



**Desarrollo de Plataforma de Monitoreo para activos críticos del Sector Eléctrico en Transformadores.**

Juan Esteban López Cardona

Informe de práctica para optar al título de Ingeniero Electricista

Asesores

José Robinsón Ortiz Castrillón, Magíster (MSc) en Ingeniería

Andrés Felipe Posada Calle, Ingeniero Eléctrico

Universidad de Antioquia  
Facultad de Ingeniería  
Ingeniería Eléctrica  
Medellín, Antioquia, Colombia  
2022

Cita	López Cardona [1]
<b>Referencia</b>	[1] J. E. López Cardona, “Desarrollo de Plataforma de Monitoreo para activos críticos del Sector Eléctrico en Transformadores”, Trabajo de grado profesional, Ingeniería Eléctrica, Universidad de Antioquia, Medellín, Antioquia, Colombia, 2022.
Estilo IEEE (2020)	



Centro de Documentación Ingeniería (CENDOI)

**Repositorio Institucional:** <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - [www.udea.edu.co](http://www.udea.edu.co)

**Rector:** John Jairo Arboleda Céspedes.

**Decano/Director:** Jesús Francisco Vargas Bonilla.

**Jefe departamento:** Noé Alejandro Mesa Quintero.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

## CONTENIDO

RESUMEN.....	10
ABSTRACT .....	11
I. INTRODUCCIÓN .....	12
II. OBJETIVOS.....	13
A. Objetivo general .....	13
B. Objetivos específicos.....	13
III. MARCO TEÓRICO.....	14
1) Transformador .....	14
2) Índice de salud .....	14
3) Análisis de gases disueltos .....	14
4) Redes Neuronales .....	14
5) Redes Neuronales Artificiales.....	15
6) Redes Neuronales Profundas.....	15
7) Librería PLOTLY .....	15
8) Librería PANDAS .....	16
9) Librería Matplotlib .....	17
10) Librería Sklearn .....	17
11. Gases generados en transformadores inmersos en aceite basado en la normativa internacional IEE STD C57.104TM-2008.....	18
12.Evaluación de la condición del transformador usando concentraciones individuales y TDGC (Gas combustible total disuelto).....	19
13. Triángulo de Duval.....	20
13.1 Comparación métodos de diagnósticos de fallas.....	21
13.2 Método del triángulo.....	21

IV. METODOLOGÍA .....	23
A. Actividad 1 Revisión Bibliográfica.....	23
B. Actividad 2. Búsqueda de bases de datos.....	23
C. Actividad 3. Preprocesamiento de datos. ....	23
D. Actividad 4. Implementación modelo de ML. ....	23
E. Actividad 5. Análisis de resultados. ....	23
F. Actividad 6. Informe final.....	23
V. RESULTADOS Y ANÁLISIS.....	24
A.1 Ejemplo de aplicación triángulo de Duval. ....	24
A.2 Tipos de fallas [13],[14],[15]. ....	25
B. Algoritmos de Machine Learning utilizando series de tiempo.....	26
B.1 Enfoque estadístico o clásico.....	27
B.2 Inteligencia artificial.....	27
B.3 Enfoques híbridos. ....	27
B.4. Análisis de resultados.....	27
B.4.1 Mapa de calor. ....	27
B.4.2 Gráficas series de tiempo consumo de Energía por año vs Ingreso de personas.....	28
B.4.4 Gráficas predicciones series de tiempo. Predicción Ingreso de personas.....	33
C. Algoritmos de Machine Learning implementados para detectar fallas en transformadores. ....	36
C.1 Medidor magnético de aceite.....	36
C.2 Árboles de decisión. ....	37
C.3 Regresión logística. ....	39
C.4 Máquinas de vectores de soporte.....	39
C.5 Análisis de resultados. ....	41
C.6 Matriz de confusión. ....	42

C.7 Métricas de medición.....44

VI. CONCLUSIONES .....46

REFERENCIAS .....47

ANEXOS.....49

## LISTA DE TABLAS

TABLA I CONCENTRACIÓN DE GASES DISUELTOS .....	20
TABLA II MÉTODOS DE DIAGNÓSTICO DE FALLAS DE GASES EN TRANSFORMADORES. ....	21
TABLA III MÉTRICAS DE MEDICIÓN DE LOS MÉTODOS UTILIZADOS .....	45

## LISTA DE FIGURAS

Fig. 1. Estructuras de datos Series y DataFrame.....	16
Fig. 2. Estructura de una red LSTM.....	18
Fig. 3. Método del triángulo de Duval. ....	22
Fig. 4. Ejemplo de aplicación Triángulo de Duval. ....	24
Fig. 5. Resultado práctico Triángulo de Duval versión 1.....	25
Fig. 6. Mapa de calor variables registradas DataSet. ....	28
Fig. 7. Ingreso de personas VS consumo de energía año 2019 centro comercial Unicentro. ....	29
Fig. 8. Ingreso de personas VS consume de energía año 2020 centro comercial Unicentro. ....	29
Fig. 9. Ingreso de personas VS consumo de energía año 2021 centro comercial Unicentro. ....	30
Fig. 10. Predicción consumo de energía año 2022 y principios 2023.....	31
Fig. 11. Predicción tendencia para el consumo de energía finales de 2022 y principios.....	31
Fig. 12. Predicción tendencia semanal para el consumo de energía a finales de 2022 y principios de 2023. ....	32
Fig. 13. Predicción tendencia anual para el consumo de energía finales de 2022 y principios 2023. ....	32
Fig. 14. Predicción tendencia diaria para el consumo de energía finales de 2022 y principios de 2023.....	33
Fig. 15. Predicción Ingreso de personas 2022 y principios 2023.....	34
Fig. 16. Predicción tendencia para el ingreso de personas finales de 2022 y principios de 2023..	34
Fig. 17. Predicción tendencia semanal para el ingreso de personas finales de 2022 y principios de 2023.....	35
Fig. 18. Predicción tendencia seminal para el ingreso de personas finales de 2022 y principios de 2023.....	35
Fig. 19. Predicción tendencia diaria para el ingreso de personas finales de 2022 y principios de 2023.....	36
Fig. 20. Imagen de medidor magnético de aceite.....	37
Fig. 21. Árboles de decisión.....	38
Fig. 22. Regresión Logística. ....	39

Fig. 23. Máquina de vectores de soporte.....	40
Fig. 24. Temperatura del aceite a largo plazo. ....	41
Fig. 25. Indicador de nivel de aceite a lo largo del tiempo. ....	41
Fig. 26. Advertencia de falla de indicador magnético de aceite a lo largo del tiempo. ....	42
Fig. 27. Matriz de confusión. ....	43
Fig. 28. Matriz de confusión árboles de decisión.....	43
Fig. 29. Matriz de confusión regresión Logística.....	44
Fig. 30. Matriz de confusión máquina de vectores de soporte.....	44
Fig. 31. Dashboard STATCOM. ....	49

## SIGLAS, ACRÓNIMOS Y ABREVIATURAS

<b>ML</b>	Machine Learning
<b>SQL</b>	Structured Query Language
<b>IEE</b>	Institution of Electrical Engineers
<b>IEC</b>	International Electrotechnical Commission
<b>ARN</b>	Artificial Neural Network
<b>DGA</b>	Diagnóstico de Gases Disueltos
<b>SVM</b>	Support Vector Machine
<b>UDEA</b>	Universidad de Antioquia

## RESUMEN

En este documento se describe el desarrollo del diseño de una plataforma de monitoreo para activos eléctricos en la nube con la ayuda de varios software de programación, como lo son Python, Google Colab y con la ayuda de bases de datos como SQL y software open source para la visualización de información en tiempo real llamado Grafana. El objetivo de este proyecto es unificar varias formas de medir el índice de salud de un transformador y sus matrices de riesgo, mediante triángulos de Duval, estado de gases, Árboles de decisión, series de tiempo y otros algoritmos de Machine Learning (ML). La idea de la plataforma de monitoreo no solamente es medir el índice de salud de un transformador, sino también aplicar ciencia de datos al consumo histórico eléctrico de los centros comerciales, que son clientes de la empresa donde se realizan las prácticas. En resumen, este trabajo unifica todos estos servicios en una plataforma en la nube.

***Palabras clave* — Python, Monitoreo, Plataforma Eléctrica, Transformador, Algoritmo, Series de tiempo, Visualización, Machine Learning.**

## ABSTRACT

This document describes the development of the design of a monitoring platform for electrical assets in the cloud with the help of various programming software, such as Python, Google Colab and with the help of databases such as SQL and open source software for the visualization of real-time information called Grafana. The objective of this project is to unify several ways to measure the health index of a transformer and its risk matrices, using Duval triangles, gas state, decision trees, time series and other Machine Learning algorithms. The idea of the monitoring platform is not only to measure the health index of a transformer, but also to apply data science to the historical electricity consumption of shopping malls, which are customers of the company where the practices are performed. In short, this work unifies all these services in a cloud platform.

***Keywords* — Python, Monitoring, Electric Platform, Electrical Transformer, Algorithm, Time Series, Visualization, Machine Learning, Dashboard.**

## I. INTRODUCCIÓN

El sector eléctrico ha venido creciendo desde sus orígenes hasta la actualidad, gracias al significativo aumento en el consumo de energía, llevando al desarrollo y a un cambio fundamental en los sistemas de potencia. Según la IEEE, un sistema de potencia es: una red formada por unidades generadoras de electricidad, cargas y líneas de transmisión de potencia, incluyendo el equipo asociado, conectado eléctrica o mecánicamente a la red [1]. Debido a este crecimiento desmesurado en el consumo energético y su desarrollo, es menester estar vigilando constantemente el comportamiento y las variables más importantes de los sistemas eléctricos para su correcto funcionamiento. Una estrategia que ha venido creciendo en su uso en la implementación de sistemas de monitoreo que son sistemas integrales telegestionados que se encargan de medir, controlar o monitorear todas las variables eléctricas en tiempo real y así medir los consumos, perfiles de carga, predecir fallas que afectan el desempeño, hacer mantenimiento predictivo o saber dónde hay fugas de energía; con el fin de aumentar la productividad del activo y prolongar su vida útil y por ende, hacer más rentables los procesos. Todas las variables pueden ser monitoreadas y analizadas en tiempo real a través de un software disponible 24/7 en internet, y generar alertas por correo electrónico y mensajes de texto SMS antes situaciones anormales o de fallas [2].

Como se mencionó anteriormente, el monitoreo más que una solución integral, es una necesidad que se tiene que implementar en la mayoría de los activos eléctricos, con esta se pueden monitorear y controlar en algunos casos variables como el consumo energético, el factor de potencia, temperatura, el estado de los gases en un transformador, etc. Para esto, es necesario tener un monitoreo lo más preciso posible. Debido a esto, se implementan algoritmos computacionales de Machine Learning (ML). Hay muchos modelos para dicho monitoreo, como lo son las redes neuronales, algoritmos de ML o una combinación de ambos. En este aspecto nos enfocaremos principalmente en el monitoreo de transformadores.

Con el fin de tener un monitoreo preciso, se propuso un modelo basado en el pronóstico de índices de salud de transformadores utilizando el método de gases disueltos en aceite combinado con modelos de ML. Este proceso será combinado con otros métodos como el triángulo de Duval, matrices de riesgo, árboles de decisión, regresión logística, máquina de vectores de soporte. La programación de este método se hizo mediante el software Python y sus librerías relacionadas con análisis de datos. Cabe resaltar que en la industria se han venido trabajando con varios modelos para el pronóstico y monitoreo del índice de salud de un transformador, y estos métodos son los que ha dado mejores resultados respecto a modelos antiguos como IEEE gas característico e IEC gases básicos [3].

---

## II. OBJETIVOS

### A. *Objetivo general*

Desarrollar una plataforma de monitoreo de activos eléctricos (centros comerciales, subestaciones) a partir de la implementación de técnicas de aprendizaje máquina para la gestión y operación.

### B. *Objetivos específicos*

- 1) Revisar sistemáticamente trabajos, artículos y otras fuentes de información, bases de datos históricas, variables de entrada, y métricas de evaluación a la hora de monitorear activos eléctricos, al igual que sus variables correspondientes.
- 2) Crear un modelo de aprendizaje de máquina para la predicción en el consumo energético y todas las variables relacionadas con los activos eléctricos que serán monitoreados a partir de la plataforma.
- 3) Crear un modelo de aprendizaje máquina para la predicción de modelos y el estado de salud de un activo eléctrico. (Limpieza de datos, visualización y análisis de las variables de entrada). De los métodos más prometedores identificados en la revisión bibliográfica, se implementará al menos un método de aprendizaje máquina.
- 4) Validar los métodos implementados a partir de las métricas utilizadas en problemas de monitoreo y pronóstico de activos eléctricos, para esto se utilizarán bases de datos públicas o datos de subestaciones y transformadores que estén disponibles.

### III. MARCO TEÓRICO

#### 1) *Transformador*

Un transformador es una máquina eléctrica que, basándose en los principios de inducción electromagnética, transfiere energía de un circuito eléctrico a otro, sin cambiar la frecuencia. La transferencia se lleva a cabo con el cambio de voltaje y corriente. Un transformador aumenta o disminuye la corriente alterna cuando es necesario [2].

Estas máquinas ayudan a mejorar la seguridad y eficiencia de los sistemas de energía durante su distribución y regulación a través de largas distancias.

Los tres componentes más importantes de un transformador son el núcleo magnético, el devanado principal y el secundario.

#### 2) *Índice de salud*

El índice de salud es un indicador para evaluar la probabilidad de falla del activo, obtenido a partir de los resultados de las pruebas que se realizan al mismo y que dan información sobre algún aspecto de su condición, que al ser cuantitativo permite establecer un ranking de los activos en función de su condición técnica [3].

#### 3) *Análisis de gases disueltos*

El análisis del diagnóstico de gases disueltos (DGA) es un método muy popular, y probablemente el más importante, para diagnosticar e interpretar los diferentes tipos de averías en transformadores. Para minimizar las interpretaciones y emitir un diagnóstico certero, actualmente se cuenta con diferentes técnicas de interpretación basadas en bases de datos de fallos históricos [4].

#### 4) *Redes Neuronales*

Las redes neuronales son más que otra forma de emular ciertas características propias de los humanos, como la capacidad de memorizar y de asociar hechos. Si se examinan con atención aquellos problemas que no pueden expresarse a través de un algoritmo, se observará que todos ellos tienen una característica en común: la experiencia. El hombre es capaz de resolver estas situaciones acudiendo a la experiencia acumulada. Así, parece claro que una forma de aproximarse al problema consista en la construcción de sistemas que sean capaces de reproducir esta característica humana. En definitiva, las redes neuronales no son más que un modelo artificial y simplificado del cerebro humano, que es el ejemplo más perfecto del que disponemos para un sistema que es capaz de adquirir conocimiento a través de la experiencia. Las redes neuronales plantean un nuevo sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso

humano: la neurona. Todos los procesos del cuerpo humano se relacionan en alguna u otra forma con la inactividad o actividad de estas neuronas. Las mismas son un componente relativamente simple del ser humano, pero cuando millares de ellas se conectan en forma conjunta se hacen muy poderosas [5].

#### 5) *Redes Neuronales Artificiales.*

Una Red Neuronal Artificial (Artificial Neural Network, RNA) se diseña sobre la base de un mecanismo de trabajo del sistema nervioso humano<sup>6</sup>. Dicho sistema aprende a implementar una tarea considerando solo ejemplos, sin estar programado con ninguna regla de tarea específica. La red neuronal artificial se basa en una colección de nodos o unidades llamadas neuronas. Las neuronas son los componentes básicos de una red neuronal a través de la cual tiene lugar la comunicación [5].

#### 6) *Redes Neuronales Profundas*

Las Redes Neuronales Profundas (Deep neural Networks, RNP) pertenecen a la familia de las Redes Neuronales Artificiales. Consisten en múltiples capas ocultas entre las capas de entrada y salida [6], [7]. La RNP procesa la entrada con manipulación matemática para producir la salida, independientemente de si la relación de datos es lineal o no lineal. La red neuronal se entrena utilizando un conjunto de datos de entrenamiento, lo que da como resultado el cálculo de la probabilidad de cada salida. Una RNP compleja contiene más capas que otras redes neuronales.

#### 7) *Librería PLOTLY*

Plotly es una librería disponible en varios lenguajes de programación, especialmente en Python (plotly.py), en Java (plotly.js), R, Julia, Matlab, etc. Esta librería permite la creación de gráficos complejos y variados. Plotly es muy utilizada en el lenguaje Python, ya que este es el lenguaje más popular en Machine Learning, y además es uno de los más fáciles de aprender.

Plotly es una librería colaborativa e interactiva. Gracias a Chart studio, ofrece un servicio web y un modo colaborativo, que permite descargar y guardar gráficos de datos en una cuenta personal. Es la interacción con los diferentes servidores de Plotly la que permite la creación de este servicio web y el modo colaborativo. Sin embargo, se puede activar un modo, importando el paquete offline que impide difundir gráficos en la nube y hace que solo se pueda acceder a ellos en local. De este modo, se interrumpe la conexión entre los diferentes servidores de Plotly y se detiene el modo colaborativo.

Una de las grandes ventajas de Plotly es su capacidad para detectar valores atípicos en una base de datos, incluso si esta es muy grande. En efecto, con Plotly se puede beneficiar determinada información de los datos simplemente pasando el ratón sobre el gráfico. Así, mediante la lectura del gráfico, se pueden identificar y recuperar los valores atípicos [8].

## 8) Librería PANDAS

Pandas es una muy popular librería de código abierto dentro de los desarrolladores de Python, y sobre todo dentro del ámbito de Data Science y Machine Learning, ya que ofrece unas estructuras muy poderosas y flexibles que facilitan la manipulación y tratamiento de datos. Pandas surgió como una necesidad para unificar en una única librería todo lo necesario para que un analista de datos pudiese tener en una misma herramienta todas las funcionalidades que necesitaba en su día a día, como son: cargar datos, modelar, analizar, manipular y prepararlos [9].

Pandas está basada en dos bibliotecas principales de Python: matplotlib para visualización de datos y NumPy para operaciones matemáticas. Pandas actúa como un contenedor sobre estas bibliotecas, permitiéndole acceder a muchos de los métodos de matplotlib y NumPy con menos código. Por ejemplo, `pandas.plot()` combina varios métodos de matplotlib en un solo método, lo que le permite trazar un gráfico en unas pocas líneas de código [10].

Antes de que existiera pandas, la mayoría de los analistas usaban Python para la recopilación y preparación de datos, y luego cambiaban a un lenguaje más específico de dominio como R para el resto de su flujo de trabajo. Pandas introdujo dos nuevos tipos de estructuras para almacenar datos que facilitan las tareas analíticas y eliminan la necesidad de cambiar de herramienta [10].

Las dos estructuras principales que se pueden encontrar dentro del paquete de Pandas son las siguientes:

- Series: array unidimensional etiquetado capaz de almacenar cualquier tipo de dato.
- DataFrame: estructura bidimensional con columnas que pueden ser también de cualquier tipo. Estas columnas son a su vez Series.

Series			Series			DataFrame		
	apples			oranges		apples	oranges	
0	3		0	0	=	0	3	0
1	2	+	1	3		1	2	3
2	0		2	7		2	0	7
3	1		3	2		3	1	2

Fig. 1. Estructuras de datos Series y DataFrame.

### 9) *Librería Matplotlib*

Matplotlib es una biblioteca de python que se utiliza para crear gráficos y diagramas 2D mediante scripts de python. Tiene un módulo llamado pyplot que facilita el trazado al proporcionar una función para controlar los estilos de línea, las propiedades de la fuente, los ejes de formato, etc. Se utiliza junto con NumPy para proporcionar un entorno que es una alternativa eficaz de código abierto para MatLab. También se puede usar con juegos de herramientas de gráficos como PyQt y wxPython [11]. También se pueden usar las API (interfaces de programación de aplicaciones) de matplotlib para incrustar gráficos en aplicaciones GUI. Un script Python matplotlib está estructurado de modo que unas pocas líneas de código son todo lo que se requiere en la mayoría de los casos para generar un gráfico de datos visuales.

### 10) *Librería Sklearn*

La librería scikit-learn, también llamada sklearn, es un conjunto de rutinas escritas en Python para hacer análisis predictivo, que incluyen clasificadores, algoritmos de clusterización, etc. Está basada en NumPy, SciPy y matplotlib, de forma que es fácil reaprovechar el código que use estas librerías [12].

Además, incorpora varias funciones para procesar los datos:

- Normalización: ajustar las variables numéricas para que tengan media 0 y varianza 1, o bien que estén en un rango como por ejemplo [0,1]. También permite normalizar vectores para que tengan norma 1.
- Transformaciones no lineales: basadas en cuantiles y exponentes, para transformar variables con distribuciones muy sesgadas, por ejemplo. Entre otros incluye la transformada de Yeo-Johnson y la de Box-Cox.
- Discretización: se trata de convertir una variable numérica en un conjunto de valores posibles, según algún criterio. Un caso extremo es cuando una variable es convertida a solo dos valores posibles, lo que se conoce como binarización.
- Valores perdidos: cuando por algunos registros faltan datos de alguna variable (p. ej., un usuario no responde a alguna pregunta de una encuesta), es posible imputar un valor en función de algún criterio automatizable, por ejemplo, sustituirlo por la mediana.
- Creación de interacciones entre variables mediante el uso de polinomios.

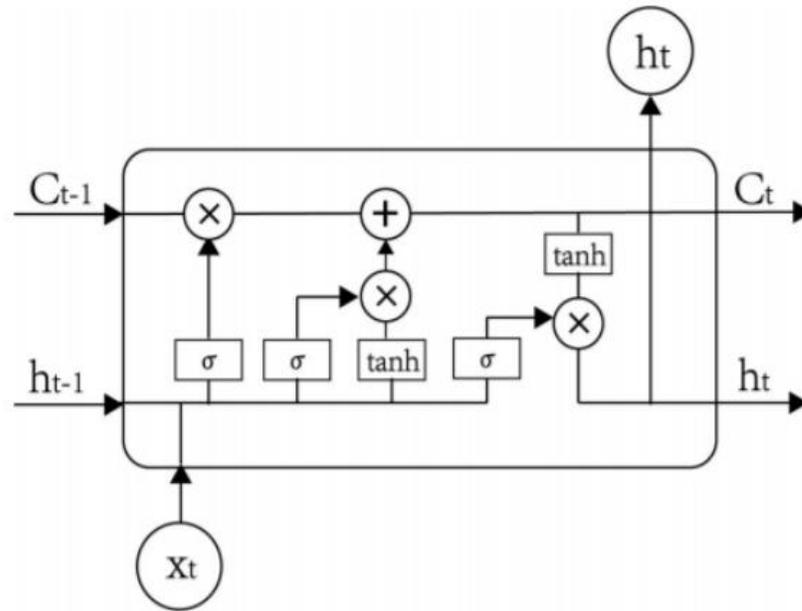


Fig. 2. Estructura de una red LSTM.

Para poder realizar la plataforma de monitoreo se cuenta con varias etapas en la elaboración, la primera es probar los modelos y códigos diseñados mediante el software de uso libre llamado Python, todos estos códigos estarán disponibles en notebooks de Google Colab. También cabe mencionar que la plataforma de monitoreo contará con dos funciones principales: monitoreo para transformadores que se encargará de evaluar su índice de salud, el estado de formación de gases, etc. Estas funciones se lograrán gracias a funciones programadas para hallar el índice de salud de transformadores, triángulos de Duval y árboles de decisión, todo esto desarrollado en Python. Otra función importante que será la de pronosticar mediante series de tiempo el consumo eléctrico de un centro comercial al igual que el ingreso del número de personas a este. Se abordará primero el tema del monitoreo de transformadores, por lo tanto, procedemos a dar una breve descripción de la generación de gases en transformadores inmersos en aceite basado en la normativa internacional IEEE STD C57.104 TM-2008 [3].

### *11. Gases generados en transformadores inmersos en aceite basado en la normativa internacional IEE STD C57.104TM-2008.*

La detección de ciertos gases generados en un transformador lleno de aceite en servicio es, con frecuencia, la primera indicación disponible de un mal funcionamiento que, si no se corrige, puede conducir a la falla. Arcos, descargas parciales, chispas de baja energía, sobrecarga severa, falla del motor de la bomba y sobrecalentamiento en el sistema de aislamiento son algunos de los posibles mecanismos. Estas condiciones que ocurren individualmente, o como varios eventos simultáneos, pueden resultar en la descomposición de los materiales aislantes y la formación de varios gases combustibles y no combustibles. El funcionamiento normal también dará como resultado la formación de algunos gases. De hecho, es posible que algunos transformadores

operen a lo largo de su vida útil con cantidades sustanciales de gases combustibles presentes. Operar un transformador con grandes cantidades de gas combustible presente no es algo normal, pero sucede, generalmente después de cierto grado de investigación y una evaluación del posible riesgo [3].

En un transformador, los gases generados se pueden encontrar disueltos en el aceite aislante, en la capa de gas sobre el aceite o en los dispositivos de recolección de gas. La detección de una condición anormal requiere una evaluación de la cantidad de gas generado presente y la tasa continua de generación. Se puede obtener alguna indicación de la fuente de los gases y el tipo de aislamiento involucrado determinando la composición de los gases generados [3].

### *12. Evaluación de la condición del transformador usando concentraciones individuales y TDGC (Gas combustible total disuelto)*

Puede ser difícil determinar si un transformador se está comportando normalmente si no tiene antecedentes de gas disuelto. Asimismo, existen considerables diferencias de opinión sobre lo que se considera un “transformador normal” con concentraciones aceptables de gases [3].

Se ha desarrollado un criterio de cuatro niveles para clasificar los riesgos de los transformadores, cuando no hay un historial previo de gas disuelto, para la operación continua a varios niveles de gas combustible. El criterio utiliza tanto las concentraciones para gases separados como la concentración total de todos los gases combustibles. Consulte la Tabla 1 [3].

- Condición 1: TDCG por debajo de este nivel indica que el transformador está funcionando satisfactoriamente. Cualquier gas combustible individual que exceda los niveles especificados debe dar lugar a una investigación adicional.
- Condición 2: TDCG dentro de este rango indica un nivel de gas combustible superior al normal. Cualquier gas combustible individual que exceda los niveles especificados debe dar lugar a una investigación adicional.
- Condición 3: TDCG dentro de este rango indica un alto nivel de descomposición. Cualquier gas combustible individual que exceda los niveles especificados debe dar lugar a una investigación adicional.
- Condición 4: TDCG que excede este valor indica una descomposición excesiva. La operación continua podría resultar en la falla del transformador. Proceda de inmediato y con precaución.

TABLA I  
CONCENTRACIÓN DE GASES DISUELTOS

Límites concentración de gases disueltos [ $\mu\text{L}/\text{L}(\text{ppm})^a$ ]								
Estado	$H_2$	$CH_4$	$C_2H_2$	$C_2H_4$	$C_2H_6$	$CO$	$CO_2$	$TDCG^b$
<b>Condición 1</b>	100	120	1	50	65	350	2500	720
<b>Condición 2</b>	101-700	121-400	2-9	51-100	66-100	351-570	2500-4000	721-1920
<b>Condición 3</b>	701-1800	401-1000	10-35	101-200	101-150	571-1400	40001-10000	1921-4630
<b>Condición 4</b>	>1800	>1000	>35	>200	>150	>1400	>10000	>4630

Los números que se muestran en la Tabla 1 están en partes de gas por millón de partes de aceite [ $\mu\text{L}/\text{L}(\text{ppm})$ ] volumétricamente y se basan en un gran transformador de potencia con varios miles de galones de aceite. Con un volumen de aceite más pequeño, el mismo volumen de gas dará una mayor concentración de gas. Los pequeños transformadores de distribución y los reguladores de voltaje pueden contener gases combustibles debido a la operación de los fusibles de expulsión internos o los interruptores de ruptura de carga [3].

La Tabla 1 enumera las concentraciones de gas disuelto para los gases individuales y TDCG para la Condición 1 a la Condición 4. Esta tabla se usa para realizar la evaluación original de una condición de gasificación en un transformador nuevo o recientemente reparado o se usa si no hay pruebas previas en el transformador para gases disueltos o si no hay antecedentes recientes [3].

### 13. Triángulo de Duval.

La técnica del triángulo de Duval se utiliza para interpretar y diagnosticar tipos de averías en transformadores de una forma gráfica más eficiente y precisa que otras técnicas utilizadas para la interpretación en la formación de gases de transformadores. Algunos de los objetivos en la cromatografía de gases que se realiza con el triángulo de Duval son [13]:

- Indicar una condición de falla incipiente.
- Determinar los materiales involucrados
- Determinar si la condición es severa y si hay cantidades anormales.
- Detectar una amplia variedad de condiciones.

- Determinar la velocidad de generación de gases.

El análisis de la muestra en el laboratorio obtiene la concentración en ppm de los siguientes gases:

- Hidrógeno(H<sub>2</sub>)
- Metano (CH<sub>4</sub>)
- Etano(C<sub>2</sub>H<sub>6</sub>)
- Etileno (C<sub>2</sub>H<sub>4</sub>)
- Acetileno (C<sub>2</sub>H<sub>2</sub>)
- Monóxido de Carbono (CO)
- Dióxido de Carbono (CO<sub>2</sub>)
- Nitrógeno (N<sub>2</sub>)
- Oxígeno (O<sub>2</sub>)

### 13.1 Comparación métodos de diagnósticos de fallas.

TABLA III [3].  
MÉTODOS DE DIAGNÓSTICO DE FALLAS DE GASES EN TRANSFORMADORES.

	%Diagnósticos Correctos	%Diagnósticos resolver	sin %Diagnósticos errados
<b>IEE gas Característico</b>	42	0	58
<b>IEE Rogers</b>	62	33	5
<b>Doernenburg</b>	71	26	3
<b>IEC Gases Básicos</b>	77	15	8
<b>IEC Triángulo de Duval</b>	96	0	4

### 13.2 Método del triángulo.

La ubicación del punto de falla se realiza por el porcentaje relativo de los tres gases en cada lado del triángulo, de 0 a 100%. Esto nos permite identificar las 6 principales fallas, más las DT (Mezcla de fallas térmicas y eléctricas). Las zonas de las fallas en el triángulo están basadas en un largo número de casos (+200) de fallas en transformadores en servicio los cuales fueron inspeccionados visualmente [14] [15].

Fórmula y esquema Triángulo de Duval # 1

$$\%C_2H_2 = 100X (X + Y + Z) \quad (1)$$

$$\%C_2H_4 = 100Y (X + Y + Z) \quad (2)$$

$$\%CH_4 = 100Z(X + Y + Z) \quad (3)$$

Siendo:  $X = C_2H_2$ ,  $Y = C_2H_4$  y  $Z = CH_4$ .

En donde:

PD : Descargas parciales.

D1 : Descargas de alta energía.

D2: Descargas de baja energía.

DT: Indeterminada, falla térmica o eléctrica.

T1: Falla térmica  $T < 300^\circ C$ .

T2: Falla térmica  $300^\circ C < T < 700^\circ C$

T3: Falla térmica  $T > 700^\circ C$

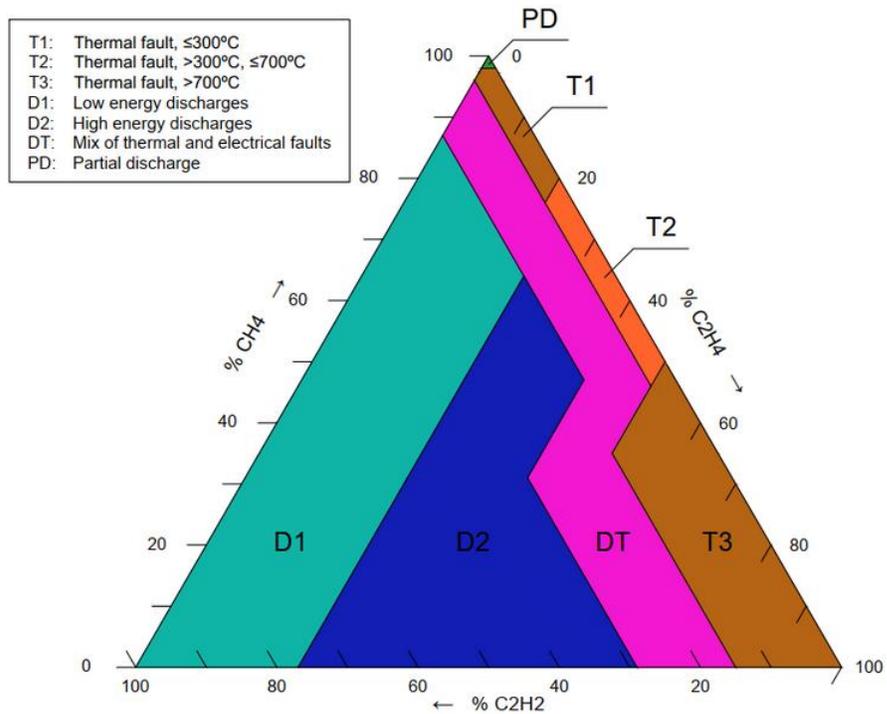


Fig. 3. Método del triángulo de Duval.

---

## IV. METODOLOGÍA

La metodología que se siguió para cumplir los objetivos propuestos son los siguientes:

*A. Actividad 1 Revisión Bibliográfica.*

De los artículos revisados, identificar los métodos y técnicas más utilizadas para el pronóstico de la vida útil de transformadores y sus matrices de riesgo en subestaciones. Al mismo tiempo, aplicar estas técnicas en el consumo energético en los centros comerciales.

*B. Actividad 2. Búsqueda de bases de datos.*

Realizar la búsqueda de bases de datos de monitoreo y predicción de activos eléctricos y variables relacionadas para el pronóstico de la vida útil y su índice de salud.

*C. Actividad 3. Preprocesamiento de datos.*

Una vez escogida la base de datos, se procesarán y eliminarán datos atípicos, ajustar la información a las entradas de los modelos. Dicho proceso se hará mediante el lenguaje de programación Python y librerías asociadas a la analítica de datos.

*D. Actividad 4. Implementación modelo de ML.*

Implementar un método de pronóstico y monitoreo, ya sea de matrices de riesgo, vida útil del transformador o su estado de salud utilizando herramientas de Machine Learning (ML).

*E. Actividad 5. Análisis de resultados.*

Analizar los resultados obtenidos evaluando el comportamiento de los diferentes modelos implementados de ML para el pronóstico y monitoreo. Para esto se utilizarán las métricas de evaluación más utilizadas, para este tipo de modelos, en la bibliografía especializada.

*F. Actividad 6. Informe final.*

Realizar un informe donde se documente el proceso de obtención y limpieza de los datos, implementación de los modelos y su evaluación. Además, se presentará un análisis de los resultados y las conclusiones pertinentes.

## V. RESULTADOS Y ANÁLISIS

En esta sección se presentan los resultados más relevantes de la práctica realizada por el estudiante. Se muestran resultados prácticos y sus correspondientes análisis. Se aplicaron varios modelos de Machine Learning y un Dashboard con actualizaciones de cada 30 segundos.

### A.1 Ejemplo de aplicación triángulo de Duval.

Si los resultados del DGA son:

Metano (CH<sub>4</sub>) = 100 ppm

Etileno (C<sub>2</sub>H<sub>4</sub>) = 100 ppm

Acetileno (C<sub>2</sub>H<sub>2</sub>) = 100 ppm.

Primer cálculo:

$$CH_4 + C_2H_4 + C_2H_2 = 300ppm \quad (4)$$

Se calcula el porcentaje relativo de cada uno de los gases:

$$\%Relativo\ de\ CH_4 = \frac{100}{300} = 33.33\% \quad (5)$$

$$\%Relativo\ de\ C_2H_4 = \frac{100}{300} = 33.33\% \quad (6)$$

$$\%Relativo\ de\ C_2H_2 = \frac{100}{300} = 33.33\% \quad (7)$$

Estos valores son las coordenadas del triángulo. Se graficará el ejercicio en Python y una macro en Excel para confirmar los resultados.

Se ingresan los resultados en el script de Python, con ingresar los valores de los gases C<sub>2</sub>H<sub>4</sub> y CH<sub>4</sub> es suficiente puesto que se trabaja con coordenadas triangulares.

```
# El orden de los gases es el siguiente C2H4 , CH4, C2H2
sample_point = np.array([33.3, 33.3, 1]) @ np.transpose(A)
print(sample_point)
```

Fig. 4. Ejemplo de aplicación Triángulo de Duval.

Resultado:

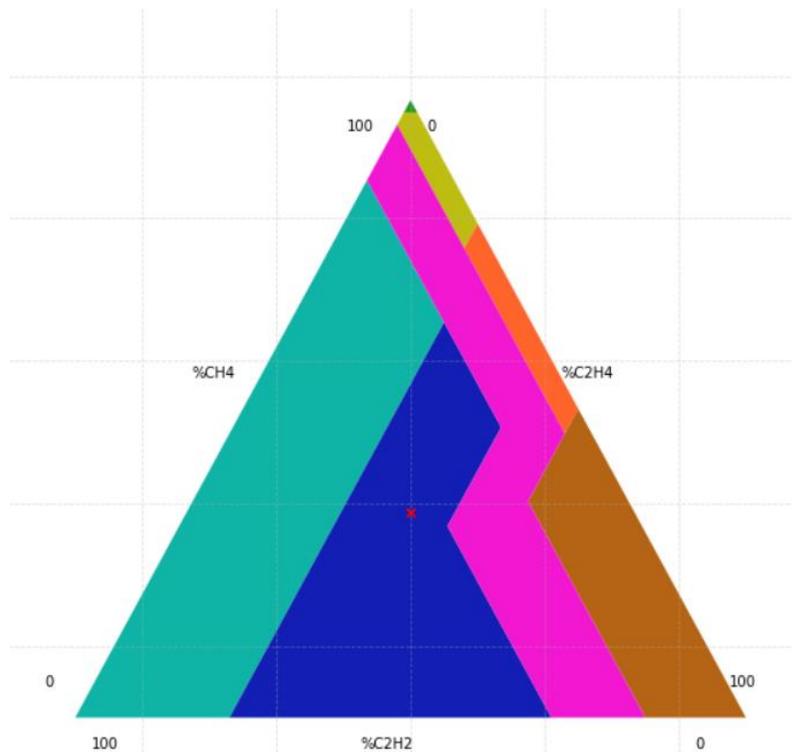


Fig. 5. Resultado práctico Triángulo de Duval versión 1.

### A.2 Tipos de fallas [13],[14],[15].

Para el triángulo 1 hay seis tipos de fallas básicas, las cuales son :

#### *Descarga Parcial de Tipo corona (PD)*

Ejemplo: descargas en burbujas de gas o papel atrapado. Resultado de un secado pobre.

#### *Descarga de baja energía (D1)*

Ejemplo: descargas parciales de tipo chispa. Arco de baja energía, inducido en la superficie rota del papel o partículas de carbón.

#### *Descarga de alta energía (D2)*

Ejemplo: flameo o salto de corriente. Cortocircuito. Esto resulta en: daño extenso del papel, gran formación de partículas de carbón en el aceite, fusión del metal, excitación del equipo o del gas.

#### *Falla térmica de temperatura < 300°C (T1)*

Ejemplo: sobrecarga, bloqueo de aceites en los ductos, insuficiencia en el enfriado, evidencia por papel enrollado (café(>200°C), negro o carbonizado (>300°C))

*Falla térmica de temperatura entre 300° C y 700° C (T2)*

Ejemplo: contacto defectuoso, Corrientes circulantes. Evidenciada por : carbonización del papel, formación de partículas de carbón.

*Falla térmica >700° C (T3)*

Ejemplo: gran circulación de corriente en el tanque y en el núcleo, cortocircuito en la lámina. Evidenciada por: extensa formación de partículas de carbón, coloración del metal (800°C) o metal fusión (>1000°C).

*B. Algoritmos de Machine Learning utilizando series de tiempo.*

Una serie de tiempo es un conjunto de datos obtenidos en puntos sucesivos en el tiempo o durante períodos sucesivos. Una secuencia de datos mensuales sobre la construcción de nuevas viviendas y una secuencia de datos semanales sobre la venta de productos son ejemplos de series de tiempo. Normalmente, los datos de una serie de tiempo se obtienen en períodos de tiempo igualmente espaciados, como por horas, días, semanas, meses o años [16].

Una de las principales preocupaciones del análisis de series de tiempo es el desarrollo de predicciones para los valores futuros de la misma. Por ejemplo, se pueden realizar series de tiempo para predecir el consumo energético de un centro comercial en base al número de personas que ingresan al mismo. La mayoría de las empresas elaboran previsiones sobre las ventas de sus productos [16].

El pronóstico de la demanda de energía eléctrica o por sus siglas en inglés llamada (EPLF) que significa Electric Power Load Forecasting se clasifica en términos de los horizontes de planeación, por ejemplo: para un horizonte de menos de un día se considera un periodo a muy corto plazo, una semana o un día de pronóstico se considera un horizonte de corto plazo, de una semana a un año mediano plazo y más de un año es un largo plazo o periodo [17].

Hablando a nivel general, una serie de tiempo puede descomponerse en tres componentes: la tendencia, su ciclo, estacionalidad y los componentes aleatorios. Aunque existen muchos métodos de pronóstico, ninguno puede generalizarse para que funcione lo suficiente bien para todos los casos, especialmente cuando se consideran muchos factores. Por lo tanto, para obtener un pronóstico adecuado, no se trata solo de adoptar un método famoso. En otras palabras, un método ideal para un caso puede funcionar mal para otro [18].

Los métodos analíticos para el pronóstico de la demanda pueden ser clasificados como métodos analíticos y no analíticos

Los métodos analíticos se subdividen en otras subclases las cuales son:

### *B.1 Enfoque estadístico o clásico.*

Son los métodos ARMA, ARIMA, regresiones lineales, etc. Aunque son los métodos más conocidos para series de tiempo y tienen una gran acogida, estos ya no son tan válidos porque tienen algunas limitaciones debido a las características no lineales de los parámetros numéricos en la predicción del clima. La aplicación de algoritmos de inteligencia artificial en donde están relacionados variables ambientales y que tienen que ver con la potencia y el voltaje pueden ser pronosticados con algoritmos de machine learning supervisados como regresiones, support vector machine, y no supervisados como K-means, clustering [18].

### *B.2 Inteligencia artificial.*

Son técnicas relativamente nuevas aplicadas a muchos campos de la ciencia. Redes neuronales y métodos de machine learning como un support vector machine son algunas de las técnicas que se utilizan. Sin embargo, algunos de estos modelos pueden sufrir sobreentrenamiento u overfitting [18].

### *B.3 Enfoques híbridos.*

Se usan combinaciones de inteligencia artificial y algoritmos de optimización como las PSO-ARMA, algoritmos genéticos [18].

Ahora que se conoce el contexto de lo que es una serie de tiempo, el objetivo de la realización de esta en mi práctica era la de predecir el consumo energético del centro comercial Unicentro a lo largo del tiempo y a su vez el ingreso de personas y si este ingreso afectaba en el consumo total del centro comercial, teniendo en cuenta variables como el día, la semana, si era feriado, etc.

Cabe resaltar que esta alternativa de predicción del consumo energético del centro comercial se busca implementar a futuro, por lo que los datos utilizados fueron simulados a excepción del ingreso de personas, que sí eran datos reales pero confidenciales. A continuación, se muestran los resultados más representativos después de analizar los datos:

### *B.4. Análisis de resultados.*

#### *B.4.1 Mapa de calor.*

Los mapas de calor son representaciones gráficas de los datos relacionadas con la matriz de correlación. Su objetivo principal es el de visualizar mejor el volumen de eventos dentro de un conjunto de datos y ayudar a dirigir a la persona que está interpretando el mapa a los datos más relevantes. Como se mencionaba anteriormente, dicho mapa está estrechamente relacionado con la matriz de correlación.

El coeficiente de correlación de Pearson oscila entre -1 y 1. Un valor menor a cero indica que hay una correlación negativa, es decir, que las dos variables asociadas son inversamente

proporcionales. Un valor mayor a cero indica que existe una correlación positiva y que las dos variables asociadas son directamente proporcionales [19].

En nuestro mapa de calor (fig.6), se puede visualizar que a simple vista no hay mucha correlación entre las variables, a excepción de ellas mismas, en el caso del consumo de energía se puede entender fácilmente porque es un valor simulado, pero en el caso del ingreso de personas respecto a los años, se puede decir que por esas épocas estuvo el fenómeno de la pandemia (COVID 19) y explica el porqué de la poca correlación.

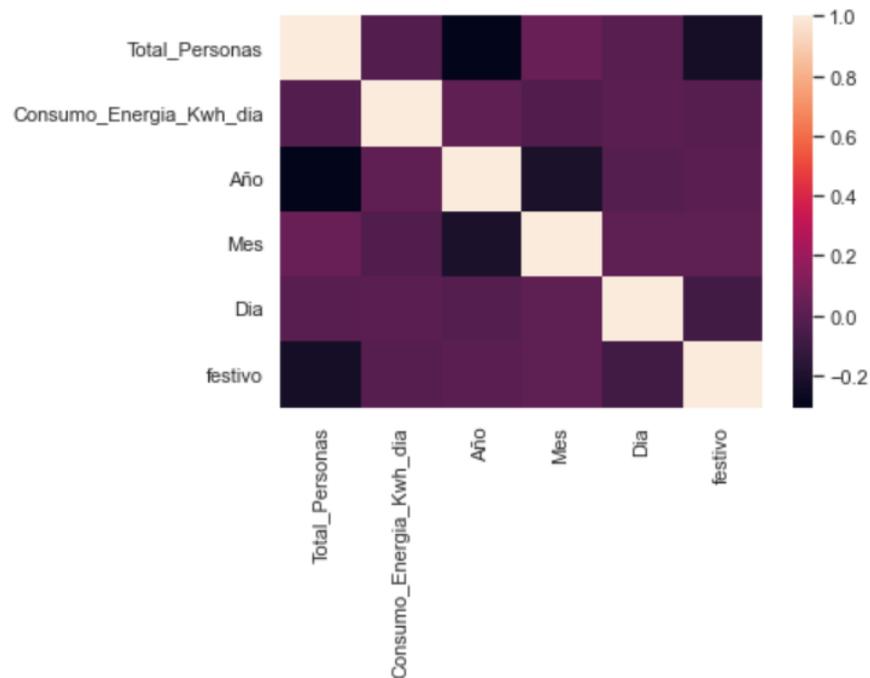


Fig. 6. Mapa de calor variables registradas DataSet.

#### B.4.2 Gráficas series de tiempo consumo de Energía por año vs Ingreso de personas.

Desde la Fig 7 hasta la 9 se puede apreciar el comportamiento del ingreso de personas respecto al consumo energético del centro comercial Unicentro a lo largo de los años, se puede ver que no hay una correlación a simple vista, y esto es explicable, ya que como se mencionó anteriormente, los datos de consumo energético fueron simulados, pero era para poder ver la aplicabilidad del uso del Software Python y a su vez de la librería Fb prophet, a la hora de realizar predicciones. La línea azul representa el consumo de energía del centro comercial respecto al día y la roja el número total de personas en el día respectivamente.

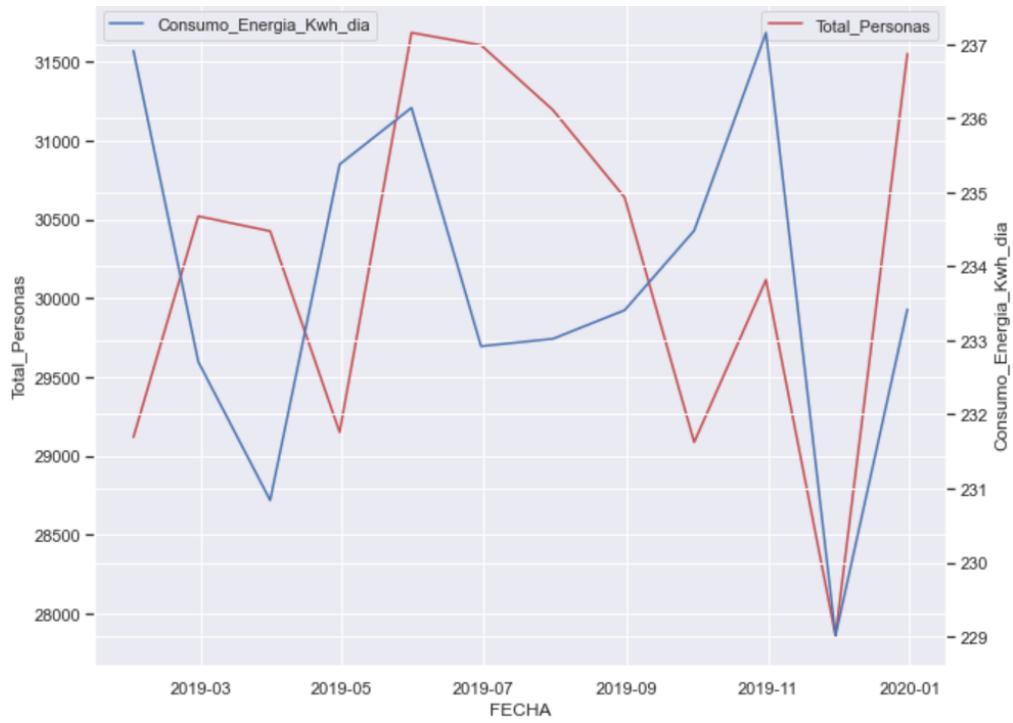


Fig. 7. Ingreso de personas VS consumo de energía año 2019 centro comercial Unicentro.

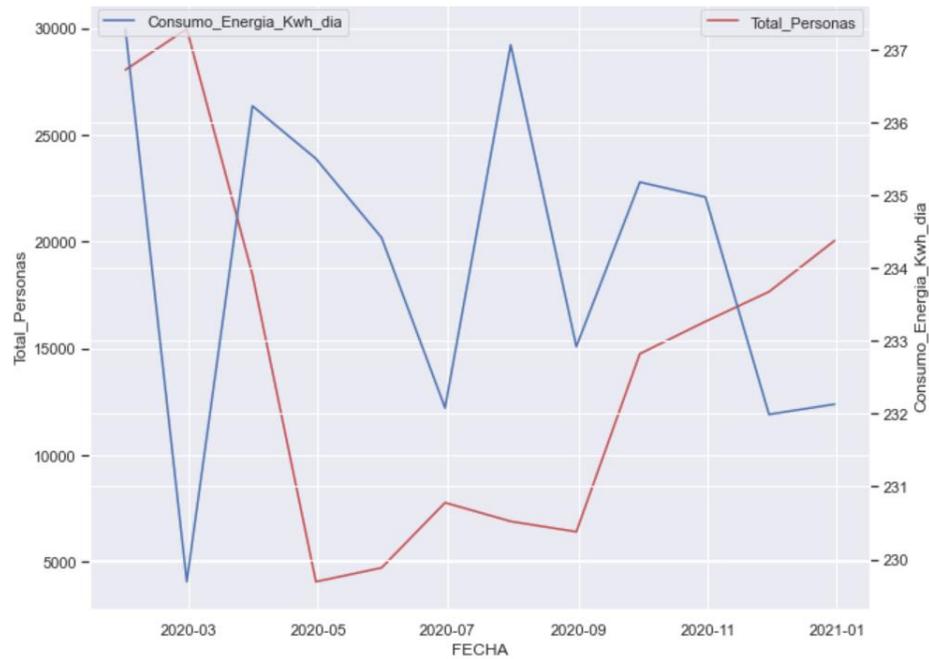


Fig. 8. Ingreso de personas VS consume de energía año 2020 centro comercial Unicentro.

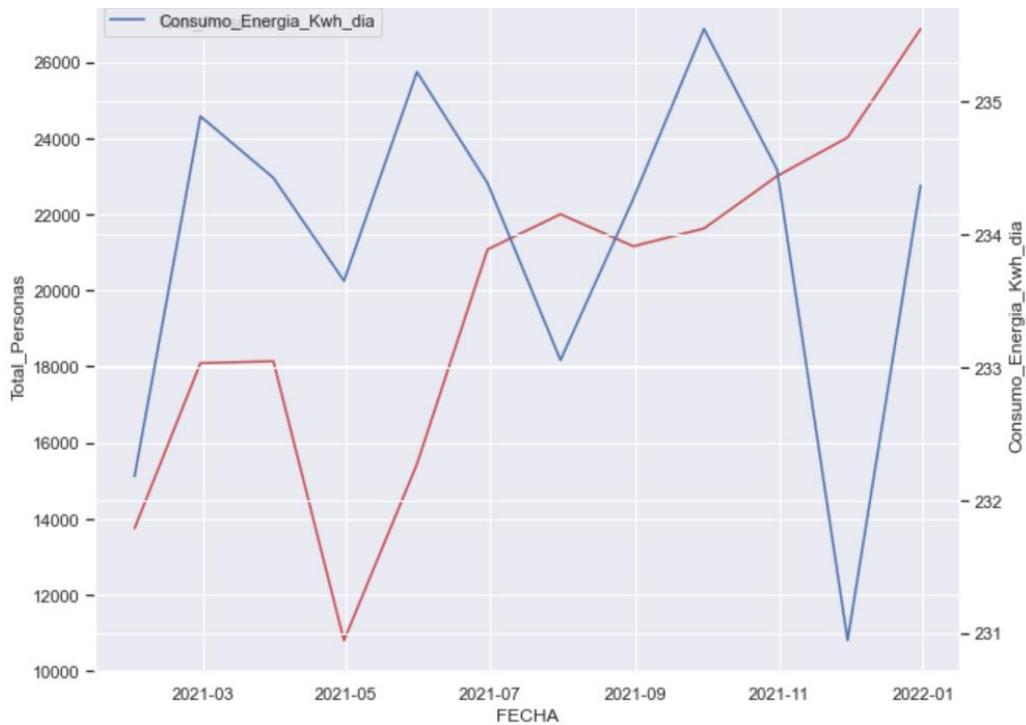


Fig. 9. Ingreso de personas VS consumo de energía año 2021 centro comercial Unicentro.

#### B.4.3 Gráficas predicciones series de tiempo. Predicción consumo de energía.

Desde la Fig 10 hasta la 14 se puede ver el potencial de la librería de inteligencia artificial Fb prophet. En la Fig 12 los puntos negros, significan los datos reales (los simulados) tomados desde el 2019 hasta principios de este 2022 y con los cuales se pretendía predecir el comportamiento del resto del 2022 y principios de 2023 respecto al consumo energético. El área azul clara representa los boundaries (límites) en donde es correcta la predicción y la línea azul oscura es la predicción como tal. Como los datos son simulados, se puede ver que hay mucho ruido y el algoritmo trata de dar su mejor pronóstico, aunque no muy acertado. Desde la Fig 11 hasta la 14 se muestran las tendencias por día, semana, etc y el comportamiento del consumo energético en el centro comercial.

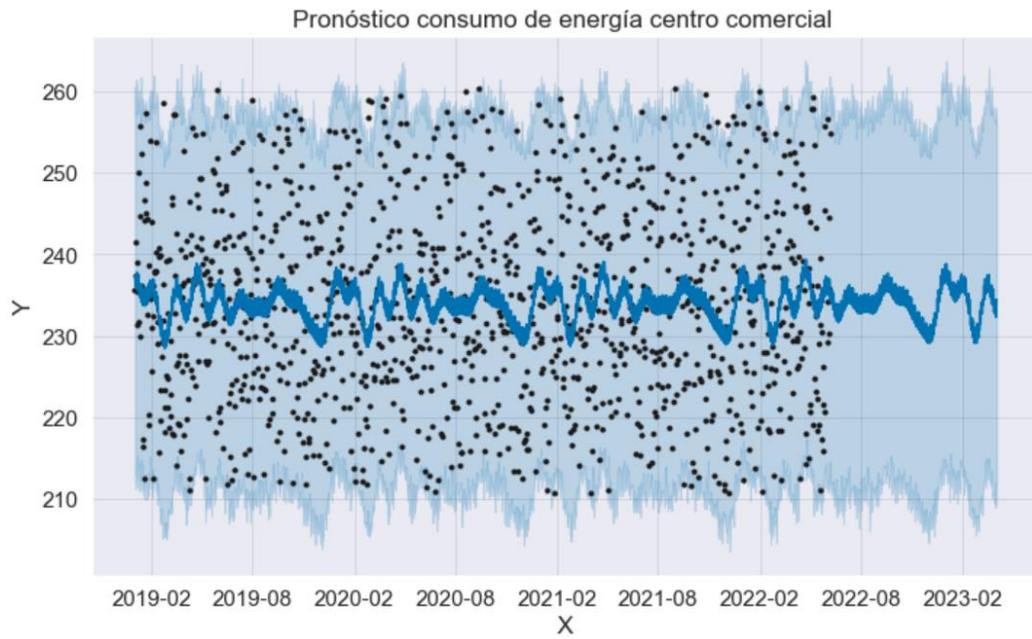


Fig. 10. Predicción consumo de energía año 2022 y principios 2023.

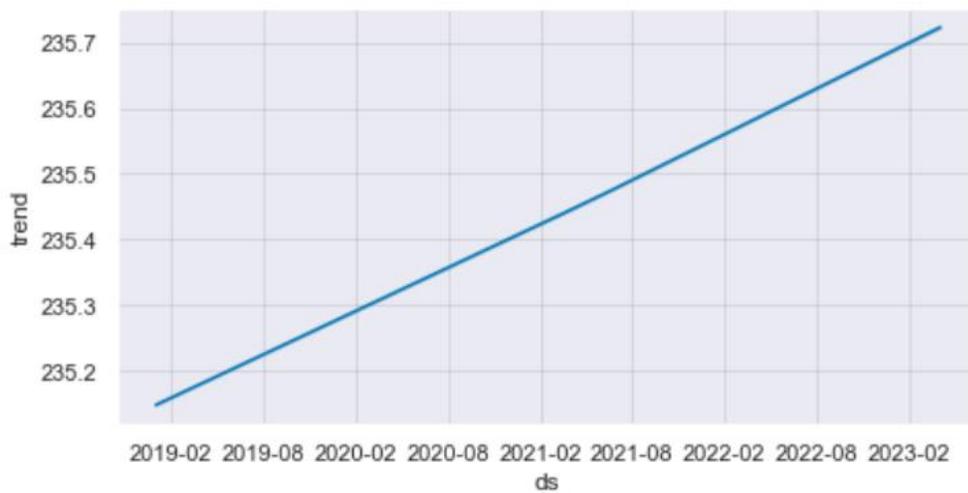


Fig. 11. Predicción tendencia para el consumo de energía finales de 2022 y principios.

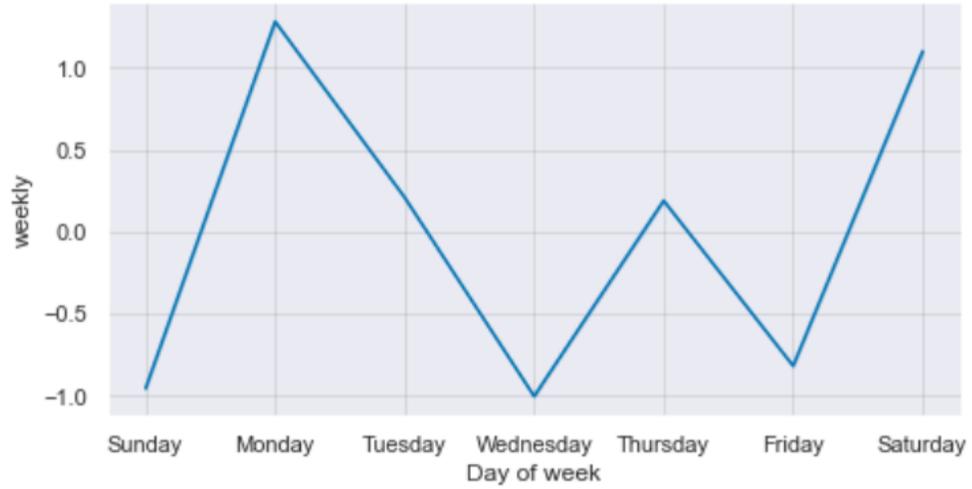


Fig. 12. Predicción tendencia semanal para el consumo de energía a finales de 2022 y principios de 2023.

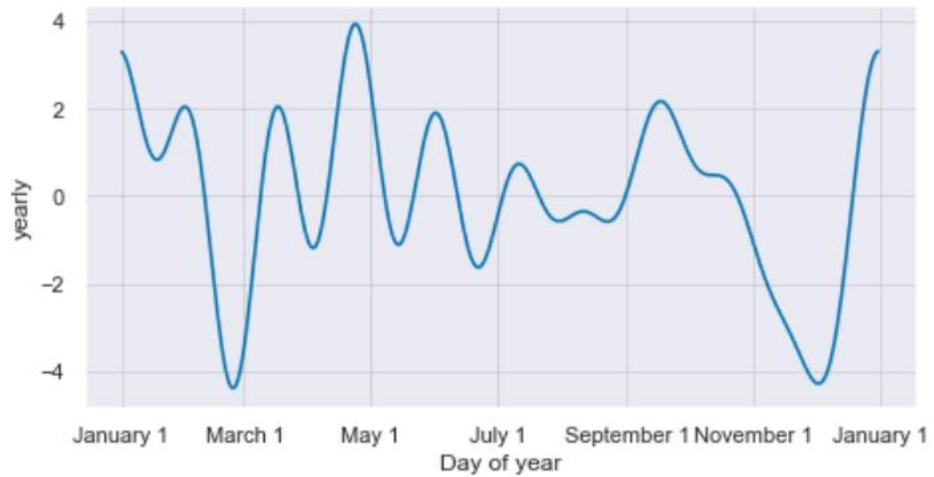


Fig. 13. Predicción tendencia anual para el consumo de energía finales de 2022 y principios 2023.

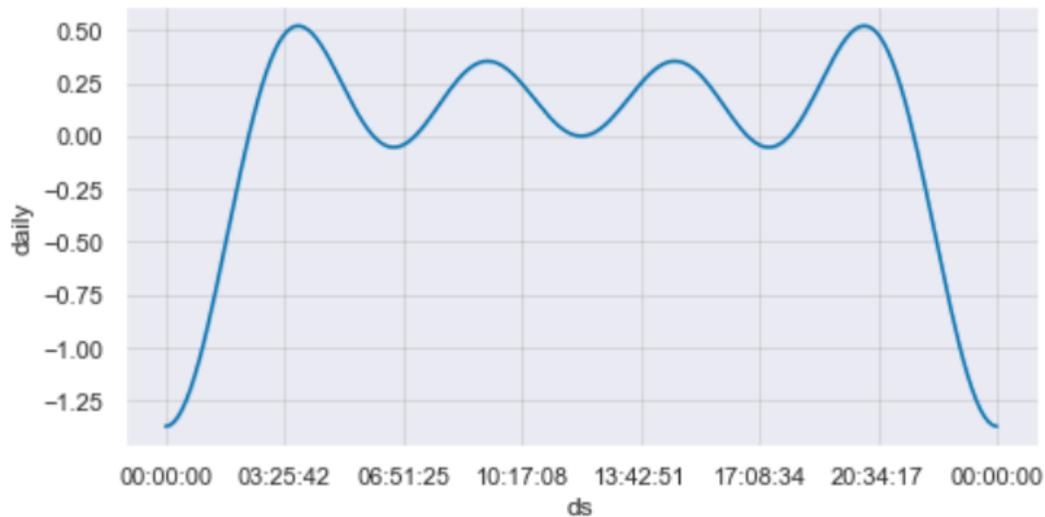


Fig. 14. Predicción tendencia diaria para el consumo de energía finales de 2022 y principios de 2023.

#### B.4.4 Gráficas predicciones series de tiempo. Predicción Ingreso de personas.

Por otro lado, desde la Fig 15 hasta la 19 se puede ver el potencial de la librería de inteligencia artificial Fb prophet. En la imagen 12 los puntos negros, significan los datos reales (los simulados) tomados desde el 2019 hasta principios de este 2022 y con los cuales se pretendía predecir el comportamiento del resto del 2022 y principios de 2023 respecto al ingreso de personas. El área azul clara representa los boundaries (límites) en donde es correcta la predicción y la línea azul oscura es la predicción como tal que el pronóstico del ingreso de personas fue más acertado y correcto, debido a que estos eran datos reales. Se aprecia un incremento en el número de personas que ingresarán al centro comercial, el cual es el comportamiento más esperado después de la pandemia.

Algo peculiar que se ve en la figura 19, es que, debido a la tendencia, se muestra que el mayor ingreso de personas se da a las 12 am, lo cual es algo incoherente. Este fenómeno se puede explicar, ya que al momento de ingresarse los datos en la base de datos se hacía a una única hora, y esta fue la escogida.

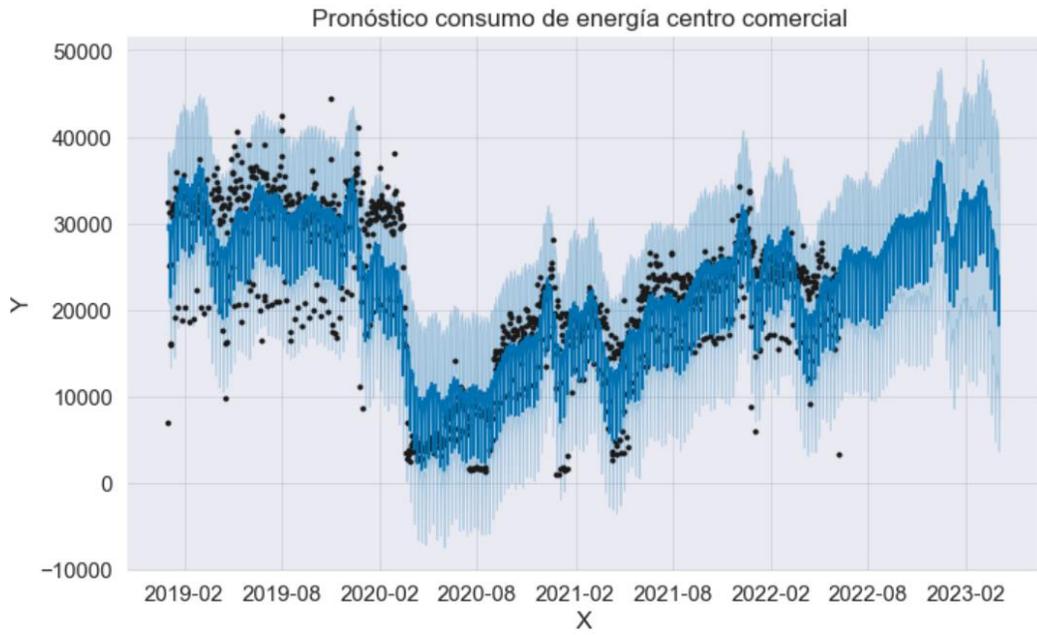


Fig. 15. Predicción Ingreso de personas 2022 y principios 2023.

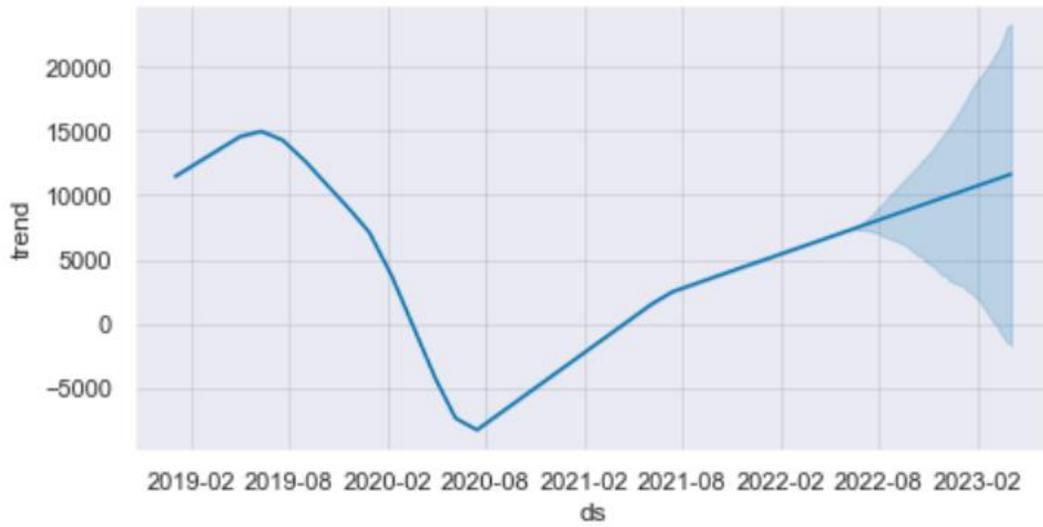


Fig. 16. Predicción tendencia para el ingreso de personas finales de 2022 y principios de 2023.



Fig. 17. Predicción tendencia semanal para el ingreso de personas finales de 2022 y principios de 2023.

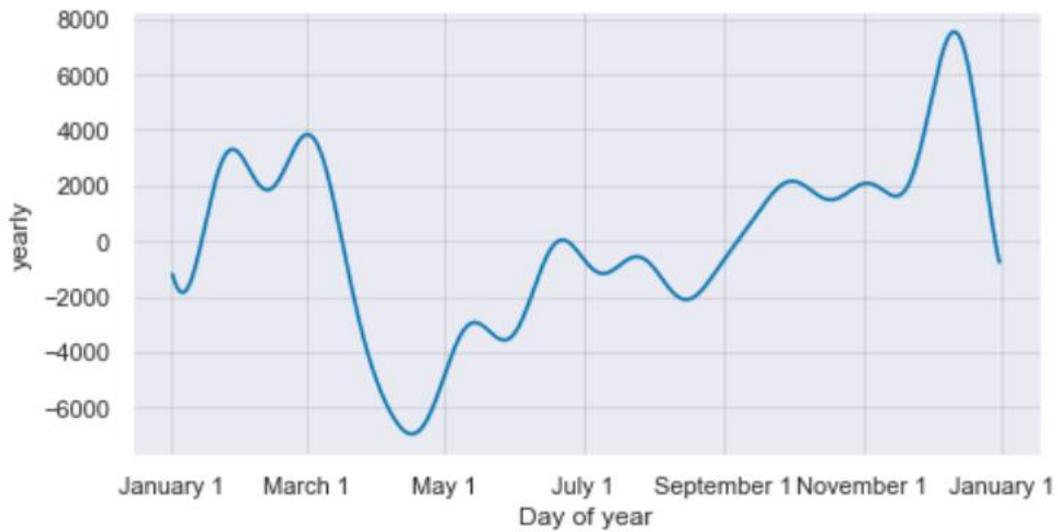


Fig. 18. Predicción tendencia semanal para el ingreso de personas finales de 2022 y principios de 2023.

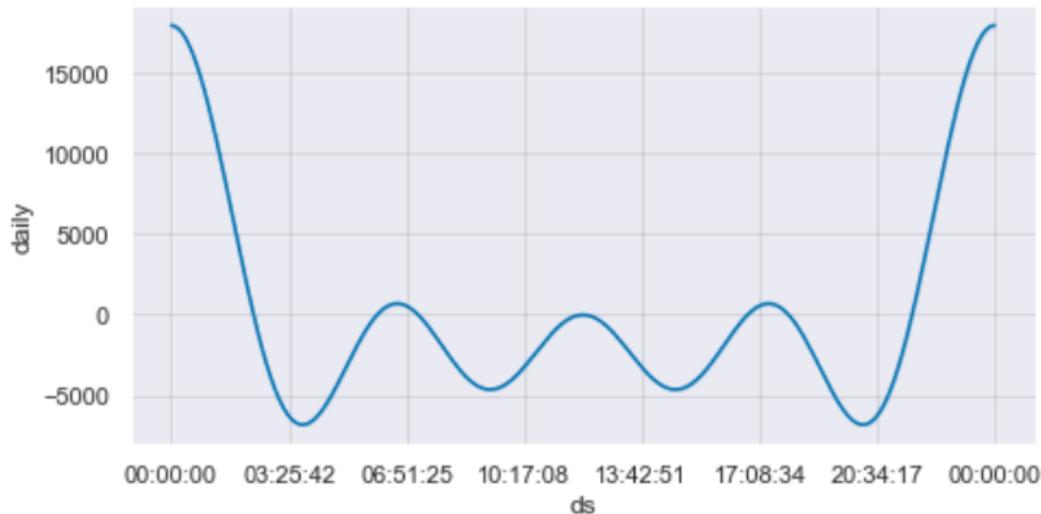


Fig. 19. Predicción tendencia diaria para el ingreso de personas finales de 2022 y principios de 2023.

### *C. Algoritmos de Machine Learning implementados para detectar fallas en transformadores.*

Los transformadores desempeñan un papel muy importante en el sistema eléctrico. Aunque son unos de los componentes más fiables de la red eléctrica, también son propensos a fallar debido a muchos factores tanto internos como externos. Puede haber muchos factores que provoquen el fallo de un transformador, pero los que pueden provocar un fallo catastrófico son los siguientes: fallas mecánicas, dieléctricas y cortocircuitos en el bobinado eléctrico.

En el notebook que quedará como anexo a este documento se explican las variables a utilizar tales como: voltajes de fase, voltajes línea a línea, corrientes de línea, corriente de neutro, temperaturas en el bobinado, nivel de aceite, temperatura del aceite y la más importante que es con la que se realiza el estudio es con un medidor magnético de aceite.

#### *C.1 Medidor magnético de aceite.*

El medidor magnético de aceite es un dispositivo mecánico que se ubica en el tanque conservador de un transformador y se utiliza para medir el nivel de aceite aislante de este mismo.

El funcionamiento de un medidor magnético de aceite se puede explicar de la siguiente manera: cuando el aceite se expande dentro del transformador debido a un aumento de temperatura, el nivel del aceite en el tanque conservador aumentará y de manera contraria cuando

disminuya la temperatura. El indicador (la aguja) sube y baja con el nivel del aceite. El indicador está conectado al engranaje del medidor que, a su vez, está conectado magnéticamente a la esfera de la pantalla que también está conectada con un interruptor. Cuando se desciende por debajo de un nivel establecido, el interruptor se cierra y se activa una alarma. De este modo, se garantiza que el nivel nunca baje demasiado sin previo aviso [20].



Fig. 20. Imagen de medidor magnético de aceite.

Ahora que se ha contextualizado de una manera precisa las variables que están involucradas en el análisis de este problema, se debe definir el objetivo el cual será: la implementación de algunos modelos de machine learning para pronosticar fallas en transformadores a través del medidor magnético de aceite. Se implementarán modelos tales como: árboles de decisión, regresión logística y máquinas de vectores de soporte [20].

Estos métodos fueron utilizados como modelos de clasificación que, a partir de varias variables como el voltaje, corriente, temperatura del aceite, voltajes de línea, de neutro etc. Se encargaba de relacionar, memorizar y analizar dichas variables para tratar de predecir cuándo se activaría la alarma del medidor magnético de aceite.

### *C.2 Árboles de decisión.*

Los árboles de decisión son técnicas de Machine en donde se construyen modelos predictivos de analítica de datos para el Big Data basados en su clasificación según ciertas

características o propiedades, o en la regresión mediante la relación entre distintas variables para predecir el valor de otra [21].

En los modelos de clasificación se predice el valor de una variable mediante la clasificación de la información en función de otras variables (tipo, pertenencia a un grupo...). Los valores a predecir son predefinidos, es decir, los resultados están definidos en un conjunto de posibles valores. El árbol de decisión es una estructura que está formada por ramas y nodos de distintos tipos:

Los nodos internos representan cada una de las características o propiedades a considerar para tomar una decisión.

Las ramas representan la decisión en función de una determinada condición (p. ej. probabilidad de ocurrencia).

Los nodos finales representan el resultado de la decisión y normalmente se les llama leafs u hojas en inglés [21].

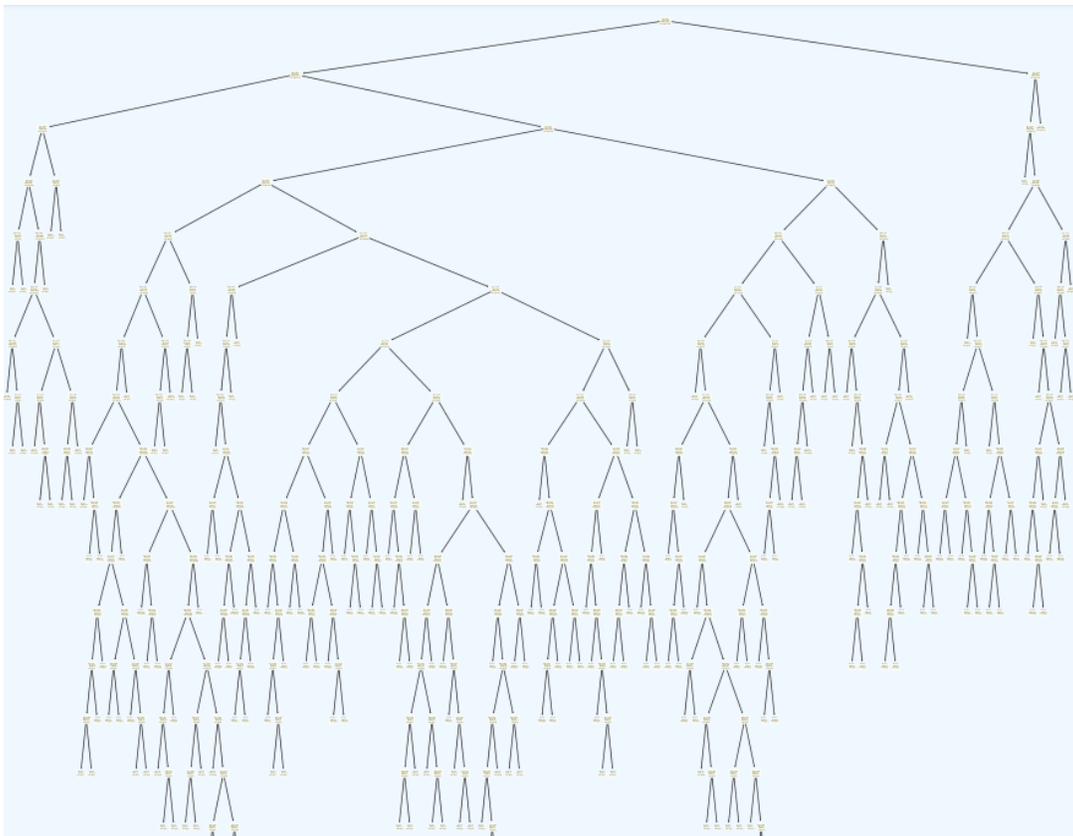


Fig. 21. Árboles de decisión.

### C.3 Regresión logística.

La regresión logística se ha convertido en una herramienta importante. Permite a los algoritmos utilizados en las aplicaciones de aprendizaje automático clasificar los datos entrantes basándose en datos históricos. A medida que llegan datos relevantes adicionales, los algoritmos mejoran la predicción de las clasificaciones dentro de los conjuntos de datos.

La regresión logística es un método de análisis estadístico para predecir un resultado binario, como un sí o un no, basado en observaciones previas de un conjunto de datos.

Un modelo de regresión logística predice una variedad de datos dependiente analizando la relación entre una o más variables independientes existentes. Por ejemplo, una regresión logística podría utilizarse para predecir si un candidato político ganará o perderá unas elecciones o si un estudiante de secundaria será admitido o no en una determinada universidad. Estos resultados binarios permiten decidir directamente entre dos alternativas [22].

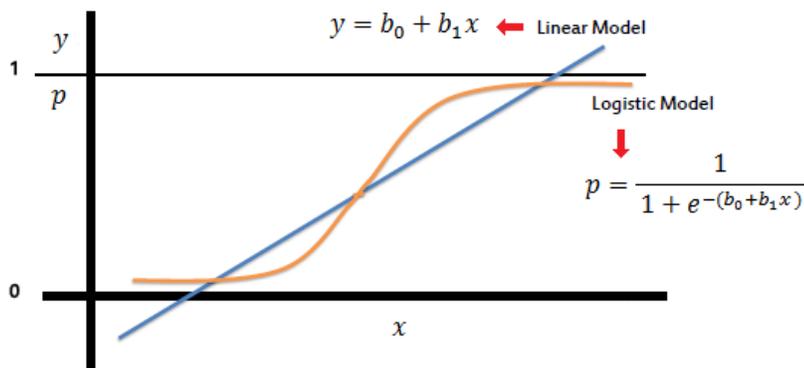


Fig. 22. Regresión Logística.

### C.4 Máquinas de vectores de soporte.

Una máquina de vectores de apoyo (SVM) es un modelo de aprendizaje automático supervisado que utiliza algoritmos de clasificación para este tipo de problemas de dos grupos. Después de suministrar a un modelo de SVM conjuntos de datos de entrenamiento etiquetados para cada categoría, son capaces de categorizar un nuevo conjunto de datos [23].

En comparación con otros algoritmos más recientes, como las redes neuronales, tienen dos ventajas principales: mayor velocidad y mejor rendimiento con un número limitado de muestras

Es un algoritmo de clasificación perteneciente al aprendizaje supervisado, puede ser aplicado para el tipo de clasificación y regresión. En este caso se usará como clasificación. Es un clasificador discriminatorio, definido formalmente por un hiperplano de separación, en otras palabras, dado los datos de entrenamiento etiquetados, el algoritmo genera un hiperplano óptimo que clasifica los nuevos ejemplos en dos espacios dimensionales. Este hiperplano es una línea que divide un plano en dos partes, en donde cada clase se encuentra en cada lado [23].

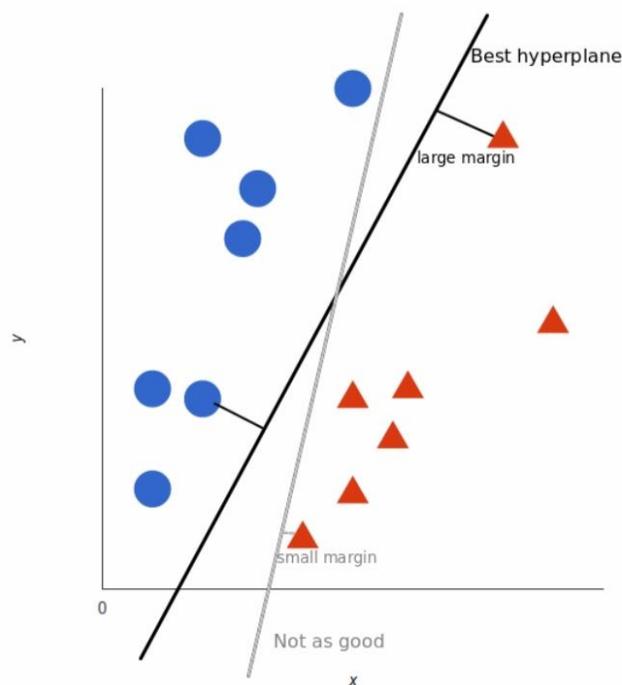


Fig. 23. Máquina de vectores de soporte.

Como se mencionaba anteriormente, el objetivo es utilizar estos tres modelos de machine learning para poder predecir a partir de un medidor magnético de aceite cuando hay un bajo nivel de aceite en el transformador o una alta temperatura y que estas dos causas lleven a que el transformador entre en una posible avería. A continuación, se adjunta un anexo en donde se podrá visualizar de mejor manera todo este estudio con sus gráficas respectivas y resultados:

### C.5 Análisis de resultados.

Las figuras 24, 25 y 26 se pueden tomar como los gráficos más representativos de los resultados encontrados, en donde se puede apreciar que cuando hay un bajo nivel de aceite o una temperatura muy alta de este, la alarma del medidor magnético se disparará. La gráfica en color negro representa la temperatura del aceite a lo largo del tiempo, la amarilla el nivel del aceite a lo largo del tiempo y la morada las veces que la alarma del indicador magnético de aceite se activa, siendo 1 activado y 0 desactivado. Todo esto entre las fechas de Julio de 2019 hasta el 05 de Mayo de 2020.

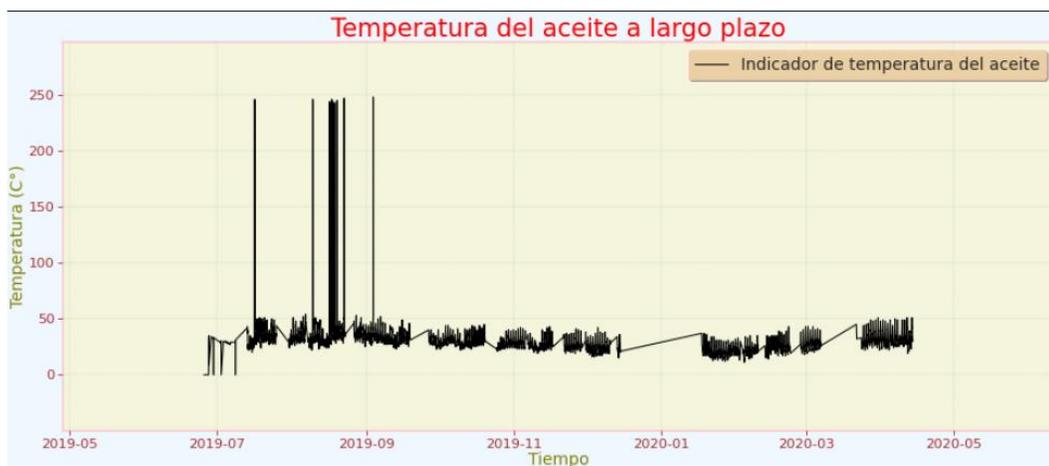


Fig. 24. Temperatura del aceite a largo plazo.

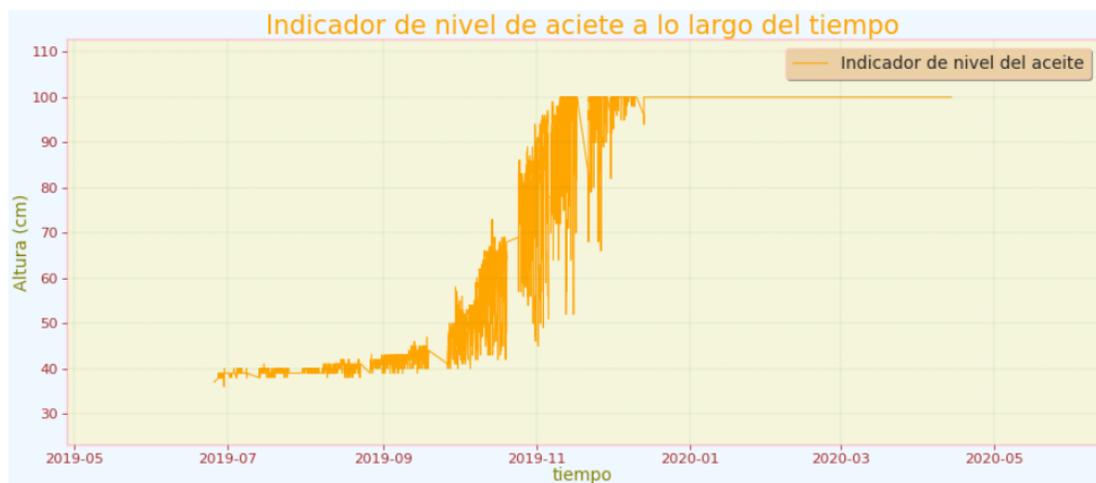


Fig. 25. Indicador de nivel de aceite a lo largo del tiempo.



Fig. 26. Advertencia de falla de indicador magnético de aceite a lo largo del tiempo.

### C.6 Matriz de confusión.

La matriz de confusión es una herramienta utilizada para visualizar el desempeño de un algoritmo de aprendizaje supervisado.

La matriz de confusión es una herramienta utilizada para visualizar el desempeño de un algoritmo de aprendizaje supervisado. Cada columna de la matriz representa el número de predicciones de cada clase y cada fila representa a las instancias en la clase real. En otras palabras, la matriz de confusión muestra qué tipos de aciertos y errores está teniendo el modelo a la hora de pasar al proceso de aprendizaje de los datos.

- Verdadero positivo: El valor real es positivo y la prueba predijo también que era positivo
- Verdadero negativo: El valor real es negativo y la prueba predijo también que el resultado era negativo
- Falso negativo: El valor real es positivo, y la prueba predijo que el resultado es negativo.
- Falso positivo: El valor real es negativo, y la prueba predijo que el resultado es positivo



Fig. 27. Matriz de confusión.

De las figuras 28, 29 y 30 se observa que, aunque los árboles de decisión no predijeron tantos verdaderos positivos como los otros dos métodos, fue más preciso a la hora de no predecir tantos falsos positivos o falsos negativos como los otros dos métodos y esto se puede observar en las métricas de medición.

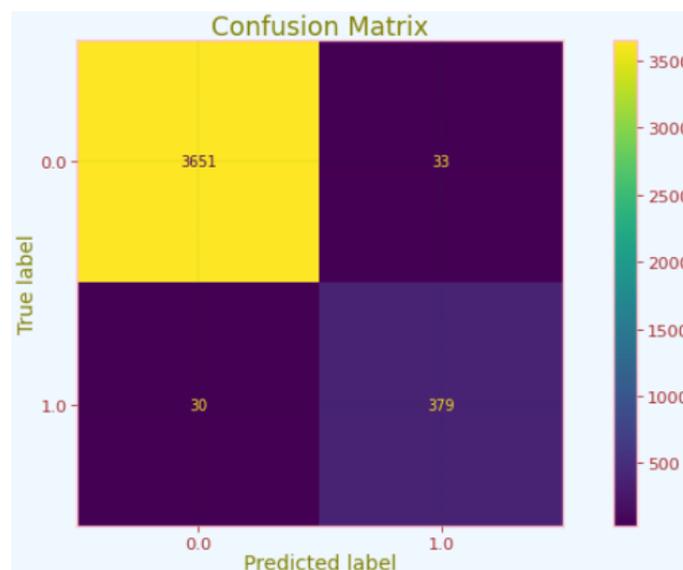


Fig. 28. Matriz de confusión árboles de decisión.

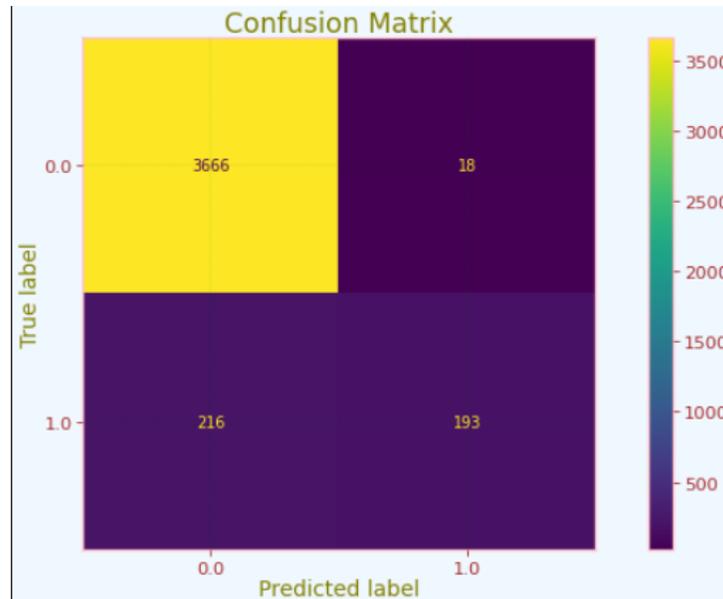


Fig. 29. Matriz de confusión regresión Logística.



Fig. 30. Matriz de confusión máquina de vectores de soporte.

### C.7 Métricas de medición.

La tabla III muestra dos métricas de medición: el “Training Accuracy ” y el “Model accuracy Score” que son las encargadas de determinar el desempeño de los modelos. Como se puede apreciar, el método de árboles de decisión es el más preciso tanto en el Training Accuracy que es la precisión que se obtiene al aplicar el modelo a los datos de entrenamiento y también se desempeña mucho mejor en el “Model Accuracy Score” que es el número de predicciones

correctas de todas las predicciones realizadas. La precisión también se puede calcular utilizando el número de predicciones correctas o todas las predicciones realizadas.

TABLA IIIII  
MÉTRICAS DE MEDICIÓN DE LOS MÉTODOS UTILIZADOS

<b>Modelo</b>	<b>Training Accuracy</b>	<b>Model Accuray Score</b>
<b>Árboles de decisión</b>	100%	98.46%
<b>Regresión Logística</b>	94.22%	94.28%
<b>Máquina de Vectores de soporte</b>	95.41%	95.55%

Según la tabla 3 se puede concluir que si se quieren detectar las fallas en el transformador mediante un medidor magnético de aceite, el mejor modelo a utilizar es el de árboles de decisión.

---

## VI. CONCLUSIONES

Desde la perspectiva académica, este trabajo contribuye en los siguientes aspectos: primero, se propusieron modelos de ML para predecir el consumo diario de potencia eléctrica en un centro comercial y para detectar fallas en un transformador. Los resultados obtenidos se validaron y/o comprobaron que el uso de modelos de ML es aplicable a la predicción del consumo de la demanda eléctrica, a la predicción de fallas de activos eléctricos y a la hora de tomar decisiones relacionadas con la parte eléctrica.

Desde el punto de vista práctico, una predicción exacta y acertada tanto en el consumo de la demanda de potencia y fallas en sistemas eléctricos, puede ayudar a los responsables en la toma de decisiones a mejorar las planificaciones relacionadas con el manejo eficiente de la energía eléctrica y de esta forma construir sistemas que ayuden a mejorar dicho manejo.

Aunque existen limitaciones relacionadas con el breve intervalo para la recopilación de datos, los modelos propuestos demuestran que el uso de series de tiempo para el pronóstico de la demanda produce un resultado preciso y exacto.

La disponibilidad de sistemas SCADA y/o Dashboards ha contribuido en gran medida a mejorar la eficiencia de sistemas eléctricos específicos, al mismo tiempo que se asegura su supervisión constante.

La detección de ciertos gases generados en un transformador lleno de aceite en servicio es, con frecuencia, la primera indicación disponible de un mal funcionamiento que, si no se corrige, puede conducir a la falla. Arcos, descargas parciales, chispas de baja energía, sobrecarga severa, falla del motor de la bomba y sobrecalentamiento en el sistema de aislamiento son algunos de los posibles mecanismos. Estas condiciones que ocurren individualmente, o como varios eventos simultáneos, pueden resultar en la descomposición de los materiales aislantes y la formación de varios gases combustibles y no combustibles. Es por esto que, el uso de métodos como los triángulos de Duval o la detección de gases en transformadores es tan importante.

## REFERENCIAS

- [1] Francisco M. Gonzalez-Longatt. "Introducción a los sistemas de Potencia". silo.tips. <https://silo.tips/download/introduccion-a-los-sistemas-de-potencia> (accedido el 8 de agosto de 2022).
- [2] A.S Karandaev, S.A Evdokimov, V.R Khramshin y O.I Karandaeva, "Information and Measuring System for Electric Arc Furnace Transformer Monitoring", APEIE, diciembre de 2014.
- [3] IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers, IEEE Std C57.104™-2008, IEEE Power & Energy Society, New York, NY 10016-5997, USA, 2019.
- [4] Chia-Hung Lin, Chien-Hsien Wu y Ping-Zan Huang, "Grey clustering analysis for incipient fault diagnosis in oil-immersed transformers", Elsevier, p. 2, noviembre de 2007
- [5] Saad ALBAWI, Tareq Abed MOHAMMED y Saad AL-ZAWI, "Understanding of a Convolutional Neural Network", ICET, p. 3, 2017.
- [6] Bengio, Y. (2009). "Learning deep architectures for AI. Foundations and trends in Machine Learning", 2(1), 1-127.
- [7] Schmidhuber, J. (2015). "Deep learning in neural networks: An overview. Neural networks", 61, 85-117.
- [8] Elias Dabbas, Interactive Dashboards and Data Apps with Plotly and Dash. Birmingham: Packt Publishing LTD, 2021.
- [9] Wes McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. 1005 Gravenstein Highway North, Sebastopol: O'Reilly, 2012.
- [10] Wes McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. 1005 Gravenstein Highway North, Sebastopol: O'Reilly, 2012.
- [11] Wes McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. 1005 Gravenstein Highway North, Sebastopol: O'Reilly, 2012.
- [12] Raúl Garreta y Guillermo Moncecchi, Learning scikit-learn: Machine Learning in Python. Birmingham, 2013.
- [13] Michel Duval, "The Duval Triangle for Load Tap Changers, Non-Mineral Oils and Low Temperature Faults in Transformers", IEEE Electrical Insulation Magazine, vol. 24, n.º 6, p. 8, julio de 2008.
- [14] Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers, IEEE Std C57.104-1991, 1991.

- 
- [15] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," IEEE Electr. Insul. Mag., vol. 18, no. 3, pp. 8–17, 2002
- [16] Gianluca Bontempi, Souhaib Ben Taieb y Yann-Aël Le Borgne, "Machine Learning Strategies for Time Series Forecasting", p. 6, 2013.
- [17] H. Marzooghi, K. Emami, P. J. Wolfs, B. Holcombe, "Short-term Electric Load Forecasting in Microgrids: Issues and Challenges", School of Engineering and Technology, Central Queensland University.
- [18] Eisa Almeshaei, Hassan Soltan, "A methodology for Electric Power Load Forecasting", Alexandria Engineering Journal, julio de 2011.
- [19] Jacob Benesty Jingdong Chen Yiteng Huang & Israel Cohen, Pearson Correlation Coefficient, 2a ed. Montreal, Canadá.
- [20] Electrical4U. "MOG in Transformer (Magnetic Oil Level Gauge)". Electrical4U. <https://www.electrical4u.com/magnetic-oil-gauge-or-mog-magnetic-oil-level-indicatortransformer/#:~:text=A%20Magnetic%20Oil%20level%20Guage,One%20float> (accedido el 6 de septiembre de 2022).
- [21] Alsabti K., Ranka S. and Singh V., CLOUDS: A Decision Tree Classifier for Large Datasets, Conference on Knowledge Discovery and Data Mining (KDD-98), August 1998
- [22] Alfred DeMaris, "A Tutorial in Logistic Regression", Journal of Marriage and Family, vol. 57, n.º 4, p. 13, noviembre de 1996.
- [23] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf. "Support vector machines." Intelligent Systems and their Applications, IEEE, 13(4), pp. 18–28, 1998.

## ANEXOS

Anexo A. Links algunos códigos de Github implementados.

- Triángulo de Duval: [https://github.com/Juanes1793/Proyectos-Porgramaci-n-/blob/main/Duval\\_Triangle.ipynb](https://github.com/Juanes1793/Proyectos-Porgramaci-n-/blob/main/Duval_Triangle.ipynb)
- Índice de salud de gases: [https://github.com/Juanes1793/Proyectos-Porgramaci-n-/blob/main/Indice\\_salud\\_gases.ipynb](https://github.com/Juanes1793/Proyectos-Porgramaci-n-/blob/main/Indice_salud_gases.ipynb)

Anexo B. Plataforma de monitoreo STATCOM.

Finalmente, también se implementó un dashboard de un STATCOM que se actualiza en tiempo real, el cual está implementando con software de uso libre, como lo es Python, Grafana y MySQL.

Ya que el STATCOM en su parte de Hardware no está implementado correctamente, se usó Python para simular los datos que luego fueron enviados a una base de datos MySQL para luego graficarse en tiempo real con el software Grafana.



Fig. 31. Dashboard STATCOM.