso el cual es igual al vector actual de medias y a la matriz de covarianzas

respectivamente, entonces el estadístico T^2 describe una distribución X^2 con m grados de libertad.

$$T^2 = X^2(m) \quad (10)$$

El conjunto $T^2 \leq T_\alpha^2$ es un coeficiente de región elíptica en el espacio de observaciones.

Cuando la matriz de covarianza actual, para el estatus en control no es conocido en vez de estimarse a partir de la matriz de covarianza, las fallas pueden ser detectadas como observaciones fuera del conjunto de entrenamiento usando el umbral obtenido con:

$$T_\alpha^2 = \frac{m(n-1)(n+1)}{n(n-m)} F_\alpha(m, n-m) \quad (11)$$

Donde $F_\alpha(m, n-m)$ es la parte superior al $100_\alpha\%$ del punto crítico de la F -distribucion con m y $n-m$ grados de libertad, para obtener el nivel de significancia de la parte superior en control. Para obtener un límite más amplio en T_α^2 se obtienen dos limites cuando la cantidad de datos incrementa ($n \rightarrow \infty$).

Para el ejemplo de la matriz de covarianza donde los valores atípicos en el conjunto de entrenamiento se pueden detectar mediante [4]:

$$T_{\alpha}^2 = \frac{(n-1)^2(m(n-m-1))F_{\alpha}(m,n-m-1)}{n\left(1+\frac{m}{n-m-1}\right)F_{\alpha}(m,n-m-1)} \quad (12)$$

IV.2.2.4. Análisis de componentes principales

El análisis de componentes principales (PCA por su nombre en inglés Principal Component Analysis), es una técnica de reducción dimensional, también es una técnica que se puede emplear para el monitoreo de procesos industriales. El PCA puede producir una representación dimensional baja o pequeña para los datos que mejor generalicen la independencia de los datos para el conjunto de entrenamiento o para un completo análisis dimensional de las fallas. La estructura abstracta del PCA puede ser usada en la identificación de cualquiera de las variables responsables de una falla o de las variables que sean más afectadas por la falla. El PCA puede separar las observaciones en espacios o subespacios para la captura de tendencias en los procesos y en un subespacio que contenga esencialmente el ruido aleatorio.

El PCA determina los vectores ortogonales, llamados vectores de carga o eigenvectores, ordenados por la cantidad de varianza explicada por la dirección de los vectores de carga, dado un conjunto de entrenamiento para n observaciones y m variables del proceso en una matriz \mathbf{X} como se muestra en la Ecuación 5 [4] y [25].

La carga de los vectores es calculada al obtener los puntos estacionarios para el problema de optimización.

$$\max_{v \neq 0} = \frac{v^T X^T X v}{v^T v} \quad (13)$$

Donde $v \in \mathbf{R}^m$. Los puntos estacionarios para la ecuación anterior pueden ser obtenidos a través de la descomposición de los valores singulares (DVS) La cual es una factorización de los vectores ortogonales.

$$\frac{1}{\sqrt{n-1}} X = U \Sigma V^T \quad (14)$$

Donde $U \in \mathbf{R}^{n \times n}$ y $V \in \mathbf{R}^{m \times m}$ son matrices unitarias y la matriz $\Sigma \times \mathbf{R}^{n \times m}$ son constantes de los valores singulares no negativos de la magnitud decreciente a lo largo de la diagonal principal ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$), y cero en los elementos fuera de la diagonal. Los vectores de peso son vectores columna ortogonales en la matriz V , y la varianza para el conjunto de entrenamiento proyectado a lo largo de i^{th} columnas de V es igual para σ_i^2 . La resolución de la ecuación anterior es equivalente a resolver una descomposición de un eigenvalor para el ejemplo de la matriz de covarianza S presentada en la Ecuación 6 y 7.

Donde la matriz diagonal $\Lambda = \Sigma^T \Sigma \in \mathbf{R}^{m \times m}$ contiene los eigenvalores no negativos de magnitud decreciente ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$) y las i^{th} iguales a los eigenvalores del cuadrado del valor singular de i^{th} (*i.e.*, $\lambda_1 = \sigma_1^2$). Con el fin de capturar de manera óptima la variación de los datos que minimicen el efecto del ruido aleatorio corruptor en el PCA, correspondiente a los

vectores de peso de un valor singular de peso típico retenido. Para la clasificación se seleccionan las columnas de la matriz de peso $\mathbf{P} \in \mathbf{R}^{m \times a}$ que corresponde a los vectores de peso asociados con los primeros valores singulares de \mathbf{a} , la proyección de las observaciones en \mathbf{X} en un espacio dimensional menor está contenida en la matriz de valores.

$$\mathbf{T} = \mathbf{X}\mathbf{P} \quad (15)$$

Y la proyección de \mathbf{T} de nuevo en el espacio m -dimensional de observación.

$$\hat{\mathbf{X}} = \mathbf{T}\mathbf{P}^T \quad (16)$$

La diferencia entre \mathbf{X} y $\hat{\mathbf{X}}$ es la matriz residual \mathbf{E} .

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}} \quad (17)$$

El residual de la matriz captura la variación en el espacio abarcado por las observaciones para los vectores de peso asociados con el valor singular más pequeño $m - a$. La expansión del subespacio para $\hat{\mathbf{X}}$ y \mathbf{E} es llamada valores de espacio y residual de espacio, respectivamente. El subespacio contenido en la matriz \mathbf{E} tiene una menor proporción de la señal de ruido, y la eliminación de este espacio para \mathbf{X} es una mayor representación de los procesos, $\hat{\mathbf{X}}$.

Reducción de orden

Es comúnmente aceptado y con cierta justificación teórica que la porción para el espacio correspondiente al PCA es la descripción de los valores singulares de pesos o variación de los estados que ocurre en el proceso, y la porción correspondiente al espacio de los valores singulares más pequeños describe el ruido aleatorio. Para una apropiada determinación de los vectores de carga \mathbf{a} , para mantener el modelo de PCA, la variación sistemática puede ser desacoplada para las variables aleatorias, y los dos tipos de variación pueden ser monitoreados por separado, existen varias técnicas para determinar el valor de la reducción de orden \mathbf{a} , pero no hay una técnica dominante, los métodos para determinar a son:

- La prueba del porcentaje de varianza.
- La prueba *SCREE*
- Análisis paralelo
- El estadístico *PRESS*.

El método de la prueba del porcentaje de varianza determina \mathbf{a} para calcular el número más pequeño para los vectores de carga necesarios para explicar un mínimo porcentaje para el total de la varianza.

La prueba *SCREE* asume que la varianza, σ_i^2 , correspondiente al ruido aleatorio forma un perfil lineal.

El análisis paralelo determina la dimensionalidad para comparar el perfil de la varianza para obtener la independencia de las variables observadas.

La dimensión para los valores de espacio pueden también ser determinados usando una validación cruzada producida con el estadístico de la predicción de la suma de residuales de los cuadrados (*PRESS*).

$$PRESS(i) = \frac{1}{mn} \|X - \hat{X}\|_F^2 \quad (18)$$

Donde *i* es el número para los vectores de carga necesarios para calcular \hat{X} y $\|\cdot\|_F$ es la norma de Frobenius (la raíz cuadrada de la suma de todos los elementos). Para la implementación de esta técnica la etapa de entrenamiento es dividida en dos grupos. El estadístico *PRESS* para un grupo es calculado basado en varias dimensiones para los valores de espacio *i*, usando todos los otros grupos. Esto es repetido para cada grupo, y el valor de *i* se asocia con el mínimo promedio del estadístico *PRESS* para determinar la dimensión de los valores de espacio [4] y [25].

IV.2.2.5. Discriminante de Fisher

Para realizar el diagnóstico de fallas, con los datos que ya han sido obtenidos directamente de la planta cuando hay presencia de fallas ya conocidas y se conoce el modo normal de operación estas se pueden clasificar en categorías. Esto cuando cada una de las clases (se determina

una clase cuando se clasifica a un individuo en un grupo con la información de un conjunto de variables que se observan previamente en un conjunto de individuos de los que se sabe que están bien clasificados en uno de los grupos) contiene datos representativos de cada tipo de falla, así como del modo normal de operación del sistema. Para realizar una clasificación en diferentes clases de los comportamientos conocidos con anterioridad de la planta o sistema, se puede utilizar el análisis discriminante de Fisher FDA (por su nombre en inglés Fisher Discriminant Analysis), el cual es una técnica de reducción lineal de la dimensionalidad de los datos. La cual maximiza de manera óptima la separación entre las clases. Esto mediante la transformación de un conjunto de vectores, ordenados en términos que maximicen la dispersión entre clases y reduzcan al mínimo la dispersión en cada clase.

Sea n el número de observaciones en un sistema, m el número de mediciones o número de variables, p el número de clases, y n_j el número de observaciones en las clases j^{th} . Para representar el vector de las variables de medición para las observaciones i^{th} para las x_i . Si los datos de entrenamiento para todas las clases en la matriz $X \in \mathbf{R}^{n \times m}$, a continuación, la transpuesta i^{th} de la fila X será un vector columna x_i .

Para poder entender el análisis discriminante de Fisher, es necesario definir las matrices que cuantifican el total de dispersión, dentro de las clases, y entre las clases. Como una matriz de dispersión total.

$$S_t = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T \quad (19)$$

Donde \bar{x} es el vector de medias totales se determina por:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (20)$$

Con X_j definida como el conjunto de vectores x_i que pertenece a la clase j , para la matriz de dispersión en cada clase j es

$$S_j = \sum_{x_i \in X_j} (x_i - \bar{x}_j)(x_i - \bar{x}_j)^T \quad (21)$$

Donde \bar{x}_j es el vector de medias para la clase j :

$$\bar{x}_j = \frac{1}{n_j} \sum_{x_i \in X_j} x_i \quad (22)$$

La matriz de dispersión entre la clase es:

$$S_w = \sum_{j=1}^p S_j \quad (23)$$

Y la matriz de dispersión entre las clases es:

$$S_b = \sum_{j=1}^p n_j (x_j - \bar{x})(x_j - \bar{x})^T \quad (24)$$

La matriz total de dispersión es igual a la suma de la matriz de dispersión entre las clases y la matriz de dispersión de la clase.

$$S_t = S_b + S_w \quad (25)$$

El objetivo para el primer vector FDA es maximizar la dispersión entre las clases mientras minimiza la dispersión entre cada una de las clases [4].

IV.2.2.6. Análisis de correlación canónica

El análisis de correlación canónica implica la colección de variables en dos conjuntos, un conjunto X y un conjunto Y . El objetivo es encontrar una combinaciones lineales $W = a'X$ y $V = b'Y$ tal que W y V tienen una posible correlación mayor. Estas combinaciones lineales pueden dar una idea de la relación entre los dos conjuntos de variables.

El análisis de correlación canónica tiene grandes propiedades similares al análisis de componentes principales. Sin embargo mientras que el PCA considera interrelaciones en un conjunto de variables. La relación canónica está enfocada al análisis de la relación entre los dos conjuntos de variables.

Una forma de ver el análisis de correlación canónica es como una extensión de regresión múltiple. En el análisis de regresión múltiple divide las variables en un conjunto de \mathbf{x} que contiene variables q y a los y conjuntos que contienen $p = 1$ variables. La solución consiste en encontrar la regresión combinación lineal $\mathbf{a}'\mathbf{X}$ que es el más altamente correlacionada con \mathbf{y} .

El análisis de correlación canónica del conjunto \mathbf{Y} contiene variables $p \geq 1$ y busca vectores \mathbf{a} y \mathbf{b} para que la correlacion entre $\mathbf{a}'\mathbf{X}$ y $\mathbf{b}'\mathbf{Y}$ sea maximizada, si \mathbf{X} se interpreta como "causa" de \mathbf{Y} , entonces $\mathbf{a}'\mathbf{X}$ puede ser llamado el predictor y $\mathbf{b}'\mathbf{Y}$ el criterio predecible. Sin embargo, no hay ninguna suposición de la asimetría causal en las matemáticas del análisis de correlación canónica; \mathbf{X} y \mathbf{Y} son tratados de forma simétrica [25].

El análisis de correlación se realiza a partir de las siguientes formulas. Dados dos conjuntos de vectores de datos en las formas [26].

$$W_m = a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mp}X_p \quad (26)$$

$$V_m = b_{m1}Y_1 + b_{m2}Y_2 + \dots + b_{mp}Y_p \quad (27)$$

De manera que los vectores \mathbf{X} y \mathbf{Y} indican la correlación máxima, es decir:

$$\max_{\mathbf{X}, \mathbf{Y}} \mathbf{a}' \Sigma \mathbf{X} \mathbf{Y} \mathbf{b} \quad (28)$$

Sujeto a

$$\max_{X,Y} \mathbf{a}' \Sigma \mathbf{X} \mathbf{X} \mathbf{a} = 1 \quad (29)$$

$$\max_{X,Y} \mathbf{b}' \Sigma \mathbf{Y} \mathbf{Y} \mathbf{b} = 1 \quad (30)$$

Dónde: $\Sigma \mathbf{X} \mathbf{Y}$ representa la matriz de covarianza entre clases y $\Sigma \mathbf{X} \mathbf{X} \mathbf{b}$, $\Sigma \mathbf{Y} \mathbf{Y} \mathbf{a}$ representan las matrices de covarianza intra-clase. El uso de los multiplicadores de Lagrange, optimiza la proyección de los vectores propios correspondientes a los mayores valores propios de las siguientes generalidades:

$$\Sigma \mathbf{X} \mathbf{X}^{-1} \Sigma \mathbf{X} \mathbf{Y} \Sigma \mathbf{Y} \mathbf{Y}^{-1} \Sigma \mathbf{Y} \mathbf{X} = \lambda_X \Sigma \mathbf{X} \mathbf{X} \mathbf{a} \quad (31)$$

$$\Sigma \mathbf{Y} \mathbf{X}^{-1} \Sigma \mathbf{Y} \mathbf{X} \Sigma \mathbf{X} \mathbf{X}^{-1} \Sigma \mathbf{X} \mathbf{Y} = \lambda_Y \Sigma \mathbf{X} \mathbf{X} \mathbf{b} \quad (32)$$

Donde \mathbf{a} y \mathbf{b} son los valores característicos de la matriz \mathbf{X} y de la matriz \mathbf{Y} .

Ahora al imponer la condición $\mathbf{a}' \Sigma \mathbf{X} \mathbf{X} \mathbf{a} = 1$ y $\mathbf{b}' \Sigma \mathbf{Y} \mathbf{Y} \mathbf{b} = 1$ se tiene.

$$\mathbf{W}_m = \frac{a_{m1}}{a' \Sigma \mathbf{X} \mathbf{X} \mathbf{a}} \mathbf{X}_1 + \frac{a_{m2}}{a' \Sigma \mathbf{X} \mathbf{X} \mathbf{a}} \mathbf{X}_2 + \dots + \frac{a_{mp}}{a' \Sigma \mathbf{X} \mathbf{X} \mathbf{a}} \mathbf{X}_p \quad (33)$$

$$\mathbf{V}_m = \frac{b_{m1}}{b' \Sigma \mathbf{Y} \mathbf{Y} \mathbf{b}} \mathbf{Y}_1 + \frac{b_{m2}}{b' \Sigma \mathbf{Y} \mathbf{Y} \mathbf{b}} \mathbf{Y}_2 + \dots + \frac{b_{mp}}{b' \Sigma \mathbf{Y} \mathbf{Y} \mathbf{b}} \mathbf{Y}_p \quad (34)$$

Y con ello se explica la relación de las variables de entrada con respecto a las variables de salida en un sistema. [25] y [27].

IV.2.2.7. Mínimos cuadrados parciales.

Los Mínimos cuadrados parciales o PLS por su nombre en inglés "Partial least squares", también conocido como proyección a las estructuras latentes, es una técnica de reducción dimensional para maximizar la covarianza entre el predictor independiente de la matriz \mathbf{X} y la matriz de predicción \mathbf{Y} para cada componente en el espacio de reducción. Una popular aplicación del PLS es cuando se selecciona una matriz \mathbf{Y} para contener sólo datos de la calidad del producto, que puede incluso incluir mediciones sin conexión de datos, y la Matriz \mathbf{X} para contener todas las otras variables del proceso. Puede ser utilizado en línea para predecir la calidad de los productos.

El PLS requiere una matriz $\mathbf{X} \in \mathbf{R}^{n \times m}$ y una matriz $\mathbf{Y} \in \mathbf{R}^{n \times p}$, donde m es el número de variables de predicción (el número de mediciones en cada observación), n es el número total de observaciones en el conjunto de entrenamiento, y p es el número de observaciones en la variable \mathbf{Y} . Cuando \mathbf{Y} se selecciona solo para contener las variables de calidad del producto. Entonces p es el número de fallas en las clases

Primero se aplica un pretratamiento de los datos, de modo que \mathbf{X} y \mathbf{Y} sean medias centrales y escaladas. La matriz \mathbf{X} es descompuesta en una matriz característica $\mathbf{T} \in \mathbf{R}^{n \times \alpha}$ y una matriz de carga $\mathbf{P} \in \mathbf{R}^{m \times \alpha}$, donde α es la

reducción de orden del componente PLS, a de más de una matriz residual $E \in R^{n \times m}$.

$$X = TP^T + E \quad (35)$$

La matriz TP^T puede ser expresada como la suma de los productos de los vector característicos t_j (de las j^{th} columnas de T) y los vectores de carga p_j (de las j^{th} columnas de P).

$$X = \sum_{j=1}^{\alpha} t_j p_j^T + E \quad (36)$$

De manera similar, Y es descompuesta en su matriz característica $U \in R^{n \times \alpha}$, y una matriz de carga $Q \in R^{p \times \alpha}$, además de un residual $\tilde{F} \in R^{n \times p}$.

$$Y = UQ^T + \tilde{F} \quad (37)$$

La matriz UQ^T puede ser expresada como la suma de los productos de los vectores característicos u_j (de las j^{th} columnas de U) y los vectores de carga q_j (de las j^{th} columnas de Q)

$$Y = \sum_{j=1}^{\alpha} u_j q_j + \tilde{F} \quad (38)$$

Las descomposiciones anteriores tienen la misma forma que la utilizada en el PCA. Las matrices \mathbf{X} y \mathbf{Y} son representadas como la suma de una serie de matrices de rango uno. Si α se iguala a un $\min(m, n)$ se reduce el ruido y la colinealidad. El objetivo de PLS es determinar los pesos y características de los vectores que están correlacionados con \mathbf{Y} mientras se describe una gran cantidad de variación en \mathbf{X} .

PLS retrocede la \mathbf{Y} del vector característico estimado $\hat{\mathbf{u}}_j$ al \mathbf{X} vector característico de \mathbf{t}_j para:

$$\hat{\mathbf{u}}_j = \mathbf{b}_j \mathbf{t}_j \quad (39)$$

Donde \mathbf{b}_j es el coeficiente de regresión. En la matriz de forma, esta relación puede ser escrita como:

$$\hat{\mathbf{U}} = \mathbf{T}\mathbf{B} \quad (40)$$

Donde $\mathbf{B} \in \mathbb{R}^{\alpha \times \alpha}$ es la diagonal de la matriz de regresión con $\mathbf{B}_{jj} = \mathbf{b}_j$, y $\hat{\mathbf{U}}$ tiene $\hat{\mathbf{u}}_j$ como sus columnas. Sustituyendo $\hat{\mathbf{U}}$ para $\hat{\mathbf{u}}_j = \mathbf{b}_j \mathbf{t}_j$ por \mathbf{U} en $\hat{\mathbf{U}} = \mathbf{T}\mathbf{B}$, y teniendo en cuenta que este va a modificar la matriz residual, se tiene:

$$\mathbf{Y} = \mathbf{T}\mathbf{B}\mathbf{Q}^T + \mathbf{F} \quad (41)$$

Donde \mathbf{F} es la predicción de la matriz de error [4].

IV.2.2.8. Distancia estadística.

Algunas herramientas ayudan a dar un pretratamiento a los datos para poder identificar el comportamiento normal de operación del comportamiento en falla son la obtención de métricas como los son:

- Distancia Euclidiana

La distancia euclidiana se define como la distancia entre dos puntos. Tal distancia se deduce con el Teorema de Pitágoras. Se trata de utilizar un espacio bidimensional, para obtener la distancia entre dos puntos P_1 y P_2 . Muestra la distancia Euclidiana en coordenadas rectangulares. La ubicación de los puntos P_1 y P_2 es (x_1, y_1) y (x_2, y_2) , respectivamente [28]:

$$d_e(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (42)$$

- Distancia de Mahalanobis

La distancia de Mahalanobis es un método estadístico para medir la similitud de dos conjuntos de datos. A diferencia de la distancia Euclidiana, la distancia de Mahalanobis considera las correlaciones entre los datos y es invariante en escala. Se define por la siguiente ecuación [28].

$$Dm(\vec{x}_1, \vec{x}_2) = \sqrt{(\vec{x}_2 - \vec{x}_1)^T S^{-1} (\vec{x}_2 - \vec{x}_1)} \quad (43)$$

Donde $(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2)$ son variables aleatorias con la misma distribución y \mathbf{S}^{-1} es la matriz de covarianzas.

IV.2.3. Computación Suave Para Realizar la Detección y Diagnóstico

La clasificación de parámetros es un problema muy importante hoy en día para el análisis de datos. La clasificación basada en los datos históricos de los procesos ya ha sido estudiada ampliamente como posible solución a la adquisición de conocimientos o la extracción del conocimiento. Uno de los principales temas a abordar en los modelos de clasificación es la mejora de la precisión de la clasificación. Para ello, muchos investigadores recientemente han puesto una atención considerable en la tarea de combinación de los métodos clasificadores. Dentro de ellos, hay tres enfoques básicos. El de un clasificador conjunto, la selección del clasificador, y la fusión paramétrica. Los algoritmos clasificadores en conjunto suelen combinar los resultados de varios clasificadores individuales. Sin embargo, la mayoría de los algoritmos de conjuntos no tienen en cuenta la experiencia local de cada clasificador. Esto es útil para descomponer un problema complejo en sub-problemas más simples y resolver cada sub-problema, uno por uno, en lugar de aprender la relación global entre las variables de entrada y la variable de destino. Numerosos enfoques relacionados con la selección de un clasificador por región local se han desarrollado [29]. En la Figura IV.2 se muestra las distintas técnicas que componen a la computación suave.

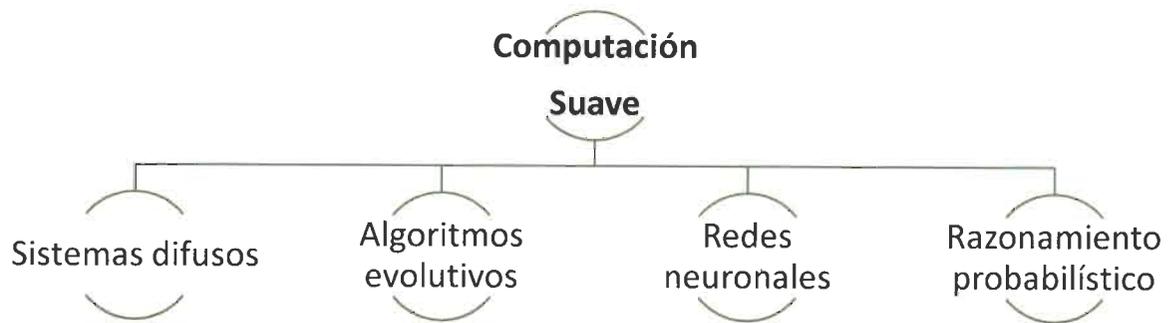


Figura IV. 2 *Métodos de computación suave para detectar y diagnosticar fallas.*

A continuación se da una breve descripción de las características de cada uno de los campos de la computación suave.

IV.2.3.1. Sistemas difusos

Un sistema difuso es una metodología que fue diseñada para tener disponibles habilidades de los expertos. Esta metodología trata de emular los patrones de razonamiento del experto. Son muy populares ya que los conceptos de la lógica difusa, son reglas del tipo *SI-ENTONCES* y métodos de inferencia difusa. Los sistemas de inferencia difusa, actualmente han encontrado diversas aplicaciones exitosas dentro de una gran variedad de áreas tales como el control automático, la clasificación de datos, el análisis de decisiones, los sistemas expertos, la predicción de series de tiempo, la

robótica y en el reconocimiento de patrones. Sin embargo, las aplicaciones más significativas de los sistemas difusos se han concentrado específicamente en el área del control automático [30].

IV.2.3.2. Algoritmos evolutivos

Los algoritmos evolutivos es una concepción de la ciencia de la computación en una forma muy convencional de la evolución como un proceso de optimización que podría ser simulado para resolver problemas de ingeniería que antes parecían intratables. La idea fundamental de este proceso consiste en evolucionar una población de posibles soluciones para un problema dado, usando operadores inspirados en la variación genética y en la selección natural. La evolución es, en efecto, un método de búsqueda entre una inmensa cantidad de soluciones posibles. En el ámbito biológico el enorme conjunto de posibilidades lo constituyen el conjunto de posibles secuencias genéticas, y las soluciones deseables son los organismos altamente aptos, es decir, los organismos mejor dotados para sobrevivir y reproducirse en su ambiente.

Actualmente existen tres vertientes que se inspiran en los principios del neodarwinismo: las estrategias evolutivas, la programación evolutiva y los algoritmos genéticos. Conjuntamente, a estas técnicas se les denomina computación evolutiva [31].

IV.2.3.3. Redes neuronales

Una red neuronal artificial (ANN por su nombre en inglés "Artificial Neuronal Network") es un término que se utilizó para describir un modelo computacional del cerebro biológico. Este es un sistema de modelado libre del cerebro humano. Está compuesto de un conjunto de unidades de procesamiento interconectadas simplemente, que se combinan para producir una señal o para resolver un problema determinado sobre una señal de entrada que se recibe. Donde la unidad de procesamiento que se encuentran interconectadas tiene ganancias ajustables, que se ajusta lentamente a través de iteraciones influenciadas por los patrones de entrada-salida dadas a la ANN [32].

Una ANN es un paradigma de procesamiento de la información que fue inspirada en las vías biológicas del sistema nervioso, tales como el cerebro, la información del proceso en una ANN es un sistema que tiene ciertas características de rendimiento en común con las redes neuronales biológicas. Básicamente, es un sistema que se encarga de muchas señales de entrada, las procesa, y la salida de ellos, es para resolver una tarea para la que ha sido entrenada. En este sentido, las ANN se utilizan como una herramienta para resolver problemas de procesamiento de patrones [8].

IV.2.3.4. Razonamiento probabilístico

El razonamiento probabilístico se puede definir como la construcción de un modelo de razonamiento que puede trabajar con incertidumbre, y se utilizan para poder capturar el conocimiento incierto de una manera natural y eficiente en muchas situaciones prácticas. En muchos casos estos modelos son representados por las redes bayesianas las cuales ayudan a obtener información de una probabilidad cuantitativa. Que es dada por los expertos y decidir más fácilmente las influencias de este conocimiento en un dominio dado. Está dado como una distribución conjunta completa, en la que muchas distribuciones condicionales se pueden representar de forma compacta por familias de las distribuciones canónicas.

En muchos casos si se tiene una gran cantidad de inferencia debido a los modelos de gran tamaño de las redes, es esencial tener en cuenta los métodos de aproximación. Los cuales se pueden describir como métodos aleatorios de muestreos que ayudan a obtener una respuesta. En donde estas mejoran su eficiencia al tener un mayor número de muestras a su disposición, es decir en casos muy particulares se realiza el uso de técnicas de muestreo en general. Estos métodos pueden ser expresados como equivalentes a las redes bayesianas [33].

IV.3. Sistemas inteligentes empleados en la metodología para realizar la detección y diagnóstico de fallas

IV.3.1. Lógica difusa

Lógica difusa es una técnica que es definida por [34] como un tipo especial de control no lineal. Esto es análogo a un sistema lineal y de control en las que la palabra "lineal" es un adjetivo técnico utilizado para especificar los sistemas de control. La palabra "difuso", esencialmente, lo que quiere destacar, es que a pesar de las características de la teoría de un sistema difuso que es caracterizada por no ser clara, la teoría en sí es precisa. Hay dos tipos de justificación de la teoría de sistemas difusos: En la primera se considera que en el mundo real, es demasiado complicado el obtener descripciones precisas, por lo tanto la aproximación a la falta de claridad se debe de obtener e introducir con el fin de obtener respuestas de manera razonable, sin embargo. A medida que se adentra en la era de la información, el conocimiento humano se hace cada vez más importante. Por lo cual se necesita una teoría para formular el conocimiento humano de forma sistemática en los sistemas de ingeniería, junto con otra información como los modelos matemáticos y mediciones sensoriales.

La segunda justificación caracteriza a las características, justifica la existencia como una rama independiente de ingeniería. Como un principio general, una buena teoría de la ingeniería debe ser capaz de hacer uso de toda la información disponible con eficacia. Para muchos sistemas prácticos, la información importante proviene de dos fuentes: una fuente proviene de los expertos humanos cuando describen sus conocimientos sobre el sistema en las lenguas naturales y la otra es mediante mediciones sensoriales y modelos matemáticos que se derivan de acuerdo a las leyes físicas. Una

tarea importante, por lo tanto, es la combinación de estos dos tipos de información en diseños de sistemas.

La esencia de los sistemas difusos es que son sistemas basados en el conocimiento y en normas. El corazón de un sistema difuso es una base de conocimientos que consiste en las denominadas reglas difusas *SI-ENTONCES*. Una regla difusa *SI-ENTONCES* es una instrucción en el lenguaje.

Por ejemplo, la siguiente es una regla difusa *SI-ENTONCES*: *SI* la velocidad de un coche es alta, *ENTONCES* se está aplicando una mayor fuerza al acelerador.

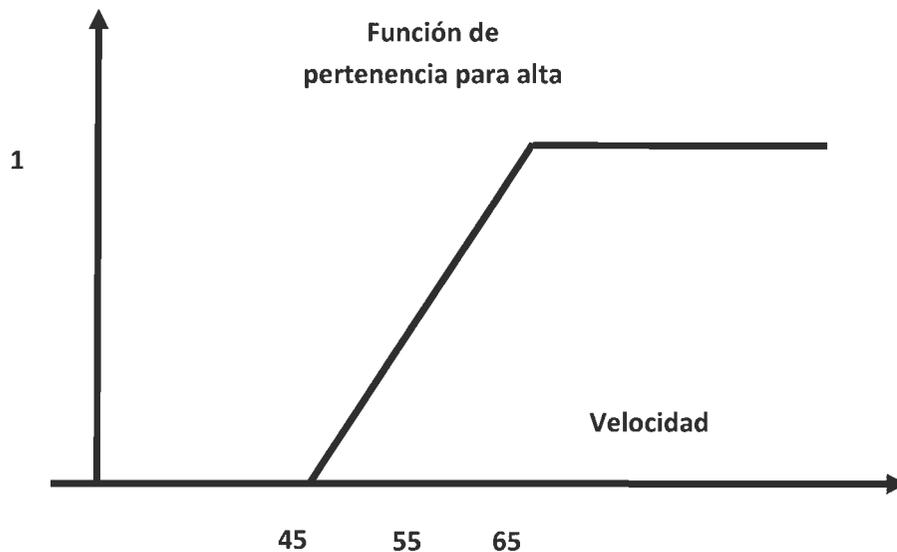


Figura IV. 3 *Función de pertenencia a la "alta", donde el eje horizontal representa la velocidad del vehículo y el eje vertical representa el valor de pertenencia a la "alta" (tomado de [34]).*

Por lo tanto un sistema difuso se describe de una manera más simple en la siguiente Figura IV.4.

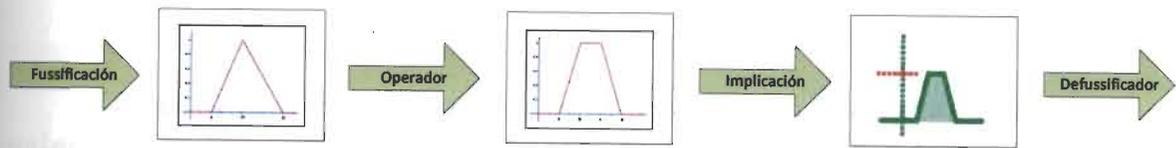


Figura IV. 4 Proceso difuso. El cual describe la metodología a seguir para ocupar esta herramienta. (tomado de Toolbox de MatLab®).

IV.3.1.1. Funciones de membresía

Los datos son fuzzificados a partir de la selección de una función de membresía. La cual se define como un conjunto difuso A sobre un universo discurso X de la forma $\mu_A: X \rightarrow [0, 1]$, donde a cada elemento de X le corresponde un valor entre 0 y 1 . Este valor, llamado valor de pertenencia o grado de pertenencia, representa el grado en el que el elemento de X pertenece al conjunto difuso A . Las funciones de membresía permiten representar gráficamente un conjunto difuso. En el eje " x " se representa el universo discurso, mientras que en el eje " y " se sitúan los grados de pertenencia en el intervalo. Para construir funciones de pertenencia se suelen utilizar funciones sencillas, ya que al estar definiendo conceptos borrosos el uso de funciones complejas no aporta mayor precisión. En muchos casos la función más utilizada es la función triangular debido a que es la más simple de utilizar.

$$\mu_{A(x)} = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{m-a} & \text{si } a < x \leq m \\ \frac{b-a}{b-m} & \text{si } a < x < m \\ 0 & \text{si } x \geq b \end{cases} \quad (44)$$

Donde a se define como el valor menor, b es el valor más alto y m es valor medio, x es cualquier valor de entrada en las funciones de pertenencia.

IV.3.1.2. Operadores difusos

Posteriormente a la fuzzificación se aplica un operador difuso, el cual es una operación básica sobre conjuntos difusos. En la secuela, que B son conjuntos difusos definidos en el mismo universo discurso que A . Por lo tanto la igualdad, el complemento, la unión e intersección de dos conjuntos difusos A y B se definen como sigue.

Donde A y B son iguales si y sólo si $\mu_{A(x)} = \mu_{B(x)}$ o todo $x \in U$. Cuando B contiene A , denotado por $A \subset B$, si y sólo si $\mu_{A(x)} \leq \mu_{B(x)}$ para todo $x \in U$. El complemento de A es un conjunto difuso A en U cuya función de pertenencia se define como:

$$\mu_{\bar{A}(x)} = 1 - \mu_{A(x)} \quad (45)$$

La unión de A y B es un conjunto difuso en U , denotado por la $A \cup B$, cuyos miembros de la función se define como:

$$\mu_{A \cup B(x)} = \max[\mu_{A(x)}, \mu_{B(x)}] \quad (46)$$

La intersección de A y B es un conjunto difuso $A \cap B$ en U con función de membrecía.

$$\mu_{A \cap B}(x) = \min[\mu_{A(x)}, \mu_{B(x)}] \quad (47)$$

IV.3.1.3. Implicaciones difusas

El conocimiento humano es representado en términos difusos con las reglas *SI-ENTONCES*. Con el fin de entender estas reglas se crea la composición difusa (una relación difusa que representa el grado de presencia o ausencia, de interacción o interconexión entre elementos de dos o más conjuntos difusos) en una interpretación de las relaciones difusas. En el cálculo proposicional clásico la expresión *SI p ENTONCES q* es escrita $p \rightarrow q$ con la implicación \rightarrow Considerándolo como un conector, donde p y q son variables proposicionales y equivalente a por lo tanto la implicación se define como un conjunto de proposiciones *SI-ENTONCES*.

$$\bar{p} \vee q \text{ y } (p \wedge q) \vee \bar{p} \quad (48)$$

Puesto que hay una amplia variedad de composiciones difusas de unión, intersección. También existe un gran número de diferentes implicaciones difusas *SI-ENTONCES* que se han propuesto en la literatura pero una de las mayormente utilizadas son las implicaciones Mamdani.

Dónde: la regla difusa *SI-ENTONCES* se interpreta como una relación difusa Q_{MM} o Q_{MP} en $U \times V$ con la función de membrecía.

$$\mu_{MM}(x, y) = \min[\mu_{FP_1}(x) + \mu_{FP_2}(y)] \quad (49)$$

$$\mu_{MP}(x, y) = \mu_{FP_1}(x)\mu_{FP_2}(y) \quad (50)$$

IV.3.1.4. Defussificador

Para obtener una solución a un problema de decisión, lo que se quiere es obtener como salida es un número y no un conjunto difuso. Por tanto, se debe de transformar el conjunto difuso obtenido en un número. Uno de los métodos más utilizados es el del centroide, que calcula el centro del área definida por el conjunto difuso obtenido. El cálculo se realiza con la siguiente formula [34].

$$g = \frac{\sum_{i=1}^n x_i \cdot u(x_i)}{\sum_{i=1}^n u(x_i)} \quad (51)$$

IV.3.2. Red Neuronal Probabilística.

La red neuronal probabilística se compone de dos capas y parte de un vector de datos \mathbf{p} expresado de la siguiente manera.

$$\{p_1, p_2, \dots, p_Q\} \quad (52)$$

Donde se tiene una matriz \mathbf{W}' de pesos transpuestos y un vector de tendencia \mathbf{b} para una capa cuando:

$$\mathbf{W}' = \begin{bmatrix} w_1^T \\ w_2^T \\ \vdots \\ w_s^T \end{bmatrix} = \begin{bmatrix} p_1^T \\ p_2^T \\ \vdots \\ p_Q^T \end{bmatrix}, \mathbf{b}' = \begin{bmatrix} R \\ \vdots \\ R \end{bmatrix} \quad (53)$$

La fila de \mathbf{W}' representa un vector prototipo que se desea reconocer y cada elemento de \mathbf{b}' es un conjunto igual para el número de elementos en cada vector de entrada (R). Así la salida para la primera capa es:

$$\mathbf{a}^1 = \mathbf{W}'\mathbf{p} + \mathbf{b}' = \begin{bmatrix} p_1^T \mathbf{p} + R \\ p_2^T \mathbf{p} + R \\ \vdots \\ p_Q^T \mathbf{p} + R \end{bmatrix} \quad (54)$$

La segunda capa es una capa competitiva. Las neuronas en esta capa, se inicializan con las salidas de la capa anterior hacia delante. Que inician el reconocimiento entre los patrones prototipo y el vector de entrada. Estas neuronas compiten con cada una de las otras para determinar un ganador. Después, una sola neurona tendrá una salida distinta a cero. La neurona ganadora indica en que categoría se está presentando la entrada para la red (cada vector prototipo representa una categoría).

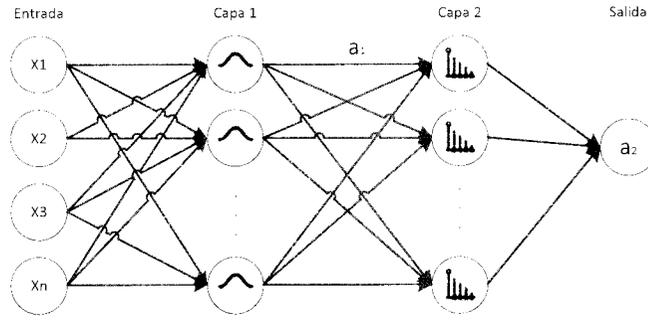


Figura IV. 5 Red Neuronal Probabilística (tomada de [35]).

La salida \mathbf{a}^1 de la primera capa es usada para iniciar la segunda capa

$$\mathbf{a}^2(0) = \mathbf{a}^1 \quad (55)$$

La salida de esta segunda capa es actualizada de acuerdo a la siguiente relación de recurrencia.

$$\mathbf{a}^2(t + 1) = \text{poslin}(W^2 \mathbf{a}^1(t)) \quad (56)$$

Los pesos de la segunda capa W^2 se establecen para que los elementos de la diagonal sean 1 y los elementos que se encuentran fuera de la diagonal sean valores negativos [35].

$$W_{ij}^2 = \left\{ \begin{array}{l} 1, \text{ si } i = j \\ -\varepsilon, \text{ en otro caso} \end{array} \right. \text{ cuando } 0 < \varepsilon < \frac{1}{s-1} \quad (57)$$

IV.3.3. Máquina de soportes vectoriales

La máquina de vectores soporte SVM (por su nombre en inglés Support Vector Machine) y sus variantes y extensiones, a menudo llamados los métodos basados en kernels, se han estudiado ampliamente y se aplica a diversos patrones de clasificación y problemas de aproximación de funciones. El objetivo de la clasificación de patrones es clasificar en una de las categorías dadas o llamadas clases. Para un problema específico de clasificación de patrones, un clasificador, se ha desarrollado de tal manera que se clasifican los objetos correctamente con razonablemente buena precisión. Donde las entradas al clasificador son llamados características, porque son determinados para que representen cada clase, o para que los datos pertenecientes a diferentes clases estén bien separados en el espacio de entrada. En general, hay dos enfoques para el desarrollo de clasificadores: un enfoque paramétrico, en el que se supone un conocimiento a priori de las distribuciones de datos, y un enfoque no paramétrico, en el que se supone que no hay un conocimiento a priori. Las redes neuronales, los sistemas difusos, y La máquina de soporte vectorial son clasificadores típicos paramétricos. La SVM es un clasificador paramétrico. A través del entrenamiento utilizando pares de entrada-salida, los clasificadores adquieren funciones de decisión que clasifican a una entrada o un punto de referencia en una de las clases dadas.

La máquina de soporte vectorial es una herramienta de aprendizaje que se emplea en muchas aplicaciones, como la clasificación, regresión, y la detección de valores atípicos. SVM fue diseñado originalmente para las tareas de clasificación de patrones. Técnicas de clasificación de patrones clasifican objetos en una de las categorías dadas llamadas clases. Clasificadores clásicos, como las redes neuronales intentan minimizar el

error en el conjunto de datos de entrenamiento que emplea la técnica de minimización del riesgo empírico. Por otro lado, la SVM se basa en el principio de minimización del riesgo estructural arraigado en la teoría del aprendizaje estadístico. El Mínimo riesgo estructural ofrece mejores capacidades de generalización a través de la minimización de la cota superior del error de generalización.

IV.3.3.1. La Máquina de soporte vectorial de dos clases

La máquina de soporte vectorial en su forma básica es un clasificador binario que aprende un hiperplano lineal que separa a un conjunto de ejemplos positivos de un conjunto de ejemplos negativos con el máximo margen. El margen se define por la distancia del hiperplano a la más cercana de los ejemplos positivos y negativos como se demuestra en la Figura IV.6. Este hiperplano de margen máximo se llama hiperplano óptimo de separación debido a las filosofías basadas en el aprendizaje estadístico que permite el máximo margen en los resultados del proceso de formación de una mejor capacidad de generalización. Los puntos de datos más cercanos se emplean para definir el margen y se conocen como vectores de soporte. El número de vectores de soporte crece con la complejidad del problema. La separación de los hiperplanos con márgenes pequeños y grandes en la clasificación binaria.

Sea M m -dimensional que contiene las entradas x_i ($i = 1, \dots, M$) pertenecientes a la clase 1 o 2 y las etiquetas asociadas sean $y_i = 1$ para la clase 1 y -1 para la Clase 2 . Si estos datos son linealmente separables, se determina la función de decisión:

$$D(X) = \mathbf{w}^T \mathbf{x} + b \quad (58)$$

$$\mathbf{W}^T \mathbf{x} + \mathbf{b} = \mathbf{0} \quad (59)$$

Donde $D(X)$ es la salida esperada y_i , \mathbf{w} es un vector m -dimensional, b es un término de sesgo, y para $i = 1, \dots, M$:

$$\mathbf{w}^T \mathbf{x} + b \begin{cases} > 0 \text{ para } y_i = 1 \\ < 0 \text{ para } y_i = -1 \end{cases} \quad (60)$$

Debido a que los datos de entrenamiento son linealmente separables, no hay datos de entrenamiento que satisfacen $\mathbf{w}^T \mathbf{x} + b = 0$. Por lo tanto, para controlar la separabilidad, en su lugar, se considera la siguiente desigualdad:

$$\mathbf{w}^T \mathbf{x} + b \begin{cases} > 1 \text{ para } y_i = 1 \\ < -1 \text{ para } y_i = -1 \end{cases} \quad (61)$$

En este caso, 1 y -1 en los lados derechos de las desigualdades puede ser una constante un $a (> 0)$ y $-a$, respectivamente. Pero al dividir ambos lados de las desigualdades por una a , se obtiene. La ecuación equivalente a:

$$y_i(\mathbf{w}^T \mathbf{x} + b) \geq 1 \text{ para } i = 1, \dots, M \quad (62)$$

El hiperplano

$$D(x) = w^T x + b = c \text{ para } -1 < c < 1 \quad (63)$$

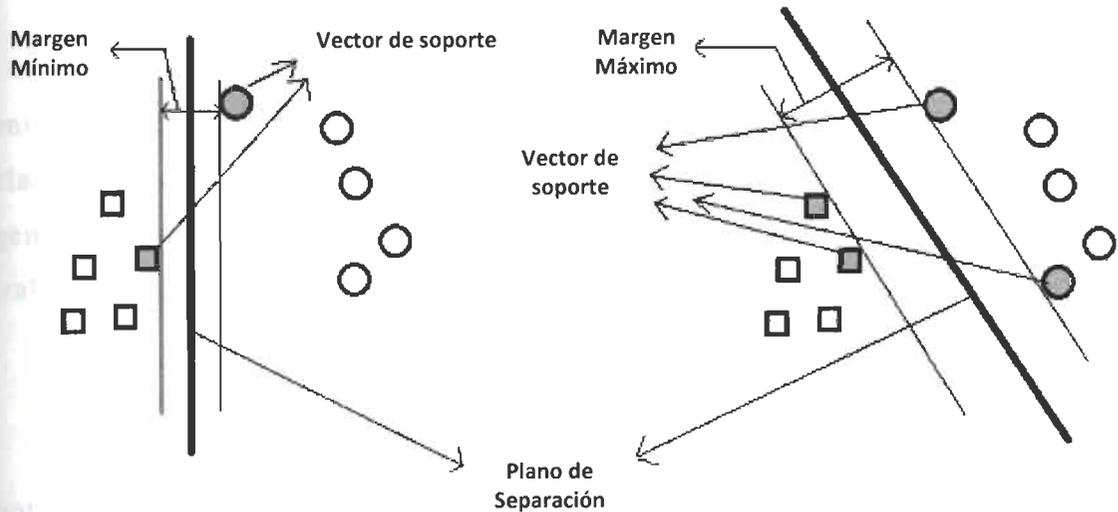


Figura IV. 6 La separación de los hiperplanos con márgenes pequeños y grandes en la clasificación binaria (tomada de [36]).

Para obtener un hiperplano con un margen mayor y mejor capacidad de generalización de una variable de holgura positiva (son variables que se agregan a la restricción para que la relación de la restricción sea de igualdad, pueden medir el ruido en los datos ya que son pequeños márgenes dados) ζ_i se define para cada muestra de entrenamiento. Esto permite que algunas de las muestras clasificadas, aun cuando ellas no tengan un máximo margen se clasifiquen correctamente. El hiperplano óptimo separa los datos y así se puede determinar como una solución para el siguiente problema de optimización cuadrática restringida por:

$$\text{Minimizar } \frac{1}{2} \|W^2\| + c \sum_{i=1}^M \zeta \quad (64)$$

$$\text{Sujeto a } y_i [Wx_i + b] \geq 1 - \zeta_i, i =, \dots, M \quad (65)$$

Donde c es el parámetro de regularización que determina el equilibrio entre la maximización del margen y la minimización del error de clasificación y se utiliza para prevenir el fenómeno de sobreajuste (se genera un margen estrecho de las variables de holgura el cual no puede trabajar con ruido y genera una mala clasificación).

Vectores W y b se determinan a través de la solución del problema de optimización en el proceso de formación. Entonces, para un dato o datos de entrada, x_i , la siguiente función de decisión discreta se utiliza para clasificar los datos de entrada en cualquiera de las clases positiva o negativa:

$$D(x_i) = \text{sign}(w^T x + b) \begin{cases} > 1 \text{ para } y_i = 1 \\ < -1 \text{ para } y_i = -1 \end{cases} \quad (66)$$

A pesar de la linealidad de la técnica básica de aprendizaje SVM, la clasificación no lineal se obtiene mediante la utilización de un mapeo del núcleo (es una generación del espacio característico de las funciones de entrada características de cada grupo, a una caracterización de una mayor dimensión) donde se utilizan funciones de núcleo (son los kernels empleados existen de distintos tipos por ejemplo Perceptron, Polinomial-homogénea, Función de base radial, Gaussiana, Sigmoide, etc.) para asignar los datos del

espacio de entrada original a un espacio de alta característica dimensional. Un hiperplano lineal está entonces capacitado para diferenciar los datos en el espacio de características que es equivalente a una clasificación no lineal en el espacio de entrada.

IV.3.3.2. Máquina de soporte vectorial multiclase

En los problemas del mundo real, se tienen problemas con más de dos clases, por ejemplo, en el diagnóstico de fallas donde varios tipos de éstas se pueden encontrar en un problema de diagnóstico. En la actualidad, se han desarrollado dos enfoques generales para la adaptación de las estrategias para el SVM multiclase. Un método construye y combina varios clasificadores binarios, mientras que el otro considera todos los datos en un problema de optimización. Dos métodos comunes, que se consideran para abordar los problemas multiclase, son los métodos uno contra todos y uno contra uno. Uno contra todos construye m clasificadores binarios donde m es el número de clases. Cada clasificador binario separa las muestras de entrenamiento de una clase de todas las otras clases. Es por eso que se llama uno contra todos. Así que en cada separador SVM binario, se utilizan las muestras de entrenamiento de todas las clases. Al clasificar un nuevo ejemplo, cada clasificador binario predice una clase y el que tiene la máxima confianza es entonces seleccionado es decir, el "ganador se lo lleva todo".

La segunda estrategia, uno contra uno construye $m(m-1)/2$ clasificadores binarios donde cada clasificador binario separa las muestras de formación de una clase de otra. Por lo tanto, en el ciclo de formación de cada SVM binario, utilizan las muestras de entrenamiento que pertenecen a

sólo dos de las clases. La estrategia de la votación por mayoría se utiliza entonces en la clasificación de un nuevo ejemplo en el que $m(m-1)/2$ SVM clasificadores binarios votarán para cada clase, y la clase ganadora será la clase que tiene los votos máximos.

IV.3.3.3. Clasificación basada en las funciones de toma continua

Una vez que los parámetros del hiperplano de \mathbf{W} y \mathbf{b} se determinan a través de la solución del problema de optimización dado por las Ecuaciones. 64 y 65, se propone la siguiente función de decisión hiperbólico en un problema de clasificación de dos clases presentada por la siguiente ecuación:

$$D(x_i) = \tanh(k(\mathbf{w}^T x + \mathbf{b})) \begin{cases} > 1 \text{ para } y_i = 1 \\ < -1 \text{ para } y_i = -1 \end{cases} \quad (67)$$

Donde el parámetro k puede ser elegido, de modo que los valores de la función de decisión muestren a que clase pertenecen los datos de entrenamiento de la clase \mathbf{I} y si están más cerca de \mathbf{I} y para la clase $-\mathbf{I}$ si están más cerca de $-\mathbf{I}$. Como el valor de k crece, la sensibilidad del clasificador se reduce frente a los cambios en las variables de proceso. Para valores bajos de k , pequeñas las variaciones en las variables defectuosas pueden ser vistas por el clasificador; Sin embargo, los resultados del sistema serán más sensibles a los ruidos del proceso. El valor de k debe ser evaluado usando las competencias del operador y el requisito previo de sensibilidad en el control del comportamiento dinámico de la falla en particular. El

enfoque propuesto de discriminación continua, también se puede emplear en el espacio de características mediante el uso de funciones de núcleo no lineales.

Las ventajas de la función de decisión continua sobre la función de decisión discreta en aplicaciones de diagnóstico de fallas en línea pueden explicarse de la siguiente manera. Considere el hiperplano de separación óptima de las funciones de toma discretas y continuas representadas en la Figura IV.7. Donde se supone que la clase marcada como -1 contiene las muestras normales de operación y la clase con la etiqueta 1 contiene las muestras de un funcionamiento defectuoso. El aumento del valor de la función de decisión representado en la Ecuación 67 de -1 a 1 se puede interpretar como la gravedad del proceso de salida de la operación normal al funcionamiento defectuoso. Por lo tanto, mediante el uso de la función de decisión continua, la información útil sobre el desarrollo de la falla se puede conseguir considerando que, en la función de decisión discreta no caracteriza el crecimiento gradual de la falla.

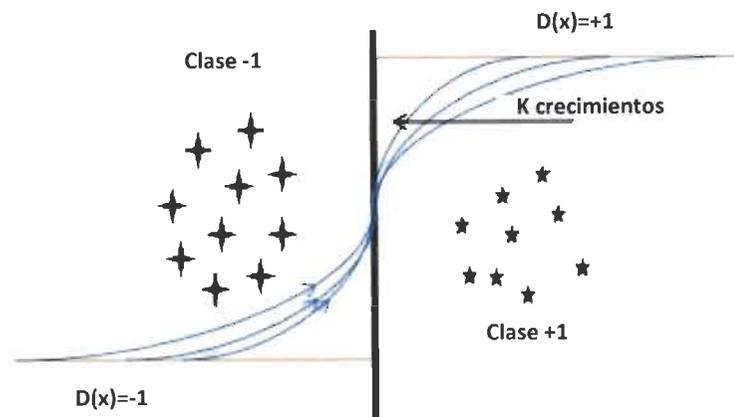


Figura IV.7 SVM-dos clase con funciones de decisión discreta y continua (tomada de [36]).

Para el diagnóstico de múltiples fallas, se emplea el SVM binario que gira en torno a la función de decisión hiperbólica y los algoritmos multiclase uno contra uno o uno contra todos. El diagrama esquemático del apoyo del SVM uno contra todos con la función de decisión continua propuesto está representado en la Figura IV.8 La función de decisión continua, $D_i(x)$, que separa la clase i con la etiqueta positiva de todas las demás clases con la etiqueta negativa, se define como

$$D_i(x) = \tanh(k_i(W_i^T x + b_i)) \quad (68)$$

Donde W_i y b_i son los parámetros de la i^{th} que separan el hiperplano. El k_i parámetro es idéntico al parámetro k utilizado en la Ecuación 67 que determina la sensibilidad del sistema. Como se muestra en la siguiente Figura, para cada vector de entrada x se tiene m valores de decisión continuas correspondientes a m clases.

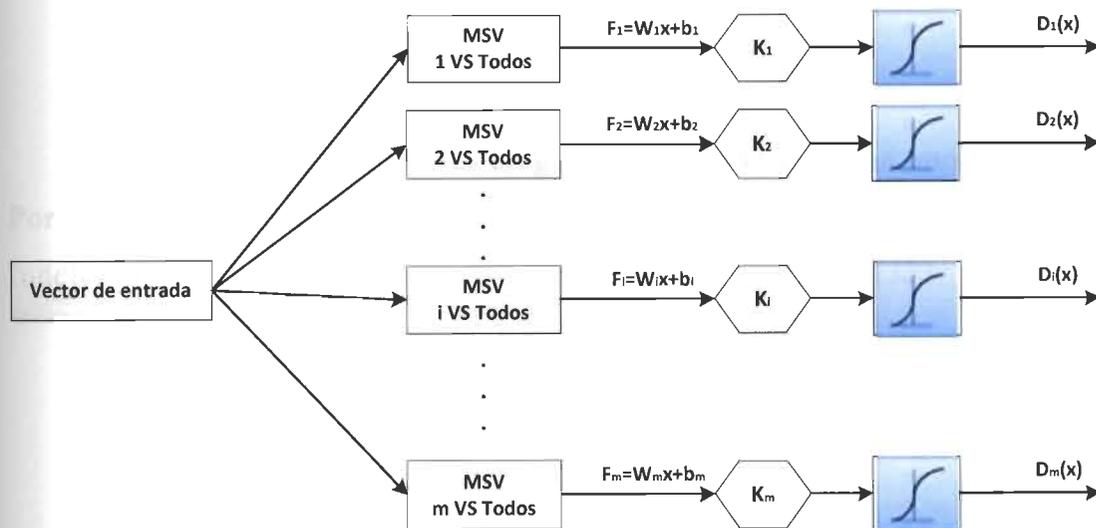


Figura IV. 8 Diagrama esquemático de UVT-SVM con funciones de decisión continuas (tomada de [36]).

La siguiente Figura IV.9 muestra el diagrama esquemático de la Máquina de soporte vectorial uno contra uno donde se propone una función de decisión continua. Para cada clasificador de dos clases, W_{ij} y b_{ij} son los parámetros del hiperplano, ($F_{ij}(x) = 0$), que separa a la clase i con la etiqueta positiva por parte de la clase j con la etiqueta negativa:

$$F_{ij}(x) = W_{ij}^T x + b_{ij} \quad (69)$$

En la Ecuación anterior se tiene $F_{ij}(x) = -F_{ij}(x)$. La función de decisión continua $D_i(x)$ para la clase i puede ser definido por una operación de minimización en F_{ij} como:

$$D_i(x) = \tanh \left(k_i \left(\min_{\substack{j \neq i \\ j = 1, \dots, m}} f_{ij}(x) \right) \right) \quad (70)$$

El k_i parámetro puede ser elegido como en el enfoque uno contra uno. Por lo tanto, es similar al enfoque de uno contra todos, y se tienen m funciones continuas para m clases correspondientes [36] y [37].

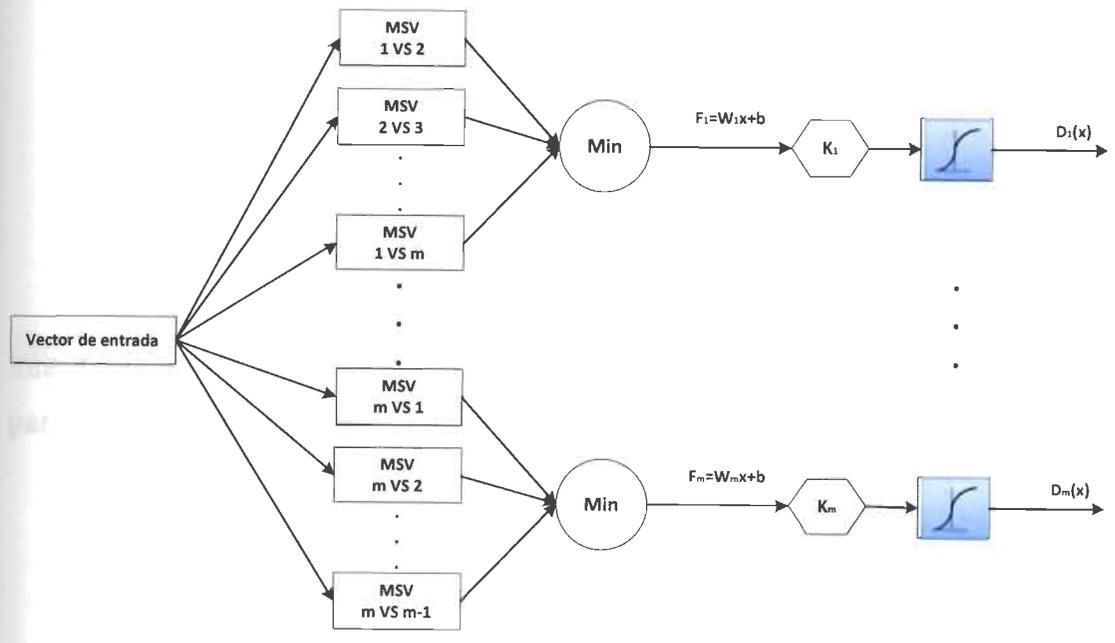


Figura IV. 9 Diagrama esquemático de UVU-SVM con funciones de decisión continuas (tomada de [36]).

V. METODOLOGÍA

En este capítulo se presenta una metodología propuesta para realizar una detección y diagnóstico, la metodología consta de los siguientes pasos para su implementación.

V.1. Pasos de la metodología

1. Generar bases de datos con los datos históricos del proceso o sistema. Las bases de datos deben presentar tanto el modo de operación normal así como las fallas que ya se han presentado en el sistema.
2. Pretratamiento de los datos. Eliminación de puntos atípicos, estandarizar los datos, análisis de componentes principales. Adicionalmente en este punto se puede determinar la correlación entre las variables y determinar cuáles de ellas son las más significantes para realizar el paso siguiente.
3. Caracterizar las fallas existentes, así como, el modo de operación normal del proceso o sistema. En este punto se determina cual es el modo normal de operación del sistema y se determinan cuáles son las fallas que ya se han presentado dentro del proceso. También se puede solo tener identificado el modo de operación normal del sistema y conforme se monitorea el sistema se pueden caracterizar la existencia de fallas.
4. Se utiliza un modelo difuso para realizar la detección de las fallas.

El modelo difuso es generado al fuzzificar los datos utilizando una función de membresía, Ecuación 44. Posteriormente se realiza una composición haciendo uso de las Ecuaciones 45 a 47. Después se

genera una implicación difusa para generar las reglas *SI- ENTONCES* con las implicaciones descritas en las Ecuaciones 49 y 50. Por ultimo defusificar los datos con el método del centroide al utilizar la Ecuación 51. Para que el modelo difuso facilite la detección, se puede tener una salida que se encuentre entre 0 y 1 donde si la salida se aproxima a uno indica que el sistema se encuentra en su modo normal de operación y si el sistema se aproxima a 0 se encuentra en modo de falla.

La ventaja de utilizar este modelo es que se pueden generar empíricamente los umbrales (al hacer uso de la experiencia de los operadores de un sistema determinado) en los cuales funciona idóneamente el proceso o sistema en su modo normal de operación, o también estos pueden ser determinados con alguna técnica estadística básica esta puede ser t de Student para datos univariados un ejemplo son las cartas de control o un estadístico t^2 de Hotelling para datos multivariados de igual forma, el ejemplo más utilizado son las cartas de control para datos multivariados (los umbrales pueden ser modificados al incrementar o disminuir el nivel de confianza que se desea obtener utilizando alguna técnica estadística) o al utilizar sus valores máximos y mínimos (utilizando el valor máximo como límite superior del umbral y como límite del umbral inferior del sistema el valor mínimo obtenido) al generar los umbrales lo que se busca es poder replicar el comportamiento del sistema en su modo normal de operación y al realizar ajustes en los umbrales se puede facilitar el trabajar con ruido dentro del sistema monitoreado.

5. El modelo difuso es capaz de detectar el modo de operación en falla en el sistema si (*SI*) se procede al siguiente paso si (*NO*) se ajustan los umbrales de entrada para sensibilizar el sistema difuso, los umbrales pueden ser ajustados a prueba y error si solo se utiliza el expertis como indicador pero si es mediante alguna técnica

estadística este se hará reduciendo o incrementando el nivel de confianza requerido.

6. Si el modelo difuso detecta que el sistema está fuera de las condiciones normales de operación la salida o los datos de origen del sistema pasaran a ser evaluados por una red neuronal y por una maquina generadora de soportes (SVM) para detectar el tipo de falla. Esta doble evaluación se realiza para hacer más robusta la metodología.
7. La SVM uno contra todos determinan si la falla es conocida o no de acuerdo a la caracterización que se elaboró con anterioridad. El SVM hace uso de las Ecuaciones 58 a 66 para caracterizar a cada uno de los diferentes grupos y como se describió dentro del marco teórico no se ha generado un SVM que pueda clasificar distintos grupos por lo cual se hace uso de distintas estrategias para poder clasificar las características de los grupos, las estrategias más usadas son el SVM uno contra todos o el SVM uno contra uno por votación esto por su sencillez, las estrategias se describen en las Ecuaciones 67 a 70. En esta metodología se hace uso del SVM uno contra todos debido a que esta estrategia puede determinar si la falla es conocida o no.

Esta clasificación se puede hacer al haber caracterizado con anterioridad los tipos de falla, y se puede determinar cuáles son las variables que están generando el comportamiento anormal del sistema o proceso.

NOTA: Si los datos del sistema no presentan un patrón el cual pueda ser clasificado por la PNN o por el SVM se puede generar un patrón que caracterice de una mejor forma a los grupos, este patrón puede ser generado con la obtención de distintas métricas como la distancia Euclidiana o de Mahalanobis entre otras, las métricas generan valores obtenidos entre las variables monitoreadas que puedan caracterizar de una mejor manera cada una de las clases, y con

ello facilitar la tarea de clasificación para la PNN o la SVM las Ecuaciones 42 y 43 describen la obtención de dichas métricas.

Si la falla no es conocida se procede a introducir los datos a la PNN para con su clasificación poder facilitar la caracterización de una nueva falla y posteriormente integrar la nueva falla como otra clase al sistema o proceso.

8. Ya que la SVM no reconoció los nuevos patrones del sistema que se introdujeron en ella, estos valores son introducidos en la PNN que ya fue entrenada previamente con los mismos datos con los cuales fue entrenada la SVM pero a diferencia de la SVM la PNN dará un diagnóstico para hacer una aproximación a las variables que pueden presentar la falla, y con este diagnóstico poder hacer más sencilla la tarea de caracterizar nuevas fallas, las Ecuaciones 52 a 57 describen la forma en que clasifica la PNN en su primera capa se realiza el entrenamiento y en la segunda capa se realiza la clasificación al dar una salida diferente a 0 se determina cual es la clase ganadora.
9. Al detectar y diagnosticar la falla, se podrán hacer los ajustes necesarios para la corrección del proceso o sistema y llevarlo nuevamente a su modo de operación normal, los ajustes serán realizados por un operador y posiblemente la automatización del sistema se centrara en trabajos futuros para cerrar el círculo de control.
10. Esta metodología puede ser replicada en distintos procesos o sistemas que se puedan monitorear y en los cuales se pueda tener acceso a los datos históricos del proceso.

V.2. Metodología de forma gráfica.

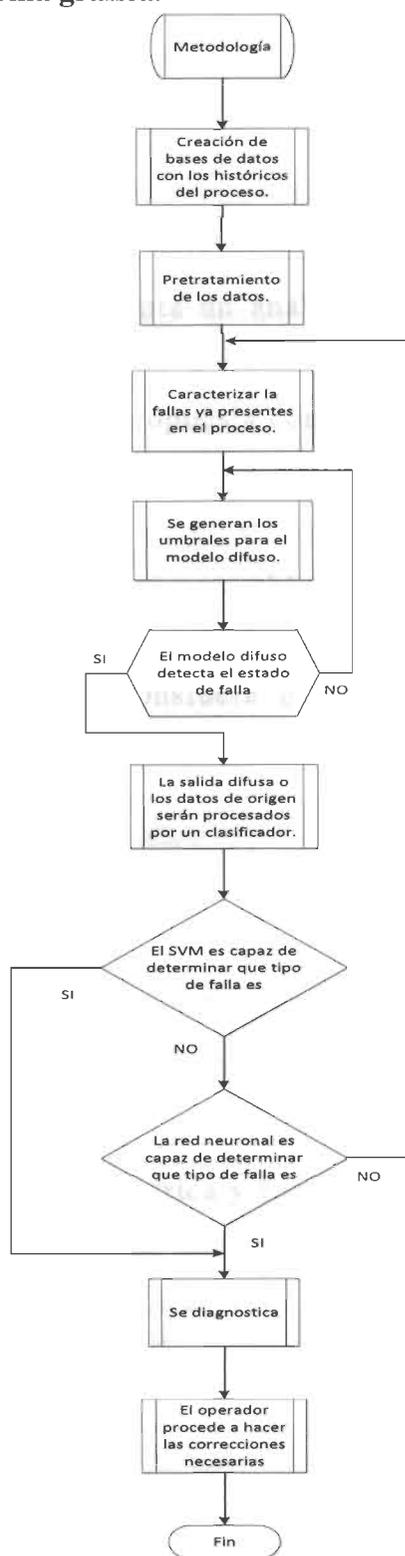


Figura V. 1 Metodología Propuesta.

VI. EXPERIMENTACIÓN

En el presente trabajo se consideraron dos casos de estudio. En el primer caso de estudio, se presenta un análisis en un sistema eléctrico de potencia, el cual tiene presencia de distintas fallas, en este caso se hace un comparativo de la metodología propuesta contra técnicas estadísticas “PCA, FDA y dos métricas estadística” para la validación de la metodología. El segundo caso de estudio, fue realizado en un proceso de mecanizado, en el cual se analizan distintas variables para obtener el acabado requerido por el cliente, este analiza los parámetros promedio provenientes de una máquina de torno y una fresadora, se considera este caso como el principal del trabajo, debido a que en él, se valida la eficiencia de la metodología propuesta por si sola.

VI.1. Caso de estudio 1

Se efectuó la experimentación para la validación de la metodología descrita en el capítulo anterior en un sistema eléctrico de potencia propuesto por El Instituto de Ingeniería Eléctrica y Electrónica IEEE por su nombre en inglés (Institute of Electrical and Electronics Engineers). El sistema está compuesto de 24 buses con 3 líneas cada uno que hacen un total de 72 variables a analizar [40]. El sistema eléctrico descrito se presenta en la Figura VI.1.

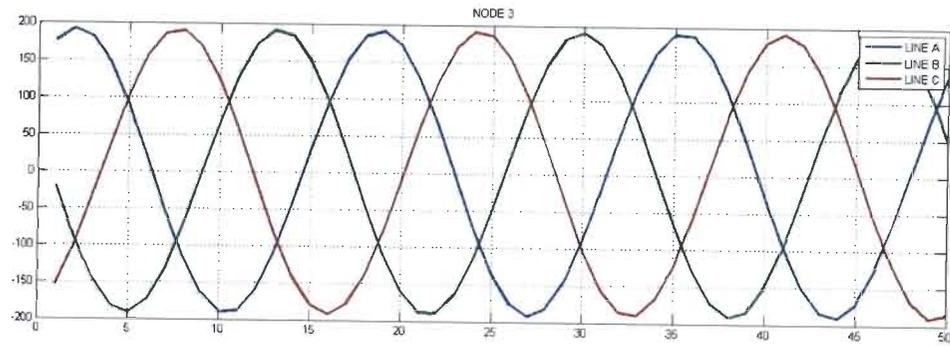


Figura VI. 2 *Modo de operación normal*

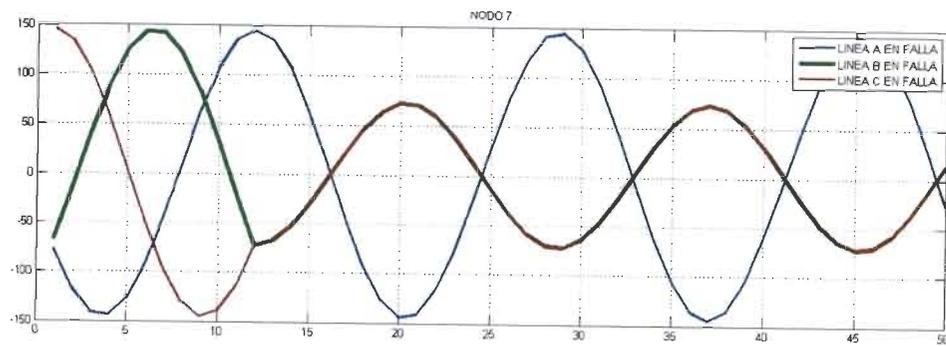


Figura VI. 3 *Falla simétrica.*

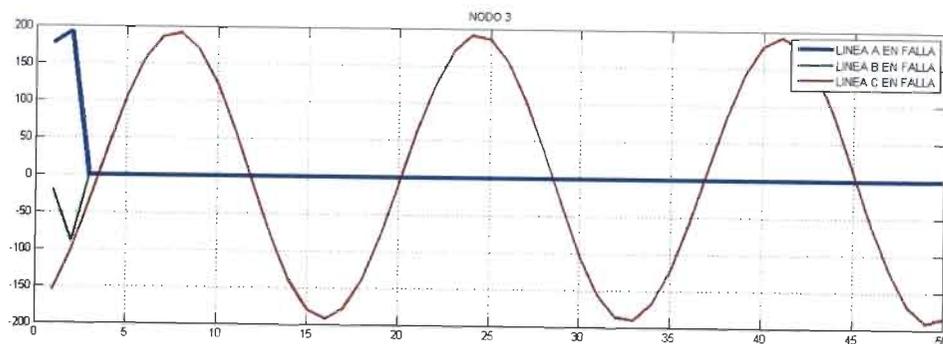


Figura VI. 4 *Falla asimétrica.*

Como se describe en la metodología el primer paso a realizar dentro del sistema, es el proceso de detección, donde se desea saber si el sistema está funcionando en un modo de operación normal o está trabajando con la

presencia de una falla, para llevar a cabo la detección, se ha propuesto la utilización de un sistema difuso para hacer más fácil y segura la detección del estado en falla. La ventaja de utilizar un sistema difuso para la etapa de detección, es que solamente se modeló uno solo de los nodos para poder evaluar el total de los 24 nodos dentro del sistema, debido a que el comportamiento de los 24 nodos, es repetitivo en todos ellos, esto hace posible que el modelado del sistema difuso de un solo de los nodos, pueda replicarse en los demás nodos.

VI.1.1. Etapa de detección para el sistema eléctrico

El sistema difuso creado, es capaz de simular el comportamiento del sistema en su modo normal de operación, en un primer paso se estableció un sistema difuso con 9 reglas, el cual solo es capaz de detectar fallas asimétricas en el sistema, para hacer más robusto el sistema difuso, se generaron 76 reglas, con lo cual, el sistema es capaz de detectar la presencia de fallas simétricas y asimétricas, ambas reglas fueron generadas haciendo uso del toolbox de MatLab®. En las siguientes Tablas se muestra cómo se generaron las reglas difusas para cada sistema.

En el primer paso se realiza la adquisición y análisis de los datos históricos del sistema eléctrico. El análisis llevó a cabo, la evaluación de las amplitudes de las tensiones de las líneas de cada nodo. Para ello se tomaron ventanas de 50 y 100 datos con muestras en modo de operación normal y en cada uno de los modos de falla. Para generar el sistema difuso se utilizaron las Ecuaciones 52-59. Estas herramientas fueron utilizadas debido a que lograron describir relativamente bien el comportamiento del sistema en su modo de operación normal.

Se realiza un comparativo al mismo grupo de datos con ventanas de datos del mismo tamaño con las técnicas estadísticas para el paso de la detección y reducción de la cantidad de variables esto con el fin de evaluar el comportamiento del sistema. La técnica utilizada para la detección fue el análisis de componentes principales PCA.

Tabla VI. 1 Reglas difusas para la detección de fallas asimétricas.

REGLA	L1	L2	L3
1	Mas	Menos	Mas
2	Mas	Menos	Menos
3	Mas	Mas	Menos
4	Menos	Mas	Mas
5	Menos	Menos	Mas
6	Menos	Mas	Menos
7	Mas	Mas	Mas
8	Menos	Menos	Menos
9	Cero	Cero	Cero

Tabla VI. 2 Reglas difusas para la detección de fallas simétricas y asimétricas.

REGLAS	LÍNEAS		
	L1	L2	L3
1	Muy positivo	El más negativo	Mínimo negativo
2	Muy positivo	El más negativo	Mínimo positivo
3	Medio positive	El más negativo	Medio positivo
⋮	⋮	⋮	⋮
74	Medio positive	Muy negativo	Medio positive
75	Mínimo positivo	Muy negativo	Mínimo positivo
76	Mínimo positivo	Muy negativo	Mínimo positivo

Al realizar el análisis con el PCA se encontró que éste solo encuentra significantes 2 variables en un total de 72 que representan el total de las líneas a analizar. Este comportamiento, se presentó tanto en los datos de los nodos en modo de operación normal, así como en los datos en modo de falla. Las siguientes Figuras muestran la cantidad de variables que el PCA considera significantes.

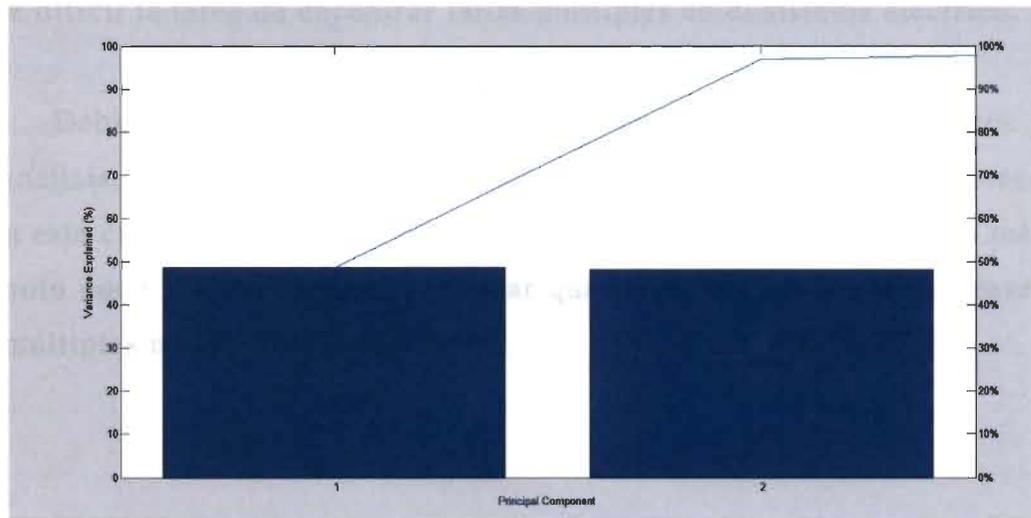


Figura VI. 5 PCA en modo de operación normal.

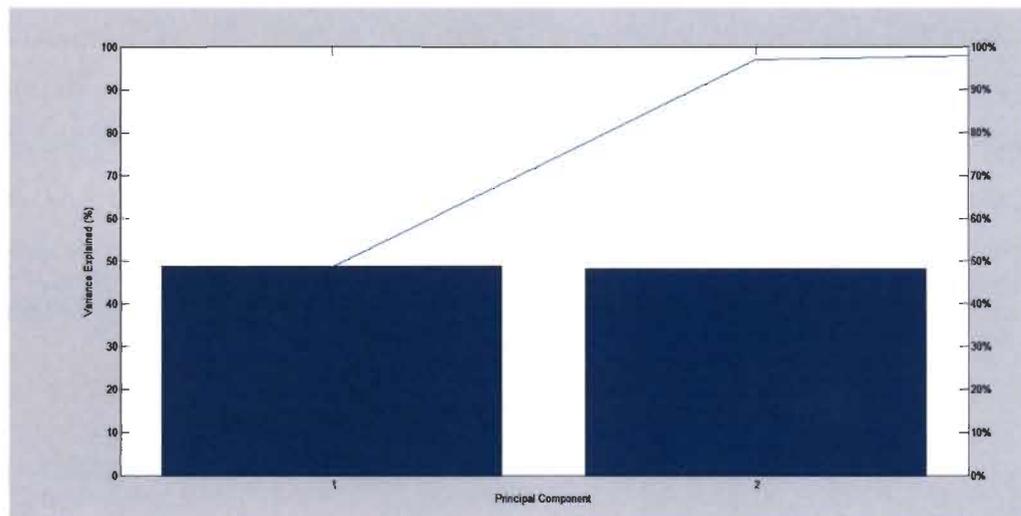


Figura VI. 6 PCA en modo de falla.

Como se muestra en las Figuras VI.5 y VI.6 para el comportamiento normal se consideran solo dos variables significantes en el sistema y aun cuando se tiene la presencia de hasta 5 fallas en diferentes nodos el comportamiento del PCA es el mismo, determinando que solo 2 variables son significantes para poder realizar la detección. Este comportamiento es debido, a que el sistema eléctrico parte del primer nodo y los demás nodos están conectados a él, por tanto, el comportamiento que presenta el PCA hace difícil la tarea de encontrar fallas múltiples en el sistema eléctrico.

Debido a que el PCA considera solo dos variables significantes para el análisis, no es una buena herramienta para realizar el paso de detección para este caso de estudio. Ya que las fallas pueden estar presentes en más de un solo nodo y se tiene que considerar que las fallas pueden estar presentes en múltiples nodos de forma similar.

VI.1.2. Etapa de diagnóstico para el sistema eléctrico.

El segundo paso de la metodología propuesta arroja el diagnóstico final, mediante el empleo de distintas métricas estadísticas y el uso de una PNN junto con una SVM. Para hacer del diagnóstico una etapa más robusta y segura.

En una primera instancia el diagnóstico fue realizado al obtener las distancias Euclidianas y posteriormente este resultado fue comparado con el resultado obtenido por las distancias de Mahalanobis. Donde para realizar el diagnóstico se obtiene un patrón anormal de operación entre las tensiones de cada uno de los 24 nodos que forman el sistema eléctrico. Por lo tanto cuando una falla está presente se compararan los patrones previamente obtenidos mediante las distancias y se determina cuál de las líneas tiene el problema.

El segundo paso o etapa de diagnóstico, el cual se lleva a cabo después de analizar la condición de operación del sistema con un sistema difuso y una vez encontrada la presencia de una falla, se procede a encontrar cuál es el tipo de falla y en qué elementos se encuentra presente. Tal búsqueda se fundamenta, en encontrar la relación presente en las distancias entre los elementos que se encuentran en modo de falla. A continuación se describen estas relaciones, primero se especifica cómo se lleva a cabo para la presencia de una falla simétrica y posteriormente para una falla asimétrica al utilizar la distancia euclidiana y la de Mahalanobis.

- Fallas simétricas

- Para el caso de la falla simétrica, la falla no se puede presentar en una sola línea, ya que ésta se presenta cuando se unen dos líneas entre sí.
- Dos líneas en falla simétrica

Si las distancias entre las líneas $L1-L2$ es diferente a $L2-L3$ y $L1-L3$ y $L2-L3$ y $L1-L3$ son iguales hay una falla a $L1$ y $L2$, o si la distancias entre $L1-L3$ es diferente a $L1-L2$ y $L2-L3$ y las distancias de $L1-L2$ y $L2-L3$ son iguales la falla se encuentra en $L1$ y $L3$, y si las distancias entre $L2-L3$ es diferente a las distancias de $L1-L2$ y $L1-L3$ y las distancias de $L1-L2$ y $L1-L3$ hay falla en $L2$ y $L3$.

- Tres líneas en falla simétrica

Si las distancias entre las tres líneas $L1-L2$, $L1-L3$ y $L2-L3$ son iguales, entonces la falla simétrica está presente en las tres líneas.

Por lo tanto, si se cumple alguna de estas condiciones propuestas en esta sección, para las distancias obtenidas por medio del cálculo de las distancias Euclidianas y de Mahalanobis, se determina que se tiene una falla simétrica presente en el sistema. La Figura VI.7 muestra el procedimiento descrito con anterioridad.

Proceso de Diagnóstico para fallas simétricas

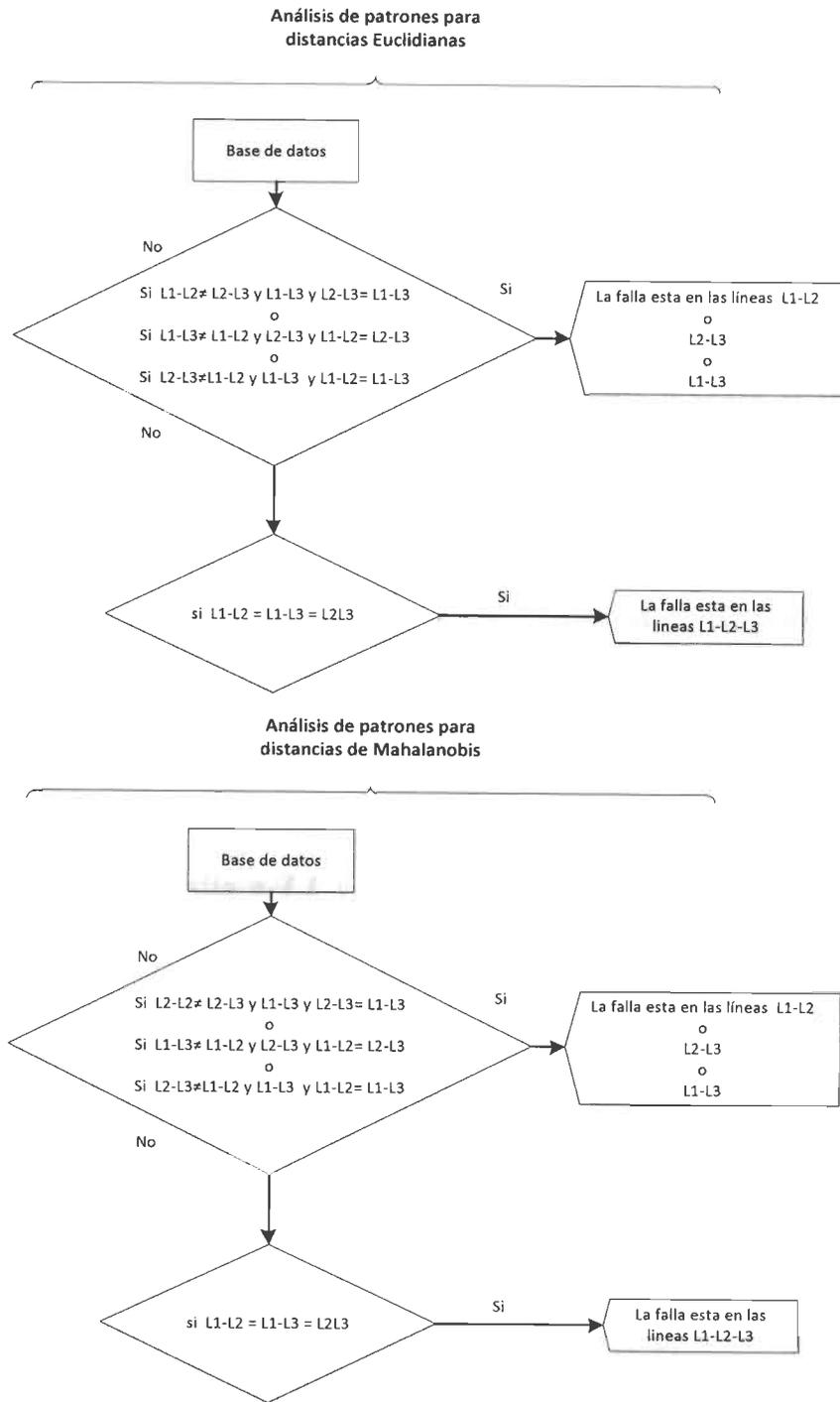


Figura VI. 7 Metodología para el diagnóstico de fallas Simétricas.

- Fallas asimétricas o a tierra

- Una de las líneas de falla a tierra

Si al comparar las distancias entre $L1-L2-L3$, las distancias de $L2-L3$ son valores iguales, entonces la falla está presente en la línea $L1$. Si las distancias de $L1-L2-L3$ son valores iguales a las distancias de $L1-L3$, entonces la falla está presente en la línea $L2$. Y de la misma manera si las distancias entre $L1-L2$ son valores iguales $L1-L2-L3$, entonces la falla está presente en la línea $L3$. Este análisis solo es aplicable para la distancia Euclidiana ya que en la distancias de Mahalanobis no se observó un patrón similar para poder determinar si la falla está presente en un solo elemento.

- Dos líneas en falla a tierra

Si las distancia entre las distancias $L1-L2$ son 0 y las distancias $L2-L3$ y $L1-L3$ son iguales hay una falla a $L1$ y $L2$, o si la distancia entre $L1-L3$ son 0 y distancias de $L1-L2$ y $L2-L3$ son iguales se tiene una falla en $L1$ y $L3$, o Si las distancias entre $L2-L3$ son iguales a 0 y distancias de $L1-L2$ y $L1-L3$ son iguales la falla está presente en $L2$ y $L3$.

- Tres líneas en falla a tierra

Si las distancias entre las tres líneas $L1-L2-L3$ es cero, entonces la falla está presente en las tres líneas.

Por lo tanto, si se cumple alguna de estas condiciones propuestas en esta sección, para las distancias obtenidas por medio del cálculo de las distancias Euclidianas y de Mahalanobis, se determina qué se tiene una falla

asimétrica presente en él sistema. La Figura VI.8 muestra el procedimiento descrito con anterioridad.

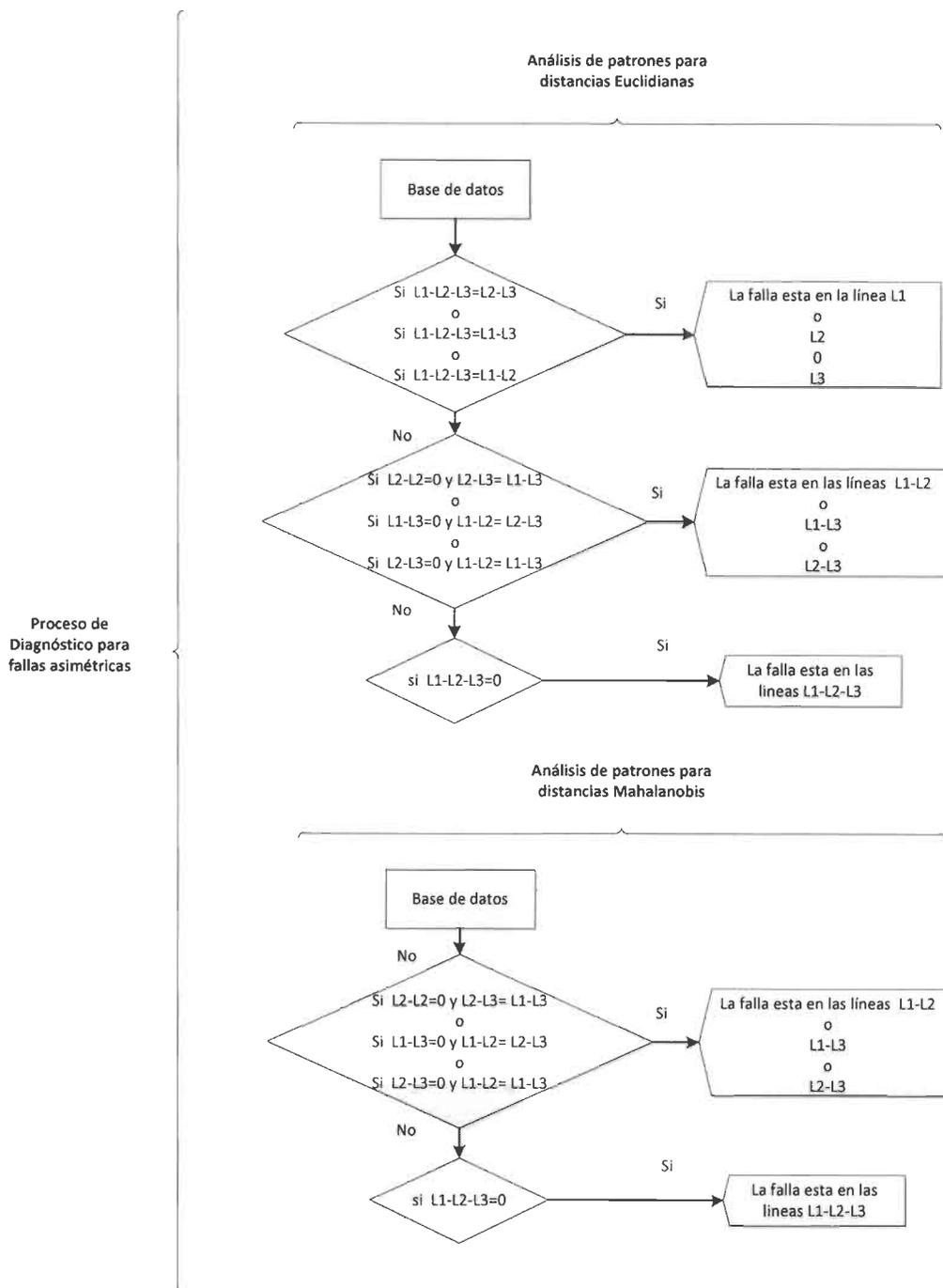


Figura VI. 8 Metodología para el diagnóstico de fallas Asimétricas.

Se procedió a realizar el diagnóstico utilizando el análisis discriminante de Fisher con la ayuda de Minitab® para hacer un comparativo del diagnóstico dado por las distancias Euclidiana y de Mahalanobis. Para realizar el diagnóstico se generaron 12 grupos en los cuales, se presentan el modo de operación normal así como la presencia de fallas simétricas y asimétricas, el resultado que se obtuvo fue el siguiente.

Análisis discriminante: g vs. L1, L2, L3

Método lineal para respuesta: g

Predictores: L1, L2, L3

Grupo	1	2	3	4
Conteo	25	25	25	25

Resumen de clasificación

Colocar en un grupo	Grupo verdadero			
	1	2	3	4
1	15	10	0	0
2	0	0	0	8
3	5	12	25	7
4	5	3	0	10
N Total	25	25	25	25

N correcta	15	0	25	10
Proporción	0.600	0.000	1.000	0.400

N = 100 N Correcta = 50 Proporción Correcta = 0.500

Distancia cuadrada entre grupos

	1	2	3	4
1	0.000000	0.263368	0.311866	0.261408
2	0.263368	0.000000	0.008252	0.000076
3	0.311866	0.008252	0.000000	0.009163
4	0.261408	0.000076	0.009163	0.000000

Función discriminativa lineal para grupos

	1	2	3	4
Constante	-0.15593	-0.00413	0.00000	-0.00458
L1	-0.00415	-0.00121	0.00000	-0.00127
L2	-0.00963	-0.00059	0.00000	-0.00062
L3	0.00014	-0.00217	0.00000	-0.00264

Como se puede observar en los datos mostrados en la parte posterior el diagnóstico utilizando el análisis discriminante de Fisher en el software de Minitab® tiene una confiabilidad de un 50%. Esto lleva a determinar que el empleo de esta herramienta, lleva a un diagnóstico erróneo en la Tabla

VI.3 se hace una comparación entre las técnicas estadísticas y el sistema difuso con el que se ha trabajado.

Tabla VI. 3 *Comparación de métricas estadísticas.*

FALLAS PRESENTES	DETECCIÓN		DIAGNÓSTICO		
	Análisis de Componentes Principales	Sistema Difuso	Análisis Discriminant e de Fisher	Distancia Euclidiana	Distancia Mahalanobis
1	No	Si	Parcial	Si	Si
2	No	Si	Parcial	Si	Si
3	No	Si	Parcial	Si	Si
4	No	Si	Parcial	Si	Si
5	No	Si	Parcial	Si	Si

Para continuar con la validación de la metodología propuesta se observó que el empleo de un sistema difuso tiene un porcentaje de eficiencia alto para realizar la detección en una primera etapa, por ello se considera una apropiada herramienta para dicha tarea. Para generar un mejor diagnóstico, se optó por utilizar una red neuronal probabilística junto con una máquina de soporte vectorial. Debido a que las métricas muestran un desempeño de entre un 70% a un 80% esto porque las distancia Euclidiana y de Mahalanobis necesitan tener en la ventana de datos adquirida todos sus datos con presencia de falla para generar el patrón descrito con anterioridad para realizar la tarea de diagnóstico.

Para el entrenamiento de la PNN y de la SVM al igual que en el análisis discriminante de Fisher se utilizaron 12 grupos en los cuales se presentan las fallas del sistema, falla simétrica y asimétrica, así como el modo de operación normal, la diferencia es que para el entrenamiento se utilizaron solamente ventanas de 25 datos para caracterizar tanto el modo de operación normal y el modo de operación en falla. La SVM utilizó 12 pares de grupos para realizar el diagnóstico y la PNN en su primera capa utilizó 12 neuronas de entrenamiento y en la segunda 12 neuronas competidoras. Tanto la SVM y la PNN se entrenó con 60 épocas ya que con este número de épocas se logró una correcta clasificación que contenga las características de cada una de las clases presentes en el sistema.

La red neuronal utilizó un 80% de los datos muestreados de uno solo de los nodos para su entrenamiento, en un principio el entrenamiento se realizó con los datos en crudo (los datos provenientes directamente del sistema), al ver que la red no brindaba una correcta clasificación al utilizar los voltajes provenientes de la simulación del SEP como datos de entrada, se procedió a utilizar las distancias Euclidianas entre los voltajes, para generar un patrón que fuera más fácil de reconocer para la red neuronal probabilística. Al hacer esto, se obtuvo una eficiencia en la red de un 99% al evaluar el 20% restante de los datos muestreados debido a que las distancias entre las líneas L1-L2, L1-L2, L2-L3 y L1-L2-L3, caracterizaban de mejor manera cada uno de los comportamientos, la eficiencia del sistema se obtuvo al analizar el 20% de los datos utilizados en la validación, donde se generó el porcentaje al evaluar el número de muestras clasificadas correctamente sobre el total de muestras utilizadas para la validación al conocer previamente a que grupo pertenecían.

En esta etapa también se utilizó la SVM para clasificar los tipos de falla que se han presentado, donde de igual manera en una primera instancia

se utilizaron ventanas de 25 datos de los voltajes provenientes directamente del SEP para realizar el diagnóstico y se observó que la SVM tiene un bajo rendimiento para hacer un correcto diagnóstico al igual que la PNN, el rendimiento presentado fue de un 60% para ambas metodologías. Por esta razón se optó por darle entrenamiento a la SVM utilizando los patrones generados por las distancias de la misma manera que con la PNN, con ellas el SVM es capaz de clasificar correctamente las fallas por lo cual hace más fácil y segura la etapa de diagnóstico y al igual que la PNN se obtuvo un porcentaje de eficiencia del 99%. De la misma forma que con la PNN se utilizó el 20% de los datos muestreados para realizar la validación, y para obtener el porcentaje se tomó en consideración los aciertos de la SVM. Para el SVM uno contra todos se generaron 12 grupos con los que se hará una correcta clasificación y para el SVM uno contra uno se generó una combinación 66 grupos para hacer una correcta clasificación, y encontrar al grupo más ganador. El SVM uno contra todos muestra un mejor desempeño debido a que el sistema puede detectar si la base de datos suministrada no pertenece a alguna de las clases.

La implementación de estas tres técnicas inteligentes es hecha con el fin de hacer de la metodología de detección y diagnóstico un proceso más robusto y más seguro con el fin de evitar la presencia de falsas alarmas al generar un diagnóstico en línea se dará una pronta respuesta a las fallas presentes en el proceso y con ello se harán los procesos más seguros para los operarios que trabajan en procesos peligrosos. Como se muestra en la Tabla VI.4. Al utilizar dichas técnicas inteligentes se evita la problemática que se tiene con las métricas de tener presente el estado de falla en todos los elementos de la ventana de datos, y la presencia de una mala clasificación al utilizar el análisis discriminante de Fisher. La Tabla VI.4 muestra el desempeño de la metodología descrita al evaluar el sistema eléctrico de potencia propuesto por la IEEE.

La metodología propuesta hace uso del SVM para realizar la clasificación, al utilizar el SVM uno contra todos se puede determinar si la falla es conocida. Si la falla es desconocida se utiliza el SVM uno contra uno por votación y la PNN para caracterizar la presencia de la falla en una nueva falla, esta caracterización se logra ya que tanto el SVM y la PNN buscan clasificar en una clase a los nuevos datos de entrada y al agruparlos en una clase con ciertas características ya determinadas con anterioridad, por este comportamiento se puede hacer una nueva y más fácil clasificación de cuáles son las variables que esta fuera de control y con ello caracterizar una nueva falla y hacer un correcto diagnóstico.

Tabla VI. 4 *Comparación de las técnicas empleadas en la detección y diagnóstico del sistema eléctrico.*

N° fallas en ventanas de 50 Datos	% en que se realizó la Detección	% para el Diagnóstico Distancia Euclidiana	% para el Diagnóstico Distancia de Mahalanobis	% para el Diagnóstico Con la Red Neuronal Probabilística	% para el Diagnóstico con la SVM	% para el Diagnóstico con la SVM y PNN
1	100	75	70	99	99	99
2	100	80	75	99	99	99
3	100	80	80	99	99	99
4	100	80	80	99	99	99
5	100	80	80	99	99	99

La Tabla VI.6 muestra en la primera columna la presencia de fallas múltiples en el SEP. La segunda columna muestra el porcentaje de detección que se tiene al utilizar la metodología de lógica difusa para dicha tarea, en la tercera y cuarta columna se presenta el desempeño de las distancia Euclidiana y de Mahalanobis respectivamente al momento de realizar el diagnóstico de fallas simétricas y asimétricas en el SEP, en la quinta y sexta columna se muestra por separado el rendimiento de la SVM y la PNN al momento de hacer el diagnóstico en el sistema eléctrico, y por último se presenta la combinación de SVM y PNN al realizar la tarea de diagnóstico. Como se describió con anterioridad el porcentaje se obtuvo al evaluar el 20% de las muestras tomadas y ver el número de acierto que se obtenían en el funcionamiento del sistema. Los datos que se obtuvieron de la simulación fueron voltajes en donde se indujeron fallas múltiples, las fallas presentaron dos comportamiento o modos de falla, las fallas simétricas y asimétricas. Se corrieron en el sistema 200 simulaciones para su validación, con un máximo de 5 de fallas simultaneas en el SEP, la metodología propuesta se modelo en uno solo de los nodos y esta fue replicada en cada uno de ellos para realizar el diagnostico, esto facilitó dicha tarea de monitoreo y demostró un buen rendimiento al realizar la tarea de detección y diagnóstico.

Al determinar el tipo y elementos dónde se encuentra la falla en un paso posterior se procederá a hacer la recuperación del sistema para llevarlo nuevamente a su modo normal de funcionamiento.

VI.2. Caso de estudio 2

En el presente caso de estudio se tiene como objetivo encontrar la variación en los parámetros de velocidad, avance y dureza del maquinado de un aluminio 6061, un acero 1080 y en un acero 4140 para poder hacer la detección y diagnóstico de las condiciones anormales de operación que no permiten que se alcance la rugosidad requerida por el cliente. Donde se analizará como variable final la rugosidad que tiene la pieza después del mecanizado y tomando como variables a analizar la velocidades de corte, el avance y la dureza del material, esto debido a que cuando se presentan variaciones dentro de los procesos de maquinado en las variables de dureza, avance y velocidad no se alcanza la rugosidad deseada, esto sucede en muchas industrias dedicadas a la fabricación de componentes. Donde algunos productos son fabricados con materiales de baja maquinabilidad.

Los rangos en que se trabaja la velocidad de corte y el avance son limitados para considerar vidas de herramientas satisfactorias. Aunado a esto la herramienta, recubrimientos y geometría de la misma influyen en la productividad, ya que se puede presentar desgaste excesivo de la herramienta lo cual puede producir errores en las tolerancias de las partes. También es conveniente conocer el grado de aportación que tiene la vibración generada en el husillo de la herramienta y la temperatura obtenida por la fricción de la pieza con la herramienta y a fin de controlarlo para tener un proceso productivo robusto y satisfactorio en cuanto a tiempo de elaboración y calidad de los productos.

Este trabajo permitirá a las empresas manufactureras, evaluar sus materiales de trabajo, mejorar la productividad, minimizar el tiempo de producción e incrementar la eficiencia de los procesos de maquinado. Con la implementación de la metodología descrita en el capítulo anterior se hará los

procesos más seguros y más simples con lo cual se alcanzarán las metas descritas anteriormente. La importancia de esta investigación en la ciencia, consiste en obtener la relación que existe entre las variables de maquinado para a través del uso de sistemas inteligentes poder dar una correcta detección y diagnóstico de las fallas presentes en el proceso.

En primer lugar, se efectuó el mecanizado de las piezas mediante una máquina de torno modelo Okuma LB15 como se muestra en la Figura VI.9. Con un tipo de herramienta DNMG 432 PG de grado RC8025 que se utiliza para el acero y un tipo de herramienta DNMG 432 GP de grado CQ23 que se utiliza para el aluminio. Los tres tipos de materiales fueron mecanizados el aluminio 6061, el acero 1018 y el acero 4140 de forma cilíndrica como se muestra en la Figura VI.10. La rugosidad se midió utilizando un rugosímetro Mitutoyo SJ-301 Figura VI.11 teniendo en cuenta la configuración mostrada en la Figura VI.12.



Figura VI.9 Máquina de torno de dos ejes CNC Okuma LB15.



Figura VI. 10 *Cilindro a mecanizar compuesto de los tres materiales, aluminio 6062, acero 1018 y acero 4140.*

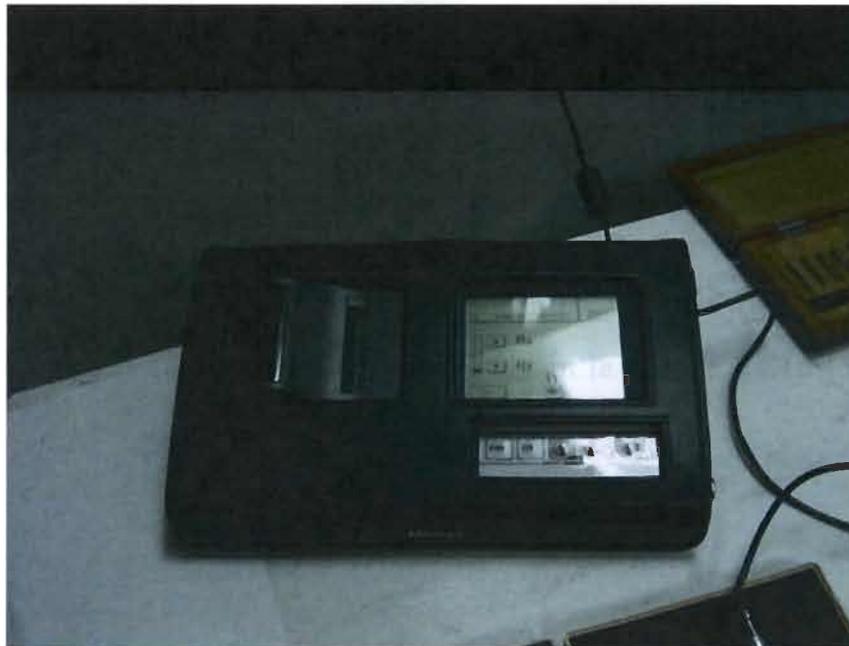


Figura VI. 11 *Rugosímetro Mitutoyo SJ-301.*

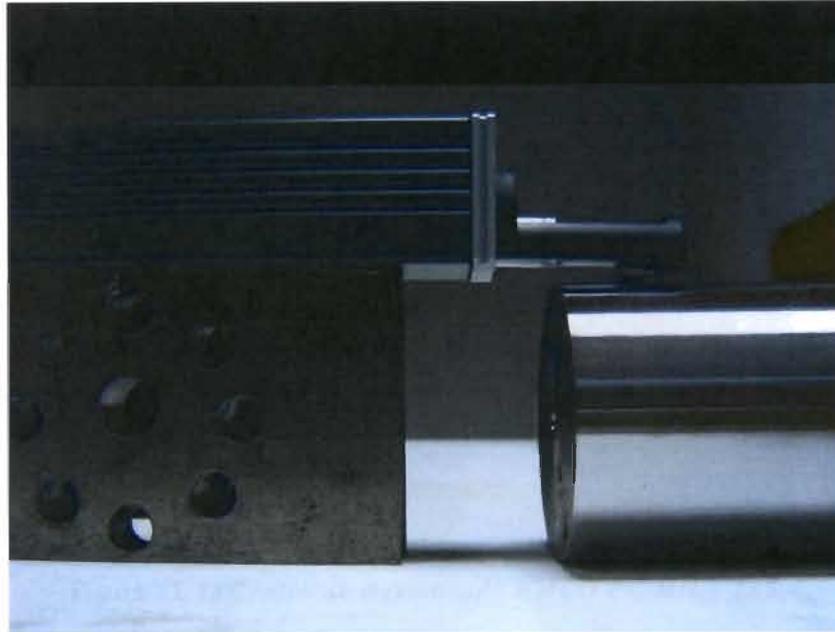


Figura VI. 12 *Uso del rugosímetro en cada cilindro mecanizado.*

Para el proceso de fresado, se hizo el mecanizado de las piezas usando un modelo de máquina de fresado EMCO PC MILL 125 Figura. VI.13. Una herramienta de acero para un fresado rápido de $9,6 \times 9,6$ mm el cual se utiliza para todos los materiales. Las piezas en forma de placa se ocupan para el mecanizado teniendo en cuenta los tres tipos de materiales, el aluminio 6062, el acero 1018 y el acero 4140 Figura VI.14. La rugosidad se midió utilizando el mismo rugosímetro Mitutoyo SJ-301 Figura VI.11 teniendo en cuenta la configuración mostrada en la Figura VI.15 [38] y [39].



Figura VI. 13 Centro de mecanizado EMCO PC MILL 125.



Figura VI. 14 Placas por ser mecanizadas hechas de los tres materiales, aluminio 6062, acero 1018 y del acero 4140.

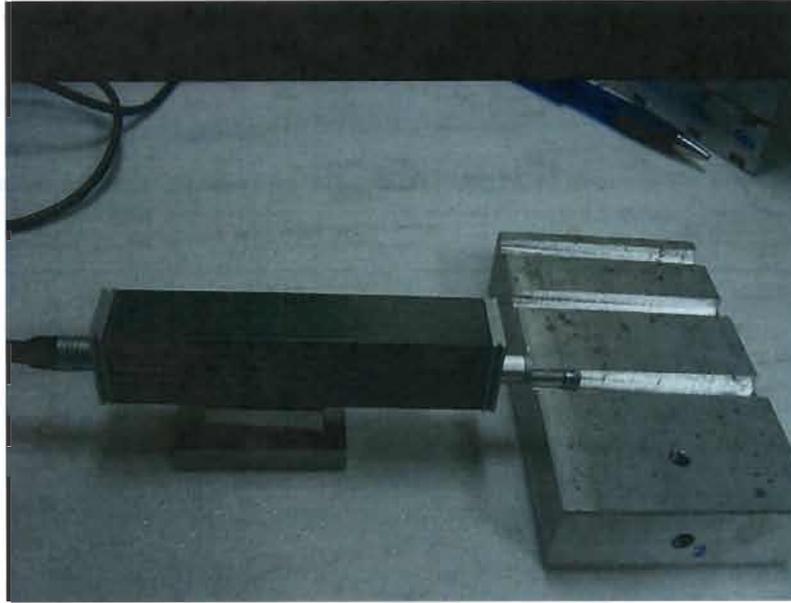


Figura VI. 15 *Utilización del rugosímetro en cada placa mecanizada.*

VI.2.1. Etapa de detección para el proceso de mecanizado.

Para la etapa de detección se generó un sistema difuso que cuenta con 12 reglas difusas las cuales describen el comportamiento en modo normal de operación. Para describir el sistema, se tomó en consideración una rugosidad máxima de $3 \mu\text{m}$ dado que para cada material varían los niveles de rugosidad requeridos por el cliente, por ello, se consideró una rugosidad máxima $3 \mu\text{m}$ debido a que esta medida está por debajo de la rugosidad máxima requerida para los tres tipos de materiales. Para realizar la detección el sistema difuso considera una salida que se encuentra entre 0 y .5 cuando el sistema está trabajando apropiadamente pero si la salida del sistema difuso se encuentra entre .51 y 1 el sistema se encuentra en falla. El sistema fue generado modelando el comportamiento del avance y velocidad en modo normal de operación, la dureza del material no se consideró debido a que no es significativo para la etapa de detección. La Tabla VI.5 muestra las reglas

difusas que se generaron para hacer la detección de un modo anormal de operación.

Tabla VI. 5 Reglas difusas para la detección en el proceso de mecanizado.

Regla	Velocidad	Avance	Rugosidad
1	Lenta	Normal	Buena
2	Lenta	Rápido	Malo
3	Rápida	Normal	Buena
4	Rápida	Rápido	Malo
5	Lenta	Normal	Buena
6	Lenta	Rápido	Malo
7	Rápida	Normal	Buena
8	Rápida	Rápido	Malo
9	Lenta	Normal	Buena
10	Lenta	Rápido	Malo
11	Rápida	Normal	Buena
12	Rápida	Rápido	Malo

El sistema difuso puede explicar con un alto porcentaje de eficiencia el comportamiento del maquinado en su modo normal de operación para los tres tipos de materiales, el sistema difuso tiene una eficiencia del 100%, esta validación se realizó utilizando el 80% de los datos muestreados para realizar el entrenamiento y un 20% de ellos para realizar la validación del sistema difuso el cual presenta un buen rendimiento para la etapa de detección.

VI.2.2. Etapa de diagnóstico para el proceso de mecanizado

En la etapa de diagnóstico se utilizó un SVM multiclase y una PNN como se describe en la metodología presentada en el capítulo anterior. El SVM hace una primera inspección para otorgarle a los datos analizados una clase la cual describe el modo de operación del proceso y con el cual se puede diagnosticar donde y cuando ocurre la falla dentro del sistema, si el SVM no logra clasificar el comportamiento en una determinada clase, la red neuronal realiza la clasificación para dar una segunda aproximación, y con la información de la PNN se pueda tener información de donde se puede encontrar la falla y si es necesario generar una nueva clase que describa ese nuevo comportamiento en falla, y con ello caracterizar un nuevo tipo de falla que no se haya presentado aún. El SVM y la PNN fueron entrenadas utilizando los datos de dureza, avance y velocidad en esta etapa de diagnóstico se utilizaron las tres variables a diferencia de la etapa de detección, esto debido a que con la dureza se logra una mejor clasificación al identificar con mayor sencillez que tipo de material es el que se está mecanizando. Para el entrenamiento de la SVM y de la PNN se generaron 12 grupos, La SVM utilizó 12 pares de grupos para realizar el diagnóstico y la PNN en su primera capa utilizó 12 neuronas de entrenamiento y en la segunda 12 neuronas competidoras. Tanto la SVM y la PNN se entrenó con 36 épocas debido a que para este caso de aplicación se utilizaron los promedios de los mecanizados del torno y la fresadora tanto para la elaboración del sistema difuso, de la SVM y la PNN se utilizó el toolbox de MatLab®. La Tabla VI.6 muestra el comportamiento de la metodología.

Tabla VI. 6 Comparación de la metodología en el proceso de mecanizado.

	Detección utilizando el Sistema Difuso	Diagnostico con la SVM	Diagnostico con la PNN	Detección y diagnostico con la combinación de un sistema difuso la SVM y la PNN
%	100	99	99	99

En la Tabla VI.6 se presenta el desempeño de la metodología propuesta, dicha metodología muestra un desempeño aceptable para la tarea de detección y diagnóstico. Al igual que en la etapa de detección la etapa de diagnóstico se entrenó con un 80% de la muestra y con un 20% se realizó la validación. La muestra es conformada por 42 datos que contienen tanto el estado de falla como el modo de operación normal, los datos son promedios del muestreo de los dos tipos de mecanizado. El porcentaje se obtuvo al analizar el 20% de los datos utilizados en la validación, donde se obtuvo al evaluar el número de muestras clasificadas correctamente sobre el total de muestras utilizadas para la validación, esto se realizó debido a que se tenía conocimiento previo de la pertenecía de dicho grupo de datos a una clase específica.

Como se muestra, tanto la SVM y la PNN muestran un rendimiento aceptable al realizar la clasificación en el paso de diagnóstico, ya que estas presentan un 99% de eficiencia, la combinación de metodologías es para hacer un sistema más robusto y evitar la presencia de falsas alarmas.

VII. CONCLUSIONES

- En el presente trabajo se desarrolló una nueva metodología para la detección y diagnóstico de fallas. Con esta nueva metodología se previene la presencia de falsas alarmas al hacer un diagnóstico seguro, con ello también se previenen riesgos de accidentes a los operadores que trabajan en procesos peligrosos.
- Con la implementación de la metodología antes descrita se logró alcanzar el objetivo deseado para este trabajo el de detectar y diagnosticar las fallas presentes en el proceso o sistema.
- La combinación de metodologías muestra un rendimiento aceptable para realizar las tareas de detección y diagnóstico. Cabe tomar en cuenta que los datos pueden ser estandarizados para mejorar el desempeño de dicha metodología o ser ayudados por alguna técnica estadística básica para realizar una mejor detección y diagnóstico.
- La técnica de lógica difusa en ambos caso de estudio demuestra su sencillez al hacer más fácil el proceso de detección que al utilizar otras técnicas.
- Tanto la SVM y la PNN muestran un gran desempeño en la clasificación de patrones pero como se describe en el siguiente punto la SVM uno contra todos fue seleccionada debido a que muestra una cualidad adicional al SVM uno contra uno y a la PNN.
- Como se describió en los dos casos de estudio se decidió por utilizar el SVM uno contra todos debido a que este puede identificar si los datos de entrada no pertenecen a ningún grupo a diferencia de la PNN y a la SVM uno contra uno.

- La metodología muestra una gran efectividad ya que a diferencia de las metodologías vistas en el estado del arte la metodología empleada aprovecha los atributos de cada una de las técnicas para realizar una correcta detección y diagnóstico en cualquier tipo de sistema o proceso en el que es empleada. La Tabla VII.1 muestra una comparación del análisis del estado del arte contra la metodología propuesta.
- La metodología propuesta se limitó solo a la detección y diagnóstico dado que la recuperación del proceso o sistema es un trabajo arduo el cual se puede desarrollar en posteriores trabajos de investigación.
- Una de las más grandes limitantes que tiene la metodología implementada es que ella no puede trabajar con datos faltantes en los históricos del proceso o sistema. Y es dependiente de la calidad y cantidad de dichos datos para realizar la detección y diagnóstico de dichas fallas. La limitante del faltante de datos puede ser manejada en trabajos futuros.

Tabla VII. *Comparación del estado del arte contra la metodología propuesta.*

AUTOR	TÉCNICAS	MEJOR	:POR QUÉ?	METODOLOGÍA PROPUESTA
Karim	SVM y ANFIS.	No	determina	Porque ocupa el SVM en La metodología propuesta es una sucesión
Salahshoor		una mejor.		ciertos casos y el ANFIS. de un sistema difuso una SVM y una PNN
et al 2010.				en otros para realizar un mejor diagnóstico. por lo cual la metodología no determina mejor diagnóstico. cual es la mejor. hace uso de los atributos individuales de cada técnica para poder trabajar en distintos casos de aplicación y poder tener un buen rendimiento.
S. Deng et al. 2011.	SVM y BPNN	No	determina	Determina que para ciertos casos, con un número menor de datos encuentra una mayor eficiencia en la SVM y cuando se cuenta con un mayor número de datos encuentra eficiente a las BPNN.
		cual es mejor.		En la metodología propuesta no es importante contar con un gran número de datos ya que SVM presenta un buen rendimiento con un bajo número de datos como se mostró en el caso de aplicación I al utilizar promedios en su entrenamiento y realizar un diagnóstico correcto. Y este diagnóstico es corroborado por la PNN.

- A. Azadeh. SVM con hiper-La SVM y los Identifica a las VSM como La metodología que se propone no et al 2012. parámetros de SVM con hiper- mejores clasificadores para optimiza pero propone un sistema optimización y parámetros de poder diagnosticar la falla. redundante para garantizar un buen ANN. optimización. diagnóstico al hacer uso de las tres metodologías.
- Ali Ajami Análisis de ICA. Selección como mejor La metodología propuesta puede trabajar et al. 2012. componente independiente herramienta el utilizar un con ruido. Y presenta una mayor (ICA) y análisis de escapes de trabajar con sistema ICA ya que este simplicidad que al implementar un ICA. componentes sistemas con ruido.
- Juan Pablo PCA. Lógica PNN. Debido a que trabaja con La metodología propuesta también utiliza un Nieto et al probabílistica y Redes procesos y mediciones con PNN pero utiliza el SVM para trabajar con un 2009 Neuronales ruido. menor número de datos de entrenamiento probabílisticas.
-

IX. REFERENCIAS

- [1] José Francisco Vilar Barrido. "Control Estadístico de los Procesos (SPC)". FC editorial. (2005), Pag. 1-30.
- [2] Douglas C. Montgomery. "Introduction to Statistical Quality Control". Sexta edición. John Wiley & Sons, Inc. (2009), Pag. 1-15.
- [3] Ali Ajami y Mahdi Daneshvar. "Data driven approach for fault detection and diagnosis of turbine in thermal power plant using Independent Component Analysis (ICA)". *Electrical Power and Energy Systems* (2012), Pag. 728–735.
- [4] L.H. Chiang, E.L. Russell and R.D. Braatz. "Fault detection and diagnosis in industrial systems". Springer (2000), Pag. 6-280.
- [5] Juan Andrés Cadena, Juan Mauricio Cadena y Sandra Milena Pérez. "Aplicación de redes neuronales probabilísticas en la detección de fallas incipientes en transformadores". *scientia et technica*, vol. xiv, núm. 39, septiembre, (2008), Pag. 48-53.
- [6] Fernando López, José Miguel Valiente, José Manuel Prats y Alberto Ferrer. "Performance evaluation of soft color texture descriptors for surface grading using experimental design and logistic regression". *Pattern Recognition* (2008), Pag. 1744 – 1755.
- [7] Juan Pablo Nieto González, Luis Garza Castañón y Rubén Morales Menéndez. "Multiple Fault Diagnosis in Electrical Power Systems with Dynamic Load Changes Using Probabilistic Neural Networks". *Computación y Sistemas* Vol. 14 No. 1, (2010), Pag. 17-30.
- [8] Ribhan Zafira Abdul Rahman, Azura Che Soh y Noor Fadzlina binti Muhammad. "Fault detection and diagnosis for continuous stirred tank reactor using neural network". *Kathmandu university journal of science, engineering and technology* vol. 6, no. ii, november, (2010), Pag. 66-74.

- [9] P. Subbaraj y B. Kannapiran, "Artificial Neural Network Approach for Fault Detection in Pneumatic Valve in Cooler Water Spray System". *International Journal of Computer Applications*, Volumen 9– No.7, November (2010), Pag. 43-51.
- [10] Karim Salahshoor, Majid Soleimani Khoshro y Mojtaba Kordestani, "Fault detection and diagnosis of an industrial steam turbine using a distributed configuration of adaptive neuro-fuzzy inference systems". *Simulation Modelling Practice and Theory* (2011), Pag. 1280–1293.
- [11] M. Arunachalam y V. Rajamani, "Distributed Fault Detection and Localization Algorithm using Artificial Neural Network in Optical WDM Networks". *European Journal of Scientific Research* ISSN 1450-216X Vol.56 No.2 (2011), Pag.194-203.
- [12] Karim Salahshoor, Majid Soleimani Khoshro y Mojtaba Kordestani, "Fault detection and diagnosis of an industrial steam turbine using a distributed configuration of adaptive neuro-fuzzy inference systems". *Simulation Modelling Practice and Theory* (2011), Pag. 1280–1293.
- [13] S. Deng, Seng-Yi Lin, We-Luan Chang, "Application of multiclass support vector machines for fault diagnosis of field air defense gun". *Expert Systems with Applications* .(2011), Pag. 6007–6013.
- [14] Tony Boutros y MingLiang, February, "Detection and diagnosis of bearing and cutting tool faults using hidden Markov models". *Mechanical Systems and Signal Processing* (2011), Pag. 2102–2124.
- [15] A. zadeh, M. Saberi, A. Kazem, V. Ebrahimipour, A. Nourmohammadzadeh y Z. Saberi, "A flexible algorithm for fault diagnosis in a centrifugal pump with corrupted data and noise based on ANN and support vector machine with hyper-parameters optimization". *Applied Soft Computing* (2013), Pag. 1478–1485.
- [16] Manjeevan Seera, Chee Peng Lim, Dahaman Ishak y Harapajan Singh "Offline and online fault detection and diagnosis of induction motors using a hybrid soft computing". *Applied Soft Computing* 13 (2013), Pag. 4493–4507.

- [17] P. Subbaraj y B. Kannapiran. "Fault detection and diagnosis of pneumatic valve using Adaptive Neuro-Fuzzy Inference System approach". *Applied Soft Computing*, (2014).
- [18] Zhimin Du, Bo Fan, Xinqiao y Jin, Jinlei Chi "Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis". *Building and Environment* 73 (2014), Pag. 1-11.
- [19] Jeevanand Seshadrinath, Bhim Singh, Fellow y B. K. Panigrahi, "Vibration analysis based interturn fault diagnosis in induction machines". *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS*, VOL. 10, NO. 1, FEBRUARY 2014, (2014), Pag. 340-350.
- [20] Yousef Shatnawi y Mahmood Al-khassaweneh, Senior Member, IEEE, "Fault Diagnosis in Internal Combustion Engines Using Extension Neural Network". *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*, VOL. 61, NO. 3, MARCH (2014), Pag. 1434-1444.
- [21] Fafa Chen, Baoping Tang, Tao Song y Li Li, "Multi-fault diagnosis study on roller bearing based on multi-kernel support vector machine with chaotic particle swarm optimization". *Measurement* 47 (2014), Pag. 576-590.
- [22] Julio César Ramírez Valenzuela. "Diagnóstico de fallos en sistemas industriales basado en razonamiento borroso y posibilístico". *Universidad politécnica de valencia Departamento de Ingeniería de Sistemas y Automática* (2007), Pag. 18-42.
- [23] Venkatasubramanian V., Rengaswamy R., Yin K. & Kavuri S. . "A review of process fault detection and diagnosis Part I. *Computers and Chemical Engineering*. 27 (2003), Pag. 293-311.
- [24] Cox, Earl, "Fuzzy fundamentals", *IEEE Spectrum* (1992), Pag. 58-61.
- [25] K. V. Mardia, J. T. Kent y J. M. Bibby, "Multivariate analysis". *ACADEMIC PRESS* (1995), Pag. 80-140.
- [26] Hongya Ge, Ivars P. Kirsteins and Xiaoli Wang, (2009). "Does canonical correlation analysis provide reliable information on data correlation in array processing", 978-1-4244-2354-5 09 (2009) IEEE, Pag. 2013-2017.

- [27] Sharma y Subbash. "Applied Multivariate Techniques". Ed.: Hohn Wiley and Sons, Inc (1996).
- [28] Elena Deza & Michel Marie Deza. "Encyclopedia of Distances". Springer (2009), Pag. 94-105.
- [29] S.Y. Sohna y H.W. Shin. "Experimental study for the comparison of classifier combination methods". Pattern Recognition (2007). Pag. 33 – 40.
- [30] Fernando Gómez Salas. "Sistemas difusos jerárquicos para modelado y control". Centro de investigación y de estudios avanzados del instituto politécnico nacional, departamento de control automático (2005), Pag. 8-18.
- [31] Antonio López Jaimes. "Diseño de un algoritmo evolutivo multiobjetivo paralelo". Centro de investigación y de estudios avanzados del instituto politécnico nacional, Departamento de Ingeniería Eléctrica Sección de Computación (2005). Pag. 12-25.
- [32] Raúl Rojas. "Neural Networks A Systematic Introduction". Springer Berlin, Heidelberg, NewYork, HongKong, London Milan, Paris y Tokyo (1996), Pag. 3-52.
- [33] Stuart J. Russell y Peter Norvig. "Inteligencia artificial un enfoque moderno". Segunda edición, Pearson educacion. S.A., Madrid. (2004), Pag. 561-600.
- [34] Li-Xin Wang. "A curse in fuzzy systems and control". International edition (1997), Pag. 1-67.
- [35] Martin T. Hagan. "Neural Network desing", Oklahoma satate University, Mark Beale MHB, Inc., (1996), pp. 14.1-14.15.
- [36] Shigeo Abe. "Support Vector Machines for Pattern Classification". Advances in Pattern Recognition ISSN 1617-7916 Springer link, (2005), Pag. 1-82.
- [37] Mehdi Namdari, Hooshang Jazayeri-Rad. "Incipient fault diagnosis using support vector machines based on monitoring continuous decision functions". Engineering Applications of Artificial Intelligence 28 (2014), 2014, Pag. 22-35.

- [38] Indira Gary Escamilla Salazar, “desarrollo de una metodología usando un sistema inteligente para identificar el desempeño de las Variables envueltas en el maquinado de alta velocidad del TI 6AL-4V” corporación mexicana de investigación en materiales División de estudios de posgrado (2011).
- [39] Luis M. Torres-Treviño, Indira G. Escamilla-Salazar, Bernardo González-Ortíz y Rolando Praga-Alejo, “An expert system for setting parameters in machining processes” *Expert Systems with Applications* 40 (2013) Pag. 6877–6884.
- [40] Juan Pablo Nieto González, “Complex Systems Fault Diagnosis Using Soft Computing and Statistical Methods”, In: Instituto Tecnológico y de Estudios Superiores de Moterrey Campus Monterrey, Escuela de Ingeniería y Tecnologías de Información (2012).