



IMPLEMENTACION DE MANTENIMIENTO PREDICTIVO PARA UNA BOMBA CENTRIFUGA UTILIZANDO MACHINE LEARNING

Daniel Felipe Soto Zabala.

Universidad Antonio Nariño
Facultad de Ingeniería Mecánica, Electrónica y Biomédica
Bogotá, Colombia
2021

IMPLEMENTACION DE MANTENIMIENTO PREDICTIVO PARA UNA BOMBA CENTRIFUGA UTILIZANDO MACHINE LEARNING

Daniel Felipe Soto Zabala.

Proyecto de grado presentado como requisito parcial para optar al título de Ingeniero
Mecatrónico.

Director:

PhD. Sergio Andrés Díaz Salas

Codirector (a):

PhD. Diego Calero Arellano

Línea de Investigación:

Mantenimiento predictivo en maquinaria industrial.

Grupo de Investigación:

GIBIO

Universidad Antonio Nariño

Facultad de Ingeniería Mecánica, Electrónica y Biomédica

Ciudad, Colombia

2021

Dedicatoria

*A mi familia, compañeros, amigos y docentes
que dieron su apoyo a lo largo de mi formación.*

Agradecimientos

Agradezco en primera instancia a mi director Sergio Andrés Díaz Salas y Codirector Diego Calero, docentes y amigos de la Universidad Antonio Nariño por aportar opiniones y correcciones sobre este documento.

Además, quiero agradecer a los jurados Mario Enrique Duarte y Diego Alejandro Pardo con los que ya he tenido el gusto de aprender de ellos en la universidad y que aportan con correcciones y opiniones para hacer de este trabajo uno de los mejores.

Resumen

En este documento se presenta una aplicación de machine learning como mantenimiento predictivo en una bomba centrífuga simulada, es una estrategia de cómo utilizar esta rama de la inteligencia artificial como técnica de mantenimiento. Por medio de simulaciones discretas se extraen datos de variables dinámicas como torque, presión y caudal de una bomba centrífuga en un sistema de suministro de agua común, se toman datos de la bomba a dos rangos de velocidades en el rotor de la bomba. Con un rango de velocidades de 1780 a 6200 RPM se simula la bomba en funcionamiento normal y a velocidades de 6201 a 9500 RPM se simula la bomba con fallas ocurridas por la alta velocidad rotacional en el eje, luego de obtener los 668 datos de la clase de falla y los 884 datos de la clase de no falla se presentan 2 modelos de clasificación como el modelo de regresión lineal de Bayes y el modelo kNN que ayudan a realizar las predicciones de datos nuevos que pueda suministrar la bomba en cualquier instante y conocer su condición. Para la evaluación de los modelos se presentan las medidas más comúnmente utilizadas en este tipo de trabajo tales como la tasa de clasificación errónea y la precisión del modelo, además se presenta un cuadro de confusión y diagrama de coordenadas paralelas en el que se observa una precisión del 100% en ambos modelos por la poca complejidad de la distribución y asignación de datos de ambas clases. El documento finalmente presenta una estrategia para el mantenimiento predictivo que, aunque se trabaja solamente con simulaciones, podría ser utilizado en un entorno real en el que los datos sean obtenidos por un monitoreo constante de la máquina.

Palabras clave: machine learning, turbo máquina, predicción, prevención, corrección.

Abstract

This document presents an application of machine learning as predictive maintenance in a simulated centrifugal pump, it is a strategy of how to use this branch of artificial intelligence as a maintenance technique. By means of discrete simulations data are extracted from dynamic variables such as torque, pressure and flow from a centrifugal pump in a common water supply system, data is taken from the pump at two speed ranges in the pump rotor. With a speed range from 1780 to 6200 RPM, the pump is simulated in normal operation and at speeds from 6201 to 9500 RPM, the pump is simulated with failures caused by high rotational speed in the shaft, after obtaining the 668 data of the failure and the 884 data of the non-failure class are presented 2 classification models such as the Bayes linear regression model and the kNN model that help to make the predictions of new data that the pump can supply at any moment and know its condition. For the evaluation of the models, the measures most commonly used in this type of work are presented, such as the rate of misclassification and the precision of the model, in addition, a confusion table and parallel coordinate diagram are presented in which a precision is observed. 100% in both models due to the low complexity of the distribution and allocation of data for both classes. The document finally presents a strategy for predictive maintenance that, although it works only with simulations, could be used in a real environment in which the data is obtained by constant monitoring of the machine.

Keywords: machine learning, turbo máquina, predicción, prevención, corrección.

Contenido

Tabla de contenido

1. Introducción	XVII
2. Planteamiento del problema.....	20
2.1 Problemática.....	20
2.2 Justificación.....	23
2.3 Objetivos.....	26
2.3.1 Objetivo general y específicos.....	26
2.4 Alcance	26
3. Marco conceptual y teórico.....	28
3.1 Uso de machine learning en la actualidad.....	28
3.2 Machine learning.....	29
3.2.1 Machine learning supervisado.....	30
3.2.2 Machine learning no supervisado.....	31
3.2.3 Machine learning reforzado.....	32
3.3 Algoritmos de machine learning.....	33
3.3.1 Algoritmos de regresión.....	33
3.3.2 Algoritmos bayesianos.....	33
3.3.3 Algoritmos de agrupación.....	34
3.3.4 Algoritmos de decisión.....	34
3.3.5 Algoritmos de redes neuronales.....	35
3.3.6 Algoritmos de aprendizaje profundo o Deep learning.....	35
3.4 Mantenimiento predictivo.....	35
3.5 Mantenimiento correctivo.....	37
3.6 Mantenimiento preventivo.....	38
3.7 Bombas centrifugas.....	39
3.7.1 Configuración básica.....	39
3.7.2 Ventajas de las bombas centrifugas.....	40
3.7.3 Tipos de bombas centrifugas.....	41
4. Planteamiento de la solución.....	45
4.1 Extracción de datos del modelo para ambas clases.....	45
4.2 Almacenamiento y pre procesado de los datos.....	51
4.3 Creación del modelo de clasificación.....	58
4.4 Realización de predicciones.....	61
4.5 Evaluación del modelo.....	64
5. Conclusiones y recomendaciones.....	69
5.1 Conclusiones.....	69

5.2	Recomendaciones.....	70
6.	Anexos.....	71
6.1	Anexo A.....	71
6.2	Anexo B.....	78
6.3	Anexo C.....	83
6.4	Anexo D.....	84

Lista de figuras

	Pág.
Figura 2-1. Bomba centrífuga convencional [15].	20
Figura 2-2. Anotación de datos en estrategias de machine learning actualmente [12].	25
Figura 3-1. Bomba centrífuga. [19].	40
Figura 3-2. Bomba centrífuga radial. [21].	41
Figura 3-3. Bomba centrífuga axial. [21].	42
Figura 3-4. Bomba centrífuga de flujo mixto. [21].	42
Figura 3-5. Bombas centrífugas en posición horizontal. [21].	43
Figura 3-6. Bombas centrífugas en posición vertical. [21].	44
Figura 4-1. Modelo de sistema de suministro de agua de Matlab. [29].	47
Figura 4-2. Bloque para propiedades del fluido.	47
Figura 4-3. Tanque de suministro.	48
Figura 4-4. Valores de volumen.	48
Figura 4-5. Estación de suministro.	49
Figura 4-6. Importación de los datos a Matlab.	52
Figura 4-7. Funcion de suavizado gaussiano de datos. [28].	53
Figura 4-8. Gráfico de presión con respecto a la velocidad.	54
Figura 4-9. Gráfico de caudal con respecto a la velocidad.	55
Figura 4-10. Gráfico de torque en el eje con respecto a la velocidad.	56
Figura 4-11. Gráfico de presión con respecto a la velocidad.	57
Figura 4-12. Gráfico de caudal con respecto a la velocidad.	57
Figura 4-13. Gráfico de torque con respecto a la velocidad.	58
Figura 4-14. Cuadro de confusión.	67
Figura 4-15. Diagrama de coordenadas paralelas.	67

Lista de tablas

	Pág.
Tabla 2-1. Artículos relacionados.....	21
Tabla 4-1. Resultados de la bomba centrífuga a 8000 rpm.....	50
Tabla 4-2. Resultados de la bomba centrífuga a 2000 rpm.....	50
Tabla 4-3. Variación estándar de la clase de falla.....	53
Tabla 4-4. Variación estándar de la clase de no falla.....	53
Tabla 6-1. Datos de clase no falla.....	71
Tabla 6-2. Datos de clase falla.	78
Tabla 6-3. Datos de prueba.	83

Lista de Símbolos y abreviaturas

Símbolos con letras latinas

Símbolo	Término	Unidad SI	Definición
P	Presión	Pa, N/m^2	$P=F/A$
F	Fuerza	F	$F=m*a$
A	Área transversal	m^2	$\pi r^2/2$
A	Aceleración	m/s^2	$\frac{\Delta v}{\Delta t}$
T	Torque	Nm	$T=F*d$
D	Distancia	m	$\int v dt$
Q	Caudal	m^3/s	$Q = V * A$

Símbolos con letras griegas

Símbolo	Término	Unidad SI	Definición
Ω	Velocidad angular	Rad/s	$\Delta\theta/\Delta t$
θ	Desplazamiento angular	rads	$\int \omega dt$

Subíndices

Subíndice Término

V	Vibración
T	Temperature

Superíndices**Superíndice Término**

E	Exponente, potencia
---	---------------------

Abreviaturas**Abreviatura Término**

-
1. *PdM* Mantenimiento predictivo
 2. *MCL* Machine Learning
 3. *GMAO* Gestion de mantenimiento asistida por ordenador
 4. *IA* Inteligencia artificial
 5. *TCE* Tasa de clasificacion erronea
 6. *kNN* K nearest neighbor

1.Introducción

Actualmente las técnicas de mantenimiento son propuestas dadas a los responsables de la gestión de una planta o cualquier tipo de maquinaria para mantenerla en operación sin ninguna interrupción . En este documento se presenta una de estas estrategias enfocada al mantenimiento predictivo y haciendo uso del machine learning o aprendizaje automático.

Debido a que la industria ha tenido un gran incremento en los últimos años es necesario comenzar a incluir nuevas tecnologías tanto en la operación de la maquinaria como en el mantenimiento, de forma que resulte más conveniente para una empresa y sus demás factores productivos.

Para poner en contexto al lector se presentará una idea que, aunque ya se utiliza en algunas industrias colombianas y es ofrecida por empresas como Preditec S.A. tiene utilidad al momento de realizar mantenimientos preventivos a varios tipos de maquinaria y también evita realizar mantenimientos correctivos que puedan generar pérdidas tanto económicas como en producción. Esta idea consiste en utilizar un modelo entrenado con datos de aprendizaje obtenidos por medio de una simulación de un sistema de suministro de agua, el sistema está compuesto de tres estaciones de bombeo que a su vez contienen 2 bombas centrífugas interconectadas.

La relación de este sistema de suministro de agua en el trabajo es únicamente para que el lector se percate del que podría ser uno de los usos de esta estrategia en un sistema real que se componga de turbo máquinas y, además, es un sistema comúnmente utilizado en la industria actual que tiene participación en desplazamiento de fluidos, generación de energía e incluso en sectores públicos como lo es el servicio de abastecimiento de agua.

El trabajo no se enfoca en el estudio del sistema de suministro sino únicamente en la bomba centrífuga. Este modelo será capaz de predecir gracias a 1552 datos de aprendizaje, cuando la bomba se encuentre próxima a fallar gracias a la medición de presión, torque y caudal será posible entrenar el modelo para que cuando se presenten datos nuevos ya sea por medio de un monitoreo constante de estas tres variables el algoritmo de una alerta al operador.

El uso de machine learning ha sido de gran utilidad en varios campos del mercado y la industria de manufactura tal como lo son, por ejemplo:

- Análisis de compradores y consumidores:

Desde la implementación del Machine Learning en la mercadotecnia, se logra hacer reconocimientos en bases de datos históricos, por ejemplo, entre que clientes está la probabilidad de no lograr una permanencia y también cuales ventas tienden a realizarse o cerrarse, para esto es necesario el aprendizaje supervisado e incluir etiquetas en la información de los datos de aprendizaje para que el sistema logre analizar de forma predeterminada el comportamiento de los clientes, que compras hicieron, que correos enviaron, el número de visitas a la página, sus llamadas etc. Todo lo anterior con el fin de establecer un patrón y en base a eso buscar en el resto de los datos cuando se repiten estos comportamientos he identificar el usuario.

- Seguridad de la empresa:

Gracias al reconocimiento de imágenes, al implementar esta función en una empresa es posible detectar los rostros y rasgos de comportamiento del equipo de trabajo, ya que el sistema es alimentado previamente con datos fotográficos y de video, estos datos sirven como datos de aprendizaje para que se establezcan patrones en las características de los rostros y así logre identificar el personal que ingresa o abandona el lugar.

Es decir, la implementación consiste en lograr que cámaras de seguridad sean capaces de identificar y dar información de quién y en donde se encuentra cada persona ya sea en un conjunto de oficinas o cerca de ellas, además, podrá identificar objetos que la empresa considere importante restringir e identificar al portador, en el caso específico de seguridad podrían ser armas u objetos prohibidos en algunas áreas como por ejemplo dispositivos electrónicos, celulares o armas de fuego.

Adicionalmente Preditec afirma que el mantenimiento predictivo, realiza evaluaciones al estado de la maquinaria y recomienda intervenir cuando sea necesario evitando así costos por mantenimientos preventivos innecesarios. El mantenimiento basado en la condición se basa en mejorar la puntualidad y optimización de procesos por medio de datos en tiempo real este es capaz de reconocer cualquier tipo de anomalía, avisando al operador para su revisión inmediata. [9]

El mantenimiento predictivo incluye estrategias, para tomar características de la condición en la que opera cualquier equipo o maquinaria por medio de mediciones y análisis de variables dinámicas, el motivo principal para implementar este tipo de mantenimiento es su fiabilidad y su manera de reducir costos ya sean de tiempo o dinero, al momento de realizar mantenimiento en la maquinaria. [9]

Básicamente la implementación como la realizada en este trabajo de grado en un ambiente real puede tener un alto impacto en industrias de generación de energía o de circulación de fluidos en las que se utilice una gran cantidad de bombas centrífugas u otro tipo de turbomáquinas.

2. Planteamiento del problema

2.1 Problemática.

Se pretende por medio de un modelo computacional, medir el comportamiento de variables dinámicas que puedan presentar fallas en una bomba centrífuga. Para que con el estudio de estos comportamientos se logre analizar las partes defectuosas, la causa de estas y su frecuencia utilizando un algoritmo de mantenimiento predictivo. Esto con el objetivo de evitar mantenimientos correctivos que puedan ser necesarios al momento de presentarse estas fallas y también evitar los mantenimientos preventivos que, aunque sean los más recomendados, estos presentan altos sobrecostos debido a su periodicidad.

El modelo estará basado en la operación de una bomba centrífuga como la que se muestra en la FIGURA 2, tiene un campo de utilidad en las industrias que se enfocan en la generación de energía, ya que, utilizan turbo máquinas en las que sus fallas las podemos definir con base a su operación y a la carga de trabajo a la que están expuestas.

Figura 2-1.Bomba centrífuga convencional [15].



Revisando los trabajos anteriores y artículos de investigación en el mantenimiento predictivo de bombas centrífugas y demás tipos de turbo maquinaria se han medido las siguientes variables:

Tabla 2-1. Artículos relacionados.

ARTICULO	VARIABLES DINAMICAS EN MEDICION.
VIBRACION EN TURBINAS DE VAPOR [1]	Vibración Torque Presión Caudal
LOT ENABLED VIBRATION MONITORING TOWARD SMART MAINTENANCE [2]	Vibración. Caudal Corriente eléctrica Volumen de flujo.
VIBRATION ANALYSIS OF ROTARY MACHINES USING MACHINE LEARNING TECHNIQUE [3]	Torque Vibración. Potencia.
DEGRADATION ASSESSMENT AND REMAINING USEFUL LIFE PROGNOSTIC OF CENTRIFUGAL PUMP USING MULTI-SENSOR.[7]	Perdidas térmicas. Vibración. Caudal. Temperatura.

TOWARDS AN INFORMATION SYSTEM FOR PROGNOSTICS AND HEALTH MANAGEMENT ANALYSIS.[8]	Vibración Temperatura Perdida de flujo Resistencia térmica caudal
--	---

Como podemos observar, las variables en medición para generar un modelo de mantenimiento predictivo para este tipo de sistemas más adecuadas serían la presión, torque y caudal conforme a la velocidad en el rotor de la bomba, ya que al medir estas variables podemos identificar por medio de datos de funcionamiento cuando se presenten, también las fallas más comunes que se suelen presentar en este tipo de máquinas tienen que ver con la vibración y la temperatura en la bomba.

La importancia de la implementación en fallas para mantenimiento predictivo en una bomba centrífuga, recae en que las fallas más usuales son ocasionadas por causa de una mala supervisión de ingenieros y operadores que ignoran las capacidades imparciales de las bombas, esto puede generar anomalías en los caudales de líquido en la bomba y así ocasionar fallos en los sistemas de la bomba, estas anomalías, pueden generar inestabilidad hidráulica y por consiguiente, desgastes acelerados y flexiones del eje además pueden existir fallas prematuras en algunas partes de la bomba pero más frecuentemente en los impulsores.

Según lo afirma Isserman, las bombas centrífugas sufren daños en sus elementos ya sean hidráulicos o mecánicos, siendo los que más frecuentan fallas los rodamientos de bolas y los sellos de anillo deslizante, aunque tampoco pueden ser ignoradas las fallas en otros componentes como las palas del impulsor, el motor impulsor y los rodamientos deslizantes. [11].

Las fallas más frecuentes en las bombas centrífugas y en la mayoría de turbo maquinas son las siguientes [11]:

- Cavitación: desarrollo de pequeñas burbujas de aire caliente dentro del fluido transportado, si la presión estática dentro de la bomba cae por debajo de la presión del aire. Las burbujas pueden colapsar repentinamente ocasionando daños en las ruedas de la hoja.
- Gas en el fluido: una caída de presión puede producir gas o vapor mezclado en el fluido. Produciendo una separación de estos y aligerando abruptamente la carga dentro de la bomba.
- Funcionamiento en seco: la falta de líquido en recorrido provoca poco enfriamiento y el sobrecalentamiento del rodamiento.
- Erosión: es el daño mecánico producido por la cavitación o presencia de partículas en el fluido.
- Corrosión: Daños por fluidos agresivos.
- Desgaste de los rodamientos: daños mecánicos por esfuerzos innecesarios y fricción del metal, esto provoca que en los rodamientos se generen picaduras o desgarros.
- Obstrucción de los orificios de alivio: conduce a la sobrecarga y al daño de los cojinetes transversos al eje de la bomba centrífuga.
- Obstrucción de los sellos de anillo deslizante: conduce a una mayor fricción y una menor eficiencia mecánica.
- Aumento de sellos divididos: conduce a una pérdida de eficiencia mecánica.

2.2 Justificación

En el mantenimiento predictivo o por condición, se evalúa el estado de los componentes mecánicos o eléctricos de la turbo máquina, mediante técnicas de seguimiento y análisis por medio del modelo computacional de Matlab, permitiéndonos programar las operaciones de mantenimiento “solamente cuando son necesarias”.

Estos modelos tienen en cuenta el estudio de ciertas variables o parámetros relacionados con el estado o condición del equipo que mantenemos, como por ejemplo la vibración, temperatura, aceites, aislamientos, etc. El estudio de estos parámetros nos suministra

información del estado de sus componentes, permitiendo no sólo detectar problemas de componentes sino también de diseño y de instalación.

En Colombia empresas como MANTENIMIENTO PREDICTIVO COMPUTARIZADO LTDA, ofrecen este tipo de estrategias, esta es una empresa de ingeniería con 23 años de experiencia especializada en instrumentación, equipos y sistemas de mantenimiento predictivo, mediante la utilización de tecnologías de punta ofrecen el desarrollo de estrategias de mantenimiento. Entre sus productos ofrecen sistemas digitales de supervisión, dispositivos de protección, vibro conmutadores, programas predictivos concertados- vibraciones y monitores de protección. [5]

Por consiguiente, prestan estos servicios para empresas de energía, ingenios azucareros, industrias alimenticias, industrias de materiales para construcción, industrias Químicas y farmacéuticas y siderúrgicas e industrias de aluminio.

También está el caso de PREDICTIVO LTDA que presta servicios de mantenimiento para el análisis de vibraciones utilizando herramientas tales como el software AMS suite: Machinery Health Manager y presta sus servicios a empresas como MAC JOHNSON COLTROLS, MAC PACIFICO, CARVAJAL EMPAQUES y ALCATEK. [6]

Hoy en día institutos como ITAINNOVA y empresas como Thermolimpic hacen uso de algoritmos de machine learning bajo el siguiente esquema:

Figura 2-2. Anotación de datos en estrategias de machine learning actualmente. [12]

La diferencia entre el modelo a tratar en este proyecto es que se trabajara solamente con un modelo evolutivo o entrenado, a partir de datos recopilados de un modelo supervisado de una bomba centrífuga en una estación de suministro acoplada a un sistema de suministro de agua.

Empresas como PREDICTIVO LTDA utilizan algoritmos parecidos al que se pretende realizar, la única diferencia es que los datos son tomados por medio de sensores físicos en la maquina como tal y con un monitoreo constante, estos datos pueden ser más precisos ya que se toman en tiempo real, además tienen en cuenta la resistencia de la maquina mientras que en nuestro caso lo ignoramos.

En thermolimpic se generan problemáticas debido a que los datos tomados por medio de simulación varían con respecto a los que se presentan en una maquina real, lo cual indica que para nuestro algoritmo también. Por los tanto es la diferencia entre tomar los datos por medio de una simulación a los tomados por medio de monitoreo constante en una maquina en tiempo real.

2.3 Objetivos.

2.3.1 Objetivo general y específicos.

Desarrollar una técnica de mantenimiento predictivo de una bomba centrífuga para la detección de fallas oportunas usando machine learning.

- Desarrollar un modelo numérico que presente el comportamiento de la presión, caudal y torque en la bomba, al variar su velocidad en el rotor.
- Recolectar los datos de comparación o aprendizaje para la realización del algoritmo.
- Implementar el algoritmo de machine learning que detecta las fallas de forma anticipada.
- Evaluar la estrategia de mantenimiento predictivo a través de simulaciones en Matlab.

2.4 Alcance

Se va a trabajar con un modelo netamente numérico, con limitados grados de libertad, donde se relacionará únicamente el comportamiento dinámico de la turbo máquina en cuestión, trabajaremos con una simulación en Matlab y no se realizará pruebas físicas con máquinas reales o algún otro dispositivo, por lo tanto, no se necesitará hacer uso de laboratorios.

El modelo numérico, será únicamente simulado por medio de la herramienta simulink y se utilizará un modelo supervisado con el que podamos obtener los datos por medio de una variación en la velocidad del rotor de la bomba.

Únicamente se simularán fallas que puedan ser ocasionadas por velocidades en el rotor muy altas, ya que esto generaría daños estructurales en los sellos mecánicos, provocaría corrosión y desgaste debido a la alta presión y una turbulencia en el caudal bastante alta.

La estrategia de mantenimiento predictivo será aplicada únicamente con los datos del modelo supervisado y para el modelo de la bomba centrífuga en SIMSCAPE.

3. Marco conceptual y teórico.

En este capítulo se presentará toda la información relacionada acerca de los temas técnicos a tratar tales como el mantenimiento predictivo y tipos de mantenimiento utilizados en la industria actual, además de lo que se trata el aprendizaje automático o Machine learning y como se ha manejado.

3.1 Uso de machine learning en la actualidad.

En la industria actual el machine learning gracias a su precisión al momento de realizar acciones que comuna operación netamente humana no se logran los mismos resultados ha tenido un alto impacto en usos comunes en varios tipos de industrias, entre estos están:

- **Conteo de piezas en video o imagen:**

Simplifica las operaciones de inventariado periódicamente de manera manual y resuelve los problemas que se puedan ocasionar por error humano, mediante reconocimiento de imágenes y la comparación de los datos conocidos es posible contar las piezas. Si se desea con análisis de bases de datos sin importar su extensión saber cuánto se deberá comprar en poco tiempo y de una manera muy acertada.

- **Optimización de tiempos:**

Este sistema es utilizado para agilizar procesos ya sean automáticos o no, dando una estimación en cuando es, o no necesario un cese en la producción. Optimiza procesos de fabricación dando avisos de mal funcionamiento o realizando predicciones de manufactura.

- Atención a clientes por medio de voz y texto:

Ya no será necesario contar con un centro de atención telefónica para ciertos servicios, debido a gracias al aprendizaje automático se puede programar con patrones de voz uno de los modelos para que este los identifique y así realizar respuestas programadas y relacionadas de forma acertada, es decir, si por ejemplo un usuario quisiera dar de alta o de baja un servicio o atención el sistema podría responder a la petición con alguna de las respuestas programadas.

- Seguridad de la información y antifraudes:

Incrementa la seguridad en conjuntos cerrados de información, ya que el sistema es capaz de detectar comportamientos indebidos o no permitidos, a diferencia de los demás sistemas que solo filtros fijos para y para una cantidad limitada de datos con antecedentes de indebidos, con ML se logra mediante patrones identificar ataques y dar aviso al personal o al sistema.

Otro uso importante es que por medio de datos históricos es capaz de identificar personas con malos antecedentes en un sitio de comercio específico, haciendo comparaciones con el comportamiento, rasgos, números de identificación e incluso la voz para que con este se logre identificar y dar aviso al comerciante evitando así posibles robos.

3.2 Machine learning.

El aprendizaje automático es una de las ramas de la inteligencia artificial, con este se logra crear sistemas que pueden aprender de una forma automática e independiente gracias a bases de cualquier tipo de dato previo de una clase en estudio. Es capaz de identificar patrones muy complejos que sin ayuda de un ordenador llevaría bastante tiempo para una persona, los patrones pueden verse en una cantidad muy grande de datos de distintos orígenes, también es posible predecir comportamientos gracias a algoritmos que con el tiempo son capaces de mejorarse de manera independiente. [25]

Actualmente los modelos que son utilizados dentro del ML ha tenido gran importancia como un recurso tecnológico computarizado y utilizado como sistemas de uso diario como por ejemplo verificadores de fraude para correos electrónicos, softwares de reconocimiento de voz, conducción automática de cualquier vehículo, herramientas para casas inteligentes.

El machine learning presenta una técnica muy interesante para la generación de un aprendizaje haciendo uso de la información que se extrae de los datos o variables analizadas por computadoras reemplazando algunas funciones de responsabilidad humana en cierta medida. Gracias a la toma de datos y al análisis de sus comportamientos de estos modelos de aprendizaje, podemos crear modelos que realicen predicciones de sucesos para la toma de decisiones con buenos resultados y un alto nivel de eficiencia.

Dentro de este campo de desarrollo de tecnología automática o IA existen diferentes tipos de machine learning. Existen tres tipos de machine learning, supervisado, no supervisado y el aprendizaje reforzado.

3.2.1 Machine learning supervisado.

El aprendizaje supervisado en primera instancia se fundamenta en incluir etiquetas en la información obtenida de variables en estudio también llamados datos de comparación o aprendizaje, para que este sea capaz de comenzar a trabajar con los datos añadidos a un algoritmo que, con el tiempo es capaz de evolucionar de una forma automática. Es capaz de realizar comparaciones, proyecciones y clasificar datos gracias a un algoritmo que sea capaz de detectar patrones en la información para así lograr compararlos con datos de aprendizaje o datos nuevos suministrados por un monitoreo constante. [26]

Los modelos supervisados suelen ser los más utilizados gracias a que consiste en utilizar Deep learning, este tiene un funcionamiento basado en redes neuronales que hace

referencia a sistemas de computación interconectados. Usa una gran cantidad de “neuronas” o unidades computacionales conectadas entre sí y clasificadas etapa por etapa, donde en etapa se encuentran redes con objetivos en común. Cada etapa es capaz de aprender de las demás etapas y así generar patrones en la información de la otra con el objetivo de que en conjunto logren desarrollar patrones de cualquier tipo de variables o comportamientos como por ejemplo colores, formas, personas, temperaturas, velocidades, resistencia, etc.

Por medio del aprendizaje desarrollado por estos modelos de ML supervisados se alimenta un conjunto de resultados que realiza predicciones adecuadas del comportamiento de datos nuevos ingresan al algoritmo y deben ser clasificados en alguna de las clases de los datos de aprendizaje. Este tipo de aprendizaje es el que se incorpora en aplicaciones tecnológicas como filtros detectores de spam en correos electrónicos, detectores de imágenes en captchas o en aplicaciones de reconocimiento de voz o escritura.

3.2.2 Machine learning no supervisado.

A diferencia de los modelos supervisados, en el ML no supervisado no se utiliza ningún tipo de etiqueta en los datos de comparación o aprendizaje para que el algoritmo comience a desarrollarse de una manera autónoma. El algoritmo se modifica a sí mismo y no contiene conocimiento previo de ninguna base de datos, es decir ningún dato ha sido etiquetado u ordenado, saca sus conclusiones de acuerdo con la evolución del algoritmo tomando factores de manera indistinta que no tienen ninguna marca que los distinga de los demás. [24]

En este tipo de aprendizaje existen dos categorías específicas conocidas como clustering y reducción dimensional. El clustering se fundamenta como una técnica de exploración para análisis de datos en la que se ordena la información por grupos o clases

desconociendo de manera predeterminada lo que los caracteriza. Esto se hace con el objetivo de obtener grupos de datos con características en común.

Este tipo de análisis de datos en estrategias de marketing se utilizan con mayor frecuencia ya que facilitan la construcción de segmentos de mercado

Por otro lado, la reducción dimensional utiliza datos de alta complejidad que demandan un mayor costo y capacidad de procesamiento ya que se manejan bases gigantescas de información no clasificada. Esta funciona determinando correlaciones entre las clases que se presentan datos, disminuyendo los datos con características repetitivas y reduciendo el tiempo de análisis para que la información con mayor prioridad sea obtenida más eficientemente.

3.2.3 Machine learning reforzado.

Otro modelo de ML se denomina aprendizaje reforzado. Este tipo de aprendizaje forma parte del Deep learning o aprendizaje de profundidad. El Aprendizaje construye modelos que mejoran su rendimiento en base a los datos de resultados obtenidos con cada interacción que realiza. Estos resultados son obtenidos a partir de una acción correcta o conjunto de datos devueltos que ingresan en una medida pre procesada.

Gracias a un agente automático el modelo utiliza los datos que van resultando de la interacción como parámetro de ajuste en su comportamiento para acciones futuras, de tal forma que la acción nueva cumpla igualmente con el objetivo o acción correcta y así obtener una recompensa más eficiente y que sea capaz de corregir los parámetros que resultaron en el dato previo no tan exacto.

El aprendizaje reforzado tiene varias aplicaciones, pero actualmente es públicamente reconocido por su aplicación en un programa conocido como AlphaZero de la industria desarrolladora tecnológica Deep Mind. Su programación logro que este agente de ML conociera a ciencia cierta la totalidad de combinaciones y jugadas posibles en un tablero de ajedrez y vencer en tan solo 4 horas de aprendizaje al motor de ajedrez computarizado conocido como Stockfish.

3.3 Algoritmos de machine learning.

Como hay varios tipos de aprendizaje automático también se debe escoger entre varios tipos de algoritmos con distintas funcionalidades y procesos, entre los más utilizados hoy en día están:

3.3.1 Algoritmos de regresión.

Este tipo de algoritmos consiste en realizar tareas de regresión, siempre y cuando el programa requiera realizar estimaciones. Es importante tener conocimiento a profundidad de las características de los datos y de cómo se clasifican, además de las relaciones que existen entre ellos.

Para este tipo de algoritmos la regresión se debe realizar enfocada en una variable independiente con una gran cantidad de datos dependientes de ella. Al ejecutarse el algoritmo este es capaz de realizar predicciones, comparaciones e incluso clasificar los datos nuevos en una de las clases conocidas.

3.3.2 Algoritmos bayesianos.

Estos algoritmos consisten en que, por medio de la aplicación del teorema de Bayes, se clasifiquen los valores de forma que no dependan de los demás datos que se encuentran en el mismo conjunto en estudio permitiendo predecir una clase en función de un conjunto predeterminado de características gracias a un índice de probabilidades.

Este tipo de algoritmo es uno de los más utilizados ya que ignorando su simplicidad permite realizar clasificaciones de datos obtenidos de bases de datos muy grandes o complejas.

3.3.3 Algoritmos de agrupación.

Los algoritmos de agrupación son los que más se utilizan en modelos de tipo no supervisado debido a su alta fiabilidad para ordenar en datos grupos, sin importar que no tengan alguna etiqueta o dirección. Este algoritmo es capaz de realizar búsquedas en grupos bastante grandes guiándose únicamente por una variable en estudio. Funcionan únicamente con la única función de asignar cada punto de datos a uno de los grupos o clases representados en la variable tomando como base únicamente todas las características con las que se haya programado para realizar la identificación de datos no identificados.

3.3.4 Algoritmos de decisión.

Los algoritmos de decisión o árboles de decisión es una estructura útil para que se elija entre varias opciones de datos o resultados teniendo en cuenta criterios previamente establecidos. Este utiliza un método en el que se representan los resultados esperados como una bifurcación en la que se elige una de ellas paso por paso y teniendo en cuenta que tengan similitud con ejecuciones anteriores. Es un diagrama muy específico, se caracteriza por incluir información de los posibles resultados esperados y los clasifica para terminar eligiendo el más conveniente.

3.3.5 Algoritmos de redes neuronales.

Los algoritmos basados en redes neuronales necesitan de un conjunto de unidades de cómputo clasificadas por etapas que a su vez están conectadas a etapas anexas. Estas redes trabajan entre sí para que aprendan las unas de las otras, son conjuntos que trabajan con datos etiquetados y suministrados por un monitoreo constante pero también dependen de los procesos realizados por las demás etapas para darle solución a problemas específicos.

3.3.6 Algoritmos de aprendizaje profundo o Deep learning.

Este tipo se caracteriza por utilizar en conjunto varios tipos de algoritmos de clasificación y pre procesamiento de datos etapa por etapa, hasta lograr simplificar las representaciones de los datos. Trabajan como una red neuronal en bases de datos muy extensas y complejas, pero al final, consideran los resultados de cada etapa para quedarse con el más acertado.

3.4 Mantenimiento predictivo.

La revolución de la industria 4.0 ha tenido un alto impacto en el sector de la fabricación ofreciendo una nueva era para las empresas en la que se comienza a implementar la comunicación maquina a máquina y mejorar la comunicación maquina a humano, además de tecnologías para analizar y fallas antes de que sea tarde. Con respecto a este tema los algoritmos de predicción dados por la GMAO son cada vez más complejos, pero a su vez efectivos. Las técnicas o estrategias de mantenimiento predictivo han facilitados de una

manera considerable la gestión y el mantenimiento de la maquinaria de la empresa, optimizando el tiempo de los mantenimientos y su costo. [27]

Se le llama mantenimiento predictivo al conjunto de técnicas que con ayuda de instrumentos de medida y análisis de variables ya sean dinámicas o no, logran identificar en términos de fallos considerables la condición operativa de los equipos o maquinaria. Su misión principal es re disponer de equipos al mínimo costo, optimizar la fiabilidad y reducir en mantenimientos preventivos. Se encarga de evaluar el estado de la maquinaria mediante un monitoreo constante y recomienda intervenir o no en función de su estado, lo cual produce grandes ahorros. [21]

El objetivo del PdM es anticipar de manera efectiva la ocurrencia de un fallo considerable o no en un equipo, que tenga antecedentes de fallar bajo ciertas condiciones en base a los datos relacionados con su estado. El mantenimiento predictivo puede incluir también actividades de mantenimiento regular, pero a diferencia del mantenimiento preventivo este prioriza que se realicen en la menor cantidad posible y en el tiempo más oportuno.

Los algoritmos de PdM se basan en datos proporcionados por una monitorización constante de la maquina o por bases de datos que contengan los datos dinámicos de un antecedente de falla, como el análisis de vibraciones, el análisis de aceite o el análisis de cualquier otra variable dinámica que se relaciones con antecedentes de fallos comunes. La diferencia entre la condición esperada de la maquinaria y como suele deteriorarse al momento de la ocurrencia de fallas hace que se puedan predecir las intervenciones de mantenimiento.

Con monitoreo constante se hace referencia a que el PdM necesita supervisar el estado de la maquinaria durante el funcionamiento normal, para reducir las interrupciones debidas a las operaciones diarias. Es necesario estar analizando los datos en tiempo real para que en caso de una medición indebida este realice el aviso inmediato.

Existen varios tipos de mantenimiento entre los cuales se encuentran el mantenimiento preventivo y el mantenimiento correctivo.

3.5 Mantenimiento correctivo.

Las medidas correctivas son las más usualmente utilizadas durante paros o interrupciones periódicas de emergencia para poner a trabajar la máquina en un corto tiempo. A diferencia del mantenimiento predictivo la acción reactiva es realizada sólo cuando ocurre la falla. Para solucionar el problema, lo que más se suele hacer es reemplazar un componente que ha fallado por un componente que sea capaz de realizar la misma operación y que tenga características similares. Es gracias a este tipo de correcciones que rara vez se trata de mejorar o incluso prevenir fallos. [22]

El mantenimiento correctivo se caracteriza por realizarse únicamente cuando ocurre una falla y sin importar que tan grave y que costo pueda generar, se examina y chequea el incidente reportado, se busca la falla y se modifica. Finalmente se documenta con cada una de sus características como causa, partes dañadas y costos. A simple vista parece ser la técnica más económica en cuanto a mano de obra y materiales, pero puede ser la más cara bajo consideraciones como: Seguridad laboral (esto considerando que la falla pudo haber ocasionado accidentes), valor de capital, certidumbre del equipo y, por último, pero no menos importantes multas por discontinuar el servicio o la producción.

Usar solamente esta técnica de mantenimiento puede ser completamente inaceptable para algunas industrias en las que su producción depende mayormente del buen estado de su maquinaria, pues los costos y pérdidas que pueda generar una falla pueden involucrar una gran cantidad de tiempo y dinero de la producción. El mantenimiento correctivo se reserva únicamente para aquellos componentes que, aunque suelen fallar periódicamente, estas fallas no involucran demás piezas.

3.6 Mantenimiento preventivo.

El objetivo del mantenimiento preventivo es evitar los fallos antes de que ocurran. Implementando este tipo de mantenimiento se resuelven las pérdidas (tiempo y costo) que se puedan ocasionar al interrumpir la operación de la maquina averiada. Al incluir este sistema se evitan daños potenciales en las máquinas, además del tiempo que tardaría la reparación de un daño grave ya que la inspección o arreglos oportunos se han realizado antes de que ocurran los fallos. [23]

El mantenimiento preventivo es una de las técnicas de mantenimiento industrial para cualquier tipo de maquinaria. Por lo tanto, podemos encontrar tres tipos de mantenimiento preventivo: programado, de oportunidad y predictivo.

- Programado: se realiza un listado con asignaciones en cierto tiempo de reparaciones, mediante un calendario, y se lleva a cabo. Este es un mantenimiento de rutina y es el que más se suele utilizar en industrias con una cantidad numerosa de maquinaria.
- De oportunidad: se realiza en momentos en los que la maquinaria entra en un largo tiempo de reposo.
- Predictivo: este realiza análisis de estadísticas de los antecedentes en los que generaron fallas para así predecir de manera oportuna cuando la maquina se encuentre próxima a fallar.

Este tipo de mantenimiento sugiere también un tiempo de interrupción en el funcionamiento de las máquinas para que se logre analizar en profundidad el funcionamiento de ordenadores, equipos y maquinarias. Es por esto por lo que al momento de realizar la programación se deben tener en cuenta fechas no muy importantes para la producción de la empresa.

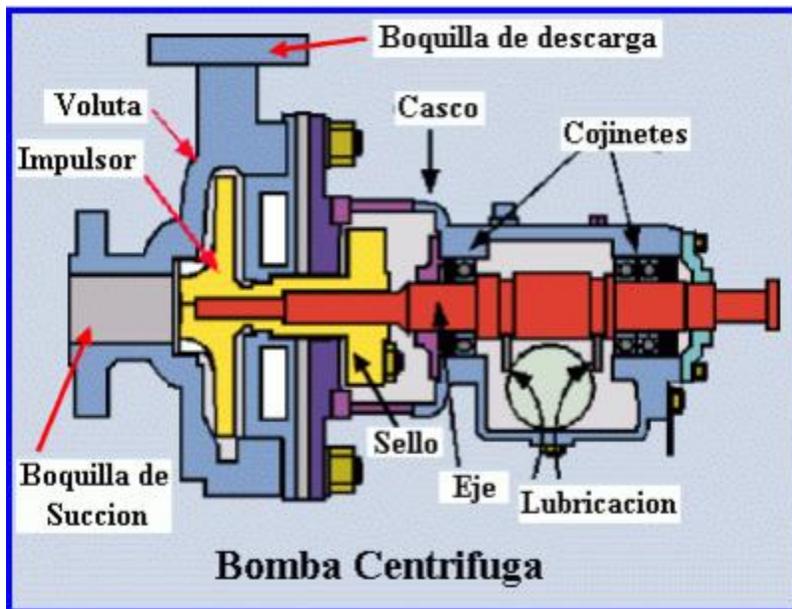
3.7 Bombas centrífugas.

Una bomba centrífuga es una de las turbomáquinas más simples de estructura, pero más esenciales en cualquier tipo de planta que maneje un movimiento de fluidos en un sistema cerrado. Su propósito es convertir energía de un motor eléctrico o turbina en velocidad o energía cinética y luego en energía de presión de un fluido que está bombeándose por canales que conforman un circuito hidráulico que comúnmente en la industria se utiliza para desplazamiento o suministro hidráulico.

La conversión de energía en la bomba ocurre gracias a dos componentes o partes fundamentales de la estructura de la bomba, el impulsor y el espiral o difusor. El impulsor es la parte que gira a cierta velocidad y convierte la energía de la máquina en energía cinética. El espiral o también llamado difusor es la parte rígida e inmóvil que convierte la energía cinética del fluido en presión. Cabe recordar que la bomba como tal no genera una presión, La presión es solamente una indicación de la resistencia al flujo (ETM, 2021)

3.7.1 Configuración básica.

La máquina de simple etapa es el tipo de turbo maquina más común y simple, la cual consiste únicamente del impulsor y un casco. El líquido es succionado hacia el centro del impulsor y puesto en movimiento gracias a las aspas. Debido a la fuerza centrífuga ocasionada por la acción rotacional del impulsor el líquido es lanzado con una considerable velocidad y presión. El casco tiene una voluta formando una especie de canal cuya área de sección transversal va aumentando esta se encarga de recoger el líquido enviado por el impulsor y aumentar su presión. Este pasaje del casco es quien conduce el fluido con mayor velocidad y presión a la conexión con las tuberías del circuito.

Figura 3-1. Bomba centrífuga. [19]

La bomba centrífuga, debido a sus características, hace parte de la clase de bombas hidráulicas más utilizadas dentro de la industria para bombear líquidos en general y permiten desplazar grandes cantidades de agua.

Hay una gran cantidad de sectores industriales que suelen usar bombas centrífugas en sus procesos mecánicos, automáticos o manuales. Algunos de los sectores que más utilizan bombas centrífugas son la industria química, el sector de la mecánica, la industria cosmética para el desarrollo de cremas, la industria alimentaria para la elaboración de todo tipo productos alimenticios o incluso en sectores públicos como lo son los sistemas de suministro de agua.

3.7.2 Ventajas de las bombas centrífugas.

Algunas ventajas de las bombas centrífugas son:

- No poseen órganos adjuntados y sus mecanismos de ajuste son reducidos en tamaño y sencillos de acoplar.
- Para un trabajo continuo y eficiente, el desgaste es constante y para la regulación puede ser suficiente con la variación de la velocidad en el rotor.
- Se pueden adaptar a cambios a cambios siempre y cuando no se realicen de manera drástica.
- Son ligeras y en su mayoría no suelen ser muy grandes lo que las hace más portables.
- El mantenimiento y el costo de sus componentes es menor, pero suele variar con respecto a la clase y la función de la turbo máquina.
- Coste ajustado.
- La incorporación de variadores de velocidad permite mayor eficiencia.
- Los sistemas de bombeo completos son suministrados según las necesidades.

3.7.3 Tipos de bombas centrifugas.

Las bombas centrifugas se clasifican según tres parámetros.

Por la dirección del flujo. Es decir, la dirección en la que el fluido atraviese la **bomba centrífuga**, según este parámetro pueden ser:

- **Radial:** En este caso el flujo atraviesa el cabezal y carcasa de forma paralela al eje de rotación. Son bombas muy eficaces y volubles y son las bombas centrifugas más comunes.

Figura 3-2. Bomba centrífuga radial. [21]



- **Axial:** En este caso el flujo atraviesa el interior de la bomba de forma paralela al eje de rotación. Son bombas muy potentes y eficaces a la hora de elevar grandes caudales siempre y cuando no sea a una altura muy grande.

Figura 3-3. Bomba centrífuga axial. [21]



- **Mixto:** es una combinación entre las bombas de flujo axial y radial.

Figura 3-4. Bomba centrífuga de flujo mixto. [21]



El segundo parámetro por el que se clasifican es por la posición del eje de rotación:

- **Horizontales.** El eje de rotación se encuentra de manera horizontal respecto a la entrada del fluido.

Figura 3-5. Bombas centrífugas en posición horizontal. [21]



- **Verticales** El eje de rotación se encuentra en posición vertical respecto a la entrada del fluido.

Figura 3-6. Bombas centrífugas en posición vertical. [21]



- **Inclinados:** El eje de rotación está inclinado.

Las demás clasificaciones son:

- **Por el diseño de la coraza (forma)** en: voluta y las de turbina.
- **Por la forma de succión** en: sencilla y doble.

4. Planteamiento de la solución.

En este capítulo se desarrollará paso por paso el modelo de machine learning para una simulación en Simscape de un modelo de suministro de agua acoplado a bombas centrífugas. Atenido a la siguiente metodología:

- Anotación de datos de aprendizaje simulando funcionamiento normal y de falla de la bomba centrífuga variando la velocidad del rotor para la creación del modelo de clasificación de machine learning.
- Análisis y pre procesado de los datos de aprendizaje de ambas clases en Matlab.
- Realizar los modelos de clasificación e ingresar datos de prueba para que por medio de predictores de regresión lineal se clasifiquen los datos en la clase que corresponda.
- Evaluar las predicciones realizadas por cada uno de los modelos sacando valores de precisión TCE y también graficando cuadro de confusión y diagrama de coordenadas paralelas para sacar conclusiones.

4.1 Extracción de datos del modelo para ambas clases.

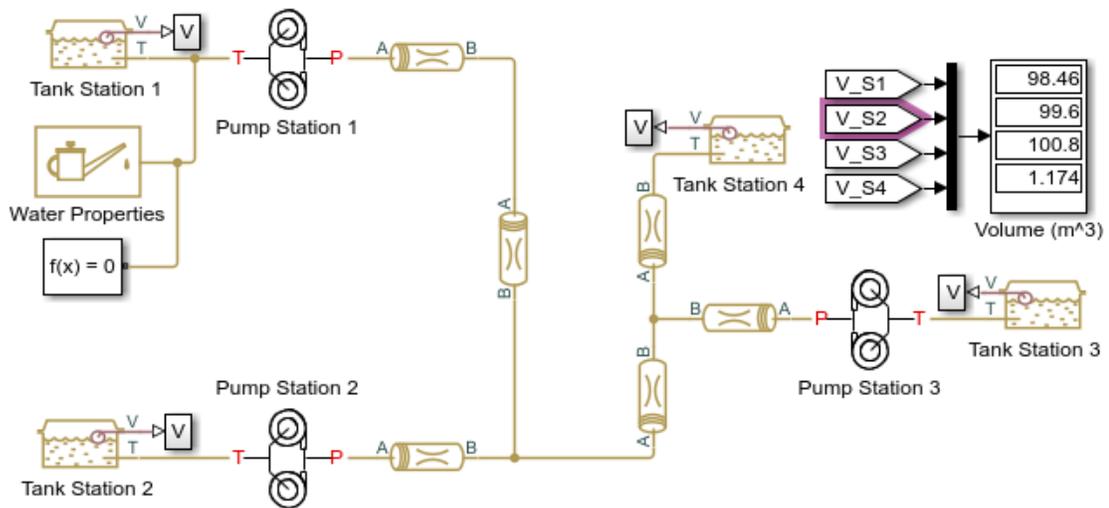
La extracción de datos se hace de una bomba centrífuga conectada a un modelo de suministro de agua. Este sistema suministra agua a cuatro tanques no presurizados por medio de tres estaciones de propulsión del líquido, en cada estación encontramos dos bombas centrífugas con características similares y velocidad de rotación distintas. El

objetivo es modificar esta velocidad de rotación en el rotor hasta valores que se conoce que generarían fallas, ya sea en las tuberías hidráulicas del sistema de suministro o en los componentes de alguna de las seis bombas centrifugas.

Al modificar esta velocidad del rotor se pueden revisar los valores de presión, la salida de la bomba o caudal, el torque en el eje y la velocidad en el eje. El modelo de suministro de agua es un ejemplo de uno de los sistemas típicos que ofrece Matlab en su plataforma, sin embargo, por su composición y su operación basada en un sistema de bombas centrifugas es útil para la obtención de los datos de aprendizaje de una de sus estaciones de suministro.

Este ejemplo muestra un sistema de suministro de agua que consta de tres estaciones de bombeo. Se espera que las tres estaciones bombeen agua en un tanque. Todos los tanques son lo suficientemente grandes como para asumir que el nivel de fluido permanece casi constante, lo que les permite ser modelados con el bloque de tanque de cabeza constante. Cada estación de bombeo consta de un tanque, dos bombas centrifugas instaladas en paralelo y un motor primario que gira a 1700 rpm (esta velocidad es la que nos interesa que sea modificable). Las características de la bomba se especifican mediante tablas de búsqueda. [29]

El estudio de este sistema de suministro de agua se ignora más adelante ya que los datos anotados serán únicamente los de la bomba centrifuga número 1 en la estación de suministro número 1.

Figura 4-1. Modelo de sistema de suministro de agua de Matlab. [29]

La **figura 4-1** será el modelo de Simscape del cual obtendremos los datos para las dos clases (falla y no-falla) que son necesarios para la implementación de Machine Learning.

Donde la **Figura 4-2** es el bloque encargado de asignar el fluido de trabajo para cada uno de los componentes ensamblados por un medio hidráulico, en él se asignan las propiedades del fluido de trabajo que para este caso será agua.

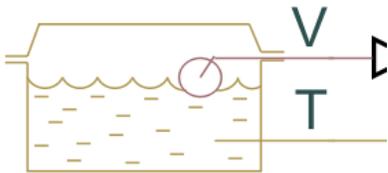
Figura 4-2. Bloque para propiedades del fluido.

Water Properties

La **Figura 4-3** representa un tanque con un fluido que se puede ajustar a presión o sin presión, que para la simulación se toma como un tanque despresurizado.

Se supone que el tamaño del tanque es lo suficientemente grande como para ignorar la presurización y el cambio de nivel de fluido debido al volumen de fluido. La conexión T es un puerto de conservación hidráulica asociado con la entrada del tanque. La conexión V es un puerto de señal físico. El caudal se considera positivo si el fluido fluye hacia el tanque.

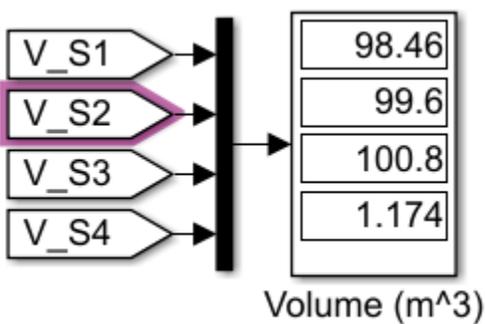
Figura 4-3. Tanque de suministro.



Tank Station 2

La **Figura 4-4** es una columna de resultados, nos representa el volumen al que se encuentra cada uno de los tanques del sistema de suministro.

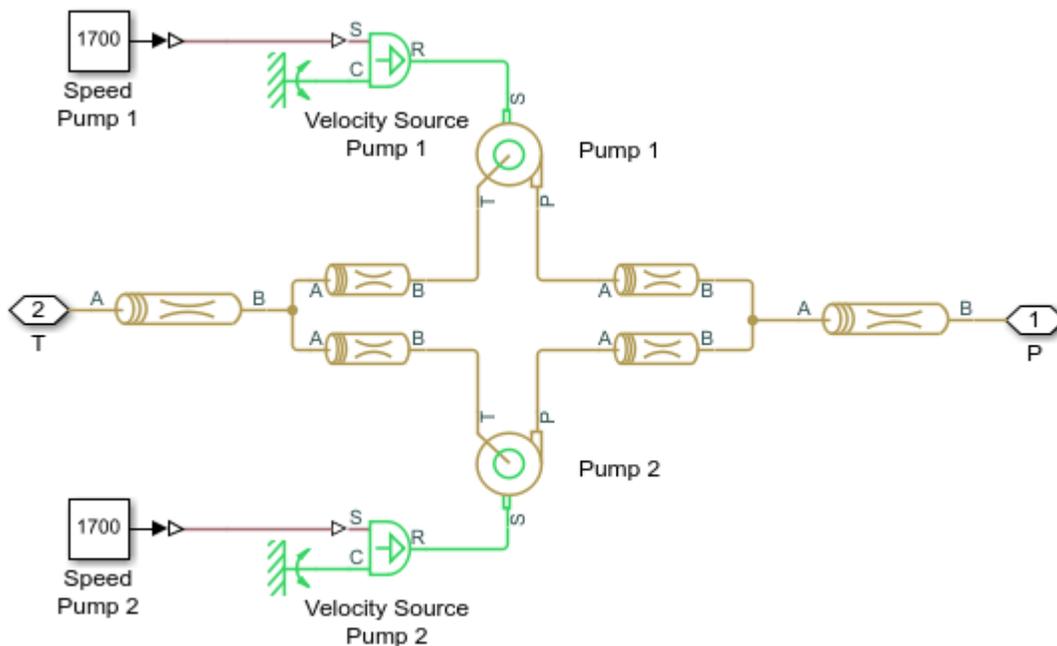
Figura 4-4. Valores de volumen.



Por ultimo las estaciones de suministro están compuestas por dos bombas centrífugas conectadas en paralelo de modo que la velocidad de rotación de una no afecte las propiedades dinámicas de la otra, por medio de tuberías hidráulicas de sección transversal circular. Estas tuberías tienen en cuenta únicamente propiedades resistivas.

Las conexiones A y B son puertos de conservación hidráulica. La dirección positiva del bloque es del puerto A al puerto B.

Figura 4-5. Estación de suministro.



El bloque de la fuente de velocidad conectado al rotor de cada bomba centrífuga es el que nos permite variar la velocidad rotacional de la bomba con respecto a un punto de referencia rotacional hasta una velocidad que ocasione fallas.

Por ejemplo, para el caso anterior tenemos la bomba 1 configurada con una velocidad rotacional de 8000 RPM lo cual es causante de falla, mientras que la bomba 2 se configura con una velocidad rotacional de 2000 RPM que es para operaciones normales.

Al realizar la simulación con esa modificación en la velocidad rotacional de la bomba 1.

Tabla 4-1. Resultados de la bomba centrífuga a 8000 rpm.

Presión (Pa)	747810
Caudal (m^3/s)	0.0866073
Torque (Nm)	337.126
Velocidad (rad/s)	837.758

Mientras que con la bomba 2:

Tabla 4-2. Resultados de la bomba centrífuga a 2000 rpm.

Presión (Pa)	705712
Caudal (m^3/s)	-0.0788994
Torque (Nm)	-346.724
Velocidad (rad/s)	209.44

Estos datos de la bomba 1 serán clasificados dentro de la clase de falla y los de la bomba 2 dentro de la clase de no-falla, este procedimiento es suficiente realizarlo con una sola bomba, ya que todas están parametrizadas de la misma manera y obtendríamos resultados similares.

Los datos de no-falla se tomaron con velocidades de 1780 a 6200 RPM, mientras que, los datos en la clase de falla fueron tomados desde las 6201 a 9500 RPM. Al final se obtuvieron por medio de simulación un total de 1552 datos de aprendizaje entre los cuales hay 668 datos de la clase de falla y 884 datos de la clase de no falla, las únicas modificaciones en la bomba centrífuga para la toma de los datos fueron con respecto a las velocidades de rotación en el eje.

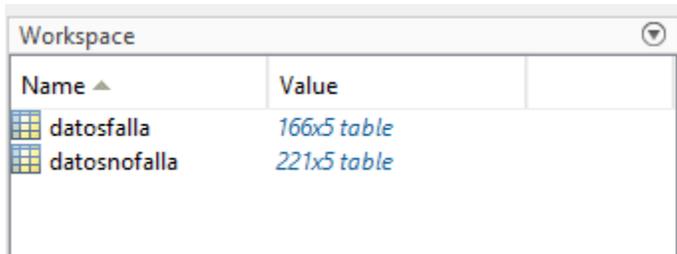
Los datos fueron tomados variando la velocidad cada 20 RPM, la cantidad de datos fue seleccionada con una variación pequeña ya que se quiere obtener algo parecido a un monitoreo constante de la maquina en estos rangos de velocidad. Con los datos totales obtenidos se observa con las gráficas que no hubo un sobreajuste o sub-ajuste de estos ya que la bomba se comportaba de manera lineal en las simulaciones, por lo tanto, no es necesario para el algoritmo “encajar” un dato forzadamente en cada clase.

Para la simulación de las fallas en la bomba centrífuga se tienen en cuenta solamente fallas que puedan ser generadas por errores del usuario ya que si se llega a ajustar una velocidad muy alta en el rotor podría ocasionar daños tales como corrosión y desgaste en los cojinetes ya que por la turbulencia y vibración que generan velocidades desde las 6200 RPM en el eje, los cojinetes y los sellos mecánicos en la bomba se llegan a desgastar y fracturar. Además, la turbulencia en el fluido podría conllevar a oscilaciones en la bomba por un desviamiento en el eje. [11]

Por lo que es conveniente que con datos de aprendizaje desde las 6200 RPM se registre la clase de falla para el modelo de clasificación ya que en condiciones normales de funcionamiento la bomba no debe llegar a este límite de velocidad. Así el modelo de clasificación avisara cuando las variables dinámicas en la bomba se relacionen con una velocidad por encima de ese límite de 6200 RPM y se pueda corregir rápidamente.

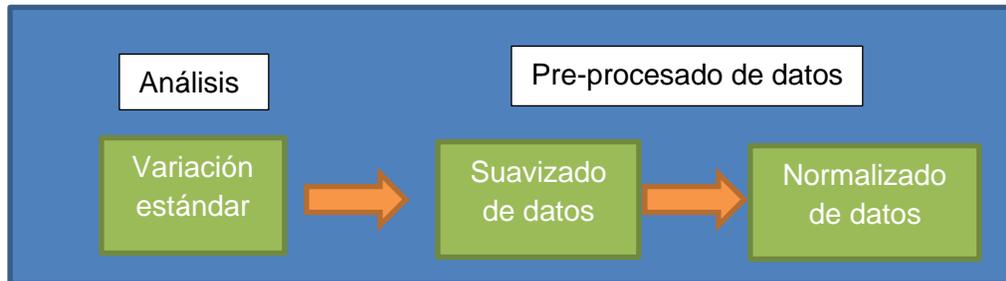
4.2 Almacenamiento y pre procesado de los datos.

Los datos se almacenaron en 2 bases de distinto nombre, pero en una ubicación conjunta, una de las bases de datos lleva por nombre **falla** y la otra lleva por nombre **no falla**. Las tablas están distribuidas con los datos arrojados por el modelo al momento de simularse con la velocidad rotacional de la bomba número 1 en la estación de bombas número 1 del modelo. Se realiza la importación de los datos de cada clase ordenados en un libro de Excel y se guardan de la siguiente manera.

Figura 4-6. Importación de los datos a Matlab.

Name ▲	Value
datosfalla	166x5 table
datosnofalla	221x5 table

Al haber obtenido todos los datos anexos al documento comenzamos a realizar un pre procesado a los mismos con el fin de analizar su comportamiento. Para esto primero se calcula su variación estándar, luego de esto se hace un suavizado y por último normalizado que se realiza con los datos brutos para transformarlos en datos que tengan formatos que sean más fáciles de utilizar.



Parte del pre procesado de datos es sacar la diferencia en cada una de las variables a medida que se va incrementando la velocidad rotacional en el eje, para esto el paso a seguir es importar los datos en Matlab para sacar así su variación estándar.

La función `std` de Matlab nos permite conocer la variación estándar de cada una de las variables medidas en ambas clases de los datos de aprendizaje a continuación, se presenta la variación estándar de cada una para ambas clases.

Tabla 4-3. Variación estándar de la clase de falla.

Variable	Variación estándar
Torque	72.2004
Presión	2.6003E04
Caudal	0.0111

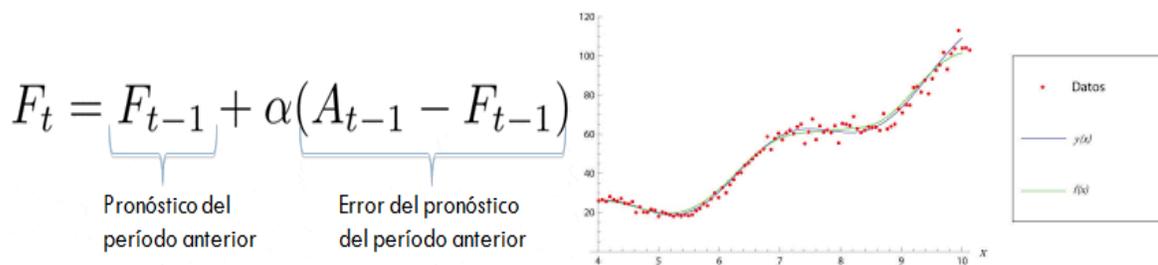
Tabla 4-4. Variación estándar de la clase de no falla.

Variable	Variación estándar
Torque	50.8112
Presión	3.0341e+04
Caudal	0.0163

Suavizado de datos

En las variables suavizado y saviavizado1 se guardan las tablas de datos obtenidos a velocidades de rotación normales, pero con una atenuación más leve.

La función smoothdata de Matlab es la encargada de realizar el suavizado de datos por método gaussiano en Matlab.

Figura 4-7. Funcion de suavizado gaussiano de datos. [28]

En la Figura 4-7 se muestra el proceso matemático que realiza la función smoothdata de Matlab con cada uno de los datos de aprendizaje y el suavizado de una muestra de ejemplo.

Normalizado y escalado de datos.

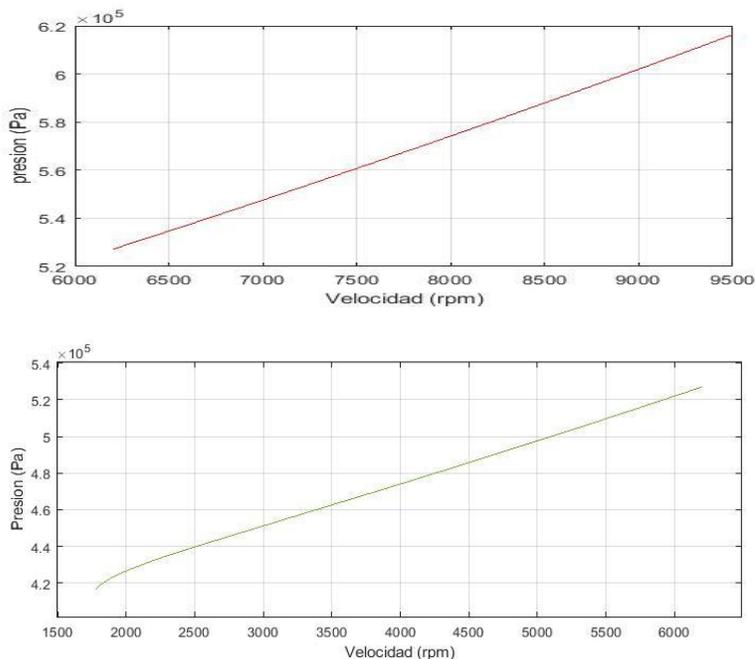
Por medio de la función “normalize” de Matlab se realiza un normalizado por puntuación estándar de los datos. Se normalizan los datos anteriormente suavizados guardados en las variables suavizado y suavizado1, esta función normaliza errores cuando los parámetros de la clase son conocidos y trabaja bien cuando se trata de datos de aprendizaje.

Luego de realizar el pre procesamiento de los datos realizamos una comparación de las gráficas que obtenemos de cada una de las variables. Las gráficas son el comportamiento de las variables con respecto a la velocidad de rotación en el eje para ambas clases, antes y después del pre procesamiento.

Datos sin procesar:

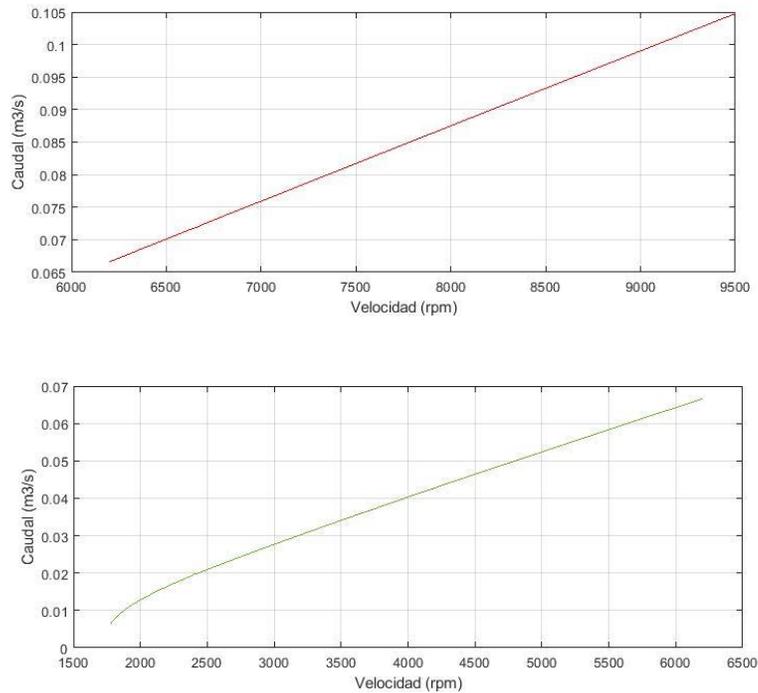
Las gráficas en color verde representan el comportamiento normal de la bomba centrífuga mientras que los datos en color rojo representan el comportamiento con falla.

Figura 4-8. Gráfico de presión con respecto a la velocidad.

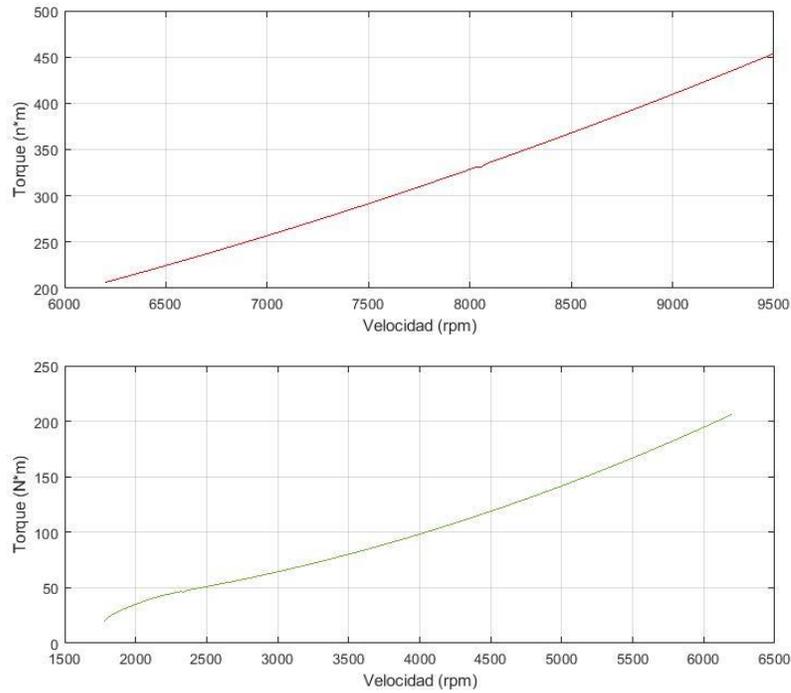


En la figura 4-18 se presenta el comportamiento de los datos de presión con respecto a la velocidad en ambas condiciones antes de su suavizado y normalizado.

Figura 4-9. Gráfico de caudal con respecto a la velocidad.



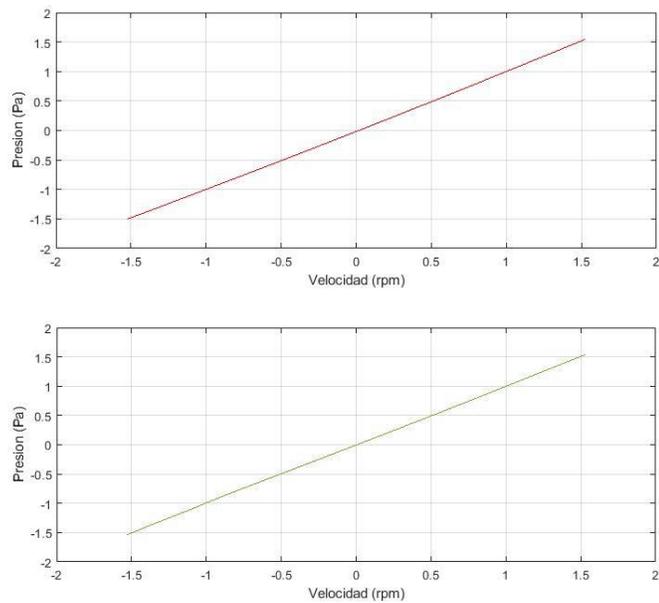
En la figura 4-9 se presenta el comportamiento de los datos de caudal con respecto a la velocidad en ambas condiciones antes de su suavizado y normalizado.

Figura 4-10. Gráfico de torque en el eje con respecto a la velocidad.

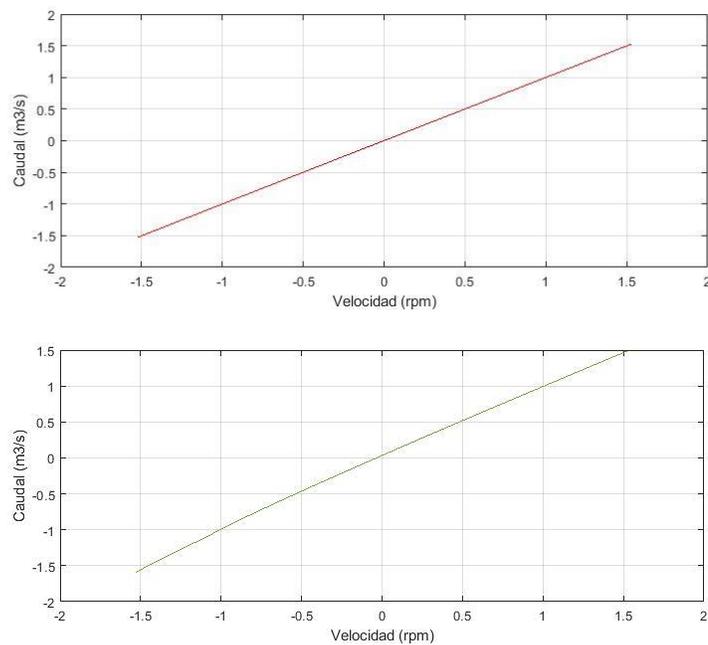
En la figura 4-10 se presenta el comportamiento de los datos de torque con respecto a la velocidad en ambas condiciones antes de su suavizado y normalizado.

Datos pre procesados:

Las gráficas en color verde representan el comportamiento normal de la bomba centrifuga mientras que los datos en color rojo representan el comportamiento con falla.

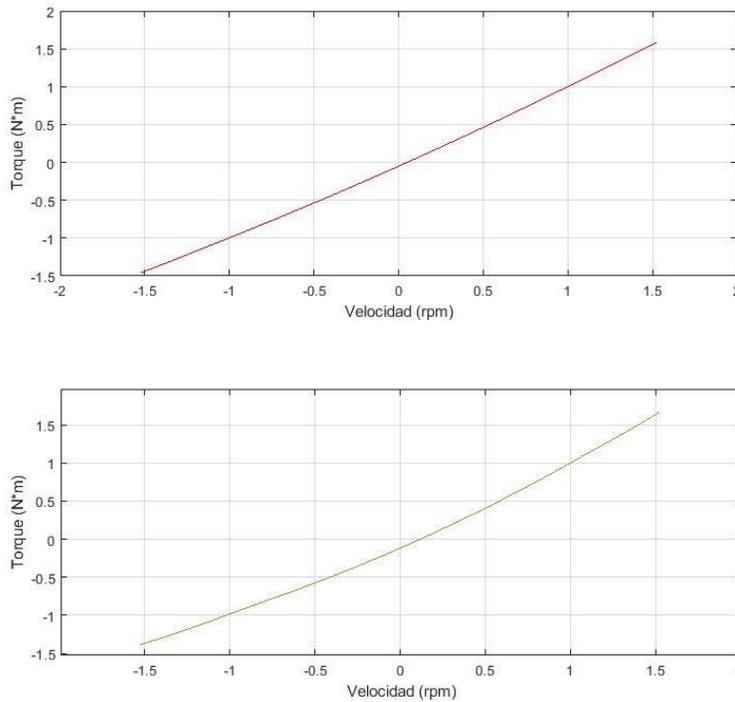
Figura 4-11. Gráfico de presión con respecto a la velocidad.

En la figura 4-11 se presenta el comportamiento de los datos de presión con respecto a la velocidad en ambas condiciones y después de su suavizado y normalizado.

Figura 4-12. Gráfico de caudal con respecto a la velocidad.

En la figura 4-12 se presenta el comportamiento de los datos de caudal con respecto a la velocidad en ambas condiciones y después de su suavizado y normalizado.

Figura 4-13. Gráfico de torque con respecto a la velocidad.



En la figura 4-13 se presenta el comportamiento de los datos de torque con respecto a la velocidad en ambas condiciones y después de su suavizado y normalizado.

4.3 Creación del modelo de clasificación.

Para realizar las predicciones de datos nuevos que se puedan obtener por medio de un monitoreo constante en la bomba centrífuga, es necesario crear un modelo de clasificación que se ajuste con los datos de aprendizaje anteriormente pre procesados.

A continuación, se realizarán dos modelos de clasificación con Matlab, el modelo de clasificación por vecino más cercano (kNN) y el modelo de clasificación Naive Bayes. Se realizarán predicciones con ambos modelos comparando así sus resultados. Todos los hiperparámetros de los modelos se toman de manera predeterminada exceptuando el número de vecino cercanos del kNN que será de $k=8$.

En primer lugar, una forma fácil de clasificar una observación es usar la misma clase que los ejemplos conocidos más cercanos. Esto se denomina modelo de k vecino más cercano (kNN).

Para este modelo se obtuvieron un total de 1154 datos de aprendizaje de cuatro variables: presión, caudal, torque y velocidad. Cada una relacionada con la clase de falla o de no falla. Para la realización del modelo se puede ajustar un modelo kNN pasando la tabla de datos de aprendizaje a la función `fitcknn` de Matlab.

Con la función `fitcknn` se ajusta el modelo de clasificación en donde, la primera entrada es la tabla de datos totales en la cual se encuentran 338 ejemplos de aprendizaje para el modelo con su respectiva clase. La segunda entrada es el nombre de la variable de respuesta en la tabla, es decir, la clase que se desea que prediga el modelo, que para este caso "condición" varía entre falla y no falla.

La salida es una variable de nombre modelo que contiene el modelo ajustado.

modelo =

ClassificationKNN

PredictorNames: {'VelocidadEnRpm' 'VelocidadEnRads' 'Presinpa' 'Caudalm3s'
'TorqueNm'}

ResponseName: 'condicion'

CategoricalPredictors: []

ClassNames: [con falla sin falla]

ScoreTransform: 'none'

NumObservations: 387

Distance: 'euclidean'

NumNeighbors: 1

Este modelo realiza una observación por cada variación de velocidad de modo que la tabla de datos totales está distribuida de esa manera con 387 velocidades diferentes y en cada observación hay 4 datos de aprendizaje.

En segundo lugar, viene el modelo de Naive Bayes o modelo ingenuo de Bayes que es capaz de utilizar predictores gaussianos, multinomiales o de Kernel. Los modelos ingenuos de Bayes asumen que las observaciones tienen alguna distribución multivariante dada la pertenencia a la clase, pero el predictor o las características que componen la observación son independientes. Este marco puede acomodar un conjunto completo de características de modo que una observación sea un conjunto de recuentos multinomiales.

Para entrenar un modelo Bayes ingenuo se puede hacer de la misma manera que con el modelo kNN ajustando la tabla de datos totales al modelo, pero esta vez con la función `fitcnb`. Después del entrenamiento, se pueden predecir etiquetas o predecir probabilidades posteriores pasando el modelo y los datos del predictor a la función `predict`.

El modelo Naive bayes fue ajustado con los siguientes parámetros:

modelo1 =

ClassificationNaiveBayes

PredictorNames: {'VelocidadEnRpm' 'VelocidadEnRads' 'Presinpa' 'Caudalm3s'
'TorqueNm'}

ResponseName: 'condicion'

CategoricalPredictors: []

ClassNames: [con falla sin falla]

ScoreTransform: 'none'

NumObservations: 387

DistributionNames: {'normal' 'normal' 'normal' 'normal' 'normal'}

DistributionParameters: {2x5 cell}

Como se puede observar dentro de la función fitcnb acomodamos el modelo de clasificación con nuestra tabla de datos totales y especificando que tome las clases de clasificación que se encuentran en la columna de condición. El modelo realiza una observación por cada velocidad de las variables medidas es decir tenemos 386 variaciones de velocidad.

4.4 Realización de predicciones.

Después de haber construido un modelo a partir de los datos, se puede usar para clasificar nuevas observaciones. Esto solo requiere calcular las características de las nuevas observaciones y determinar en qué región del espacio de predicción se encuentran.

- Predicciones con el modelo kNN.

Los predictores de Matlab determinan la clase predicha de nuevas observaciones utilizando métodos de regresión lineal. Los predictores utilizados son los que por medio de regresión lineal hacen una comparación entre datos nuevos.

Las entradas de la función son el modelo entrenado y una tabla de observaciones, con las mismas variables predictoras que se utilizaron para entrenar el modelo. El resultado es una matriz categórica de la clase predicha para cada observación en nuevos datos. Devuelve un vector de etiquetas de clase predichas para los datos del predictor en la tabla o matriz predicción, según el modelo de clasificación de la máquina de vectores de soporte entrenado compacto (modelo).

La base de datos 'datosdeprueba' contiene una tabla que se encuentra en la sección del Anexo C que tiene las mismas variables que las características. Sin embargo, las observaciones son nuevas y no se encuentran en los datos de aprendizaje.

En 'datosdeprueba' tenemos contenidas observaciones para las que se conoce la clase correcta (almacenadas en la variable condición). Esto nos da una manera de probar el modelo comparando las clases predichas por el modelo con las clases verdaderas. La función de predicción ignorará la variable Condición al realizar predicciones a partir del modelo. Se realiza una predicción de 9 datos de prueba:

predicciones =

9x1 categorical array

sin falla

sin falla

con falla

con falla

sin falla

con falla

sin falla

con falla

sin falla

Para tener un mayor grado de precisión al momento de realizar predicciones futuras se puede especificar el valor de k para nuestro modelo de clasificación kNN estableciendo la propiedad "NumNeighbors" al llamar a fitcknn. Se ajusta un k=8 para que el modelo compare un dato nuevo con 8 datos cercanos y lo ubica en la misma clase que tengan la mayoría de esos 8 datos.:

modelo =

```
ClassificationKNN
```

```
    PredictorNames: {'VelocidadEnRpm' 'VelocidadEnRads' 'Presinpa' 'Caudalm3s'  
'TorqueNm'}
```

```
    ResponseName: 'condicion'
```

```
    CategoricalPredictors: []
```

```
    ClassNames: [con falla  sin falla]
```

```
    ScoreTransform: 'none'
```

```
    NumObservations: 387
```

```
    Distance: 'euclidean'
```

```
    NumNeighbors: 8
```

- **Predicciones con el modelo Naive Bayes.**

Se realiza el mismo proceso para el modelo anterior con este modelo para observar cómo se comporta al momento de predecir datos nuevos que no se incluyen en los datos de aprendizaje, pero que pertenecen a las mismas variables medidas.

Esta predicción se realiza con los mismos datos de prueba puestos en el modelo anterior para comparar los resultados de ambos modelos:

predicciones =

9x1 categorical array

sin falla

sin falla

con falla

con falla

sin falla

con falla

con falla

con falla

sin falla

Como se observa, al realizar la predicción de las clases con las 9 velocidades de prueba se observa que el modelo predice de manera correcta todos los datos y acomoda correctamente los datos que pertenecen a una falla y los que no.

4.5 Evaluación del modelo.

Para la evaluación del modelo el primer paso es comparar las clases conocidas con las predicciones del modelo kNN para observar que tan bien se desempeña el modelo con datos nuevos.

La evaluación del kNN se guarda en una variable llamada evaluación1, una comparación realizada entre las predicciones obtenidas con el modelo y las clases conocidas en los datos de prueba:

evaluacion1 =

9x1 logical array

1

1

1

1

1

1

1

1

1

Y también realizamos el mismo proceso con el modelo de clasificación de Bayes guardando el array en una variable llamada evaluación2:

evaluacion2 =

9x1 logical array

1
1
1
1
1
1
1
1
1

Como se observa en la imagen anterior con estos datos de prueba se logra una precisión en la predicción del 100 % en ambos modelos y no será necesario buscar un modelo nuevo ya que se obtuvo, gracias a la gran cantidad de datos de aprendizaje un modelo con la precisión deseada. Como la precisión de estos dos modelos fue del 100% no se tienen en cuenta algunos otros modelos de clasificación porque el resultado sería el mismo.

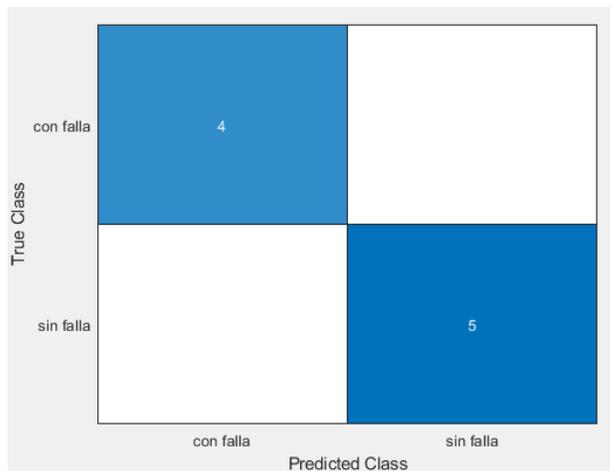
En lugar de precisión (la proporción de predicciones correctas), una métrica de uso común para evaluar un modelo es la tasa de clasificación errónea (la proporción de predicciones incorrectas). Por lo que tener una precisión del 100 % nos indica a su vez que la tasa de clasificación errónea es de 0.

La precisión y la tasa de clasificación errónea logran dar un valor único para el rendimiento general del modelo, pero puede ser útil ver de manera más detallada las clases que suele confundir el modelo. Una matriz de confusión muestra el número de observaciones para cada combinación de clase verdadera y predicha.

Una matriz de confusión se visualiza comúnmente sombreando los elementos según su valor. Comúnmente, los elementos diagonales (las clasificaciones correctas) están sombreados en un color y los otros elementos (las clasificaciones incorrectas) en otro color.

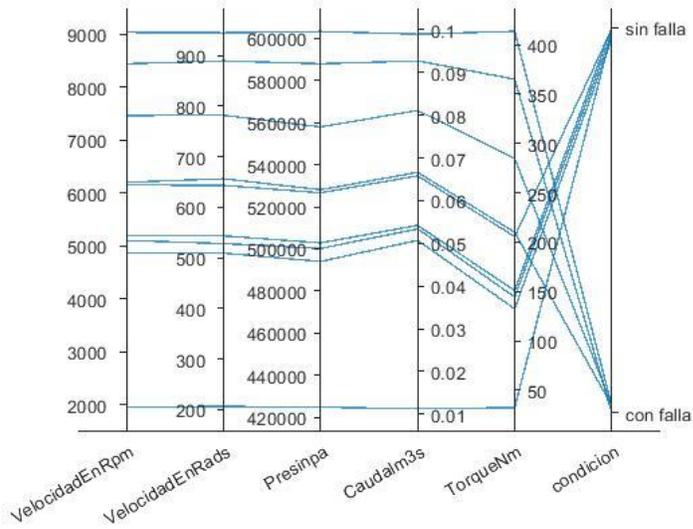
Con la función `confusionchart` de Matlab generamos un cuadro de confusión comparando las predicciones realizadas con las clases conocidas en los datos de prueba. Lo mismo se realiza con el diagrama de coordenadas paralelas que está a continuación que hace una ilustración de los datos de prueba con su respectiva clase predicha.

Figura 4-14. Cuadro de confusión.



El cuadro de confusión nos muestra las clases que suele confundir el modelo con respecto a las clases reales de los datos de prueba.

Figura 4-15. Diagrama de coordenadas paralelas.



El diagrama de coordenadas paralelas nos muestra cada uno de los datos de prueba relacionados con la clase predicha por el algoritmo. Se presentan las variables medidas de cada dato de prueba tales como velocidad, presión, caudal y torque. Por cada línea tenemos un dato de prueba.

5. Conclusiones y recomendaciones.

5.1 Conclusiones.

- Una precisión del 100% se debe a que como los datos de aprendizaje y de prueba fueron tomados por simulación, las variables de la bomba presentan un comportamiento lineal en ambas clases, por esto no es muy complejo para el algoritmo realizar predicciones con los modelos de clasificación, estos son capaces de ubicar los datos en la clase correcta. De modo que la precisión del modelo puede reducirse si se aplica en un entorno real.
- Fue posible utilizar un modelo que simule los datos dinámicos de una bomba centrífuga en un sistema de uso común en cualquier industria para extraer los datos suficientes para el entrenamiento de un modelo de machine learning.
- Gracias a la evaluación del modelo de machine learning en Matlab se puede concluir que con cualquier dato de prueba que se le suministre al modelo este realizará una predicción de manera correcta y alertará en caso de que su predicción lo ubique en la clase de falla.
- El uso de este algoritmo de machine learning en un ambiente real sería de gran ayuda y tendría una gran variación con respecto a su precisión ya que los datos de aprendizaje ya no serían tomados por simulación sino por medio de sensores, lo que puede hacer que las variables no se comporten de una manera tan lineal en ambas clases. Las señales tomadas de los sensores se deben procesar para sacar picos que puedan generar fallas y filtrar valores de presión, caudal y torque que ocasionan fallas
- Con ambos los modelos de clasificación kNN y de Bayes obtuvimos resultados similares y una tasa de clasificación errónea del 0%

5.2 Recomendaciones.

- Si se toman los datos por medio de simulación se deben tomar los datos de una forma detallada y completa ya que de estos datos de aprendizaje depende todo el algoritmo.

6.Anexos

6.1 Anexo A

Tabla 6-1. Datos de clase no falla.

Velocidad en rpm	Velocidad en rad/s	Presión [Pa]	Caudal [m ³ /s]	Torque [N*m]
1780	186,4	416908	0,00655	20,2837
1800	188,5	418177	0,0074018	22,8018
1820	190,6	419289	0,0081407	24,5234
1840	192,7	420311	0,0088161	26,0617
1860	194,8	421278	0,0094507	27,4686
1880	196,9	422207	0,0100573	28,7718
1900	199	423062	0,0106131	29,9772
1920	201,1	423827	0,0111076	31,1042
1940	203,2	424537	0,0115649	32,1785
1960	205,3	425210	0,0119972	33,2125
1980	207,3	425858	0,0124115	34,213
2000	209,4	426486	0,0128122	35,1844
2020	211,5339	427100	0,0132025	36,1293
2050	214,7	428001	0,013773	37,5003
2070	216,77	428592	0,0141458	38,3846
2100	219,9	429469	0,0146966	39,6671
2120	222,005881	430048	0,0150496	40,4926
2140	224,100276	430620	0,015417	41,2776
2160	226,194671	431177	0,0157638	42,0027

2180	228,289066	431722	0,0161023	42,6798
2200	230,383461	432256	0,0164334	43,318
2220	232,477856	432781	0,0167583	43,9556
2240	234,572251	433299	0,0170777	44,5072
2260	236,666647	433810	0,0173923	45,0679
2280	238,761042	434316	0,0177027	45,6114
2300	240,855437	434816	0,0180092	46,1409
2320	242,949832	435312	0,0183123	46,6589
2340	245,044227	435804	0,0186123	46,1677
2360	247,138622	436292	0,0189095	47,6691
2380	249,233017	436777	0,0192041	48,1649
2400	251,327412	437260	0,094964	48,6563
2420	253,421807	437739	0,0197864	49,1446
2440	255,516202	438216	0,0200745	49,631
2460	257,610598	438691	0,0203606	50,1162
2480	259,704993	439164	0,0206451	50,6012
2500	261,799388	439636	0,0209279	51,0867
2520	263,893783	440105	0,0212092	51,5734
2540	265,988178	440574	0,0214891	52,0619
2560	268,082573	441041	0,0217676	52,5527
2580	270,176968	441506	0,0220449	53,0462
2600	272,271363	441971	0,0223211	53,543
2620	274,365758	442435	0,0225961	54,0434
2640	276,460154	442897	0,02287	54,5478
2660	278,554549	443359	0,023143	55,0565
2680	280,648944	443820	0,023415	55,5698
2700	282,743339	444281	0,0236861	56,088
2720	284,837734	444740	0,0239563	56,6113
2740	286,932129	445199	0,0242258	57,14
2760	289,026524	445658	0,0244945	57,6743
2780	291,120919	446116	0,0247624	58,2144
2800	293,215314	446574	0,0250296	58,7602
2820	295,309709	447032	0,0252962	59,3102

2840	297,404105	447489	0,0255623	59,8639
2860	299,4985	447947	0,0258278	60,4214
2880	301,592895	448404	0,0260928	60,9826
2900	303,68729	448861	0,0263572	61,5475
2920	305,781685	449318	0,0266211	62,1163
2940	307,87608	449774	0,0268845	62,6887
2960	309,970475	450231	0,0271474	63,2649
2980	312,06487	450688	0,0274099	63,8449
3000	314,159265	451144	0,0276718	64,4286
3020	316,25366	451600	0,0279332	65,0161
3040	318,348056	452057	0,0281942	65,6073
3060	320,442451	452513	0,0284548	66,2022
3080	322,536846	452969	0,0287149	66,801
3100	324,631241	453426	0,0289746	67,4034
3120	326,725636	453882	0,0292338	68,0097
3140	328,820031	454338	0,0294926	68,6197
3160	330,914426	454795	0,029751	69,2334
3180	333,008821	455251	0,030009	69,8509
3200	335,103216	455707	0,0302666	70,4721
3220	337,197611	456163	0,0305238	71,0971
3240	339,292007	456620	0,0307806	71,7259
3260	341,386402	457076	0,0310371	72,3584
3280	343,480797	457533	0,0312932	72,9947
3300	345,575192	457989	0,0315489	73,6347
3320	347,669587	458446	0,0318043	74,2785
3340	349,763982	458903	0,0320593	74,926
3360	351,858377	459359	0,0323139	75,5773
3380	353,952772	459816	0,0325683	76,2324
3400	356,047167	460273	0,0328223	76,8912
3420	358,141563	460730	0,033076	77,5538
3440	360,235958	461188	0,333243	78,2201
3460	362,330353	461645	0,0335824	78,8902
3480	364,424748	462102	0,0338351	79,5641

3500	366,519143	462560	0,0340875	80,2417
3520	368,613538	463018	0,0343397	80,9231
3540	370,707933	463476	0,0345915	81,6083
3560	372,802328	463934	0,034843	82,2972
3580	374,896723	464392	0,0350943	82,9898
3600	376,991118	464850	0,0353453	83,6863
3620	379,085514	465309	0,035596	84,3865
3640	381,179909	465768	0,0358464	85,0904
3660	383,274304	466226	0,0360966	85,7981
3680	385,368699	466686	0,0363465	86,5096
3700	387,463094	467145	0,0365961	87,2249
3720	389,557489	467604	0,0368455	87,9439
3740	391,651884	468064	0,0370947	88,6667
3760	393,746279	468524	0,0373436	89,3932
3780	395,840674	468984	0,0375922	90,1235
3800	397,935069	469444	0,0378407	90,8576
3820	400,029465	469905	0,0380889	91,5954
3840	402,12386	470366	0,0383368	92,337
3860	404,218255	470827	0,0385845	93,0824
3880	406,31265	471288	0,038832	93,8315
3900	408,407045	471749	0,0390793	94,5844
3920	410,50144	472211	0,0393264	95,3411
3940	412,595835	472673	0,0395732	96,1015
3960	414,69023	473135	0,0398199	96,8657
3980	416,784625	473598	0,0400663	97,6337
4000	418,87902	474060	0,0403125	98,4054
4020	420,973416	474523	0,0405586	99,1809
4040	423,067811	474987	0,0408044	99,9602
4060	425,162206	475450	0,04105	100,743
4080	427,256601	475914	0,0412955	101,53
4100	429,350996	476378	0,0415407	102,321
4120	431,445391	476842	0,0417858	103,115
4140	433,539786	477307	0,0420307	103,913

4160	435,634181	477772	0,0422754	104,715
4180	437,728576	478237	0,0425199	105,52
4200	439,822972	478702	0,0427642	106,33
4220	441,917367	479168	0,0430084	107,143
4240	444,011762	479634	0,0432924	107,96
4260	446,106157	480100	0,0434962	108,781
4280	448,200552	480567	0,0437399	109,605
4300	450,294947	481034	0,0439834	110,433
4320	452,389342	481501	0,0442267	111,265
4340	454,483737	481969	0,0444699	112,101
4360	456,578132	482437	0,0447129	112,941
4380	458,672527	482905	0,0449557	113,784
4400	460,766923	483374	0,0451984	114,631
4420	462,861318	483842	0,045441	115,482
4440	464,955713	484311	0,0456834	116,337
4460	467,050108	484781	0,0459257	117,195
4480	469,144503	485251	0,0461678	118,057
4500	471,238898	485721	0,0464097	118,923
4520	473,333293	486191	0,0466515	119,793
4540	475,427688	486662	0,0468932	120,667
4560	477,522083	487133	0,0471348	121,544
4580	479,616478	487605	0,0473762	122,425
4600	481,710874	488077	0,0476174	123,31
4620	483,805269	488549	0,0478586	124,198
4640	485,899664	489021	0,0480996	125,091
4660	487,994059	489494	0,0483405	125,987
4680	490,088454	489967	0,0485812	126,887
4700	492,182849	490441	0,0488218	127,79
4720	494,277244	490915	0,0490623	128,698
4740	496,371639	491389	0,0493027	129,609
4760	498,466034	497864	0,049543	130,524
4780	500,560429	492339	0,0497831	131,443
4800	502,654825	492814	0,0500231	132,365

4820	504,74922	493290	0,050263	133,291
4840	506,843615	493766	0,0505028	134,221
4860	508,93801	494242	0,0507425	135,155
4880	511,032405	494719	0,050982	136,093
4900	513,1268	495196	0,0512215	137,034
4920	515,221195	495673	0,0514608	137,979
4940	517,31559	496151	0,0517	138,928
4960	519,409985	496630	0,0519391	139,881
4980	521,50438	497108	0,0521781	140,838
5000	523,598776	497587	0,052417	141,798
5020	525,693171	498067	0,0526558	142,762
5040	527,787566	498546	0,0528945	143,73
5060	529,881961	499027	0,0531331	144,701
5080	531,976356	499507	0,0533715	145,676
5100	534,070751	499988	0,05361	146,656
5120	536,165146	500469	0,0538483	147,638
5140	538,259541	500951	0,0540865	148,625
5160	540,353936	501433	0,0543246	149,616
5180	542,448332	501915	0,0545626	150,61
5200	544,542727	502398	0,0548006	151,608
5220	546,637122	502881	0,0550384	152,609
5240	548,731517	503365	0,0552761	153,615
5260	550,825912	503849	0,0555138	154,624
5280	552,920307	504333	0,0557513	155,637
5300	555,014702	504818	0,0559888	156,654
5320	557,109097	505303	0,0562261	157,675
5340	559,203492	505789	0,0564634	158,699
5360	561,297887	506275	0,0567006	159,727
5380	563,392283	506761	0,0569378	160,759
5400	565,486678	507248	0,0571748	161,795
5420	567,581073	507735	0,0574118	162,834
5440	569,675468	508223	0,0576486	163,878
5460	571,769863	508711	0,0578854	164,925

5480	573,864258	509199	0,0581221	165,976
5500	575,958653	509688	0,0583588	167,03
5520	578,053048	510177	0,0585953	168,088
5540	580,147443	510667	0,0588318	169,151
5560	582,241838	511157	0,0590682	170,217
5580	584,336234	511647	0,0593045	171,286
5600	586,430629	512138	0,0595408	172,36
5620	588,525024	512630	0,059777	173,437
5640	590,619419	513121	0,0600131	174,518
5660	592,713814	513613	0,0602491	175,603
5680	594,808209	514106	0,0604851	176,691
5700	596,902604	514599	0,0607209	177,784
5720	598,996999	515092	0,0609568	178,88
5740	601,091394	515586	0,0611925	179,979
5760	603,185789	516080	0,0614282	181,083
5780	605,280185	516574	0,0616638	182,191
5800	607,37458	517069	0,0618993	183,302
5820	609,468975	517565	0,0621348	184,417
5840	611,56337	518061	0,0623702	185,535
5860	613,657765	518557	0,0626056	186,658
5880	615,75216	519054	0,0628408	187,784
5900	617,846555	519551	0,0630761	188,914
5920	619,94095	520048	0,0633112	190,048
5940	622,035345	520546	0,0635463	191,186
5960	624,129741	521044	0,0637813	192,327
5980	626,224136	521543	0,0640163	193,472
6000	628,318531	522043	0,0642512	194,621
6020	630,412926	522542	0,064486	195,774
6040	632,507321	523042	0,0647208	196,931
6060	634,601716	523543	0,0649555	198,091
6080	636,696111	524044	0,0651902	199,255
6100	638,790506	524545	0,0654248	200,423
6120	640,884901	525047	0,0656594	201,594

6140	642,979296	525549	0,0658939	202,77
6160	645,073692	526052	0,0661283	203,949
6180	647,168087	526555	0,0663627	205,132
6200	649,262482	527058	0,066597	206,318

6.2 Anexo B.

Tabla 6-2. Datos de clase falla.

Velocidad (rpm)	velocidad (rad/s)	Presión [Pa]	Caudal (m ³ /s)	Torque (N*m)
6201	649,367	527083	0,066609	206,378
6220	651,357	527587	0,066843	207,569
6240	653,451	528092	0,067077	208,763
6260	655,546	528597	0,067311	209,961
6280	657,640	529102	0,067546	211,163
6300	659,734	529608	0,067780	212,369
6320	661,829	530114	0,068014	213,579
6340	663,923	530620	0,068248	214,792
6360	666,018	531127	0,068481	216,009
6380	668,112	531635	0,068715	217,230
6400	670,206	532143	0,068949	218,454
6420	672,301	532626	0,069171	219,621
6440	674,395	533134	0,069405	220,853
6460	676,490	533643	0,069638	222,089
6480	678,584	534153	0,069872	223,329
6500	680,678	534663	0,070106	224,572
6520	682,773	535174	0,070339	225,819
6540	684,867	535685	0,070573	227,070
6560	686,962	536196	0,070806	228,325
6580	689,056	536708	0,071039	229,583

6600	691,150	537220	0,071273	230,845
6620	693,245	537733	0,071506	232,111
6640	695,339	538246	0,071739	233,381
6660	697,434	538760	0,071912	234,655
6680	699,528	539274	0,072206	235,922
6700	701,622	539788	0,072439	237,213
6720	703,717	540303	0,072672	238,498
6740	705,811	540818	0,072905	239,787
6760	707,906	541334	0,073138	241,079
6780	710,000	541850	0,073371	242,375
6800	712,094	542367	0,073604	243,675
6820	714,189	542884	0,073894	244,979
6840	716,283	543402	0,074069	246,287
6860	718,378	543920	0,074302	247,598
6880	720,472	544439	0,074535	248,913
6900	722,566	544957	0,074767	250,232
6920	724,661	545477	0,075000	251,554
6940	726,755	545997	0,075233	252,881
6960	728,849	546517	0,075465	254,211
6980	730,944	547038	0,075698	255,545
7000	733,038	547559	0,075930	256,883
7020	735,133	548081	0,076163	258,224
7040	737,227	548603	0,076395	259,570
7060	739,321	549125	0,076627	260,919
7080	741,416	549648	0,076860	262,271
7100	743,510	550172	0,077092	263,628
7120	745,605	550696	0,077324	264,988
7140	747,699	551220	0,077556	266,353
7160	749,793	551745	0,077789	267,721
7180	751,888	552270	0,078021	269,092
7200	753,982	552796	0,078253	270,468
7220	756,077	553322	0,078485	271,847
7240	758,171	553849	0,078717	273,230

7260	760,265	554376	0,078949	274,617
7280	762,360	554903	0,079181	276,008
7300	764,454	555431	0,079413	277,402
7320	766,549	555960	0,079645	278,800
7340	768,643	556489	0,079877	280,202
7360	770,737	557018	0,080109	281,608
7380	772,832	557548	0,080340	283,017
7400	774,926	558078	0,080572	284,431
7420	777,021	558609	0,080804	285,848
7440	779,115	559140	0,081036	287,268
7460	781,209	559672	0,081267	288,693
7480	783,304	560204	0,081499	290,121
7500	785,398	560736	0,081731	291,554
7520	787,493	561269	0,081962	292,990
7540	789,587	561803	0,082194	294,429
7560	791,681	562337	0,072425	295,873
7580	793,776	562871	0,082657	297,320
7600	795,870	563406	0,082888	298,771
7620	797,965	563941	0,083120	300,226
7640	800,059	564477	0,083351	301,684
7660	802,153	565013	0,083582	303,147
7680	804,248	565550	0,083814	304,613
7700	806,342	566087	0,084045	306,083
7720	808,437	566625	0,084276	307,557
7740	810,531	567163	0,084508	309,034
7760	812,625	567701	0,084739	310,515
7780	814,720	568240	0,084970	312,000
7800	816,814	568780	0,085201	313,489
7820	818,908	569320	0,085432	314,982
7840	821,003	569860	0,085664	316,478
7860	823,097	570401	0,085895	317,978
7880	825,192	570942	0,086126	319,482
7900	827,286	571484	0,086357	320,990

7920	829,380	572026	0,086588	322,501
7940	831,475	572569	0,086819	324,017
7960	833,569	573112	0,087050	325,536
7980	835,664	573656	0,087281	327,058
8000	837,758	574200	0,087512	328,585
8020	839,852	574744	0,087743	330,115
8040	841,947	575289	0,087973	331,649
8060	844,041	575835	0,088204	331,187
8080	846,136	576381	0,088435	334,729
8100	848,230	576927	0,088666	336,274
8120	850,324	577474	0,088897	337,824
8140	852,419	578021	0,089127	339,377
8160	854,513	578569	0,089358	340,934
8180	856,608	579117	0,089589	342,494
8200	858,702	579666	0,089819	344,058
8220	860,796	580215	0,090050	345,627
8240	862,891	580765	0,090281	347,198
8260	864,985	581315	0,090511	348,774
8280	867,080	581866	0,090742	350,354
8300	869,174	582417	0,090972	351,937
8320	871,268	582968	0,091203	353,524
8340	873,363	583520	0,091433	355,115
8360	875,457	584073	0,091664	356,709
8380	877,552	584625	0,091894	358,308
8400	879,646	585179	0,092125	359,910
8420	881,740	585733	0,092355	361,516
8440	883,835	586287	0,092586	363,125
8460	885,929	586842	0,092816	364,739
8480	888,024	587397	0,093046	366,356
8500	890,118	587953	0,093277	367,977
8520	892,212	588509	0,093507	369,602
8540	894,307	589065	0,093737	371,230
8560	896,401	589623	0,093967	372,863

8580	898,495	590180	0,094198	374,499
8600	900,590	590738	0,094428	376,139
8620	902,684	591297	0,094658	377,782
8640	904,779	591856	0,094888	379,430
8660	906,873	592415	0,095118	381,081
8680	908,967	592975	0,095349	382,736
8700	911,062	593535	0,095579	384,395
8720	913,156	594096	0,095809	386,057
8740	915,251	594658	0,096039	387,724
8760	917,345	595219	0,096269	389,394
8780	919,439	595782	0,096499	391,068
8800	921,534	596344	0,096729	392,745
8820	923,628	596908	0,096959	394,427
8840	925,723	597471	0,097189	396,112
8860	927,817	598035	0,097419	397,801
8880	929,911	598600	0,097649	399,494
8900	932,006	599165	0,097879	401,190
8920	934,100	599731	0,098109	402,891
8940	936,195	600297	0,098339	404,595
8960	938,289	600863	0,098568	406,303
8980	940,383	601430	0,098798	408,014
9000	942,478	601997	0,099028	409,730
9020	944,572	602565	0,099258	411,449
9040	946,667	603134	0,099488	413,172
9060	948,761	603702	0,099717	414,899
9080	950,855	604272	0,099947	416,629
9100	952,950	604841	0,100177	418,363
9120	955,044	605412	0,100407	420,102
9140	957,139	605982	0,100636	421,843
9160	959,233	606554	0,100866	423,589
9180	961,327	607125	0,101096	425,339
9200	963,422	607697	0,101325	427,092
9220	965,516	608270	0,101555	428,849

9240	967,611	608843	0,101785	430,610
9260	969,705	609417	0,102014	432,374
9280	971,799	609991	0,102244	434,142
9300	973,894	610565	0,102473	435,914
9320	975,988	611140	0,102703	437,690
9340	978,083	611715	0,102933	439,470
9360	980,177	612291	0,103162	441,253
9380	982,271	612868	0,103392	443,041
9400	984,366	613444	0,103621	444,832
9420	986,460	614022	0,103851	446,626
9440	988,554	614600	0,104080	448,425
9460	990,649	615178	0,104309	450,227
9480	992,743	615756	0,104539	452,033
9500	994,838	616336	0,104768	453,843

6.3 Anexo C.

Tabla 6-3. Datos de prueba.

Velocidad en rpm	Velocidad en rad/s	Presión [Pa]	Caudal [m ³ /s]	Torque [N*m]	condición
1930	202,109127	424187	0,01134	31,647	sin falla
5070	530,929158	499267	0,053235	145,188	sin falla
6230	652,404074	527814	0,066948	208,106	con falla
7430	778,067781	558874	0,080920	286,588	con falla
5190	543,495529	502157	0,0546816	151,108	sin falla
8470	886,976326	587119	0,092931	365,547	con falla
6170	646,120889	526303	0,0662455	204,54	sin falla
9030	945,619389	602849	0,099373	412,310	con falla
4830	505,796417	493528	0,0503829	133,756	sin falla

6.4 Anexo D.

Código Matlab.

```
>> variacion_torque_falla=std(datosfalla.torque);
>> variacion_caudal_falla=std(datosfalla.caudal);
>> variacion_presion_falla=std(datosfalla.presion);
>> variacion_torque_nofalla=std(datosnofalla.torque);
>> variacion_caudal_nofalla=std(datosnofalla.caudal);
>> variacion_presion_nofalla=std(datosnofalla.presion);
>> suavizado=smoothdata(datosfalla);
>> suavizado1=smoothdata(datosnofalla);
>> normalizado=normalize(suavizado);
>> normalizado1=normalize(suavizado1);
>> modelo_knn=fitcknn(datostotales, datostotales.condicion);
>> modelo_cnb=fitcnb(datostotales, datostotales.condicion);
>> prediccion=predict(modelo_knn, datosdeprueba);
>> prediccion1=predict(modelo_cnb, datosdeprueba);
>> precision=prediccion==datosdeprueba.condicion;
>> precision1=prediccion1==datosdeprueba.condicion;
>> TCE=prediccion~datosdeprueba.condicion;
>> TCE1=prediccion1~datosdeprueba.condicion;
>> confusionchart (datosdeprueba. condicion, prediccion)
>> parallelcoords(datosdeprueba.condicion, prediccion)
```

Bibliografía

- [1] GARCÍA GARRIDO, Santiago. Vibración en turbinas de vapor. [En línea]. <http://www.renovetec.com/590-mantenimiento-industrial/110-mantenimiento-industrial/309-vibracion-en-turbinas-de-vapor>. 01/10/2020.J.
- [2] KHADEMI, A., RAJI, F. y SADEGHI, M. (2019, abril). LoT Enabled Vibration Monitoring Toward Smart Maintenance. In 2019, Third International Conference on Internet of Things and Applications, University of Isfahan, Isfahan, Iran p. 1-6.
- [3] PINHEIRO ALVES, Pinheiro, BRANDAO MODESTO, Lago y DA COSTA, Cesar. (2019). Vibration Analysis of Rotary Machines Using Machine Learning Technique. European Journal of Engineering Research and Science, Vol. 2, p. 12-16.
- [4] EMPRESITE, Mantenimiento predictivo. [En línea]. Recuperado de <https://empresite.eleconomistaamerica.co/Actividad/MANTENIMIENTO-PREDICTIVO/> Fecha de consulta 19 de octubre de 2020.
- [5] MPC. Mantenimiento predictivo computarizado Ltda. [En línea]. Recuperado de <http://mpc-predictivo.com/representaciones.htm>. Fecha de consulta 19 de octubre de 2020.
- [6] PREDICTIVO. Análisis de vibraciones balanceo dinámico - alineación laser. [En línea]. Recuperado de <http://www.predictivolta.com/> Fecha de consulta 19 de octubre de 2020.
- [7] ZHOU, Taotao y LI Yan-fu. (2020). Degradation Assessment and Remaining Useful Life Prognostic of Centrifugal Pump using Multi-Sensor Process Monitoring The 9th Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling (APARM 2020 – Vancouver).
- [8] ARDILA, Alfonso, MARTÍNEZ, Felipe, GARCÉS, Kelli, BARBIERI, Giacomo, SÁNCHEZ-LONDONO, David, CAIELLI, Andrea, y FUMAGALLI, Laura. (2020). XRepo - Towards an information system for prognostics and health management analysis. Procedia Manufacturing, 42, p.146-153.
- [9] PREDITEC. Mantenimiento predictivo. [en línea]. <http://www.preditec.com/mantenimiento-predictivo/>. 01/02/2021.
- [10] AUTOMATION FÓRUM. ¿qué es el mantenimiento correctivo? [página web]. Fecha de consulta. recuperado de <https://www.tecsaqro.com.mx/blog/mantenimiento-correctivo/>
- [11] ROLF, Isserman. Fault-Diagnosis Applications. [libro]. Alemania: Technische Universität Darmstadt, 2011. Cap 6 fault diagnosis of pumps, pag 143.
- [12] ITAINNOVA. [imagen]. Funcionamiento de IA y machine learning. Consultado enero 20 de 2021]. Disponible en internet: <https://www.itainnova.es/blog/big-data-y-sistemas-cognitivos/mantenimiento-predictivo-mediante-inteligencia-artificial-y-algoritmos-de-deep-learning/>
- [13] MATHWORKS. [Imagen]. Onramp de aprendizaje automático. [Consultado noviembre 25 de 2020]. Disponible en internet: <https://matlabacademy.mathworks.com/R2020a/portal.html?course=machinelearning>
- [14] LEZAMA, Jean. [Imagen]. Conceptos básicos bombas centrifugas. SlidePlayer. [consultado noviembre 20 de 2020]. Disponible en internet <https://slideplayer.es/slide/13971578/>

- [15] : ANONIMO. [imagen].Bombas Centrifugas y uso en instalaciones eléctricas.[Consultado noviembre 20 de 2020]. Disponible en internet: <https://www.seguas.com/bombas-centrifugas-instalaciones-hidraulicas/>
- [16] MISTRAL. Gabriela.[fotografía]. Bombas y válvulas: Apuntando a un mejor desempeño. [Consultado noviembre 25 de 2020]. Disponible en internet:<https://www.mch.cl/informes-tecnicos/bombas-y-valvulas-apuntando-a-un-mejor-desempeno/#>
- [17] MATHWORKS. [Imagen]. Centrifugal Pump (TL) . [Consultado ENERO 25 de 2020]. Disponible en internet: <https://www.mathworks.com/help/physmod/hydro/ref/centrifugalpumt1.html>.
- [18] Anonimo. Leyes de Afinidad de las Bombas centrífugas. [En línea]. Consultado [febrero 03 de 2021]. Recuperado de <https://www.bombascentrifugas.net/p/leyes-de-afinidad.html>
- [19] Características de las Bombas Centrifugas. [En línea]. E-Learning Universidad del Norte. [Citado el 3 de mayo de 2021]. Disponible https://www2.javerianacali.edu.co/sites/ujc/files/normas_icontec.pdf
- [20] Bombas centrífugas y su uso en instalaciones hidráulicas. [En línea]. Seguas aire comprimido y frio industrial. [Citado el 3 de mayo de 2021]. Disponible <https://www.seguas.com/bombas-centrifugas-instalaciones-hidraulicas/>
- [21] O. Motaghare, A. S. Pillai and K. I. Ramachandran, "Predictive Maintenance Architecture," 2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2018, pp. 1-4, doi: 10.1109/ICCIC.2018.8782406.
- [22] L. Guo, B. Xiao, R. Shi y C. Lv, "La tarea de mantenimiento correctivo determina la investigación del método basada en FMECA", *2009 8th International Conference on Reliability, Maintainability and Safety*, 2009, págs. 665-669, doi: 10.1109/ICRMS.2009.5270104.
- [23] A. Jezzini, M. Ayache, L. Elkhansa, B. Makki and M. Zein, "Effects of predictive maintenance(PdM), Proactive maintenace(PoM) & Preventive maintenance(PM) on minimizing the faults in medical instruments," 2013 2nd International Conference on Advances in Biomedical Engineering, 2013, pp. 53-56, doi: 10.1109/ICABME.2013.6648845.
- [24] K. Pahwa and N. Agarwal, "Stock Market Analysis using Supervised Machine Learning," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 2019, pp. 197-200, doi: 10.1109/COMITCon.2019.8862225.
- [25] M. G. Gramajo, L. Ballejos and M. Ale, "Software Requirements Engineering through Machine Learning Techniques: A Literature Review," 2018 IEEE Biennial Congress of Argentina (ARGENCON), 2018, pp. 1-7, doi: 10.1109/ARGENCON.2018.8646010.
- [26] F. Khomh, B. Adams, J. Cheng, M. Fokaefs and G. Antoniol, "Software Engineering for Machine-Learning Applications: The Road Ahead," in *IEEE Software*, vol. 35, no. 5, pp. 81-84, September/October 2018, doi: 10.1109/MS.2018.3571224.
- [27] DP Wangoo, "Técnicas de inteligencia artificial en ingeniería de software para la reutilización y el diseño automatizados de software", *4a Conferencia Internacional sobre Comunicación y Automatización Informática (ICCCA)* 2018, págs. 1-4, doi: 10.1109/CCAA.2018.8777584.
- [28] PTC Mathcad. [imagen]. Cómo utilizar los datos de suavizado en el análisis predictivo. Consultado mayo 20 de 2021. Disponible en internet: <https://maniqui.ru/computadoras-y-software/big-data/la-ciencia-de-datos/1409-cmo-utilizar-los-datos-de-suavizado-en-el-analisis.html>.
- [29] The MathWorks, Inc. Water Supply System. [En línea]. fecha de consulta mayo del 2021. disponible en <https://la.mathworks.com/help/physmod/hydro/ug/water-supply-system.html>