

Proceso de análisis de datos para predecir consumo horario eléctrico en tres transformadores en Salento Quindío: enfoque para decisiones en tiempo real

Paulo César Albis Ramos

Leidy Andrea Raigoza Palacio

Jeidy Alexandra Villanueva Torres

Asesor

Rafael Gaitán Ospina

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Especialización en Ciencia de Datos y Analítica

2024

Dedicatoria

Paulo César Albis Ramos: Dedico este trabajo a mi familia, cuyo amor, apoyo y sacrificio incondicional me han dado la fuerza para superar los desafíos a lo largo de este camino. A mis amigos, que siempre creyeron en mí, y a todos aquellos que me inspiraron a seguir adelante con dedicación y perseverancia. Agradezco especialmente a mis compañeros de equipo, Jeidy y Andrea, por su compromiso, ideas y apoyo constante, que hicieron de este proyecto una experiencia enriquecedora y memorable. Este logro es un reflejo de todo el esfuerzo compartido a lo largo de esta travesía.

Jeidy Alexandra Villanueva Torres: Dedico este logro a mis padres, que con su ejemplo de esfuerzo y valores me han enseñado a luchar por mis sueños, y a mi esposo, cuya paciencia, apoyo incondicional y amor han sido mi mayor fortaleza durante este camino. A mis compañeros de equipo, Paulo y Andrea, cuya compañía y colaboración hicieron de este proyecto una experiencia significativa y enriquecedora, y a Dios, por guiarme y darme la sabiduría necesaria para superar cada desafío.

Leidy Andrea Raigoza Palacio: Dedico este proyecto a mi hijo Emmanuel y a mi esposo Paulo, quienes han sido mi motor y mi razón para seguir adelante. A mis profesores y mentores, cuyo conocimiento y paciencia me han permitido crecer tanto académica como personalmente.

También quiero expresar mi gratitud a mis compañeros de equipo por su esfuerzo, compañerismo y confianza, que hicieron de este trabajo una experiencia llena de aprendizaje y logros compartidos. Este proyecto es un reflejo de todo lo que pudimos alcanzar juntos.

Agradecimientos

Expresamos nuestro más sincero agradecimiento a Rafael Gaitán Ospina, por su valioso apoyo técnico en el desarrollo de los modelos de predicción que sustentan este trabajo de grado, y a Luis Ángel Anillo Arrieta, cuya orientación y colaboración en la preparación y revisión del documento final fueron fundamentales para la culminación exitosa de esta investigación. Sus aportes fueron esenciales en el desarrollo del proyecto titulado “*Proceso de análisis de datos para predecir consumo horario eléctrico en tres transformadores en Salento Quindío: enfoque para decisiones en tiempo real*”. Este trabajo no habría sido posible sin su generosidad al compartir conocimientos, su disposición para resolver dudas técnicas y su compromiso con la excelencia académica. Su contribución refleja el impacto del trabajo en equipo y la colaboración interdisciplinaria en la construcción de nuevos conocimientos.

Resumen

La Resolución 40072 de 2018 del Ministerio de Minas y Energía de Colombia establece directrices para implementar la Infraestructura de Medición Avanzada (AMI), cuyo propósito es mejorar la gestión energética, reducir pérdidas técnicas y no técnicas, y garantizar un servicio eléctrico más eficiente. En este contexto, el trabajo se enfoca en analizar datos obtenidos de medidores inteligentes para predecir el consumo horario de energía en transformadores de Salento, Quindío. La metodología CRISP-DM guía este análisis, desde la preparación de los datos hasta la creación y evaluación de modelos predictivos utilizando técnicas de aprendizaje automático, lo que mejora la precisión de las predicciones y la toma de decisiones. Este enfoque integra análisis en tiempo real con redes inteligentes, optimizando la gestión del recurso eléctrico, apoyando la sostenibilidad y cumpliendo con los estándares técnicos y regulatorios. Así, se mejora la eficiencia y la calidad del servicio energético en Colombia.

Palabras clave: Medidores Inteligentes, Infraestructura de Medición Avanzada (AMI), Demanda de Energía, Predicción, Eficiencia Energética, Decisiones, Análisis, Datos, Energía, Modelos.

Abstract

Resolution 40072 of 2018 of the Colombian Ministry of Mines and Energy establishes guidelines to implement the Advanced Metering Infrastructure (AMI), whose purpose is to improve energy management, reduce technical and non-technical losses, and guarantee a more efficient electric service. In this context, the work focuses on analyzing data obtained from smart meters to predict hourly energy consumption in transformers in Salento, Quindío. The CRISP-DM methodology guides this analysis, from data preparation to the creation and evaluation of predictive models using machine learning techniques, which improves prediction accuracy and decision making. This approach integrates real-time analytics with smart grids, optimizing electricity resource management, supporting sustainability and complying with technical and regulatory standards. This improves the efficiency and quality of energy service in Colombia.

Keywords: Smart Meters, Advanced Metering Infrastructure (AMI), Energy Demand, Prediction, Energy Efficiency, Decisions, Analytics, Data, Energy, Modeling.

Tabla de Contenido

Glosario.....	13
Introducción	16
Planteamiento del Problema	17
Justificación	19
Objetivos	21
Objetivo General	21
Objetivos Específicos	21
Marco Conceptual.....	22
Análisis de la Variabilidad del Consumo	24
Modelos de Regresión para la Predicción	25
Técnicas de Clustering para el Perfilado de Consumidores.....	25
Análisis de Series Temporales para la Predicción Horaria	25
Marco Teórico.....	26
Metodología	28
Entendimiento del Negocio.....	28
Identificación del Problema.....	28
Determinación de Objetivos	29
Evaluación de la Situación Actual.....	29
Entendimiento de los Datos	29
Recolección de Datos	29
Descripción de Datos.....	30
Exploración de Datos	30

Preparación de Datos.....	30
Limpieza de Datos.....	30
Creación de Indicadores	31
Transformación de Datos.....	31
Modelado.....	31
Selección de Técnicas de Modelado	31
Construcción del Modelo	31
Optimización de Parámetros	31
Evaluación del Modelo	31
Validación del Modelo	31
Interpretación de Resultados	32
Pruebas de Robustez.....	32
Ajuste del Modelo	32
Implementación del Modelo en Producción	32
Integrar el Modelo en una Aplicación o Sistema Existente	32
Automatizar la Recolección de Datos y las Predicciones	32
Monitorear el Rendimiento del Modelo de Forma Continua	32
Análisis de Variables Relacionadas con el Consumo Energético.....	33
Conjunto de Datos.....	33
Actividades de Pre-Procesamiento.....	34
Análisis Exploratorio	35
Visualización Original de la Serie Temporal (Sin Remuestreo).....	36
Modelos de Predicción de Consumo Energético Diario	39

Visualización Datos con Remuestreo.....	39
Descomposición de la Serie Temporal.....	41
Modelo Exponential Smoothing	44
Transformador TRANS001	46
Transformador TRANS002.....	48
Transformador TRANS003.....	50
Modelo ARIMA	52
Transformador TRANS001	53
Transformador TRANS002.....	57
Transformador TRANS003.....	60
Modelo Redes Neuronales Recurrentes (<i>RNN</i>).....	64
Transformador TRANS001	66
Transformador TRANS002.....	68
Transformador TRANS003.....	69
Modelo Prophet.....	71
Transformador TRANS001	73
Transformador TRANS002.....	75
Transformador TRANS003.....	77
Sistema de Análisis y Predicción en Tiempo Real.....	79
Backend (Python)	80
Frontend (Flask y Dash).....	80
Conclusiones	83
Recomendación.....	88

Referencias..... 89

Lista de Tablas

Tabla 1 <i>Dataframe Consolidado de Consumo Eléctrico y Demanda</i>	35
Tabla 2 <i>Comparación de Métricas de Modelos para el Transformador TRANS001</i>	84
Tabla 3 <i>Comparación de Métricas de Modelos para el Transformador TRANS002</i>	85
Tabla 4 <i>Comparación de Métricas de Modelos para el Transformador TRANS003</i>	86

Lista de Figuras

Figura 1 <i>Metodología CRISP-DM</i>	28
Figura 2 <i>Agrupación de Clientes por Transformador</i>	36
Figura 3 <i>Consumo Energía Activa Importada Mes a Mes TRANS001</i>	37
Figura 4 <i>Consumo Energía Activa Importada Mes a Mes TRANS002</i>	38
Figura 5 <i>Consumo Energía Activa Importada Mes a Mes TRANS003</i>	38
Figura 6 <i>Consumo Energía Activa Importada Mes a Mes con Remuestreo TRANS001</i>	40
Figura 7 <i>Consumo Energía Activa Importada Mes a Mes con Remuestreo TRANS002</i>	41
Figura 8 <i>Consumo Energía Activa Importada Mes a Mes con Remuestreo TRANS003</i>	41
Figura 9 <i>Descomposición de la Serie Temporal para el Transformador TRANS001</i>	42
Figura 10 <i>Descomposición de la Serie Temporal para el Transformador TRANS002</i>	43
Figura 11 <i>Descomposición de la Serie Temporal para el Transformador TRANS003</i>	44
Figura 12 <i>Predicción del Consumo Energético Diario para TRANS001 - Suavizamiento Exponencial Triple</i>	47
Figura 13 <i>Predicción del Consumo Energético Diario para TRANS002 - Suavizamiento Exponencial Triple</i>	49
Figura 14 <i>Predicción del Consumo Energético Diario para TRANS003 - Suavizamiento Exponencial Triple</i>	51
Figura 15 <i>Predicciones del Modelo ARIMA (3, 1, 3) – [TRANS001]</i>	54
Figura 16 <i>Resumen de Resultados del Modelo ARIMA (3, 1, 3) – [TRANS001]</i>	55
Figura 17 <i>Predicciones del Modelo ARIMA (3, 1, 3) – [TRANS002]</i>	58
Figura 18 <i>Resumen de Resultados del Modelo ARIMA (3, 1, 3) – [TRANS002]</i>	59
Figura 19 <i>Predicciones del Modelo ARIMA (3, 1, 3) – [TRANS003]</i>	61

Figura 20 <i>Resumen de Resultados del Modelo ARIMA (3, 1, 3) – [TRANS003]</i>	63
Figura 21 <i>Detalles del Modelo LSTM: Capas, Parámetros y Dimensiones de Salida</i>	65
Figura 22 <i>Predicción del Consumo Energético Diario para TRANS001 utilizando RNN</i>	67
Figura 23 <i>Predicción del Consumo Energético Diario para TRANS002 utilizando RNN</i>	68
Figura 24 <i>Predicción del Consumo Energético Diario para TRANS003 utilizando RNN</i>	70
Figura 25 <i>Predicción del Consumo Energético Diario para TRANS001 utilizando PROPHET</i> 74	
Figura 26 <i>Predicción del Consumo Energético Diario para TRANS002 utilizando PROPHET</i> 76	
Figura 27 <i>Predicción del Consumo Energético Diario para TRANS003 utilizando PROPHET</i> 78	
Figura 28 <i>Interfaz Gráfica Inicial de Modelos de Predicción</i>	81
Figura 29 <i>Interfaz Gráfica de la Descomposición Inicial de la Serie Temporal</i>	81
Figura 30 <i>Interfaz Gráfica de Modelos de Predicción</i>	82
Figura 31 <i>Interfaz Gráfica de las Métricas</i>	82

Glosario

AMI (Infraestructura de Medición Avanzada): Sistema que permite la medición, recopilación y análisis remoto de datos sobre el consumo de energía, a través de medidores inteligentes. Facilita la gestión eficiente de la energía y la toma de decisiones en tiempo real.

ARIMA (Autoregressive Integrated Moving Average): Modelo estadístico utilizado para la predicción de series temporales. Es ampliamente usado en la predicción del consumo energético debido a su capacidad para modelar datos que muestran patrones estacionales y de tendencia.

Consumo Horario: Cantidad de energía eléctrica consumida en un intervalo de una hora. Este dato es clave para analizar patrones de consumo y desarrollar predicciones más precisas.

Demanda de Energía: Cantidad total de energía requerida por los consumidores en un período determinado. La predicción de la demanda es crucial para la gestión eficiente de los recursos energéticos.

Eficiencia Energética: Uso eficiente de la energía, que implica la reducción de pérdidas y la optimización de recursos energéticos sin comprometer los servicios prestados. Es uno de los principales objetivos de la implementación de AMI.

Estacionalidad: Componente de una serie temporal que refleja patrones repetitivos a intervalos regulares, como los cambios en el consumo de energía en diferentes estaciones del año o días de la semana.

Machine Learning (Aprendizaje Automático): Rama de la inteligencia artificial que se centra en la creación de algoritmos que permiten a las computadoras aprender de los datos. Se utiliza para mejorar la precisión de los modelos de predicción del consumo energético.

Medidor Inteligente: Dispositivo electrónico que mide el consumo de energía de un hogar o empresa y transmite esta información de manera remota a la empresa distribuidora. Permite la monitorización en tiempo real y contribuye a una mejor gestión de la demanda energética.

Predicción de Consumo: Proceso mediante el cual se anticipa la demanda futura de energía utilizando datos históricos, modelos estadísticos y técnicas de aprendizaje automático. Este enfoque permite mejorar la planificación y la operación del sistema eléctrico.

Pérdidas Técnicas y no Técnicas: Las pérdidas técnicas se refieren a la energía que se disipa en los componentes del sistema eléctrico (cables, transformadores, etc.), mientras que las pérdidas no técnicas son aquellas causadas por factores como el fraude o errores en la medición.

Prophet: Algoritmo de predicción desarrollado por Facebook, diseñado para modelar datos de series temporales con componentes estacionales y no lineales. Es útil para hacer pronósticos a largo plazo en escenarios con datos irregulares.

Redes Neuronales Recurrentes (RNN): Se utiliza para procesar datos secuenciales, como series temporales, siendo particularmente útil para la predicción a lo largo del tiempo.

Regresión Logística: Técnica estadística utilizada para predecir la probabilidad de un evento binario, como en la clasificación de un servicio energético como "excelente" o "regular" en función de ciertas variables independientes.

Series Temporales: Conjunto de datos recolectados en intervalos sucesivos, que permiten analizar patrones a lo largo del tiempo. El análisis de series temporales es esencial en la predicción del consumo energético.

Suavizamiento Exponencial: Técnica estadística que permite hacer predicciones basadas en la ponderación de los datos pasados, dándole más importancia a los datos más recientes. Se usa en modelos de predicción de series temporales.

Transformador: Dispositivo eléctrico que transforma el voltaje de la energía eléctrica para su distribución eficiente. Los datos de consumo por transformador son clave en el análisis y la predicción del consumo energético.

MAE (Error Absoluto Medio): Mide la magnitud promedio del error en las predicciones. Cuánta diferencia hay, en promedio, entre las predicciones del modelo y los valores reales.

MSE (Error Cuadrático Medio): Mide el promedio del cuadrado de las diferencias entre los valores reales y los valores predichos por el modelo. Un valor más bajo de MSE indica que el modelo predice con mayor precisión.

Introducción

La gestión eficiente de la energía eléctrica representa uno de los mayores desafíos para los sistemas de distribución modernos, especialmente en un contexto donde el consumo energético y las pérdidas técnicas y no técnicas crecen continuamente. En Colombia, la Resolución 40072 de 2018 del Ministerio de Minas y Energía impulsa la adopción de la Infraestructura de Medición Avanzada (AMI) como estrategia para mejorar la calidad del servicio y optimizar los recursos energéticos mediante el análisis de datos en tiempo real. Los medidores inteligentes, pieza clave de esta infraestructura, generan volúmenes significativos de información útil para anticipar el consumo energético y gestionar la demanda de manera eficiente.

Este trabajo se centra en el uso de datos de medidores inteligentes instalados en tres transformadores ubicados en Salento, Quindío, con el objetivo de predecir el consumo horario de energía. A través de metodologías como *CRISP-DM* y el uso de modelos avanzados de aprendizaje automático, se analizan variables críticas del consumo energético, se desarrollan modelos predictivos y se implementan sistemas de análisis en tiempo real.

En el Capítulo 1 se aborda la identificación de las variables que influyen en el consumo energético, apoyándose en una revisión sistemática del marco teórico. El Capítulo 2 detalla el desarrollo de modelos de predicción utilizando técnicas como Suavizado exponencial, ARIMA, *RNN* y *Prophet*, evaluando su desempeño en cada transformador. Finalmente, el Capítulo 3 presenta la integración de los modelos en un sistema de análisis en tiempo real para facilitar la toma de decisiones operativas y estratégicas.

Planteamiento del Problema

El desafío de mejorar la eficiencia energética y reducir las pérdidas en el sector eléctrico colombiano ha impulsado la necesidad de adoptar tecnologías avanzadas. En respuesta a esta necesidad, la Resolución 40072 de 2018 emitida por el Ministerio de Minas y Energía, establece un marco normativo para la implementación de la Infraestructura de Medición Avanzada (AMI). La resolución aplica a todos los agentes involucrados en la prestación del servicio de energía eléctrica y en la promoción de la gestión eficiente de la energía, tanto en el Sistema Interconectado Nacional (SIN) como en las Zonas No Interconectadas (ZNI) (CREG, 2018).

La implementación de AMI tiene como objetivo facilitar esquemas de eficiencia energética, mejorar la calidad del servicio, dinamizar la competencia en la comercialización minorista de energía eléctrica, y gestionar la reducción de pérdidas técnicas y no técnicas (Rodríguez, 2021).

La Comisión de Regulación de Energía y Gas (CREG) establecerá las condiciones para la implementación de AMI y determinará los responsables de su instalación, administración, operación, mantenimiento y reposición. La meta para el año 2030 es que al menos el 75% de los usuarios conectados en un mercado de comercialización en el SIN cuenten con AMI (CREG, 2018).

Teniendo en cuenta lo anterior, se evidencia que el despliegue creciente de la infraestructura de medición avanzada (AMI) ha hecho que los datos de consumo de energía deban ser tratados y analizados de manera diferente (Mitra et al, 2024), con fines de apoyar la toma de decisiones en compañías, ya que actualmente no se cuenta las herramientas que permitan gestionar adecuadamente estos grandes volúmenes de datos (Liu et al, 2015). La problemática que se abordará en el trabajo de grado es la necesidad de mejorar la gestión de la

demanda de energía eléctrica mediante la predicción del consumo horario del cliente, utilizando información de medidores inteligentes (Ketchledge, 2024). Este análisis supone un amplio desafío, con miras a abordar enfoques de pronósticos y aplicación de diferentes modelos estadísticos y de aprendizaje automático (Ignatiadis et al, 2019).

Ante este contexto, surge la pregunta de investigación central del trabajo de grado: "¿Cómo se puede mejorar la gestión de la demanda de energía eléctrica y facilitar la toma de decisiones operativas y estratégicas en tiempo real, mediante la predicción del consumo horario de los clientes de tres transformadores ubicados en el municipio de Salento Quindío, utilizando la información recopilada por medidores inteligentes?". Esta pregunta busca explorar métodos eficaces para la utilización de datos de AMI en modelos predictivos que no solo anticipen el consumo energético, sino que también optimicen la gestión y la respuesta de las infraestructuras eléctricas en tiempo real (Ketchledge, 2024).

Este enfoque permitirá no solo responder a los requisitos técnicos y regulatorios establecidos por la Comisión de Regulación de Energía y Gas (CREG) y la legislación colombiana, sino también avanzar hacia una gestión más inteligente y eficiente del recurso eléctrico en áreas críticas, apoyando la sostenibilidad y la eficiencia en el servicio de energía eléctrica.

Justificación

La implementación de un proceso de análisis de datos para predecir el consumo horario de los transformadores utilizando información de medidores inteligentes ofrece una oportunidad para mejorar la eficiencia operativa y la calidad del servicio. Estos dispositivos, a diferencia de los medidores tradicionales, pueden transferir información de consumo a sistemas de procesamiento de datos remotos, lo que permite identificar variables clave para el análisis. Esto facilita la retroalimentación sobre el consumo de electricidad, el reconocimiento de patrones y anomalías, el pronóstico de la demanda de energía y generación, y la creación de perfiles de carga, habilitando así la toma de decisiones informadas en tiempo real para gestionar de manera más eficiente la demanda de energía eléctrica (Völker et al, 2021).

El objetivo de seleccionar las variables más relevantes se basa en la necesidad de comprender mejor los factores que afectan el consumo horario del cliente que permite mejorar la precisión del modelo de predicción y, por lo tanto, la calidad de las decisiones tomadas en función de estas predicciones. Para lograr una amplia adopción de estos servicios basados en datos de medidores inteligentes, se deben abordar varios desafíos. Estos incluyen la estandarización del hardware y los formatos de datos, el desarrollo de algoritmos de procesamiento de datos (Völker et al, 2021).

El desarrollo de un modelo de predicción mediante técnicas de aprendizaje automático no supervisado permite explorar patrones de demanda y respuesta de los clientes en diferentes contextos, lo que proporciona una ventaja competitiva al permitir una gestión más eficiente de los recursos energéticos y una mejor planificación de la capacidad de la red eléctrica (Kang & Reiner, 2022). Además, al implementar un sistema de análisis de datos en tiempo real, se

garantiza que las predicciones sean actualizadas continuamente, facilitando la toma de decisiones informadas y ágiles para optimizar la operación.

Objetivos

Objetivo General

Predecir el consumo diario de los clientes de tres transformadores ubicados en el municipio de Salento Quindío, utilizando información de medidores inteligentes, que faciliten la toma de decisiones en tiempo real.

Objetivos Específicos

Identificar las variables que influyen en el consumo diario de tres transformadores ubicados en el municipio de Salento Quindío, usando los datos de medidores inteligentes, basados en la revisión sistemática del marco teórico.

Desarrollar modelos de predicción del consumo diario de cada transformador utilizando técnicas de aprendizaje automático.

Visualizar un modelo de análisis de datos en tiempo real que basados en los datos de medidores inteligentes, genere predicciones actualizadas del consumo diario de cada transformador, para apoyar la toma de decisiones operativas y estratégicas.

Marco Conceptual

Un medidor inteligente de energía es un dispositivo avanzado que registra y transmite datos detallados sobre el consumo eléctrico de un hogar o negocio. A diferencia de los medidores tradicionales, los medidores inteligentes pueden comunicarse de manera bidireccional con la red eléctrica, permitiendo la recopilación remota de datos y la implementación de servicios como la facturación precisa, la gestión de la demanda y el monitoreo en tiempo real del consumo de energía (Rashid, 2018). Estos dispositivos también pueden proporcionar información útil a los usuarios para ayudarles a optimizar su consumo y reducir costos. Lo que permite implementar estrategias de precios dinámicos y mejorar la eficiencia y confiabilidad de los sistemas de distribución de energía (Oprea & Bâra, 2019).

Los medidores inteligentes desempeñan un papel fundamental en la gestión moderna de la energía eléctrica, proporcionando una gran cantidad de información detallada (Arias Martín, 2019). Esto resalta su importancia tanto desde una perspectiva técnica como económica, siendo instrumentos clave para soportar modelos predictivos avanzados que mejoran la toma de decisiones y la planificación a largo plazo en las redes de distribución y gestión de la demanda energética (Athanasiadis et al, 2024).

En el campo de la predicción del consumo energético utilizando infraestructuras de medición avanzada (AMI), el estudio realizado por Mathumitha et al (2023) aborda el uso de técnicas de aprendizaje automático para predecir el consumo en edificaciones a través de datos obtenidos de medidores inteligentes y datos históricos de cargas. Este estudio compara dos algoritmos de predicción: la regresión basada en árboles de decisión y la regresión basada en Máquinas de Vectores de Soporte (*SVM*), donde demuestran cómo la regresión *SVM* puede ser

utilizada efectivamente para pronosticar el consumo de energía con una precisión notablemente alta.

Igualmente, Burnett et al (2022) exploran el uso de diversas técnicas de machine learning para estimar la demanda de energía en hogares estadounidenses, donde emplean un conjunto de datos que incluyen consumo de energía eléctrica, hora del día, día de la semana, estacionalidad, ubicación geográfica, características de vivienda, socioeconómicas y comportamentales. La investigación destaca el algoritmo de bosque aleatorio (Random Forest) como el más efectivo entre varios modelos evaluados, incluyendo regresiones lineales, k-nearest neighbors, y métodos de gradiente. Además, subrayan la potencialidad del aprendizaje automático para prever el consumo energético, ofreciendo perspectivas sobre la gestión eficiente y la implementación de nuevas tecnologías en los sistemas residenciales.

Por su lado, Czétány et al (2021) desarrolló un estudio utilizando datos de medidores inteligentes de casi mil hogares de Hungría, con el objetivo de establecer perfiles de consumo de electricidad empleando métodos de agrupamiento como k-means, fuzzy k-means y jerárquico aglomerativo; examinando el impacto de varios parámetros como el tipo de medidor, el día de la semana, la estacionalidad, alineando así la práctica de medición avanzada con las necesidades de una gestión energética inteligente y sostenible, marcando un paso significativo hacia la optimización de recursos en entornos residenciales.

Entre las variables utilizadas por los medidores inteligentes de energía encontramos las siguientes, según (González-Briones et al, 2019):

Fecha: Representa el día de la lectura programada.

Hora: Representa la hora de la lectura programada.

Import kWh: Cantidad de energía eléctrica importada medida en kilovatios-hora.

Import kVarh: Cantidad de energía reactiva importada en kilovars-hora.

Interval-LP Import kWh y Interval-LP Import kVarh: Mediciones de energía y energía reactiva importadas durante un intervalo específico, en kilovatios-hora y kilovars-hora, respectivamente.

Export kWh y Export kVarh: Cantidad de energía y energía reactiva exportadas, en kilovatios-hora y kilovars-hora, respectivamente.

Interval-LP Export kWh y Interval-LP Export kVarh: Similar a las variables de intervalo de importación, pero para la energía exportada.

DemandForActiveImportPower y DemandForActiveExportPower: Demanda de potencia activa para la importación y exportación, respectivamente, en kilovatios.

DemandForReactiveImportPower y DemandForReactiveExportPower: Demanda de potencia reactiva para la importación y exportación, respectivamente, en kilovatios.

L1, L2, L3 Last average voltage: Últimas lecturas promedio de voltaje para cada una de las tres fases de suministro de energía, medidas en voltios.

L1, L2, L3 Last average current: Últimas lecturas promedio de corriente para cada una de las tres fases de suministro de energía, medidas en amperios.

Análisis de la Variabilidad del Consumo

La variabilidad en el consumo de energía es crucial para identificar patrones de uso y desarrollar modelos predictivos eficientes. El análisis de variabilidad puede incluir el cálculo de histogramas de consumo y la evaluación de la desviación estándar para entender las fluctuaciones en el consumo energético (Azaza et al, 2017).

Modelos de Regresión para la Predicción

Los modelos de regresión, como los modelos lineales y no lineales, son ampliamente utilizados para correlacionar el consumo de energía con variables influyentes como la temperatura exterior y la hora del día. Estos modelos ayudan a prever el consumo basado en condiciones climáticas y patrones diarios (Liu et al, 2015).

Técnicas de Clustering para el Perfilado de Consumidores

El *clustering* de datos de medidores inteligentes permite agrupar consumidores con patrones similares de consumo. Este agrupamiento es fundamental para diseñar programas de respuesta a la demanda y para la gestión eficiente de la carga en la red. Las técnicas de *clustering* utilizan características como el factor de responsabilidad y la variabilidad del consumo para categorizar a los consumidores (Azaza et al, 2017).

Análisis de Series Temporales para la Predicción Horaria

El análisis de series temporales es esencial para la predicción del consumo eléctrico horario. Algoritmos como ARIMA y modelos de aprendizaje automático pueden predecir el consumo futuro basándose en datos históricos, lo cual es vital para la planificación operativa y la toma de decisiones en tiempo real (Liu et al, 2015).

Marco Teórico

La gestión eficiente de la energía eléctrica ha sido un desafío recurrente para los sistemas de distribución modernos. En el contexto colombiano, la implementación de la Infraestructura de Medición Avanzada (AMI), establecida por la Resolución 40072 de 2018 del Ministerio de Minas y Energía, representa un paso significativo hacia la optimización de recursos energéticos. Este marco normativo fomenta la incorporación de tecnologías avanzadas que permiten gestionar grandes volúmenes de datos y facilitar decisiones en tiempo real, especialmente a través de los medidores inteligentes.

Los medidores inteligentes son dispositivos capaces de registrar y transmitir datos de consumo eléctrico de forma bidireccional, lo que habilita una comunicación continua con las redes eléctricas. Según Oprea & Bâra (2019), estos dispositivos no solo ofrecen monitoreo en tiempo real, sino que también facilitan la implementación de modelos de precios dinámicos y estrategias de eficiencia energética. Desde una perspectiva técnica, los datos capturados por los medidores inteligentes permiten un análisis detallado del comportamiento energético, identificando patrones y tendencias que son esenciales para la predicción del consumo.

En el campo de la predicción energética, las series temporales han jugado un papel fundamental. Modelos estadísticos como ARIMA y técnicas más avanzadas de aprendizaje automático, como las Redes Neuronales Recurrentes (*RNN*) y *Prophet*, son herramientas clave para anticipar patrones de consumo. Según Burnett et al. (2022), estos métodos permiten modelar variaciones estacionales y tendencias en diferentes contextos, mejorando la precisión de las predicciones. En particular, los algoritmos de aprendizaje automático han demostrado ser efectivos para manejar datos complejos y no lineales, lo que resulta crítico en la gestión de redes inteligentes.

La metodología *CRISP-DM*, utilizada como base para este trabajo, integra el análisis de datos desde la recopilación hasta la evaluación de los modelos predictivos. Este enfoque asegura que las decisiones estén respaldadas por un proceso riguroso y replicable. Kang & Reiner (2022) destacan la importancia de combinar esta metodología con herramientas de aprendizaje automático para obtener *insights* precisos y accionables en tiempo real.

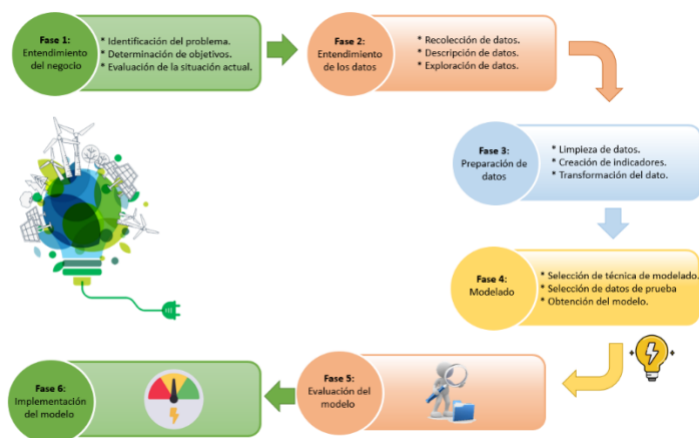
Desde una perspectiva más amplia, los estudios de González-Briones et al. (2019) y Mathumitha et al. (2023) subrayan cómo las técnicas de *clustering* y los modelos de regresión han permitido analizar grandes volúmenes de datos de consumo, agrupando clientes según patrones de uso y optimizando la distribución de energía. Estos enfoques refuerzan la necesidad de combinar métodos tradicionales con herramientas modernas de inteligencia artificial.

Metodología

En el desarrollo del proyecto de grado "Proceso de análisis de datos para predecir consumo horario eléctrico en tres transformadores en Salento Quindío: enfoque para decisiones en tiempo real", presenta la metodología *CRISP-DM*, la cual se adoptará y comprende las siguientes seis fases, como se visualiza en la Figura 1, descrito por (Espinosa-Zúñiga, 2020).

Figura 1

Metodología *CRISP-DM*



Entendimiento del Negocio

Identificación del Problema

La Resolución 40072 de 2018 del Ministerio de Minas y Energía de Colombia establece lineamientos para implementar la Infraestructura de Medición Avanzada (AMI) en el servicio público de energía eléctrica, con el objetivo de promover la gestión eficiente de la energía y la incorporación de nuevas tecnologías. Para 2030, se espera que al menos el 75% de los usuarios tengan AMI, mejorando así la eficiencia energética, la calidad del servicio, y reduciendo pérdidas técnicas y no técnicas. El creciente uso de AMI requiere un nuevo enfoque en el análisis de datos de consumo para apoyar decisiones empresariales, usando modelos estadísticos y de

aprendizaje automático para mejorar la gestión de la demanda de energía mediante la predicción del consumo horario con medidores inteligentes.

Determinación de Objetivos

El objetivo general de este estudio es predecir el consumo horario de los clientes de tres transformadores en Salento Quindío, utilizando datos de medidores inteligentes para facilitar decisiones en tiempo real. Los objetivos específicos incluyen identificar las variables clave que influyen en el consumo, desarrollar un modelo predictivo utilizando técnicas de aprendizaje automático, e implementar un sistema de análisis de datos en tiempo real que permita actualizar continuamente las predicciones para mejorar la gestión operativa y estratégica en la empresa de distribución eléctrica.

Evaluación de la Situación Actual

Actualmente, se enfrenta una ineficiencia significativa en la administración de los datos provenientes de los medidores, lo que afecta la gestión de la demanda de energía eléctrica. Esta limitación en el procesamiento y análisis de los datos impide una óptima toma de decisiones y la ejecución de estrategias efectivas de gestión energética. La implementación de la solución de minería de datos propuesta será crucial para superar estos desafíos, permitiendo establecer un punto de comparación claro que facilitará la medición del éxito del proyecto en términos de mejora en la eficiencia y efectividad de las decisiones operativas y estratégicas.

Entendimiento de los Datos

Recolección de Datos

En este proyecto, la recolección de datos se realiza mediante el uso de medidores inteligentes instalados en tres transformadores ubicados en el municipio de Salento, Quindío. Estos dispositivos son las fuentes principales de datos y están diseñados para capturar

información detallada sobre el consumo horario de energía eléctrica. Durante la fase de recolección, se emplean técnicas automatizadas para extraer los datos directamente de los medidores, asegurando así la precisión y la actualización continua de la información.

Descripción de Datos

Las variables registradas en los medidores inteligentes son esenciales para monitorear y analizar el flujo de energía eléctrica y la calidad de la misma en una instalación específica o a través de la red eléctrica.

Exploración de Datos

La exploración de datos en el contexto de la Infraestructura de Medición Avanzada (AMI) implica la realización de pruebas estadísticas elementales para discernir las características fundamentales de los datos obtenidos de los medidores inteligentes. Este análisis preliminar es esencial para entender la calidad y estructura de los datos de consumo energético, detectar irregularidades, comprender la variabilidad del consumo e identificar tendencias y patrones temporales. Al examinar detalladamente estos aspectos, se pueden apreciar las dinámicas del consumo energético y establecer una base sólida para el desarrollo de modelos analíticos y predictivos en la gestión eficiente de la demanda de energía.

Preparación de Datos

Limpieza de Datos

La limpieza de datos en el ámbito de los medidores inteligentes de AMI es un paso crucial que involucra varias técnicas para garantizar la integridad y calidad de los datos antes del análisis. Esto incluye la normalización de los datos para uniformar escalas y formatos, la discretización de campos numéricos para categorizar datos continuos, el manejo adecuado de

valores ausentes para evitar sesgos, la eliminación de registros duplicados para asegurar la precisión y la imputación de datos para rellenar espacios donde la información es incompleta.

Creación de Indicadores. Permite ayudar a encontrar patrones que incidan en el comportamiento del modelo.

Transformación de Datos. La transformación de datos es un paso crucial en la preparación de datos para el modelado, que permite que las técnicas analíticas trabajen más eficazmente y que los modelos aprendan de manera más eficiente.

Modelado

Selección de Técnicas de Modelado

Basado en los objetivos específicos y las características de los datos, se podría elegir entre varios modelos como regresión lineal para relaciones directas, árboles de decisión para clasificaciones basadas en características, o modelos más complejos como las redes neuronales para patrones no lineales y complejos.

Construcción del Modelo

Utilizar los datos transformados y los indicadores creados para entrenar el modelo.

Optimización de Parámetros

Ajustar los parámetros del modelo para mejorar la precisión (Wang et al, 2019).

Evaluación del Modelo

Validación del Modelo

Utilizar el conjunto de datos de prueba para evaluar la precisión del modelo. Métricas comunes incluyen el error cuadrático medio (MSE), el coeficiente de determinación (R^2), y para clasificación, la precisión y la matriz de confusión.

Interpretación de Resultados

Analizar cómo las predicciones del modelo se alinean con los datos reales.

Pruebas de Robustez

Realizar pruebas de estrés en el modelo cambiando condiciones o introduciendo datos anómalos para ver cómo responde.

Ajuste del Modelo

Basado en los resultados de las pruebas, ajustar el modelo según sea necesario. Esto puede implicar volver a la fase de modelado para refinar los algoritmos o para incluir más datos.

Implementación del Modelo en Producción

Integrar el Modelo en una Aplicación o Sistema Existente

Esto puede incluir la implementación del modelo como un servicio de *backend*, que puede ser consultado mediante APIs por otras aplicaciones de la empresa, o integrarse directamente en el software de gestión energética para proporcionar recomendaciones en tiempo real basadas en los datos recibidos.

Automatizar la Recolección de Datos y las Predicciones

Esto asegura que las predicciones se basan en los datos más recientes, aumentando la relevancia y utilidad de las mismas en la toma de decisiones operativas y estratégicas.

Monitorear el Rendimiento del Modelo de Forma Continua

Consiste en supervisar y evaluar el rendimiento del modelo de predicción de manera sistemática para asegurar que sigue siendo preciso y relevante a medida que cambian las condiciones de operación y los patrones de datos.

Análisis de Variables Relacionadas con el Consumo Energético

En este capítulo se presenta el análisis de las variables que influyen en el consumo diario de energía en tres transformadores ubicados en Salento, Quindío, utilizando datos provenientes de medidores inteligentes. A través de una revisión sistemática del marco teórico, se identificaron las variables más relevantes relacionadas con el consumo energético, incluyendo el comportamiento horario, estacionalidades y patrones de uso. Además, se describe el proceso de preprocesamiento de datos, que incluye la consolidación, limpieza y transformación de los registros para garantizar su integridad y utilidad en el análisis.

Conjunto de Datos

Se realiza un análisis del conjunto de datos de consumo eléctrico de medidores horarios. Este conjunto de datos se compone de múltiples archivos en formato .csv, correspondientes al consumo mensual. Para integrarlos en un único conjunto de datos, se configura un directorio en Python y se gestionan listas y *dataframes* que permiten agregar cada archivo .csv. Finalmente, todos los *dataframes* se combinan en un solo conjunto, organizados cronológicamente.

El *dataframe* resultante contiene 11 variables. A partir de marzo de 2024 y hasta octubre de 2024, se registraron datos horarios, lo que representa 24 mediciones diarias de energía activa. Posteriormente, se realizó un pivot en la columna *CHANNEL* para generar columnas adicionales que facilitaran el análisis y la exploración de los datos. Al finalizar este proceso, se obtuvo un *dataframe* con 25 columnas. Para el desarrollo del trabajo, se seleccionaron 3 atributos principales: Fecha, Transformador, Energía Activa Importada llamada en el dataframe inicial Interval-LP Import kWh. Este volumen de información requiere el uso de técnicas específicas para su tratamiento y análisis.

Actividades de Pre-Procesamiento

Para iniciar el análisis del consumo eléctrico, se realiza la lectura de múltiples archivos CSV que contienen datos horarios de medidores inteligentes correspondientes a distintos meses. Estos archivos se consolidan en un único *DataFrame* utilizando la librería pandas, obteniendo un total de 461.538 registros. Se identifican las columnas relevantes del conjunto de datos, que incluyen el número del medidor, la fecha y hora de la lectura, el valor de la energía importada y exportada, y otras variables relacionadas con la demanda de potencia activa y reactiva.

En primer lugar, se lleva a cabo una revisión de los nombres de las columnas para estandarizarlos, eliminando espacios, tildes y caracteres no deseados. Posteriormente, se verifica la existencia de valores faltantes en las columnas, encontrando que algunos campos como *CHANNEL* y *UNIT_MEASUREMENT* contienen valores nulos. Para facilitar el análisis, se crea una copia del *DataFrame* original y se reemplazan los valores faltantes con el texto "Null" cuando es necesario.

Se realiza un pivot del *DataFrame*, utilizando las columnas *METER_NO* y *DATE_TIME* como índices, y la columna *CHANNEL* para crear nuevas columnas con los valores de consumo correspondientes. Este proceso permite generar un *DataFrame* de 459.589 registros con 6 columnas las cuales se pueden observar en la Tabla 1 que incluye información detallada sobre el consumo de energía.

Además, se procede a renombrar las columnas para hacerlas más representativas del contenido, utilizando un diccionario de mapeo. Entre las columnas clave se incluyen medidor, fecha, transformador, Intervalo de energía activa importada (*LP_act_imp*). Las columnas relacionadas con la energía se convierten a formato numérico para asegurar un tratamiento adecuado durante el análisis.

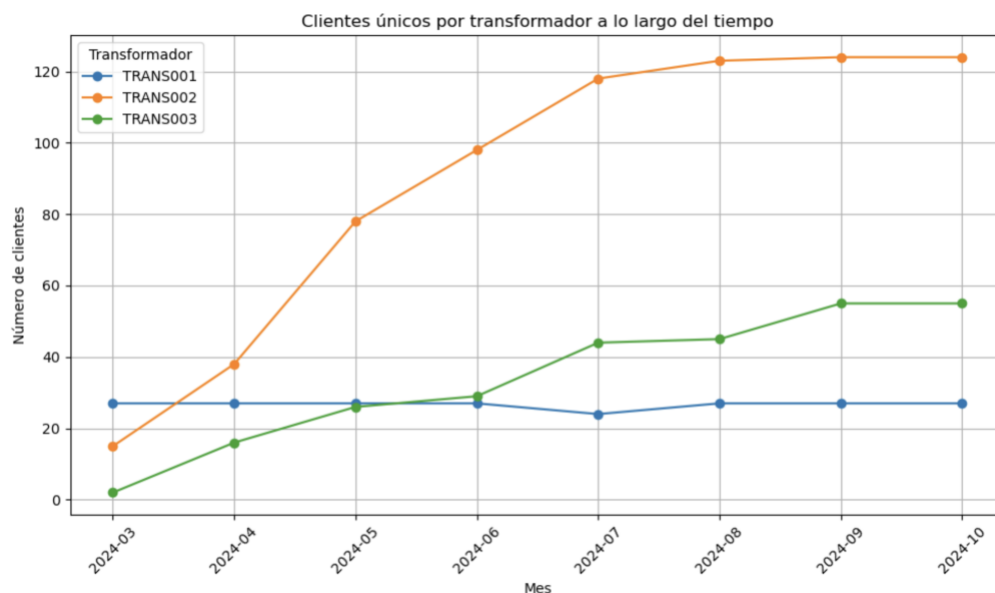
Tabla 1*Dataframe Consolidado de Consumo Eléctrico y Demanda*

Fecha	Medidor	LP_act_imp	Status	Punto_medida	Transformer
1/03/24	MT2222172868NUM	0.02	InstallWithActive	MS319593PT	TRANS001
1/03/24	MT2222172901NUM	0.20	InstallWithActive	MS319596PT	TRANS001
1/03/24	MT2222172902NUM	0.11	InstallWithActive	MS319600PT	TRANS001
1/03/24	MT2222172862NUM	0.42	InstallWithActive	MS319599PT	TRANS001
1/03/24	MT2222172861NUM	0.00	InstallWithActive	MS319592PT	TRANS001
1/03/24	MT2222172860NUM	0.00	InstallWithActive	MS319591PT	TRANS001
1/03/24	MT2222172859NUM	0.02	InstallWithActive	MS319590PT	TRANS001
1/03/24	MT2222172858NUM	0.19	InstallWithActive	MS319589PT	TRANS001
1/03/24	MT2222172857NUM	0.13	InstallWithActive	MS319588PT	TRANS001
1/03/24	MT2223003526NUM	0.99	InstallWithActive	MS319597PT	TRANS001

Nota. Tabla que ilustra el dataframe de las variables para el análisis de consumos horarios de energía.

Análisis Exploratorio

Se agrupan los datos por transformador como se observa en la Figura 2 para obtener un resumen del intervalo de consumo de energía activa importada (LP_act_imp) a lo largo del tiempo, y se realizan remuestreo diarios, semanales y mensuales. Los datos se visualizan utilizando gráficos para evaluar el comportamiento del consumo y se identifican patrones temporales, tales como tendencias y estacionalidades. El análisis de los datos remuestreados proporciona una visión clara del consumo de energía en diferentes frecuencias temporales, lo que permite una mejor comprensión de los patrones de uso.

Figura 2*Agrupación de Clientes por Transformador***Visualización Original de la Serie Temporal (Sin Remuestreo)**

Se realiza la visualización original de la serie temporal del consumo de energía activa para los transformadores seleccionados (TRANS001, TRANS002 y TRANS003). Esta visualización muestra los datos sin remuestreo, reflejando las fluctuaciones diarias en el consumo energético desde marzo de 2024 hasta octubre de 2024. Para obtener esta visualización, los datos de consumo energético se agruparon y sumaron por día, consolidando los valores de energía activa importada en un intervalo diario. Como resultado, se generó un nuevo *dataframe* diario para cada transformador, lo que permite observar los patrones y variaciones en el consumo de energía en una escala temporal más manejable y adecuada para el análisis.

Cada gráfico ilustra la variabilidad y los patrones de uso característicos de cada transformador, proporcionando una perspectiva inicial de la tendencia y los picos de consumo. El

transformador TRANS001 muestra una serie de variaciones constantes a lo largo del periodo, con incrementos notables hacia finales de octubre como se ve en la

Figura 3. Por otro lado, el TRANS002 presenta un comportamiento más estable en los primeros meses, seguido de un aumento considerable y fluctuaciones marcadas en la segunda mitad del año como se ve en la Figura 4. Finalmente, el TRANS003 tiene patrones de consumo bajos al inicio del periodo, con un aumento pronunciado desde septiembre y variaciones intensas en ciertos días como se ve en la Figura 5. Esta visualización inicial es importante para comprender los patrones de consumo diario y para guiar los pasos posteriores en el análisis, como el ajuste de modelos predictivos y la descomposición de la serie en componentes de tendencia y estacionalidad.

Figura 3

Consumo Energía Activa Importada Mes a Mes TRANS001

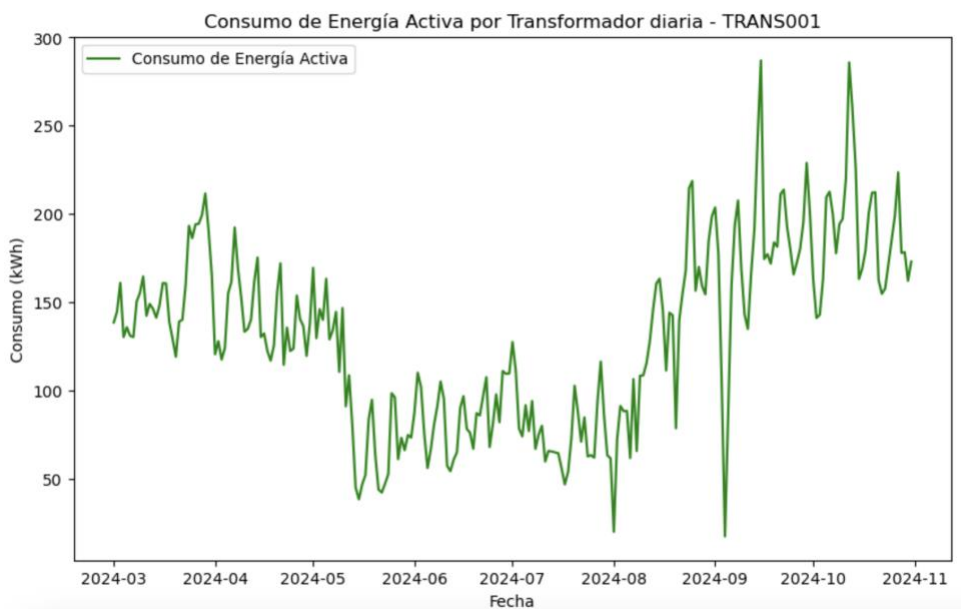
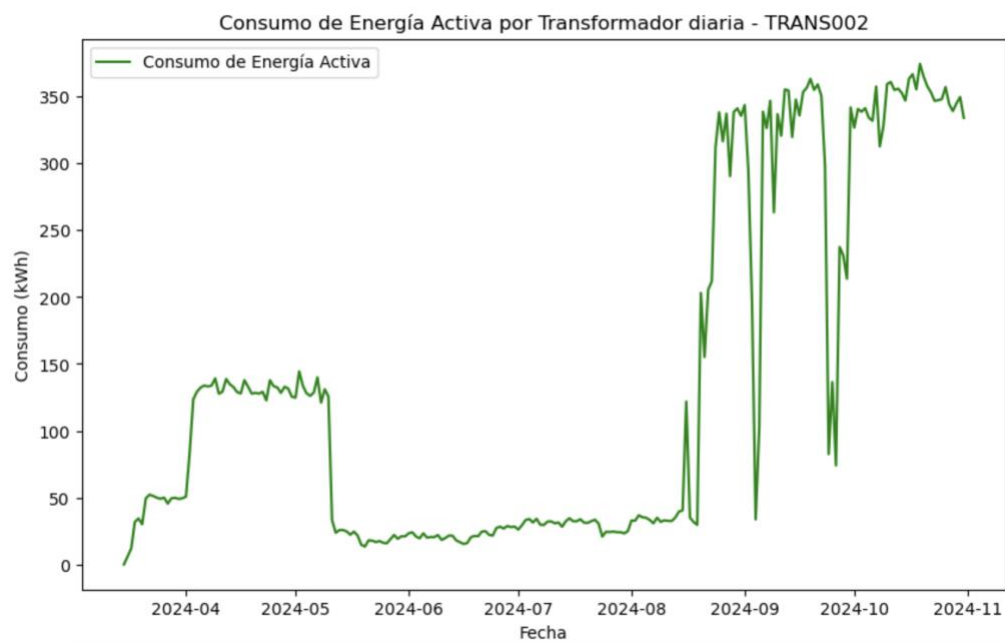
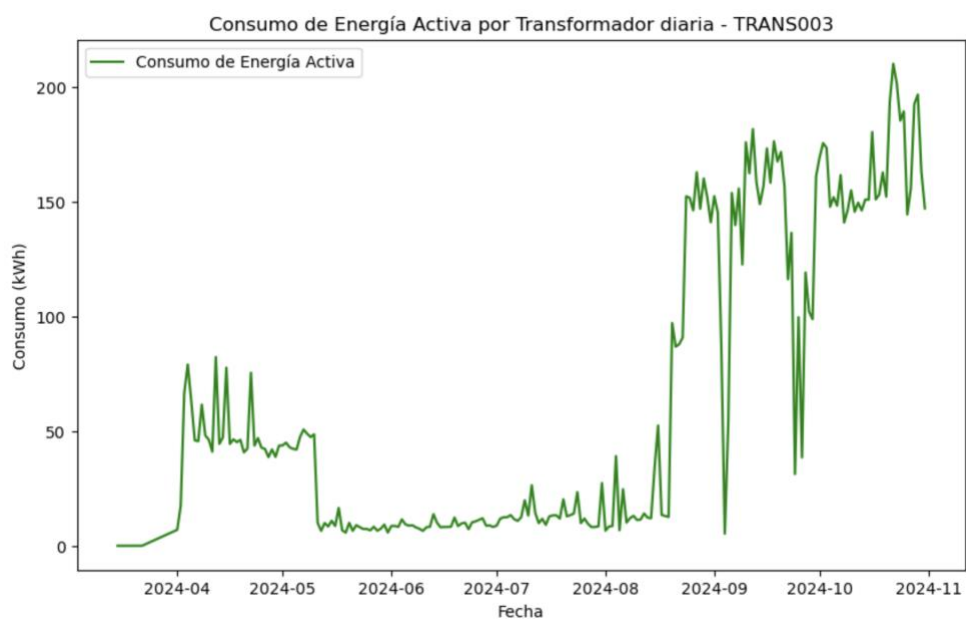


Figura 4

Consumo Energía Activa Importada Mes a Mes TRANS002

**Figura 5**

Consumo Energía Activa Importada Mes a Mes TRANS003



Modelos de Predicción de Consumo Energético Diario

En este capítulo se detalla el desarrollo de modelos de predicción del consumo diario de energía para los tres transformadores ubicados en Salento, Quindío, utilizando técnicas de aprendizaje automático y métodos estadísticos. Entre los enfoques implementados se encuentran ARIMA, Redes Neuronales Recurrentes (*RNN*), *Prophet* y el modelo de Suavizamiento Exponencial (*Exponential Smoothing*). Cada técnica fue aplicada tras un preprocesamiento de los datos, que incluyó remuestreo diario y selección de hiperparámetros. Se analizaron métricas como MAE, MSE y R^2 para evaluar la precisión y robustez de los modelos.

Para abordar la problemática planteada, se aplica un remuestreo diario con los datos del consumo energético, agrupando y sumando los valores del intervalo de la energía activa importada para cada día. Con lo anterior, se cambia la frecuencia original de los datos a un intervalo diario y se genera un nuevo *dataframe* para cada transformador.

Posteriormente, se implementan los modelos en el lenguaje de programación python *Exponential Smoothing*, ARIMA, Redes Neuronales Recurrentes (*RNN*) y *Prophet* sobre tres subconjuntos de datos asociados a los transformadores: TRANS001, TRANS002 y TRANS003 para un periodo comprendido entre el 01/03/2024 al 31/10/2024. Se realiza conversión del índice Fecha al formato *Datetime* para asegurar un manejo adecuado de los modelos de series temporales.

Visualización Datos con Remuestreo

Se presenta la visualización de los datos de consumo energético para los transformadores seleccionados (TRANS001, TRANS002 y TRANS003) con un remuestreo diario. En esta visualización, los datos horarios originales han sido agrupados y sumados para cada día, permitiendo observar las fluctuaciones y patrones en el consumo energético diario desde marzo

hasta noviembre de 2024. Este proceso de remuestreo ayuda a simplificar el análisis, proporcionando una perspectiva clara de las variaciones de consumo a lo largo del tiempo.

En los gráficos, se puede observar cómo cada transformador muestra patrones distintos. TRANS001 presenta un comportamiento relativamente estable con un leve aumento en el consumo a partir de septiembre de 2024 como se observa en la Figura 6. TRANS002 y TRANS003, por su parte, exhiben incrementos más significativos y periodos de alta variabilidad, especialmente en los últimos meses, lo cual podría estar relacionado con cambios en la base de clientes o con condiciones de demanda específicas como se observa en la Figura 7 y la Figura 8. Esta visualización es útil para identificar tendencias generales, picos de consumo y posibles eventos anómalos en la serie temporal con remuestreo de cada transformador.

Figura 6

Consumo Energía Activa Importada Mes a Mes con Remuestreo TRANS001

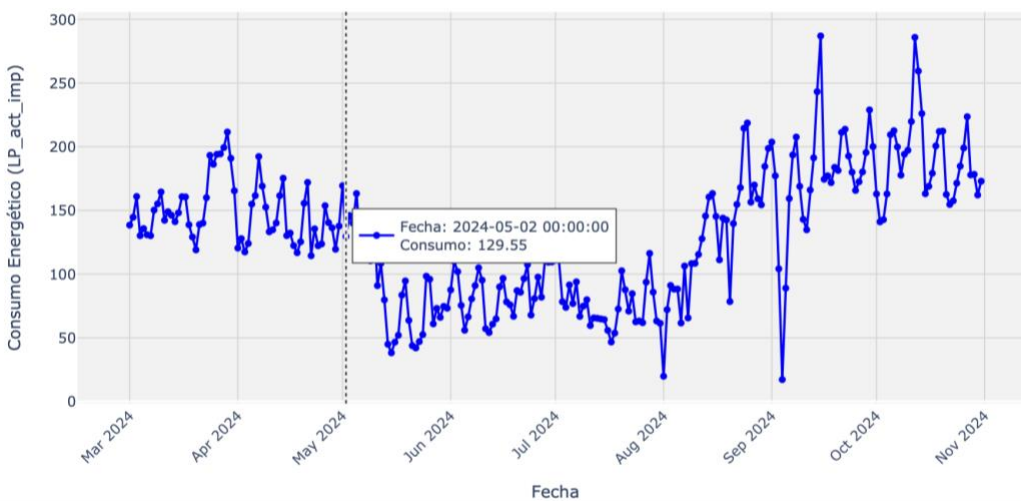
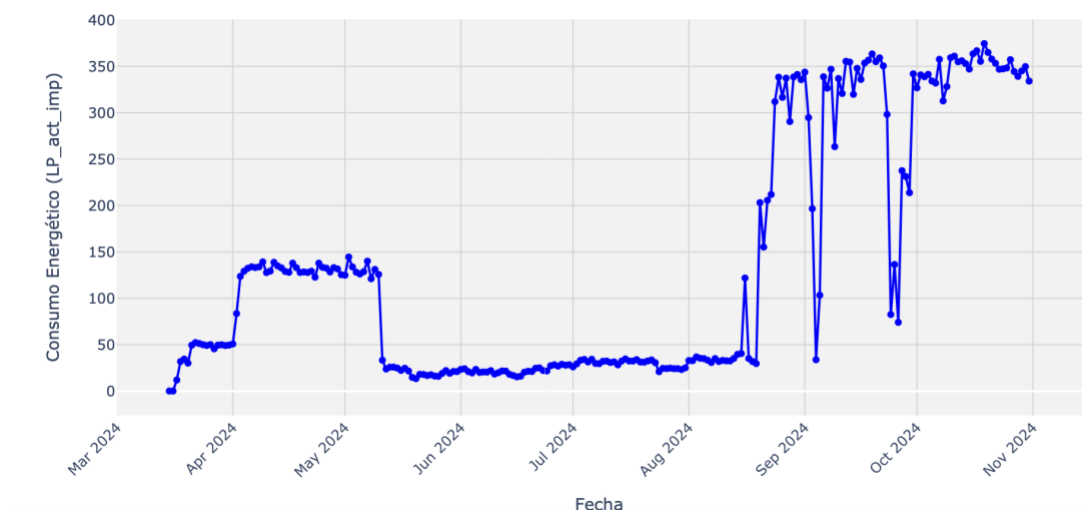
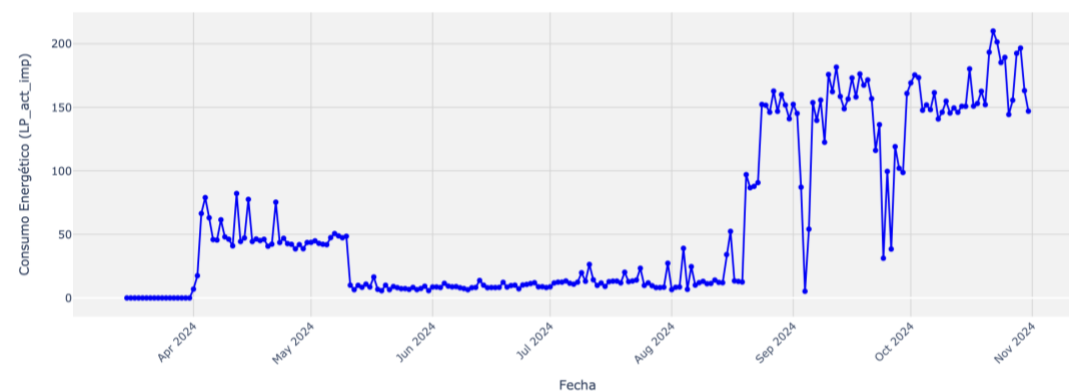


Figura 7

Consumo Energía Activa Importada Mes a Mes con Remuestreo TRANS002

**Figura 8**

Consumo Energía Activa Importada Mes a Mes con Remuestreo TRANS003



Descomposición de la Serie Temporal

Se lleva a cabo una descomposición de la serie temporal del consumo energético, analizando los componentes de tendencia, estacionalidad y residuos para cada transformador

(TRANS001, TRANS002 y TRANS003). Este desglose proporciona información para el ajuste de modelos predictivos en etapas posteriores del análisis.

En la Figura 9, se observa un análisis detallado de los componentes de la serie temporal para este transformador. La serie original muestra un consumo relativamente estable a lo largo del periodo, sin un aumento marcado en la segunda mitad de 2024, lo cual es consistente con el hecho de que TRANS001 mantuvo una cantidad constante de clientes. El componente de tendencia refleja una leve variabilidad, pero sin un crecimiento sostenido. El componente estacional muestra fluctuaciones regulares y repetitivas en el consumo diario, indicando patrones que se mantienen estables en el tiempo. Finalmente, los residuos presentan pocas fluctuaciones, lo que sugiere que el comportamiento de TRANS001 es bastante consistente y predecible debido a la estabilidad en su base de clientes.

Figura 9

Descomposición de la Serie Temporal para el Transformador TRANS001



En la Figura 10, se observa un análisis detallado de los componentes de la serie temporal para este transformador. La serie original muestra un aumento significativo en el consumo energético a partir de la segunda mitad de 2024, lo cual está relacionado con un incremento en la cantidad de clientes conectados a este transformador. El componente de tendencia evidencia un aumento sostenido, reforzando la idea de una mayor demanda energética de más clientes en este periodo. La estacionalidad muestra patrones regulares y consistentes, aunque con una intensidad que parece haber aumentado en la segunda mitad del año. Los residuos en TRANS002 presentan mayores fluctuaciones, especialmente después de septiembre de 2024, lo que podría estar asociado con la demanda adicional de sus nuevos clientes.

Figura 10

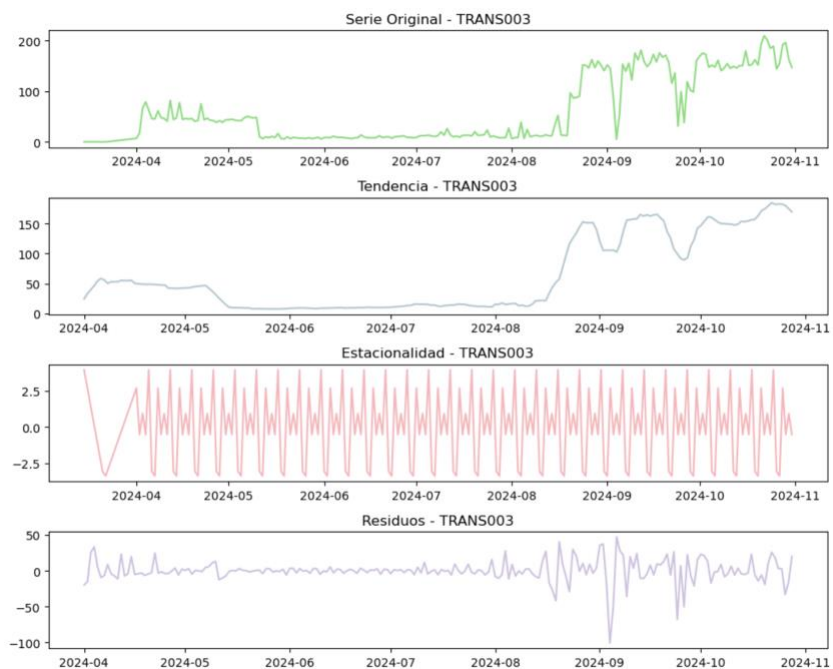
Descomposición de la Serie Temporal para el Transformador TRANS002



En la Figura 11, se observa un análisis detallado de los componentes de la serie temporal para este transformador. La serie original muestra un patrón similar al de TRANS002, con un aumento notable en el consumo energético en la segunda mitad de 2024, posiblemente debido a la expansión en la cantidad de clientes. El componente de tendencia también evidencia este incremento, reflejando una mayor demanda de energía. El componente estacional revela fluctuaciones regulares, con un patrón que se ha vuelto más constante y pronunciado hacia finales de 2024. En cuanto a los residuos, se observan mayores fluctuaciones después de la incorporación de nuevos clientes.

Figura 11

Descomposición de la Serie Temporal para el Transformador TRANS003



Modelo Exponential Smoothing

El modelo de Suavizamiento Exponencial Triple (también conocido como el método de *Holt-Winters*) es una técnica estadística de pronóstico utilizada para predecir series temporales

que presentan patrones tanto de tendencia como de estacionalidad. Este modelo es una extensión del suavizamiento exponencial simple y del suavizamiento doble, ya que incorpora tres componentes principales: nivel, tendencia y estacionalidad. La característica clave del Suavizamiento Exponencial Triple es que asigna un peso exponencialmente decreciente a los valores más antiguos, de modo que los datos recientes tienen mayor influencia en las predicciones, mientras que los datos más antiguos tienen menor peso. Esta propiedad es útil para capturar patrones recientes en el comportamiento de la serie, especialmente en series con cambios estacionales o tendencia variable (Hyndman & Athanasopoulos, 2021).

En este proyecto, el Suavizamiento Exponencial Triple fue implementado para pronosticar el consumo energético diario de tres transformadores, con el objetivo de identificar tanto las fluctuaciones estacionales diarias como las tendencias generales en el consumo. Se configuró el modelo para capturar la tendencia aditiva y una estacionalidad multiplicativa diaria, lo que permite que el modelo se adapte a cambios en los patrones de consumo cada día. Esta configuración es especialmente útil en contextos donde el consumo de energía tiene patrones predecibles que se repiten diariamente.

Este modelo es adecuado para analizar el consumo de energía en series temporales diarias, ya que permite descomponer la serie en sus componentes y realizar pronósticos basados en la tendencia y estacionalidad observadas. La técnica resulta útil en el análisis de datos de consumo energético, ya que muchos sistemas de energía presentan ciclos de consumo diarios, como patrones de uso diurno y nocturno. La capacidad de capturar estos patrones con el Suavizamiento Exponencial Triple facilita el ajuste de modelos que ayudan a predecir la demanda de energía y a gestionar eficientemente el consumo diario en redes eléctricas.

El modelo fue implementado con la librería statsmodels en Python, dividiendo los datos en un conjunto de entrenamiento (80%) y un conjunto de prueba (20%). Tras ajustar el modelo con los datos de entrenamiento, se evaluaron las predicciones en el conjunto de prueba utilizando métricas como el error absoluto medio (MAE), el error cuadrático medio (MSE), la raíz del error cuadrático medio (RMSE) y el coeficiente de determinación (R^2). Estas métricas permitieron evaluar el desempeño del modelo en cada transformador y su capacidad para aproximarse a los valores reales.

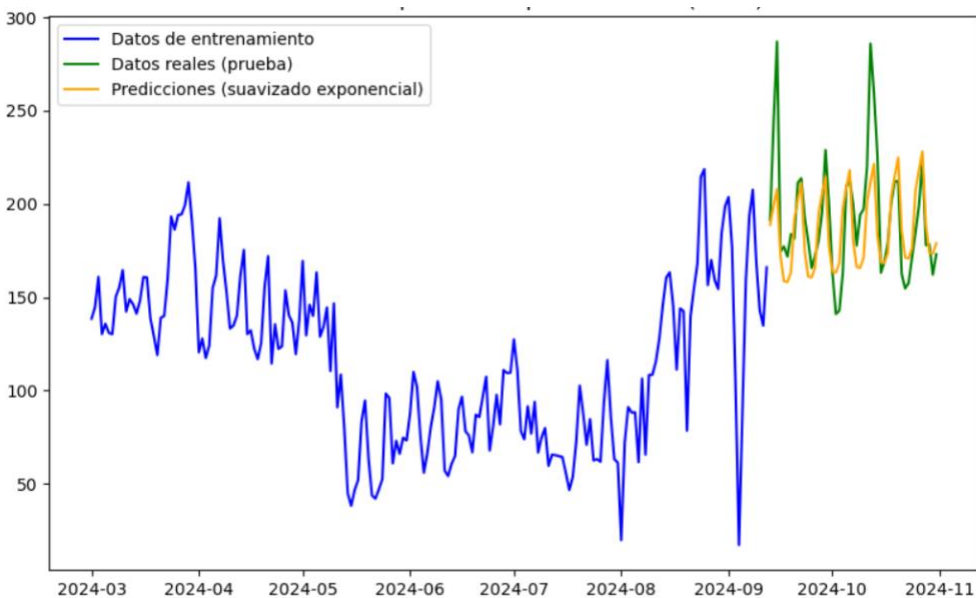
A continuación, se presenta un análisis de los resultados para cada transformador:

Transformador TRANS001

En la Figura 12 se destaca los datos de entrenamiento en azul, los datos de prueba en verde y las predicciones generadas mediante el modelo de Suavizamiento Exponencial Triple (línea naranja). Durante el periodo de entrenamiento, se observa una fluctuación moderada en los valores, seguida de un comportamiento más estable antes del incremento brusco alrededor de septiembre de 2024. En los datos de prueba, se aprecian picos que reflejan una tendencia ascendente en el consumo. El modelo capta esta tendencia general y algunos de los patrones estacionales presentes, como se observa en la cercanía de la línea naranja a la verde. Sin embargo, el modelo presenta ligeros desfases en la captura de los picos más pronunciados, reflejando una precisión moderada en las predicciones para TRANS001. Este resultado es consistente con las métricas de rendimiento obtenidas, que indican que el modelo sigue adecuadamente la tendencia general, aunque con limitaciones para capturar fluctuaciones abruptas en el consumo diario.

Figura 12

Predicción del Consumo Energético Diario para TRANS001 - Suavizamiento Exponencial Triple



Los valores obtenidos para las métricas de evaluación son:

- MAE: 16.95
- MSE: 554.79
- R²: 0.44

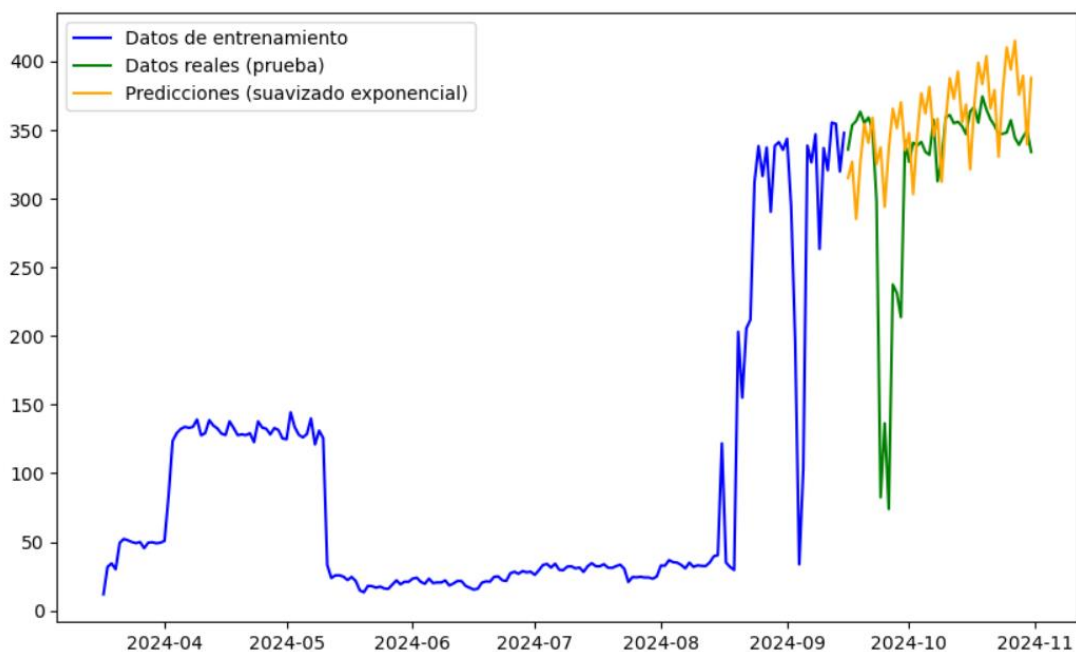
El modelo mostró un desempeño aceptable para TRANS001, con un R² positivo que sugiere una capacidad moderada para capturar la variabilidad en el consumo diario. Las métricas MAE y MSE indican que las predicciones están cercanas a los valores reales. Esto se debe a que el consumo en TRANS001 presenta un patrón diario estable que se adapta bien al modelo de suavizamiento exponencial.

Transformador TRANS002

En la Figura 13, se destacan los datos de entrenamiento en azul, los datos de prueba en verde y las predicciones generadas mediante el modelo de Suavizamiento Exponencial Triple (línea naranja). Durante el periodo de entrenamiento, se observa un aumento inicial en el consumo hasta alcanzar un nivel estable alrededor de abril de 2024, seguido de un periodo de estabilidad con pocas fluctuaciones significativas. Sin embargo, hacia septiembre de 2024, se produce un aumento brusco en el consumo, que continúa en el periodo de prueba. En los datos de prueba, los picos y variaciones reflejan una tendencia ascendente en el consumo diario. El modelo de Suavizamiento Exponencial Triple capta esta tendencia general y algunos patrones estacionales, evidenciado por la cercanía de la línea naranja a la verde en varias secciones. No obstante, el modelo presenta desfases en algunos de los picos más pronunciados, lo que indica una precisión moderada en la captura de las fluctuaciones más abruptas para TRANS002. Este desempeño es consistente con las métricas de evaluación obtenidas, que muestran que el modelo sigue adecuadamente la tendencia general, aunque presenta limitaciones al capturar variaciones extremas en el consumo diario.

Figura 13

Predicción del Consumo Energético Diario para TRANS002 - Suavizamiento Exponencial Triple



Los valores obtenidos para las métricas de evaluación son:

- MAE: 47.52
- MSE: 5612.35
- R^2 : -0.20

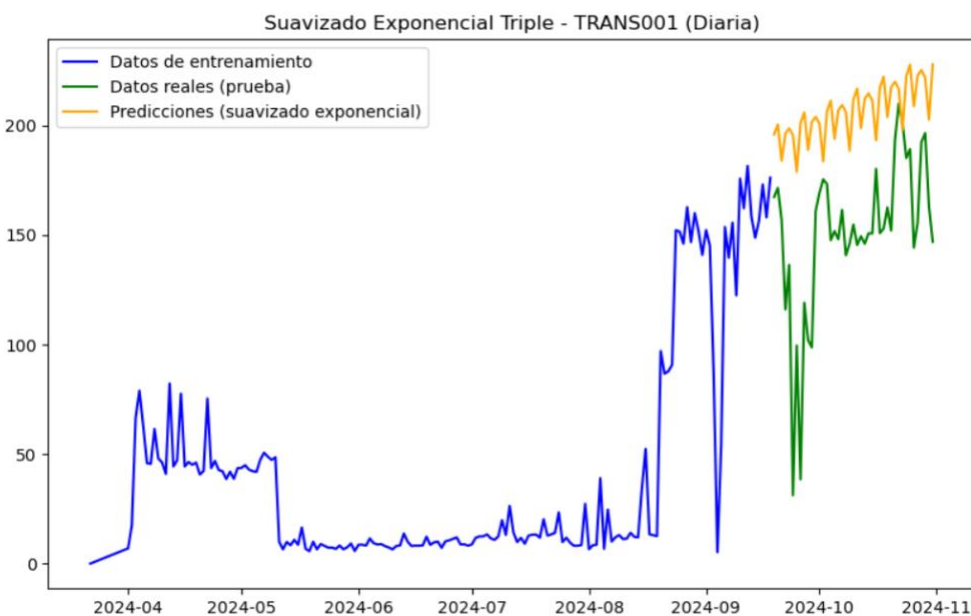
En TRANS002, el modelo mostró limitaciones significativas, reflejadas en un R^2 negativo que sugiere una capacidad limitada para capturar la variabilidad diaria. Este resultado puede deberse a que TRANS002 presenta cambios abruptos en el consumo diario debido a la variabilidad de los clientes en el transformador, los cuales el modelo de suavizamiento exponencial no logra capturar eficazmente. Este transformador presenta una variabilidad más alta, lo que limita la eficacia de este modelo en datos con fluctuaciones irregulares.

Transformador TRANS003

En la Figura 14, muestra los datos de entrenamiento en azul, los datos de prueba en verde y las predicciones generadas mediante el modelo de Suavizamiento Exponencial Triple (línea naranja). Durante el periodo de entrenamiento, se observa un consumo relativamente bajo y estable con algunas fluctuaciones iniciales, seguido de un incremento más notable hacia septiembre de 2024. En el periodo de prueba, los datos reflejan un aumento marcado en el consumo, evidenciado por picos y variaciones ascendentes. El modelo de Suavizamiento Exponencial Triple logra capturar la tendencia general de crecimiento y algunos patrones estacionales en el consumo, como se observa en la cercanía de la línea naranja a la línea verde en múltiples puntos. Sin embargo, presenta desfases en los picos más pronunciados, lo que indica una precisión limitada en la captura de las fluctuaciones más extremas para TRANS003. Este desempeño es consistente con las métricas de evaluación obtenidas, que reflejan que el modelo sigue adecuadamente la tendencia general, aunque tiene dificultades para representar cambios abruptos en el consumo diario.

Figura 14

Predicción del Consumo Energético Diario para TRANS003 - Suavizamiento Exponencial Triple



Los valores obtenidos para las métricas de evaluación son:

- MAE: 55.70
- MSE: 4198.43
- R^2 : -2.20

En el caso de TRANS003, el modelo también tuvo dificultades, con un R^2 negativo que indica que el suavizamiento exponencial no logra capturar la variabilidad en el consumo diario de este transformador. La alta fluctuación en el consumo diario de TRANS003, posiblemente influenciada por factores externos no capturados en el modelo, limita la precisión del Suavizamiento Exponencial Triple en este set de datos.

El modelo de Suavizamiento Exponencial Triple fue útil para capturar los patrones de consumo diario en TRANS001, donde la variabilidad era más estable. Sin embargo, en los transformadores TRANS002 y TRANS003, que presentaron variaciones más abruptas y patrones

menos regulares, el modelo no fue capaz de capturar adecuadamente la variabilidad en el consumo diario.

Modelo ARIMA

Se realiza implementación del modelo ARIMA para analizar el consumo energético diario para tres transformadores, aprovechando el enfoque estadístico de las series temporales para capturar patrones de estacionalidad y tendencia en los datos históricos disponibles.

En el análisis inicial se aplicó la prueba de *Dickey-Fuller* Aumentada (ADF) para verificar si la serie temporal de consumo energético cumple con el requisito de estacionalidad, condición fundamental para este modelo. Se realiza diferenciación para eliminar tendencias y estacionalidades, resultando en una serie estacionaria apta para modelado e identificando finalmente el parámetro d – diferenciación para aplicar el algoritmo (Hyndman & Athanasopoulos, 2021).

Posteriormente, mediante los métodos ACF - Función de Autocorrelación, PACF - Autocorrelación Parcial y Criterio de Información de Akaike (AIC), se identifican los valores más óptimos de los parámetros p [Parámetro Autorregresivo (AR)] cuántos valores pasados de la serie (o lags) se usan para predecir el valor actual y q [Parámetro de Promedio Móvil (MA)] número de términos de error pasados incluidos en el modelo. Lo anterior, verificando los gráficos obtenidos y diversas combinaciones para evaluar y comparar modelos en función de su ajuste y complejidad.

Se utilizó el método ARIMA (*AutoRegressive Integrated Moving Average*) de la librería statsmodels de Python, la cual facilita el ajuste y diagnóstico. Los modelos fueron optimizados a través de la selección de los parámetros (p, d, q) , ajustando el término autorregresivo, de diferenciación y de promedio móvil para minimizar el criterio de información de Akaike (AIC).

El modelo se entrenó utilizando el 80% de los datos disponibles, reservando el 20% restante para la validación, lo cual permitió monitorear el rendimiento y ajustar hiperparámetros para evitar el sobreajuste.

Para evaluar la precisión de las predicciones, se utilizaron métricas del modelo ARIMA como AIC (Criterio de Información de Akaike) que es utilizada para evaluar la calidad de un modelo estadístico al balancear el ajuste del modelo y la complejidad del mismo. Un valor de AIC más bajo indica un modelo mejor, ya que implica un ajuste adecuado con menos complejidad y la métrica BIC (Criterio de Información Bayesiano) que penaliza aún más la cantidad de parámetros del modelo, lo que lo hace más conservador. Las otras métricas usadas fueron el error absoluto medio (MAE), el error cuadrático medio (MSE), la raíz del error cuadrático medio (RMSE), el coeficiente de determinación (R^2). Estas métricas permiten evaluar la capacidad predictiva del modelo y compararla con el comportamiento real del consumo energético en cada transformador.

A continuación, se presenta un análisis de los resultados para cada transformador:

Transformador TRANS001

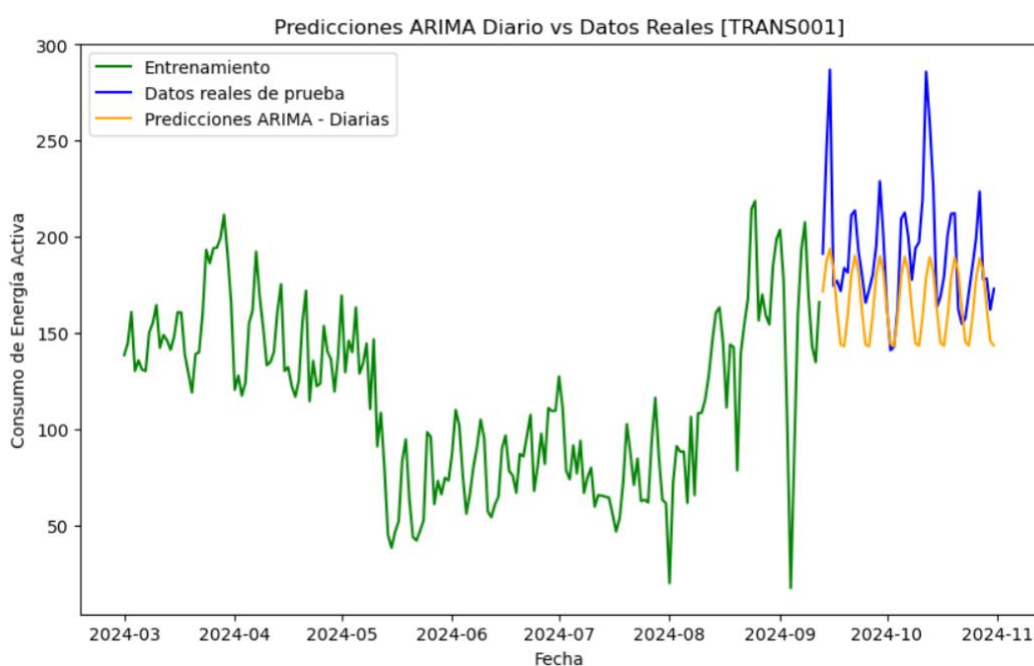
Para el transformador TRANS001, el modelo ARIMA (3,1,3) mostró ser la configuración más óptima, logrando un buen ajuste a los datos de consumo energético diario. Los parámetros seleccionados permiten captar adecuadamente las dependencias temporales en la serie y contribuyen a minimizar el error de predicción.

De la Figura 15 se puede observar para la línea de entrenamiento (línea verde) una serie con patrones de fluctuación significativos y una tendencia estacionaria en ciertas secciones, con variabilidad considerable en algunos intervalos de tiempo.

Aunque las predicciones siguen una trayectoria general parecida a los datos de prueba, se observa una tendencia conservadora y diferencias en los picos más altos; mostrando que el modelo ARIMA identifica bien patrones generales, pero podría haber un sesgo o falta de capacidad del modelo a adaptarse a cambios bruscos en la serie de tiempo sin que se especifique un componente estacional.

Figura 15

Predicciones del Modelo ARIMA (3, 1, 3) – [TRANS001]



Al ajustar un modelo ARIMA (3, 1, 3) sobre los datos de consumo energético diario, se obtuvieron los siguientes resultados clave según la Figura 16:

Los coeficientes AR (L1, L2, L3) y MA (L1, L2, L3) tienen valores significativos (p -valor = 0.000), por lo que son relevantes y contribuyen al modelo. AR.L1 y MA.L2 también tienen un impacto considerable, con coeficientes de 1.7664 y 1.9760, respectivamente.

AIC (Criterio de información de Akaike): Medida de la calidad del modelo en función de la complejidad y la bondad del ajuste. Cuanto menor sea el valor de AIC, mejor es el equilibrio entre ajuste y complejidad del modelo. El modelo tiene AIC de 1762.619 es moderadamente bajo, lo que indica que el modelo es aceptable.

BIC (Criterio de Información Bayesiano): Penaliza más a los modelos con más parámetros.

Un valor de BIC más bajo indica un modelo más adecuado. Para este modelo el BIC de 1785.530 es ligeramente superior al AIC, lo que indica que el modelo penaliza la complejidad un poco más en comparación con el AIC.

Figura 16

Resumen de Resultados del Modelo ARIMA (3, 1, 3) – [TRANS001]

SARIMAX Results						
=====						
Dep. Variable:	LP_act_imp	No. Observations:	196			
Model:	ARIMA(3, 1, 3)	Log Likelihood	-874.309			
Date:	Sun, 17 Nov 2024	AIC	1762.619			
Time:	13:05:20	BIC	1785.530			
Sample:	03-01-2024	HQIC	1771.895			
	- 09-12-2024					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

ar.L1	1.7664	0.104	17.039	0.000	1.563	1.970
ar.L2	-1.6440	0.132	-12.424	0.000	-1.903	-1.385
ar.L3	0.5168	0.106	4.888	0.000	0.310	0.724
ma.L1	-2.0377	0.094	-21.622	0.000	-2.222	-1.853
ma.L2	1.9760	0.134	14.780	0.000	1.714	2.238
ma.L3	-0.8081	0.090	-9.014	0.000	-0.984	-0.632
sigma2	449.0927	37.427	11.999	0.000	375.736	522.449
=====						
Ljung-Box (L1) (Q):		0.00	Jarque-Bera (JB):	43.84		
Prob(Q):		0.99	Prob(JB):	0.00		
Heteroskedasticity (H):		2.45	Skew:	-0.60		
Prob(H) (two-sided):		0.00	Kurtosis:	4.99		
=====						

La diferencia entre el AIC y el BIC es relativamente pequeña, mostrando que el modelo es razonablemente eficiente en términos de complejidad, pero aun así podría explorarse un modelo más simple y/o comparar con otros modelos para confirmar si es el mejor ajuste posible.

La prueba de diagnóstico Ljung-Box con un p-valor de 0.99 sugiere que no hay autocorrelación significativa en los residuos, lo cual es positivo y significa que el modelo está capturando adecuadamente las dependencias temporales de la serie. Sin embargo, la prueba Jarque-Bera con un valor p-valor de 0.00) muestra que los residuos no siguen una distribución normal, lo cual no es ideal en un ajuste de ARIMA y podría sugerir que hay aspectos de los datos que el modelo no está capturando completamente.

Las métricas de rendimiento fueron las siguientes:

- MAE: 27.95
- MSE: 1247.97
- RMSE: 35.37
- R²: -0.2620

El modelo ARIMA para el transformador TRANS001 presenta un MAE (Error Absoluto Medio) de 27.95, lo que indica que, en promedio, las predicciones se desvían en esta magnitud de las unidades de energía activa al momento de hacer las predicciones. Este valor refleja que el modelo no está capturando adecuadamente las fluctuaciones en los datos. Además, el MSE (Error Cuadrático Medio) es de 1247.97, un valor alto que existen grandes diferencias entre las predicciones y los valores reales, con algunos errores particularmente significativos. Esto refuerza la conclusión de la incapacidad del modelo para captar los picos y caídas pronunciados en el consumo de energía. Finalmente, el coeficiente de determinación (R²) de -0.2620 indica que el modelo ARIMA tiene un rendimiento muy bajo al intentar capturar la variabilidad en los datos de prueba.

Transformador TRANS002

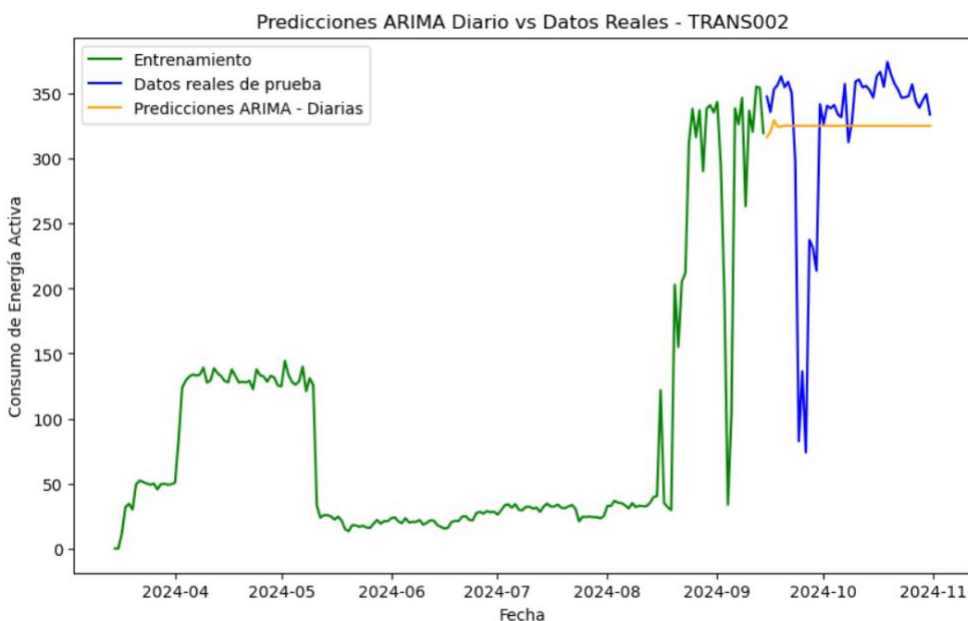
Para el transformador TRANS002, el modelo ARIMA(3,1,3) mostró ser la configuración más óptima, logrando un ajuste aceptable de los datos de consumo energético diario. Los parámetros seleccionados permiten captar adecuadamente las dependencias temporales en la serie y contribuyen a minimizar el error de predicción.

En la Figura 17 se puede observar para la línea de entrenamiento una serie con patrones de fluctuación significativos. Se observa un aumento sostenido en el consumo hasta abril/2024, seguido de una caída abrupta que se mantiene estable hasta agosto/2024. A partir de este periodo hay un incremento muy abrupto en el consumo, con variaciones significativas mostrando cambios en el patrón de consumo de energía.

Las predicciones muestran una tendencia más estable en comparación con los datos reales, pero se mantienen en un nivel muy constante, sin capturar adecuadamente los picos que ocurren en los datos reales. Por lo anterior, el modelo no está capturando adecuadamente la dinámica del consumo de energía en el transformador TRANS002.

Figura 17

Predicciones del Modelo ARIMA (3, 1, 3) – [TRANS002]



Al ajustar un modelo ARIMA (3, 1, 3) sobre los datos de consumo energético diario, se obtuvieron los siguientes resultados clave como se ve en la Figura 18.

El coeficiente AR (L1) con ($p < 0.05$) tiene un impacto negativo y considerable en el primer rezago. Los coeficientes L2, L3 no son tan significativos. Por su parte, MA (L1, L2, L3) son significativos ($p < 0.05$) por lo que los errores de predicción de los rezagos anteriores tienen un efecto notable en el ajuste del modelo, contribuyendo en la captura de la estructura de la serie temporal.

AIC (Criterio de información de Akaike): Medida de la calidad del modelo en función de la complejidad y la bondad del ajuste. Cuanto menor sea el valor de AIC, mejor es el equilibrio entre ajuste y complejidad del modelo. El modelo tiene AIC de 1770.268, es moderadamente bajo, lo que indica que el modelo es aceptable.

Figura 18*Resumen de Resultados del Modelo ARIMA (3, 1, 3) – [TRANS002]*

SARIMAX Results						
=====						
Dep. Variable:	LP_act_imp	No. Observations:	184			
Model:	ARIMA(3, 1, 3)	Log Likelihood	-878.134			
Date:	Sun, 17 Nov 2024	AIC	1770.268			
Time:	15:32:13	BIC	1792.735			
Sample:	03-15-2024	HQIC	1779.375			
	- 09-14-2024					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

ar.L1	-0.4420	0.113	-3.894	0.000	-0.664	-0.219
ar.L2	-0.2371	0.135	-1.755	0.079	-0.502	0.028
ar.L3	-0.0369	0.125	-0.295	0.768	-0.282	0.208
ma.L1	0.4833	0.113	4.287	0.000	0.262	0.704
ma.L2	0.2212	0.148	1.494	0.135	-0.069	0.511
ma.L3	-0.4047	0.125	-3.245	0.001	-0.649	-0.160
sigma2	857.2839	45.648	18.780	0.000	767.816	946.752
=====						
Ljung-Box (L1) (Q):		0.01	Jarque-Bera (JB):	1578.81		
Prob(Q):		0.92	Prob(JB):	0.00		
Heteroskedasticity (H):		6.90	Skew:	0.69		
Prob(H) (two-sided):		0.00	Kurtosis:	17.32		
=====						

BIC (Criterio de Información Bayesiano): Penaliza más a los modelos con más parámetros.

Un valor de BIC más bajo indica un modelo más adecuado. Para este modelo el BIC de 1792.735 es ligeramente superior al AIC, lo que indica que el modelo penaliza la complejidad un poco más en comparación con el AIC.

La diferencia entre el AIC y el BIC es relativamente pequeña, mostrando que el modelo es razonablemente eficiente en términos de complejidad, pero aun así podría explorarse un modelo más simple y/o comparar con otros modelos para confirmar si es el mejor ajuste posible.

La prueba de diagnóstico Ljung-Box con un p-valor de 0.92 sugiere que no hay autocorrelación significativa en los residuos, lo cual es positivo y significa que el modelo está capturando adecuadamente las dependencias temporales de la serie. Sin embargo, la prueba Jarque-Bera con un valor p-valor de 0.00) muestra que los residuos no siguen una distribución

normal, lo cual no es ideal en un ajuste de ARIMA y podría sugerir que hay aspectos de los datos que el modelo no está capturando completamente.

Las métricas de rendimiento fueron las siguientes:

- MAE: 42.15
- MSE: 4599.99
- RMSE: 67.82
- R^2 : -0.00235

El modelo ARIMA presenta un MAE (Error Absoluto Medio) de 42.15, lo que indica un promedio de desviación significativo en las predicciones de consumo de energía activa diaria para el transformador TRANS002. Este valor indica que el modelo no logra capturar adecuadamente las fluctuaciones de los datos. Además, el MSE (Error Cuadrático Medio) es considerable 4599.99, mostrando que existen grandes diferencias entre las predicciones y los valores reales, y señalando errores altos en algunos puntos. Esto se observa en la gráfica XX, donde las predicciones tienden a estabilizarse en un valor promedio y no logran capturar la variabilidad de los datos de prueba. El coeficiente de determinación (R^2) de -0.00235 confirma que el modelo ARIMA tiene un rendimiento muy bajo al intentar capturar la variabilidad o patrones en los datos de prueba y se resalta la necesidad de aplicar ajustes adicionales o validar otro modelo para la predicción del consumo en este transformador.

Transformador TRANS003

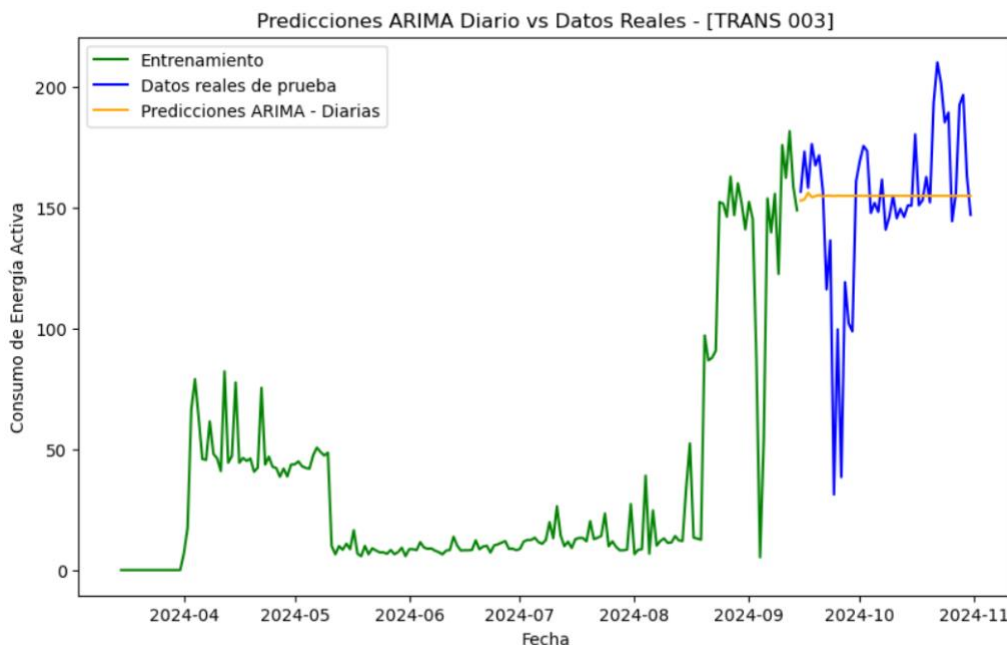
Para el transformador TRANS003, el modelo ARIMA (3,1,3) mostró ser la configuración óptima, logrando un ajuste aceptable de los datos de consumo energético diario. Los parámetros seleccionados permiten captar adecuadamente las dependencias temporales en la serie y contribuyen a minimizar el error de predicción.

En la Figura 19 se puede observar para la línea de entrenamiento (línea verde) un comportamiento variable, con picos esporádicos en el consumo de energía seguidos de períodos de estabilidad en niveles bajos. En la segunda mitad del periodo se observa un cambio brusco en el patrón consumo, que representa un desafío para este tipo de modelo.

Las predicciones muestran una tendencia más estable en comparación con los datos reales, pero se mantienen en un nivel muy constante, sin capturar adecuadamente los picos que ocurren en los datos reales. Por lo anterior, el modelo tiende a suavizar los valores y no se ajusta a la dinámica del consumo de energía en el TRANS003.

Figura 19

Predicciones del Modelo ARIMA (3, 1, 3) – [TRANS003]



Al ajustar un modelo ARIMA (3, 1, 3) sobre los datos de consumo energético diario, se obtuvieron los siguientes resultados clave como se ve en la Figura 20:

El coeficiente AR (L1) con ($p < 0.05$) tiene un impacto negativo significativo sobre la predicción actual. Los coeficientes L2, L3 no son tan significativos. Por su parte, MA (L1, L3)

son significativos ($p < 0.05$), por lo que los errores pasados tienen un efecto notable en la predicción, ayudando al modelo a corregir sus errores de pronóstico de manera más precisa.

AIC (Criterio de información de Akaike): Medida de la calidad del modelo en función de la complejidad y la bondad del ajuste. Cuanto menor sea el valor de AIC, mejor es el equilibrio entre ajuste y complejidad del modelo. El modelo tiene AIC de 1563.325, es moderadamente bajo, lo que indica que el modelo es aceptable.

BIC (Criterio de Información Bayesiano): Penaliza más a los modelos con más parámetros.

Un valor de BIC más bajo indica un modelo más adecuado. Para este modelo el BIC de 1585.791 es ligeramente superior al AIC, lo que indica que el modelo penaliza la complejidad un poco más en comparación con el AIC.

Figura 20*Resumen de Resultados del Modelo ARIMA (3, 1, 3) – [TRANS003]*

SARIMAX Results						
=====						
Dep. Variable:	LP_act_imp	No. Observations:	184			
Model:	ARIMA(3, 1, 3)	Log Likelihood	-774.662			
Date:	Sun, 17 Nov 2024	AIC	1563.325			
Time:	15:59:36	BIC	1585.791			
Sample:	03-15-2024	HQIC	1572.432			
	- 09-14-2024					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

ar.L1	-0.7217	0.125	-5.769	0.000	-0.967	-0.477
ar.L2	-0.3246	0.247	-1.316	0.188	-0.808	0.159
ar.L3	0.0740	0.148	0.501	0.616	-0.216	0.364
ma.L1	0.6080	0.099	6.170	0.000	0.415	0.801
ma.L2	0.1230	0.210	0.587	0.557	-0.288	0.534
ma.L3	-0.4729	0.152	-3.113	0.002	-0.771	-0.175
sigma2	277.0595	16.907	16.387	0.000	243.922	310.197
=====						
Ljung-Box (L1) (Q):		0.00	Jarque-Bera (JB):	555.69		
Prob(Q):		0.96	Prob(JB):	0.00		
Heteroskedasticity (H):		2.83	Skew:	0.49		
Prob(H) (two-sided):		0.00	Kurtosis:	11.48		
=====						

La diferencia entre el AIC y el BIC es relativamente pequeña, mostrando que el modelo es razonablemente eficiente en términos de complejidad, pero aun así podría explorarse un modelo más simple y/o comparar con otros modelos para confirmar si es el mejor ajuste posible.

La prueba de diagnóstico Ljung-Box con un p-valor de 0.96 sugiere que no hay autocorrelación significativa en los residuos, lo cual es positivo y significa que el modelo está capturando adecuadamente las dependencias temporales de la serie. Sin embargo, la prueba Jarque-Bera con un valor p-valor de 0.00) muestra que los residuos no siguen una distribución normal, lo cual no es ideal en un ajuste de ARIMA y podría sugerir que hay aspectos de los datos que el modelo no está capturando completamente.

Las métricas de rendimiento fueron las siguientes:

- MAE: 22.69

- MSE: 1208.87
- RMSE: 34.77
- R^2 : -0.00706

El modelo ARIMA presenta un MAE (Error Absoluto Medio) de 22.69, indicando que las predicciones de consumo energético diario se desvían, en promedio, en esta magnitud de las unidades de energía activa, por lo que se infiere que el modelo tiene dificultades para capturar con precisión las fluctuaciones en los datos. Además, el MSE (Error Cuadrático Medio) es alto 1208.87, mostrando grandes diferencias entre las predicciones y los valores reales, con algunos errores particularmente grandes. Finalmente, el Coeficiente de Determinación (R^2) es de -0.0070, lo que indica un rendimiento muy bajo en la captura de la variabilidad o patrones en los datos de prueba y no logra captar adecuadamente los picos y caídas pronunciados en el consumo de energía, resaltando la necesidad de ajustes adicionales para mejorar la precisión del modelo.

Modelo Redes Neuronales Recurrentes (RNN)

El modelo *RNN* fue diseñado utilizando capas de *LSTM* (*Long Short-Term Memory*) con el objetivo de capturar las dependencias a largo plazo presentes en la serie temporal de datos de consumo energético. Las capas *LSTM* son adecuadas para modelar series temporales complejas debido a su capacidad de retener información pasada y adaptarse a patrones secuenciales. Para evitar el sobreajuste, se añadieron capas de *Dropout* después de cada capa *LSTM*, eliminando aleatoriamente un porcentaje de las neuronas durante el entrenamiento (Goodfellow et al, 2016). Esta estructura, mostrada en la Figura 21, proporciona un modelo robusto capaz de aprender de grandes cantidades de datos sin perder su capacidad de generalización.

Figura 21

Detalles del Modelo LSTM: Capas, Parámetros y Dimensiones de Salida

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 10, 50)	10,400
dropout (Dropout)	(None, 10, 50)	0
lstm_1 (LSTM)	(None, 50)	20,200
dropout_1 (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51

Total params: 30,651 (119.73 KB)

Trainable params: 30,651 (119.73 KB)

Non-trainable params: 0 (0.00 B)

Este modelo fue implementado utilizando la librería *TensorFlow* y su API de alto nivel, *Keras*, la cual facilita la construcción y entrenamiento de redes neuronales profundas.

TensorFlow, desarrollado por Google, es ampliamente utilizado en aplicaciones de aprendizaje profundo debido a su capacidad para manejar grandes volúmenes de datos y optimizar redes neuronales complejas, como las *RNN* y *LSTM*. En este proyecto, *TensorFlow* proporcionó la flexibilidad necesaria para definir una arquitectura de red adecuada y optimizar los parámetros del modelo de forma eficiente, aprovechando también su capacidad para ejecutarse en GPU y mejorar el rendimiento en el entrenamiento.

El proceso de preprocesamiento comenzó con la normalización de los datos, escalando el rango de valores entre 0 y 1. Esto asegura que todas las características tengan el mismo peso durante el entrenamiento. A continuación, se crearon secuencias de datos de entrada utilizando una ventana de 10 periodos, es decir, cada entrada al modelo corresponde a los valores de consumo de los últimos 10 días. El modelo se entrenó utilizando el 80% de los datos disponibles,

reservando el 20% restante para la validación, lo cual permitió monitorear el rendimiento y ajustar hiperparámetros para evitar el sobreajuste.

La optimización del modelo se realizó minimizando el error cuadrático medio (MSE), ajustando los pesos de la red en función de los errores de predicción observados durante el entrenamiento. La elección de MSE como función de pérdida es adecuada para problemas de regresión, ya que penaliza los errores grandes, forzando al modelo a mejorar la precisión.

Los resultados obtenidos con el modelo *RNN* demostraron un desempeño variable entre los diferentes transformadores. Para evaluar la precisión de las predicciones, se utilizaron métricas como el error absoluto medio (MAE), el error cuadrático medio (MSE), la raíz del error cuadrático medio (RMSE), el coeficiente de determinación (R^2), y el error porcentual absoluto medio (MAPE). Estas métricas permiten evaluar la capacidad predictiva del modelo y compararla con el comportamiento real del consumo energético en cada transformador.

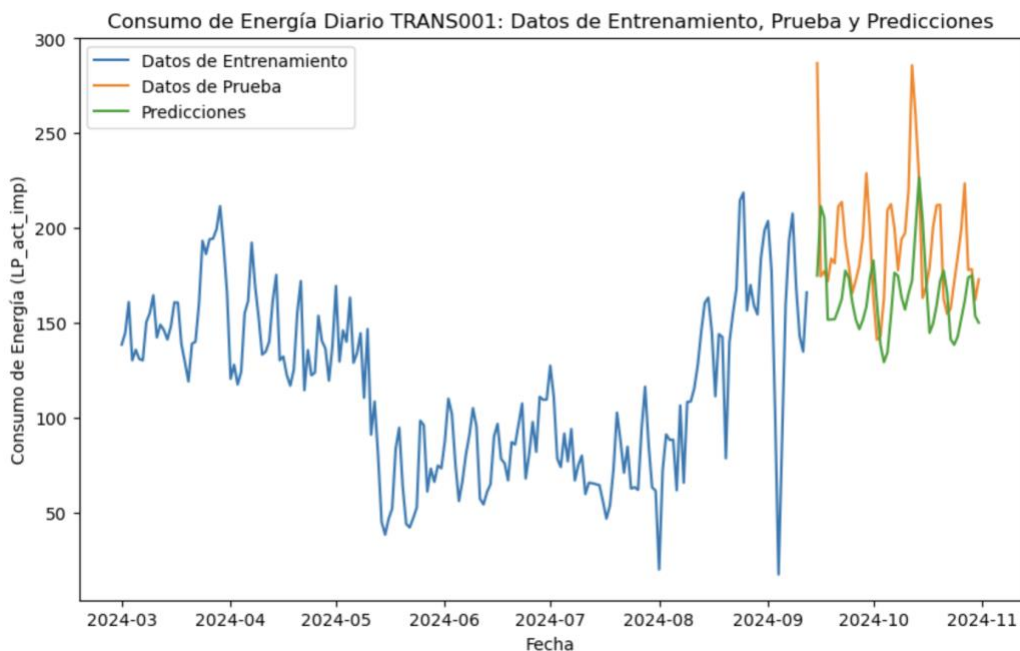
A continuación, se presenta un análisis de los resultados para cada transformador:

Transformador TRANS001

En la Figura 22, se muestra el consumo de energía diario para el transformador TRANS001, destacando los datos de entrenamiento (línea azul), los datos de prueba (línea naranja) y las predicciones del modelo (línea verde). Los datos de entrenamiento reflejan un consumo estable con fluctuaciones moderadas, mientras que los datos de prueba exhiben mayor variabilidad, con picos y caídas más pronunciados. Las predicciones del modelo siguen la tendencia general, pero presentan ciertos desfases en los valores extremos, lo que indica una precisión moderada en la captura de las fluctuaciones abrupta.

Figura 22

Predicción del Consumo Energético Diario para TRANS001 utilizando RNN



Los valores obtenidos para las métricas de evaluación son:

- MAE: 32.54
- MSE: 1662.03
- RMSE: 40.77
- R²: -0.7053
- MAPE: 16.04%

Estos resultados indican una precisión moderada en las predicciones para TRANS001. La baja puntuación en R² sugiere que el modelo no logra capturar completamente la variabilidad del consumo de energía en este transformador, posiblemente debido a que este transformador tiene un patrón de consumo más estable y menos fluctuante en comparación con los otros. Sin

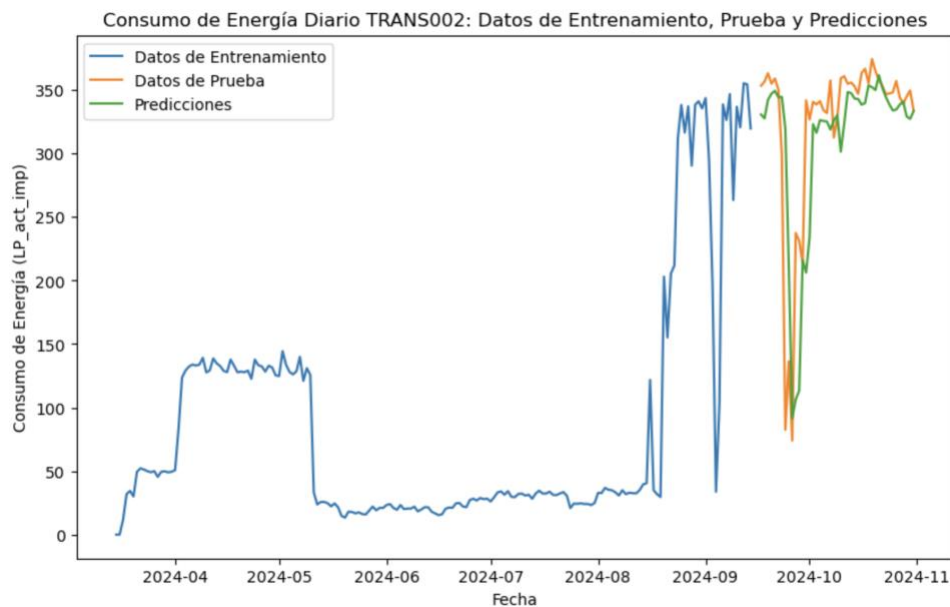
embargo, los valores de MAE y MAPE muestran que las predicciones están relativamente cerca de los valores reales en promedio, aunque existen ciertas limitaciones en la precisión.

Transformador TRANS002

En Figura 23, se presenta el consumo de energía diario para el transformador TRANS002, mostrando los datos de entrenamiento (línea azul), los datos de prueba (línea naranja) y las predicciones del modelo (línea verde). Los datos de entrenamiento muestran un incremento gradual hasta alcanzar un nivel estable, mientras que los datos de prueba reflejan fluctuaciones más pronunciadas y una serie de picos elevados. Las predicciones del modelo logran capturar la tendencia general y siguen de cerca las variaciones de los datos de prueba, aunque con algunos desfases en los cambios abruptos.

Figura 23

Predicción del Consumo Energético Diario para TRANS002 utilizando RNN



Las métricas de rendimiento fueron las siguientes:

- MAE: 30.87

- MSE: 3018.50
- RMSE: 54.94
- R²: 0.3681
- MAPE: 15.73%

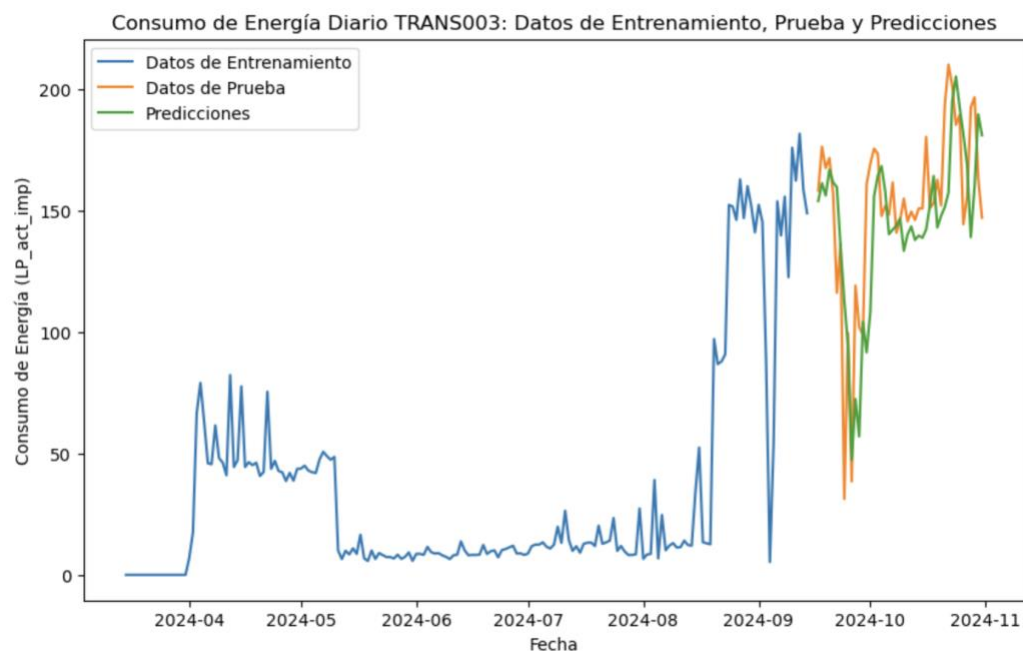
El valor de R² de 0.3681 indica que el modelo captura mejor la variabilidad en el consumo de energía para TRANS002 en comparación con TRANS001. Este transformador muestra una mayor fluctuación en el consumo, lo que permite al modelo ajustar mejor los patrones de consumo. Las métricas MAE y MAPE reflejan una precisión adecuada tanto en términos absolutos como relativos, lo que sugiere que el modelo es más preciso en la predicción del consumo para este transformador, posiblemente debido a la mayor variabilidad y fluctuación en los datos de este transformador.

Transformador TRANS003

En la Figura 24, se muestra el consumo de energía diario para el transformador TRANS003, con los datos de entrenamiento (línea azul), los datos de prueba (línea naranja) y las predicciones del modelo (línea verde). Durante el periodo de entrenamiento, se observa una fase inicial de variabilidad seguida de un periodo estable de bajo consumo. En los datos de prueba, se presentan incrementos significativos en el consumo, reflejando una tendencia ascendente. El modelo logra seguir de cerca esta tendencia en sus predicciones, aunque con ligeros desfases en algunos picos.

Figura 24

Predicción del Consumo Energético Diario para TRANS003 utilizando RNN



Las métricas de rendimiento fueron las siguientes:

- MAE: 21.29
- MSE: 857.99
- RMSE: 29.29
- R^2 : 0.3099
- MAPE: 18.73%

Para el TRANS003, el modelo logró una precisión aceptable, con un RMSE de 29.29 y un MAE relativamente bajo. El valor de R^2 (0.3099) muestra una capacidad moderada del modelo para capturar la variabilidad en los datos. Aunque el valor de MAPE es un poco mayor que el de los otros transformadores, indicando un mayor error porcentual en promedio, las

métricas globales muestran que las predicciones están bastante alineadas con los datos reales, aunque sigue existiendo espacio para mejorar la capacidad predictiva del modelo.

El modelo *RNN*, basado en *LSTM*, mostró una capacidad razonable para predecir el consumo de energía en los transformadores seleccionados. Aunque el desempeño varió entre transformadores, los resultados globales indican que el modelo puede capturar la tendencia general y algunos patrones específicos de consumo. Las gráficas de comparación entre los datos reales y las predicciones sugieren que el modelo tiene una precisión aceptable, especialmente para TRANS002 y TRANS003, que presentan una mayor variabilidad en el consumo.

Para TRANS001, que mantiene un consumo relativamente estable, el modelo tuvo más dificultades para capturar variaciones, lo que se refleja en un valor de R^2 negativo. Esto sugiere que otros modelos pueden adaptarse mejor a patrones de consumo más estables.

Modelo Prophet

El modelo *Prophet* es una herramienta de pronóstico de series temporales desarrollada por Facebook, diseñada para captar patrones complejos de estacionalidad y tendencia en series de datos de alta frecuencia y de crecimiento constante. *Prophet* permite manejar de forma intuitiva factores como estacionalidades diarias, semanales y anuales, así como tendencias de largo plazo y días festivos o eventos especiales que puedan afectar los datos. Este modelo es robusto frente a datos faltantes y valores atípicos, lo cual lo hace especialmente útil para series temporales de consumo energético, donde puede existir variabilidad significativa en patrones de uso (Taylor & Letham, 2018).

En este proyecto, el modelo *Prophet* fue aplicado para pronosticar el consumo diario de energía de tres transformadores, capturando tanto la tendencia de fondo como los patrones estacionales presentes en los datos. *Prophet* permite descomponer la serie temporal en

componentes de tendencia, estacionalidad y eventos adicionales, proporcionando una visualización clara de los factores que influyen en el consumo energético diario. Esta capacidad de ajuste es ideal para series con estacionalidades variables o múltiples ciclos en diferentes horizontes de tiempo.

Para este análisis, *Prophet* se utilizó con el fin de predecir el consumo energético diario de tres transformadores. Se partió de un conjunto de datos que incluía la fecha y los valores de consumo energético, y se adaptó el formato de acuerdo con los requerimientos de *Prophet*, renombrando las columnas como ds (para la fecha) e y (para el consumo energético).

En la implementación, se configuraron los componentes estacionales diarios y semanales de *Prophet*, con el objetivo de capturar las variaciones en el consumo energético en función de los ciclos temporales de cada transformador. *Prophet* permite realizar una predicción diaria, y se generó un horizonte de pronóstico de 30 días adicionales para evaluar la precisión en datos no vistos.

El modelo fue implementado en Python utilizando la librería *Prophet*, que facilitó la construcción y ajuste del modelo. Pandas se empleó para la manipulación y limpieza de los datos, asegurando que cumplieran con los requisitos de *Prophet*, mientras que Matplotlib se utilizó para graficar tanto las predicciones del modelo como los componentes de tendencia y estacionalidad. Esto proporcionó una visualización clara de los patrones capturados y permitió un análisis de los resultados.

Para evaluar la precisión de *Prophet*, se calcularon las métricas de error absoluto medio (MAE), error cuadrático medio (MSE), y el coeficiente de determinación (R^2). También se utilizó el estadístico *Durbin-Watson* para evaluar la autocorrelación en los residuos, una medida clave para entender la independencia de los errores del modelo. Estas métricas permiten evaluar

la capacidad predictiva del modelo y determinar qué tan bien se alinea con el comportamiento real de consumo de energía en cada transformador.

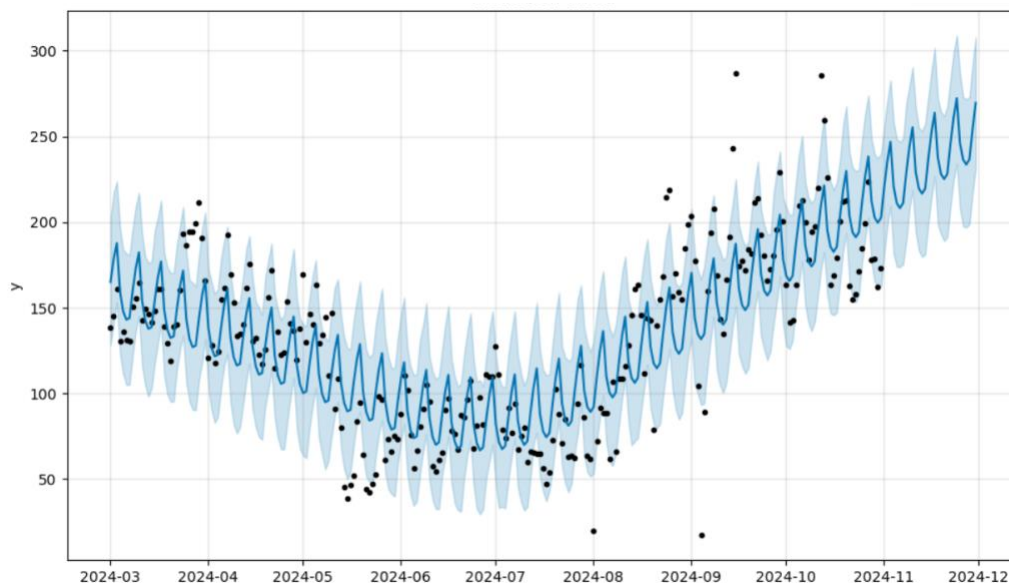
Transformador TRANS001

En la Figura 25, se observa la tendencia diaria del consumo energético en una línea azul continua, con un intervalo de confianza sombreado en azul claro que representa la incertidumbre en la predicción. Los puntos negros corresponden a los valores reales de consumo energético diario. A lo largo del periodo de marzo a diciembre de 2024, se puede identificar una tendencia general de aumento en el consumo. Inicialmente, el consumo desciende levemente hasta mediados de año (junio-julio), momento en el que empieza a incrementarse de manera sostenida. Este patrón sugiere la presencia de estacionalidad diaria, marcada por picos y valles cíclicos en el consumo, que el modelo captura adecuadamente a través de los ciclos regulares de la línea de predicción.

El intervalo de confianza refleja la capacidad del modelo para manejar variaciones en los datos. Aunque en algunos puntos los valores reales (puntos negros) caen fuera del intervalo de predicción, especialmente en los momentos de mayor variabilidad, la línea de tendencia general sigue de cerca el comportamiento real, mostrando la habilidad de *Prophet* para capturar patrones estacionales y la tendencia ascendente en el consumo diario de TRANS001 hacia finales de 2024.

Figura 25

Predicción del Consumo Energético Diario para TRANS001 utilizando PROPHET



Las métricas de evaluación para TRANS001 son:

- MAE: 23.37
- MSE: 29.26
- R^2 : 0.681
- Durbin-Watson: 0.531

El MAE de 23.37 indica que el modelo presenta un error promedio moderado, sugiriendo que las predicciones están generalmente cerca de los valores reales en términos absolutos. El MSE de 29.26, por otro lado, refuerza este desempeño al reflejar que los errores grandes son relativamente infrecuentes, penalizando los desvíos extremos de manera significativa y subrayando la consistencia del modelo.

El coeficiente de determinación (R^2) de 0.681 sugiere que *Prophet* es capaz de explicar aproximadamente el 68.1% de la variabilidad en el consumo de energía de TRANS001, lo cual

es un indicador positivo de ajuste al capturar tanto la tendencia general como los patrones estacionales. No obstante, el valor *Durbin-Watson* de 0.531 señala una ligera autocorrelación positiva en los residuos, lo que implica que existen aún algunos patrones recurrentes en el consumo que no fueron completamente capturados, especialmente en las variaciones más abruptas.

Transformador TRANS002

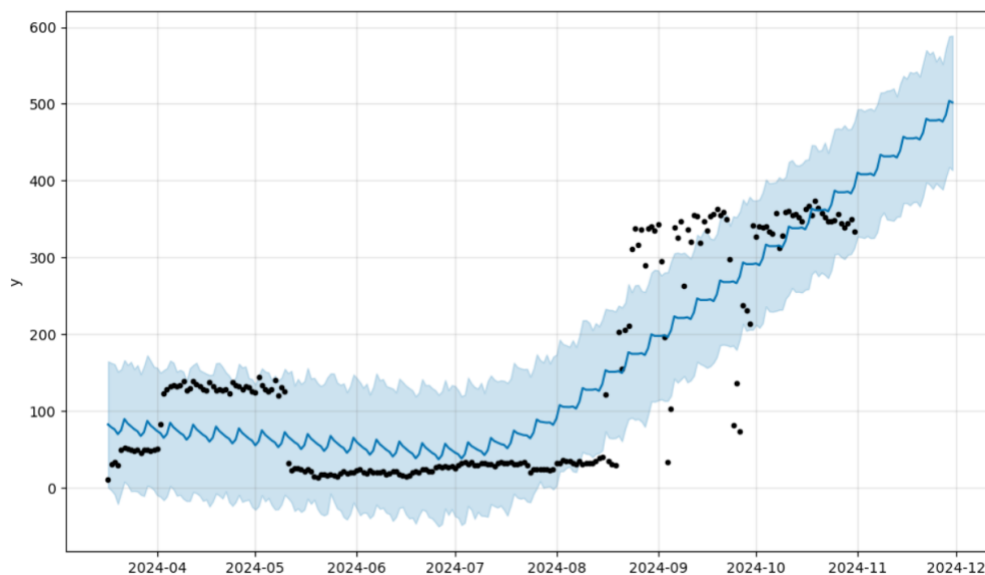
En la Figura 26, se observa la tendencia diaria del consumo energético representada por una línea azul continua, mientras que el intervalo de confianza, sombreado en azul claro, muestra la incertidumbre en la predicción del modelo. Los puntos negros corresponden a los valores reales de consumo energético diario.

A lo largo del periodo de marzo a diciembre de 2024, se puede identificar una tendencia general de aumento en el consumo. Inicialmente, el consumo experimenta una ligera disminución hasta mediados de año (junio-julio), momento en el cual empieza a incrementarse de manera constante. Este patrón sugiere la presencia de estacionalidad diaria, marcada por picos y valles cíclicos en el consumo, que el modelo captura de manera adecuada mediante los ciclos regulares de la línea de predicción.

El intervalo de confianza refleja la capacidad del modelo para manejar variaciones en los datos. Aunque en algunos puntos los valores reales (puntos negros) se sitúan fuera del intervalo de predicción, especialmente en los momentos de mayor variabilidad, la línea de tendencia general sigue de cerca el comportamiento observado. Esto muestra la habilidad de *Prophet* para capturar tanto los patrones estacionales como la tendencia ascendente en el consumo diario de TRANS002 hacia finales de 2024, reflejando una precisión adecuada en el seguimiento de los cambios diarios en el consumo energético.

Figura 26

Predicción del Consumo Energético Diario para TRANS002 utilizando PROPHET



Las métricas de evaluación para TRANS002 son:

- MAE: 53.57
- MSE: 65.12
- R^2 : 0.754
- Durbin-Watson: 0.300

El MAE de 53.57 indica que el modelo tiene un error promedio moderado, lo cual sugiere que las predicciones de *Prophet* están, en general, relativamente cerca de los valores reales en términos absolutos. El MSE de 65.12 refuerza esta interpretación al penalizar los errores grandes de manera significativa, lo que subraya que los desvíos extremos en las predicciones son limitados, reflejando una consistencia razonable en el desempeño del modelo.

El coeficiente de determinación (R^2) de 0.754 sugiere que *Prophet* es capaz de explicar aproximadamente el 75.4% de la variabilidad en el consumo de energía de TRANS002, un

indicador positivo de ajuste al capturar tanto la tendencia general como los patrones estacionales de consumo. Sin embargo, el valor *Durbin-Watson* de 0.300 señala una autocorrelación positiva considerable en los residuos, lo cual implica que el modelo no logra capturar completamente algunos patrones recurrentes, especialmente en las fluctuaciones más abruptas del consumo energético.

Transformador TRANS003

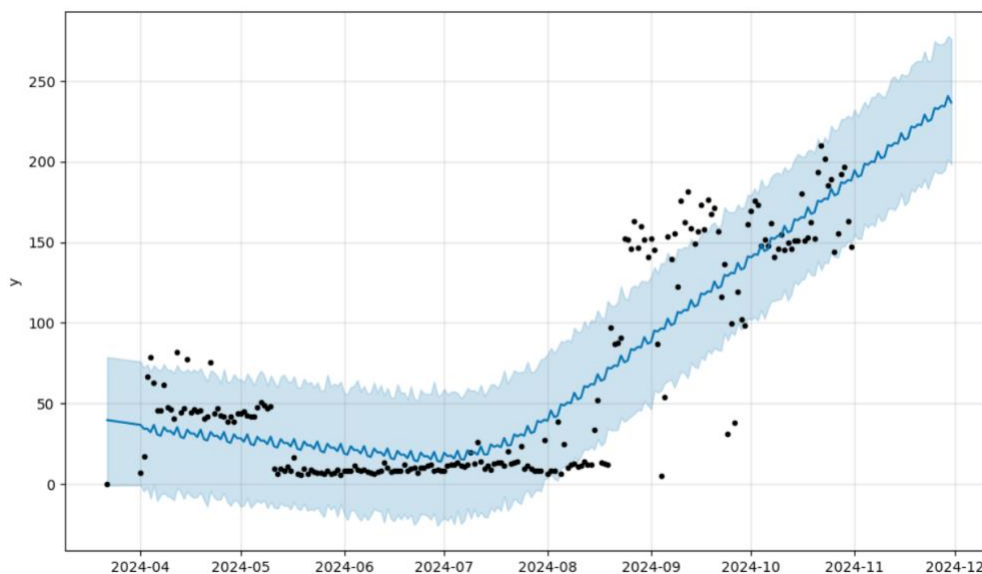
En la Figura 27, se muestra la tendencia diaria del consumo energético representada por una línea azul continua, mientras que el intervalo de confianza, sombreado en azul claro, representa la incertidumbre en las predicciones del modelo. Los puntos negros corresponden a los valores reales de consumo energético diario.

A lo largo del periodo de marzo a diciembre de 2024, se observa una tendencia general de incremento en el consumo. Inicialmente, el consumo desciende levemente hasta mitad de año (alrededor de julio), momento a partir del cual comienza un aumento progresivo y sostenido. Este patrón sugiere una estacionalidad diaria, caracterizada por picos y valles regulares, que el modelo logra capturar adecuadamente en la predicción.

El intervalo de confianza refleja la capacidad del modelo para manejar las fluctuaciones de los datos, mostrando que, aunque algunos valores reales (puntos negros) se sitúan fuera del rango de predicción en momentos de alta variabilidad, la línea de tendencia general sigue de cerca el comportamiento observado. Esto demuestra la habilidad de *Prophet* para capturar tanto la tendencia ascendente como los patrones estacionales en el consumo diario de TRANS003 hacia finales de 2024, ofreciendo una precisión adecuada en el seguimiento de las variaciones diarias del consumo energético en este transformador.

Figura 27

Predicción del Consumo Energético Diario para TRANS003 utilizando PROPHET



Las métricas de evaluación para TRANS003 son:

- MAE: 23.28
- MSE: 30.20
- R^2 : 0.77
- Durbin-Watson: 0.511

El MAE de 23.28 indica que el modelo *Prophet* presenta un error promedio relativamente bajo, lo cual sugiere que las predicciones están bastante cerca de los valores reales en términos absolutos. El MSE de 30.20 refuerza esta interpretación al penalizar de manera significativa los errores grandes, lo que sugiere que las desviaciones extremas en las predicciones son limitadas y que el modelo mantiene una consistencia adecuada en su desempeño.

El coeficiente de determinación (R^2) de 0.77 sugiere que *Prophet* logra explicar aproximadamente el 77% de la variabilidad en el consumo de energía de TRANS003, lo cual es

un indicador positivo de ajuste, ya que captura tanto la tendencia general como los patrones estacionales de consumo. No obstante, el valor de *Durbin-Watson* de 0.511 indica una ligera autocorrelación positiva en los residuos, lo que implica que aún existen patrones recurrentes en el consumo que el modelo no captura completamente, especialmente en los picos y caídas más pronunciados.

Sistema de Análisis y Predicción en Tiempo Real

En este capítulo se presenta la implementación y visualización de un modelo de análisis de datos en tiempo real, basado en la información proveniente de medidores inteligentes. Este modelo permite generar predicciones actualizadas del consumo diario para cada uno de los transformadores analizados. Se integraron las predicciones de los modelos desarrollados, como *ARIMA*, *RNN*, *Prophet* y Suavizamiento Exponencial. Esta visualización facilita la identificación de tendencias y comportamientos de consumo, apoyando la toma de decisiones operativas y estratégicas. Adicionalmente, se diseñó una interfaz gráfica intuitiva que permite a los usuarios interactuar con los datos y los resultados predictivos.

El desarrollo de un sistema robusto de análisis y predicción en tiempo real para el consumo diario de energía de transformadores se constituye como un pilar en la gestión operativa y estratégica de recursos energéticos. Utilizando datos obtenidos de medidores inteligentes, nuestro objetivo fue implementar una solución que no solo facilitara la visualización de los patrones de consumo energético, sino que también permitiera predecir futuras demandas con precisión.

Para alcanzar este objetivo, se optó por una arquitectura dividida en dos componentes clave: un backend desarrollado con python y un frontend utilizando framework Flask y Dash, para la aplicación web. Esta estructura permitió una segregación clara de funciones, donde Flask

se encargaba del procesamiento de datos y ejecución de modelos predictivos, y Dash proporcionaba una interfaz interactiva para la visualización y manipulación de datos.

Backend (Python)

Se implementaron varios endpoints para procesar los datos en tiempo real, aplicando transformaciones, ejecutando cálculos estadísticos y modelos de predicción mediante el uso de librerías especializadas como pandas y statsmodels.

Frontend (Flask y Dash)

La interactividad del usuario fue centralizada mediante la creación de gráficos dinámicos y tablas actualizables. Utilizando Plotly para las visualizaciones gráficas y dash.DataTable para las tablas. Los *'callbacks'* de Dash facilitaron la actualización de contenidos en tiempo real, respondiendo a las interacciones del usuario como la selección de transformadores y modelos de análisis.

El sistema fue diseñado para funcionar de manera local, con Flask operando en el puerto 5000 y Dash interactuando con estos servicios. Este entorno local permitió un desarrollo y pruebas eficientes.

La interacción inicia con la selección de un transformador desde un menú desplegable, seguido por la visualización de datos específicos al dar clic en el botón 'Ver Datos' como se observa en la Figura 28, Esto carga la información de consumos por día en una tabla y prepara el aplicativo para la predicción. Los usuarios pueden entonces elegir entre los diferentes modelos de predicción desarrollados en la lista de selección como los son: *Prophet*, *RNN*, *ARIMA*, *Smoothing Triple*, para el análisis de descomposición de series temporales se debe dar clic en el botón “Descomposición” y para la respectiva predicción se debe dar clic en el botón

“Predicción”. Cada elección se refleja inmediatamente en las visualizaciones gráficas correspondientes como se ve en la Figura 29 y la Figura 30.

Figura 28

Interfaz Gráfica Inicial de Modelos de Predicción

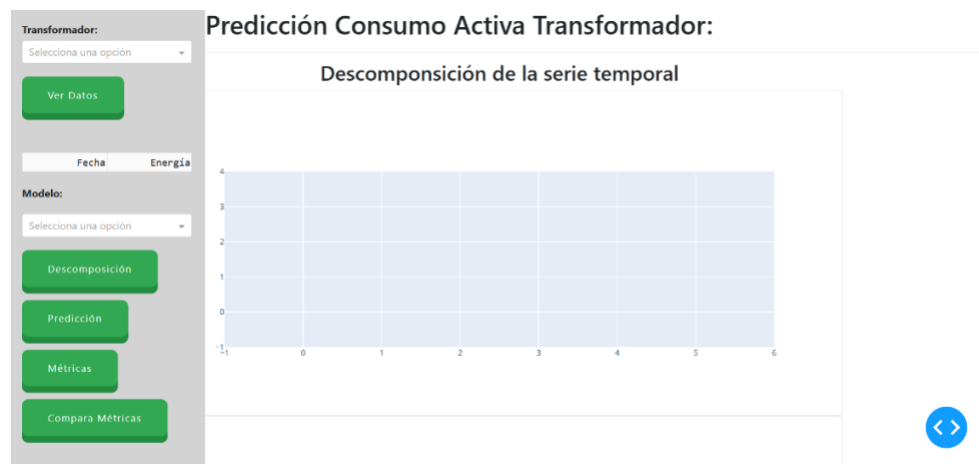
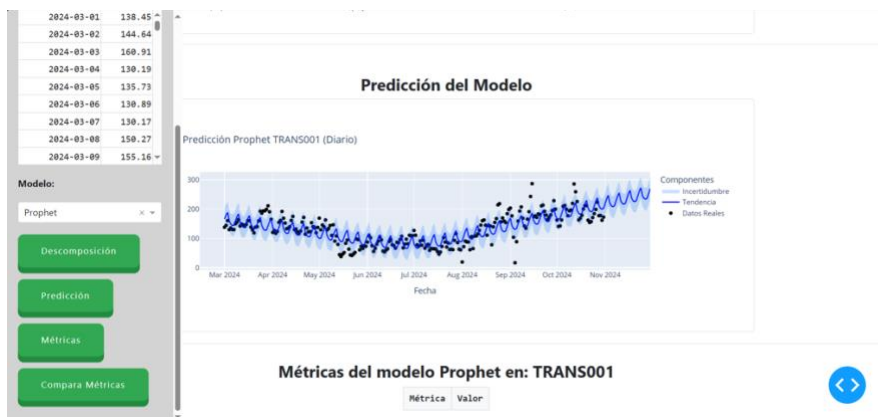


Figura 29

Interfaz Gráfica de la Descomposición Inicial de la Serie Temporal



Figura 30*Interfaz Gráfica de Modelos de Predicción*

Finalmente, los usuarios pueden comparar las métricas de los diferentes modelos desarrollados dando click en el botón “Compara Métricas” como se ve en la Figura 31, facilitando así la evaluación de la efectividad de cada método en la predicción del consumo energético.

Figura 31*Interfaz Gráfica de las Métricas*

Conclusiones

Después de desarrollar el trabajo y analizar el consumo energético diario del transformador TRANS001 utilizando diversos modelos, se puede concluir que el modelo *Prophet* demostró ser el más adecuado para este escenario, caracterizado por un patrón de consumo relativamente estable con menores fluctuaciones. *Prophet* capturó de manera efectiva tanto la tendencia como la estacionalidad, logrando un R^2 de 0.681 y un MAE de 23.37, lo que indica que sus predicciones se aproximan de manera consistente a los valores reales. Aunque presenta ligeros desfases en los picos de consumo, su capacidad para modelar patrones complejos lo posiciona como la mejor herramienta para este caso, como se observa en la Tabla 2.

El modelo *RNN*, diseñado para manejar dependencias a largo plazo en series temporales, no logró adaptarse bien a los datos de TRANS001. Esto se refleja en su R^2 negativo de -0.705 y un MAE más alto de 32.54, indicando que este enfoque no es ideal para patrones de consumo menos dinámicos. Por el contrario, el modelo ARIMA, con un AIC de 1762.619, mostró un ajuste moderado a los datos, pero su desempeño general fue limitado, como lo demuestra su R^2 negativo de -0.262, sugiriendo que no logra capturar de manera adecuada las fluctuaciones observadas en los datos reales.

Finalmente, el método de Suavizamiento Exponencial presentó un desempeño aceptable, con un MAE de 16.95, posicionándose como una alternativa válida para capturar patrones simples. Sin embargo, este enfoque es menos robusto en comparación con *Prophet* y se limita a modelar datos con menor complejidad.

Tabla 2

Comparación de Métricas de Modelos para el Transformador TRANS001

Métricas	Smoothing	ARIMA	RNN	Prophet
MAE	16,950	27,950	32,546	23,374
MSE	554,791	1247,970	1662,028	
RMSE		35,370	40,768	29,264
R2	0,439	-0,262	-0,705	0,681
Durbin-Watson				0,531
MAPE			16,035	
AIC		1762,619		
BIC		1785,530		

Nota. Tabla donde se ilustra las métricas correspondientes a los modelos para el TRANS001.

En el caso del transformador TRANS002, con un patrón de consumo más dinámico y fluctuante, se observa un cambio en la eficacia de los modelos utilizados. El modelo *Prophet* nuevamente destacó con un R^2 de 0.754 y un MAE de 53.57, capturando de forma adecuada tanto la tendencia general como las estacionalidades presentes en los datos, como se detalla en la Tabla 3. Sin embargo, la alta variabilidad y los picos abruptos en el consumo energético limitaron su capacidad de predicción en ciertos periodos, reflejándose en un ligero aumento del error en comparación con TRANS001.

El modelo *RNN*, diseñado para manejar patrones complejos y fluctuaciones más pronunciadas, mostró un desempeño notablemente mejor en este transformador, logrando un R^2 de 0.368 y un MAE de 30.87, evidenciando que es una opción viable para escenarios con mayor dinámica. En este caso, el *RNN* fue capaz de capturar las variaciones significativas en los datos, superando a *Prophet* en términos de precisión en los valores absolutos.

Por otro lado, el modelo ARIMA, aunque presentó un AIC moderadamente bajo de 1770.268, no logró capturar completamente las fluctuaciones más pronunciadas, reflejando un R^2 de -0.002 que sugiere limitaciones en su capacidad de modelado para este tipo de comportamiento energético.

Finalmente, el método de Suavizamiento Exponencial, con un MAE de 47.52, mostró un desempeño más limitado al no adaptarse adecuadamente a la naturaleza dinámica del consumo en TRANS002, quedando rezagado frente a los demás modelos evaluados.

Tabla 3

Comparación de Métricas de Modelos para el Transformador TRANS002

Métricas	Smoothing	ARIMA	RNN	Prophet
MAE	47,529	42,150	30,867	53,575
MSE	5612,355	4599,990	3018,504	
RMSE		67,820	54,941	65,120
R2	-0,200	-0,002	0,368	0,755
Durbin-Watson				0,300
MAPE			15,728	
AIC		1770,268		
BIC		1792,735		

Nota. Tabla donde se ilustra las métricas correspondientes a los modelos para el TRANS002.

Para el transformador TRANS003, caracterizado por un consumo inicial bajo y estable seguido de un aumento abrupto en la segunda mitad del periodo analizado, los resultados resaltan la importancia de seleccionar modelos que se adapten a cambios significativos en los patrones de consumo. En este caso, el modelo *Prophet* demostró un excelente desempeño, logrando un R^2 de 0.778 y un MAE de 23.28, destacándose como el mejor modelo para capturar tanto la tendencia como la estacionalidad presentes en los datos, como se muestra en la Tabla 4. Sin embargo, se

identificaron ligeros desfases en los picos de mayor variabilidad, lo que refleja ciertas limitaciones en la predicción de valores extremos.

El modelo RNN, con un R^2 de 0.310 y un MAE de 21.29, mostró un rendimiento sólido en la captura de la dinámica del transformador TRANS003, siendo especialmente eficaz en contextos donde predominan fluctuaciones abruptas y cambios repentinos en el patrón de consumo. Este resultado evidencia que el RNN es una opción robusta para escenarios con alta variabilidad.

En contraste, el modelo ARIMA, con un AIC de 1563.325, presentó un ajuste aceptable, pero su desempeño general fue limitado, reflejado en un R^2 de -0.007, lo que indica que no logra adaptarse adecuadamente a los cambios pronunciados en el consumo energético. Por su parte, el Suavizamiento Exponencial, con un MAE de 55.70, fue el menos efectivo, siendo incapaz de modelar los patrones complejos y dinámicos observados en TRANS003.

Tabla 4

Comparación de Métricas de Modelos para el Transformador TRANS003

Métricas	Smoothing	ARIMA	RNN	Prophet
MAE	55,703	22,690	21,285	23,288
MSE	4198,432	1208,870	857,990	
RMSE		34,770	29,291	30,201
R2	-2,266	-0,007	0,310	0,778
Durbin-Watson				0,511
MAPE			18,725	
AIC		1563,325		
BIC		1585,791		

Nota. Tabla donde se ilustra las métricas correspondientes a los modelos para el TRANS003.

Los resultados obtenidos indican que *Prophet* es la herramienta más adecuada para modelar el consumo energético diario de los transformadores analizados, debido a su capacidad para equilibrar precisión y flexibilidad. Este modelo demostró ser el más robusto al capturar tanto tendencias como estacionalidades en los tres transformadores, logrando métricas consistentes en MAE, MSE y R^2 , lo que lo posiciona como una opción confiable para escenarios con características diversas, desde patrones más estables (TRANS001) hasta consumos más dinámicos y fluctuantes (TRANS002 y TRANS003).

Por su parte, el modelo RNN destacó en transformadores con alta variabilidad, como TRANS002 y TRANS003, al capturar de manera efectiva las fluctuaciones a corto plazo. Sin embargo, en transformadores con menor dinamismo, como TRANS001, este modelo tuvo dificultades, reflejadas en métricas menos favorables, lo que sugiere que su fortaleza radica en manejar patrones complejos y dinámicos.

En el caso de ARIMA, aunque mostró utilidad para capturar patrones generales en series más estables, su capacidad predictiva fue limitada frente a los transformadores con alta variabilidad. Si bien los valores de AIC y BIC indicaron un ajuste moderado, su desempeño global resalta la necesidad de ajustes adicionales en sus parámetros para mejorar la captura de fluctuaciones más abruptas.

Finalmente, el Suavizamiento Exponencial fue adecuado únicamente para el transformador más estable (TRANS001), donde logró un desempeño aceptable. Sin embargo, no logró adaptarse a datos más dinámicos, como los de TRANS002 y TRANS003, obteniendo métricas desfavorables y quedando rezagado frente a los demás modelos.

Recomendación

Un trabajo futuro se proyecta en mejorar el rendimiento de la aplicación mediante la incorporación y procesamiento eficiente de datos reales obtenidos directamente de clientes individuales, en lugar de datos simulados o agrupados por transformadores. Actualmente, la aplicación funciona con volúmenes reducidos de datos, pero se planea fortalecer su arquitectura para manejar grandes volúmenes de datos históricos de manera eficiente. Este enfoque permitirá personalizar las predicciones del consumo energético y optimizar la gestión de la demanda en tiempo real. Además, se buscará adaptar la plataforma desarrollada en Python, Dash y Flask para mantener la precisión de las predicciones y la agilidad en la visualización y toma de decisiones, garantizando que sea escalable y robusta para escenarios más complejos.

Referencias

- Arias Martín, J. (2019). *Predicción de consumos eléctricos en Inglaterra, Gales y Escocia a través de datos de medidores inteligentes. Trabajo Final de Máster, Universitat Oberta de Catalunya*. <http://hdl.handle.net/10609/102266>
- Athanasiadis et al, C. L. (2024). *A review of distribution network applications based on smart meter data analytics. Renewable and Sustainable Energy Reviews*. <https://www-sciencedirect-com.bibliotecavirtual.unad.edu.co/search?qs=A%20review%20of%20distribution%20network%20applications%20based%20on%20smart%20meter%20data%20analytics>
- Azaza et al, M. W. (2017). *Smart meter data clustering using consumption indicators: responsibility factor and consumption variability. Energy Procedia*. <https://www.sciencedirect.com/science/article/pii/S1876610217308648>
- Burnett et al, J. W. (2022). *How do machines predict energy use? Comparing machine learning approaches for modeling household energy demand in the United States. Energy Research & Social Science*. <https://www-sciencedirect-com.bibliotecavirtual.unad.edu.co/science/article/pii/S2214629622002195>
- CREG. (2018). *Gestor normativo alejandria. Resolución 40072*. https://gestornormativo.creg.gov.co/gestor/entorno/docs/resolucion_minminas_40072_2018.htm
- Czétány et al, L. V.-B.-B. (2021). *Development of electricity consumption profiles of residential buildings based on smart meter data clustering. Energy & Buildings*. <https://doi.org/10.1016/j.enbuild.2021.111376>

- Espinosa-Zúñiga, J. J. (2020). *Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. Ingeniería, Investigación y Tecnología.* <https://doi.org/10.22201/fi.25940732e.2020.21n1.008>
- González-Briones et al, A. H. (2019). *Machine learning models for electricity consumption forecasting: A Review.* <https://ieeexplore-ieee-org.bibliotecavirtual.unad.edu.co/document/8769508>
- Goodfellow et al, I. B. (2016). *Deep learning. MIT Press.* <https://www.deeplearningbook.org>
- Hyndman, R., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice (3rd ed.). OTexts.* <https://otexts.com/fpp2/>
- Ignatiadis et al, D. G. (2019). *Forecasting Residential Monthly Electricity Consumption using Smart Meter Data. North American Power Symposium (NAPS), Wichita, KS, USA.* <https://ieeexplore-ieeeorg.bibliotecavirtual.unad.edu.co/document/9000285>
- kang, J., & Reiner, D. (2022). *Off seasons holidays and extreme weather events: Using data-mining techniques on smart meter and energy consumption data from China. Energy Research & Social Science, 89, 102637.* <https://doi.org/10.1016/j.erss.2022.102637>
- Ketchledge, J. (2024). *Challenges of Implementing AMI. Electric Energy Online.* https://electricenergyonline.com/EE_magazine_article.php?ID=414.
- Liu et al, X. G. (2015). *Benchmarking Smart Meter Data Analytics. Proceedings of the 18th International Conference on Extending Database Technology (EDBT).* <https://openproceedings.org/2015/conf/edbt/paper-55.pdf>
- Mathumitha et al, R. R. (2023). *SVM-based regression for forecasting building power energy consumption using smart meter data. 14th International Conference on Computing*

- Communication and Networking Technologies (ICCCNT)*. <https://ieeexplore-ieee-org.bibliotecavirtual.unad.edu.co/document/10307674>
- Mitra et al, S. C. (2024). *Energy*, 289, 129920. <https://doi.org/10.1016/j.energy.2023.129920>
- Oprea, S.-V., & Bâra, A. (2019). *Machine Learning Algorithms for Short-Term Load Forecast in Residential Buildings Using Smart Meters, Sensors, and Big Data Solutions*. *IEEE Access*. <https://ieeexplore.ieee.org/document/8928601?denied=>
- Rashid, M. H. (2018). *AMI Smart Meter Big Data Analytics for Time Series of Electricity Consumption*. *17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering*, <https://ieeexplore-ieee-org.bibliotecavirtual.unad.edu.co/document/8456135>
- Rodríguez, V. Á. (2021). *Energía: Infraestructura de Medición Avanzada (AMI)*. *Castro Nieto Abogados*. <https://castronieto.co/energia-la-estructura-de-medicion-avanzada-ami/>
- Taylor, S., & Letham, B. (2018). *Forecasting at scale*. *The American Statistician*, 72(1), 37-45. <http://lethalletham.com/ForecastingAtScale.pdf>
- Völker et al, B. R. (2021). *Watt's up at Home? Smart Meter Data Analytics from a Consumer-Centric Perspective*. *Energies*, 14(3), 719. <https://doi.org/10.3390/en14030719>
- Wang et al, Y. C. (2019). *Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges*. *IEEE Transactions on Smart Grid*, 10(3), 3125-3148. [https://ieeexplore-ieee-org.bibliotecavirtual.unad.edu.co/document/8322199\[RG1\]](https://ieeexplore-ieee-org.bibliotecavirtual.unad.edu.co/document/8322199[RG1])