Proyecto Fin de Máster

Máster en Sistemas de Energía Eléctrica

Técnicas de predicción aplicada a la demanda desagregada con variables exógenas

Autor: Jaime Fernando Jácome García

Tutor: Jesús Manuel Riquelme Santos

Dpto. Ingeniería Eléctrica
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2023

Proyecto Fin de Máster
Máster en Sistemas de Energía Eléctrica

Técnicas de predicción aplicada a la demanda desagregada con variables exógenas

Autor:
Jaime Fernando Jácome García

Tutor:
Jesús Manuel Riquelme Santos
Profesor titular

Dpto. Ingeniería Eléctrica
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla
Sevilla, 2023

Proyecto Fin de Máster: Técnicas de predicción aplicada a la demanda desagregada con variables exógenas

Autor: Jaime Fernando Jácome García

Tutor: Jesús Manuel Riquelme Santos

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, 2023

El Secretario del Tribunal

A mi familia

A mis amigos

Agradecimientos

Agradezco a Dios por guiarme y otorgarme la sabiduría necesaria para superar obstáculos y debilidades en mi camino.

Ivon y Patricio, mis queridos padres, merecen un agradecimiento especial por su apoyo constante, confianza inquebrantable y por proporcionarme las herramientas necesarias para alcanzar cualquier meta que me proponga.

Cristian, mi primo, ha sido tanto mi ejemplo a seguir como mi apoyo incondicional en la vida, y por eso le estoy profundamente agradecido.

Mis hermanos, Esteban e Ivonne, han sido una fuente constante de apoyo, compañía y amor incondicional a lo largo de mi vida.

No puedo dejar de mencionar a mi sobrina Maite, quien, a pesar de ser una niña, es la persona más fuerte y valiente que conozco. Su fuerza y alegría han sido una inspiración constante para mí.

A mis amigos, que contribuyeron de manera significativa a este logro en mi vida profesional, les dedico un sincero agradecimiento. Cada uno de ustedes ha sido un pilar fundamental en mi camino, y este logro no habría sido posible sin su presencia y apoyo.

Jaime Jácome

Sevilla, 2023

Resumen

El progreso de los mercados eléctricos liberalizados ha revolucionado la industria energética en los últimos años. La necesidad de predecir la demanda desagregada de electricidad se ha vuelto crucial para la toma de decisiones y la ventaja competitiva. A pesar de los enfoques tradicionales, la predicción de demanda desagregada ha emergido como una solución para abordar patrones complejos en el consumo eléctrico.

Este estudio se enfoca en la predicción de la demanda eléctrica desagregada en centros de transformación. Se ha aplicado un riguroso proceso de tratamiento de datos destinado a abordar la presencia de valores faltantes y anomalías en los conjuntos de datos. Se implementa modelos, como ARIMA y redes neuronales de memoria a largo plazo (LSTM), con el fin de realizar una comparación de su capacidad predictiva.

La selección del modelo óptimo desempeña un papel fundamental en la mejora de la toma de decisiones en el sector energético. Con miras a mejorar la robustez y la precisión de las predicciones, se ha propuesto la utilización de modelos ensambladores. Estos ensambladores tienen la finalidad de aprovechar y fusionar las mejores características de los modelos individuales, permitiendo así una respuesta predictiva aún más precisa y confiable.

El objetivo de este estudio es proporcionar un análisis detallado de la predicción de demanda desagregada en centros de transformación, lo que puede contribuir significativamente a la planificación estratégica en la industria energética

Abstract

The progress of liberalized electricity markets has revolutionized the energy industry in recent years. The need to forecast unbundled electricity demand has become crucial for decision making and competitive advantage. Despite traditional approaches, disaggregated demand forecasting has emerged as a solution to address complex patterns in electricity consumption.

This study focuses on disaggregated electricity demand forecasting at transformer substations. A rigorous data treatment process has been applied to address the presence of missing values and anomalies in the data sets. Then, models, such as ARIMA and long-term memory neural networks (LSTM), have been implemented in order to perform a comparison of their predictive capability.

The selection of the optimal model plays a key role in improving decision making in the energy sector. In order to improve the robustness and accuracy of predictions, the use of ensemble models has been proposed. These assemblers are intended to exploit and merge the best features of individual models, thus enabling an even more accurate and reliable predictive response.

The objective of this study is to provide a detailed analysis of disaggregated demand forecasting at transformation centers, which can significantly contribute to strategic planning in the energy industry.

Índice

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

1 INTRODUCCIÓN

Quien no espera vencer, ya está vencido

- José Joaquín de Olmedo-

En los últimos 15 años, el progreso de los mercados eléctricos liberalizados, también conocidos como mercados eléctricos desregulados o mercados eléctricos abiertos, ha revolucionado la forma en que se genera, distribuye y comercializa la electricidad. Estos sistemas son diseñados para fomentar la competencia y permitir la participación de múltiples actores en la industria energética. Al mismo tiempo, el rápido crecimiento de las fuentes de energía renovables variables ha impulsado la necesidad de desarrollar modelos de mercado eléctrico personalizados en diferentes regiones, como Estados Unidos y Europa[1]–[3].

En un entorno global cada vez más orientado hacia el análisis de datos y la toma de decisiones basados en datos, la predicción de la demanda desagregada ha adquirido una relevancia fundamental para empresas y organizaciones. Esta herramienta se ha convertido en imprescindible en la optimización de operaciones, en la mejora de la planificación estratégica y en la capacidad de brindar productos y servicios altamente personalizados a sus clientes. La demanda desagregada implica prever patrones de consumo a niveles más detallados, como productos individuales, categorías de productos, segmentos de clientes o ubicaciones geográficas específicas. Esto permite optimizar operaciones, mejorar la planificación y ofrecer productos y servicios personalizados a los clientes, lo que se ha vuelto crucial para mantener una ventaja competitiva en el mercado[4].

A lo largo de las últimas décadas, la predicción de la demanda eléctrica ha sido objeto de intensa investigación y desarrollo, debido a la necesidad de abordar los desafíos asociados con la generación, distribución y consumo de electricidad. Aunque los enfoques tradicionales basados en técnicas estadísticas son ampliamente utilizados en la predicción de la demanda agregada, han mostrado limitaciones en la representación de patrones complejos y cambios abruptos en el consumo. Esto ha impulsado el interés en la predicción de demanda desagregada, que busca ofrecer una visión más detallada y precisa de los patrones de consumo, lo que resulta esencial para la toma de decisiones estratégicas y operativas en el sector eléctrico y otras industrias[4].

La predicción de la demanda desagregada en centros de transformación (CT) es esencial para anticipar con precisión los patrones de consumo de energía eléctrica. En este trabajo, se aborda la tarea de predecir la demanda desagregada de dos centros de transformación específicos, cada uno con su propia base de datos. Para lograrlo, se emplea diversas técnicas de tratamiento de datos, especialmente diseñadas para abordar la presencia de datos faltantes y anómalos.

Es importante destacar que, en este estudio, además de considerar la limpieza y procesamiento de datos, se valora la influencia de variables exógenas en la predicción de la demanda de energía eléctrica en una serie temporal. En particular, se analizan las variables de temperatura y humedad, cuyo impacto en la demanda energética es fundamental y ha sido ampliamente reconocido en la literatura científica. La temperatura y la humedad son factores críticos que afectan directamente los patrones de consumo eléctrico, ya que inciden en la necesidad de climatización y refrigeración, así como en otros aspectos relevantes para la demanda energética.

Una vez que los datos se procesan y limpian, la tarea de predicción se lleva a cabo mediante la implementación de modelos como ARIMA (autoregressive integrated moving average), RNN (Recurrent Neural Networks), y LSTM (long short-term memory). Estos modelos se seleccionan cuidadosamente y se evalúan en función de su capacidad para capturar las relaciones temporales y las influencias de las variables exógenas, como la temperatura y la humedad. La elección del mejor modelo desempeña un papel crítico en la mejora de la toma de decisiones y la planificación estratégica en estos centros de transformación.

La información obtenida a través de este estudio tiene un valor significativo para la planificación estratégica en el sector energético, ya que permite comprender cómo las variables exógenas, en particular la temperatura y la humedad, influyen en la demanda de energía eléctrica en una serie temporal. Esto proporciona una base sólida para la toma de decisiones informadas y la optimización de recursos en el suministro de energía eléctrica.

La organización de este estudio se compone de siete capítulos. El primero, la introducción, establece la relevancia de la predicción de la demanda eléctrica. El segundo capítulo, 'Objetivos', define el objetivo general y los objetivos específicos que guían la investigación. En el tercer capítulo, 'Marco Teórico', se profundiza en conceptos fundamentales, como demanda desagregada y técnicas de machine learning. El cuarto capítulo, 'Metodología', detalla la metodología empleada, que incluye procesamiento de datos y modelos de predicción. El quinto capítulo, 'Resultados y Discusiones', presenta y analiza los hallazgos obtenidos. El sexto capítulo, 'Conclusiones', resume las conclusiones del estudio y sugiere direcciones futuras. El séptimo capítulo, 'Referencias', proporciona una lista exhaustiva de fuentes bibliográficas utilizadas para respaldar la investigación.

2 OBJETIVOS

Cada día sabemos más y entendemos menos

- Albert Einstein-

2.1. Objetivo General

Este trabajo tiene como objetivo realizar un análisis exhaustivo de la predicción de la demanda desagregada en centros de transformación, investigando y aplicando diversas metodologías y técnicas de pronóstico con el propósito de obtener resultados robustos y significativos desde una perspectiva científica y técnica

2.1 Objetivos Específicos

- Exponer de manera sistemática y detallada la metodología de investigación empleada, incluyendo la descripción de los procedimientos utilizados para la recopilación de datos, las herramientas y técnicas específicas utilizadas para el análisis, así como cualquier enfoque metodológico pertinente.
- Llevar a cabo un análisis minucioso de los datos recopilados, empleando herramientas y técnicas apropiadas que estén respaldadas por la literatura científica, y presentar los resultados de manera clara y comprensible, incorporando gráficos, estadísticas descriptivas y visualizaciones pertinentes.
- Realizar una interpretación exhaustiva de los resultados del análisis de datos, estableciendo conexiones con los hallazgos dentro del marco teórico proporcionado por la revisión de la literatura.
- Llevar a cabo una discusión crítica y fundamentada sobre las implicaciones de los resultados obtenidos, estableciendo relaciones con teorías relevantes en el campo de estudio, y destacando cualquier contribución novedosa o perspicaz al avance del conocimiento.
- Sintetizar de manera concisa y precisa las conclusiones fundamentales derivadas de la investigación, enfocándose en los hallazgos más significativos que aborden directamente a los resultados de la investigación

3 MARCO TEÓRICO

En realidad, no me preocupa que quieran robar mis ideas, me preocupa que ellos no las tengan.

- Nikola Tesla-

En este capítulo, se explora técnicas para la predicción de la demanda de energía eléctrica desagregada. Se destaca la importancia del tratamiento de datos como base fundamental, se aborda el aprendizaje automático (machine learning) en redes neuronales y la aplicación de modelos tradicionales como ARIMA. Sin embargo, es esencial destacar que, en este estudio, es necesario incorporar variables exógenas como la temperatura y la humedad. La inclusión de estas variables permite evaluar si hay una mejora significativa en la precisión de las predicciones energéticas.

La temperatura y la humedad son factores exógenos que pueden tener un impacto considerable en la demanda de energía eléctrica. Por ejemplo, en días calurosos, es probable que la demanda de energía para sistemas de aire acondicionado aumente significativamente. De manera similar, la humedad puede afectar la eficiencia de equipos y sistemas de climatización, lo que influye en la cantidad de energía requerida

Además, en este contexto, también es relevante mencionar la utilización de un modelo ensamblador. Un modelo ensamblador combina múltiples modelos de predicción en uno solo, aprovechando sus fortalezas individuales para obtener una predicción más precisa y robusta de la demanda de energía eléctrica. La combinación de modelos ARIMA o ARIMAX (AutoRegressive Integrated Moving Average with eXogenous variables) y LSTM puede ser especialmente efectiva para abordar los desafíos únicos en la predicción de la demanda eléctrica, y, como resultado, contribuir a una gestión energética más eficiente y sostenible

3.1 Predicción de la Demanda

La predicción de la demanda de energía eléctrica es esencial para garantizar un suministro confiable y eficiente de electricidad. Predecir cuánta energía se requerirá en el futuro es crucial para planificar la generación y distribución de energía, así como para evitar cortes de energía y desperdicio de recursos. Este marco teórico abordará métodos y técnicas utilizados en la predicción de la demanda eléctrica, destacando su importancia en la gestión de sistemas eléctricos.

3.1.1 Demanda Desagregada

La demanda desagregada en energía eléctrica se refiere a la subdivisión o descomposición de la demanda total de electricidad en sus componentes individuales o segmentos. En otras palabras, implica analizar y comprender el consumo de electricidad en un nivel más detallado, desglosando la demanda en categorías específicas, como sectores industriales, comerciales, residenciales, o incluso por regiones geográficas y horas del día.

La demanda desagregada de energía eléctrica en España se refiere a la subdivisión o descomposición de la demanda total de electricidad en el país en sus distintos componentes o sectores específicos de consumo[5]. Estos componentes se dividen generalmente en categorías clave, que pueden incluir:

- **Sector Residencial:** Esta categoría abarca la demanda de electricidad de hogares y viviendas.
- **Sector Industrial:** Incluye la demanda de electricidad de las instalaciones y procesos industriales.
- **Sector Comercial:** Se refiere a la demanda eléctrica de comercios, oficinas y otros establecimientos comerciales.
- **Sector de Servicios:** Engloba la electricidad utilizada en servicios públicos, como hospitales, escuelas y edificios gubernamentales.
- **Agricultura:** La demanda de energía eléctrica en actividades agrícolas.

- **Transporte eléctrico:** Esto abarca la carga de vehículos eléctricos y otros sistemas de transporte eléctrico.
- **Pérdidas en la Red:** Se refiere a la electricidad que se pierde en la transmisión y distribución de la energía

El análisis de la demanda desagregada es una herramienta crucial para entender y gestionar la demanda de energía eléctrica, especialmente en contextos donde se controla la Media Tensión y la baja tensión en comunidades de energía. La Media Tensión se refiere al monitoreo y regulación de la electricidad en niveles intermedios de tensión, que típicamente suministra energía a áreas residenciales, comerciales e industriales. Por otro lado, la baja tensión se enfoca en la distribución local de electricidad, como en hogares y pequeñas empresas en comunidades energéticas.

Al descomponer la demanda en segmentos más pequeños, se obtiene una visión detallada de cómo se consume energía eléctrica en estos sectores específicos. Esta aproximación permite a los reguladores y empresas de servicios públicos crear estrategias adaptadas a cada sector. Esto puede incluir medidas para mejorar la eficiencia energética en comunidades energéticas, reducir la carga durante picos de demanda o diseñar tarifas de electricidad que promuevan un uso más responsable de la energía.

Además, este análisis minucioso facilita la identificación de posibles mejoras en la infraestructura eléctrica, como la expansión de redes o la instalación de sistemas de gestión de carga. Estas acciones contribuyen a asegurar un suministro eléctrico confiable y adaptado a las necesidades específicas de las comunidades de energía, lo cual es esencial para mantener el funcionamiento eficiente de la sociedad en su conjunto.

La Figura 3-1 representa el consumo de energía final por sectores durante el período 2000-2020, ofreciendo una visión general de esta práctica esencial para controlar y gestionar la demanda de energía eléctrica en comunidades energéticas.

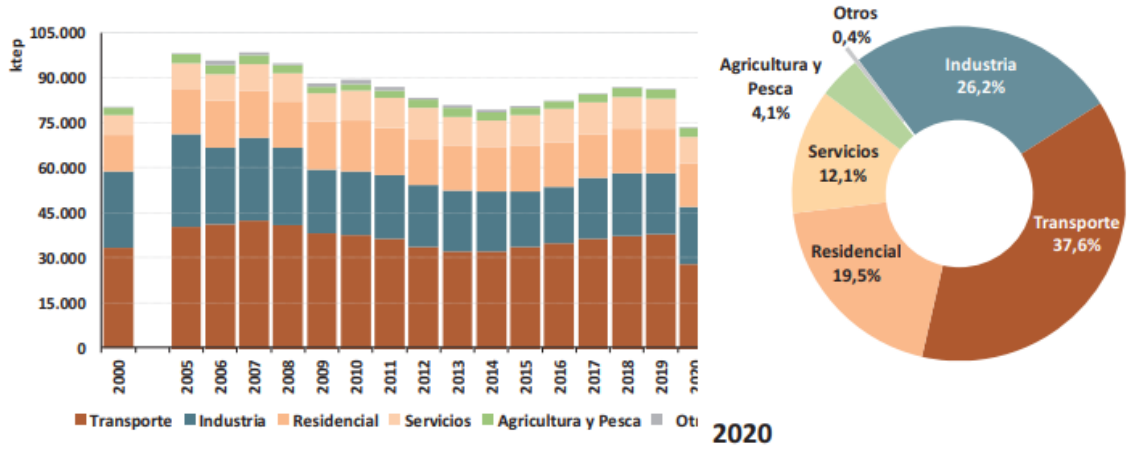


Figura 3-1. Consumo de energía final por sectores, 2000-2020. **Fuente MITERD/IDEA**.

3.1.2 Variables Exógenas

Las variables exógenas, como la temperatura y la humedad, desempeñan un papel crítico en la demanda desagregada de energía eléctrica[6]. La temperatura, por ejemplo, incide directamente en los patrones de consumo eléctrico: su aumento se traduce en una mayor demanda eléctrica destinada a la refrigeración, mientras que su disminución conlleva un incremento en el consumo eléctrico para la calefacción, especialmente en sectores residenciales y comerciales. Asimismo, la humedad afecta significativamente la demanda eléctrica al influir en la percepción térmica de las personas y en los procesos de transferencia de calor. En climas cálidos y húmedos, la necesidad de sistemas de aire acondicionado se intensifica, resultando en un aumento sustancial en el consumo de electricidad.

3.2 Machine Learning

El aprendizaje automático, comúnmente conocido como "machine learning" en inglés, constituye una rama de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos computacionales[7]. Estos sistemas poseen la capacidad de adquirir conocimiento y mejorar su desempeño en tareas específicas mediante el análisis de datos y la acumulación de experiencia, en contraposición a depender exclusivamente de una programación explícita. En su esencia, el aprendizaje automático capacita a las computadoras para identificar patrones, tomar decisiones y realizar predicciones o acciones basadas en la información extraída de los datos. Esta versatilidad lo convierte en una herramienta de amplia aplicabilidad en diversas áreas, desde el procesamiento del lenguaje natural y la recomendación de contenido hasta la anticipación de tendencias y el diagnóstico médico[7].

La elección del algoritmo adecuado puede resultar abrumadora, ya que existen numerosos algoritmos de aprendizaje automático supervisados y no supervisados, cada uno con un enfoque distinto para el aprendizaje[8], [9].

No existe un método definitivo como el mejor porque la elección depende de los datos, los objetivos, los ajustes específicos y la evolución constante de las nuevas técnicas que van surgiendo con el tiempo. La selección se basa en un análisis cuidadoso de cada problema. Encontrar el algoritmo adecuado, en parte, implica ensayo y error; incluso los profesionales con mucha experiencia en ciencia de datos no pueden determinar si un algoritmo funcionará sin probarlo primero. Sin embargo, la elección del algoritmo también depende del tamaño y tipo de datos con los que se esté trabajando, los conocimientos que se deseen obtener de los datos y cómo se utilizar esos conocimientos. La Figura 3-2 presenta los algoritmos disponibles para el aprendizaje automático y su clasificación.

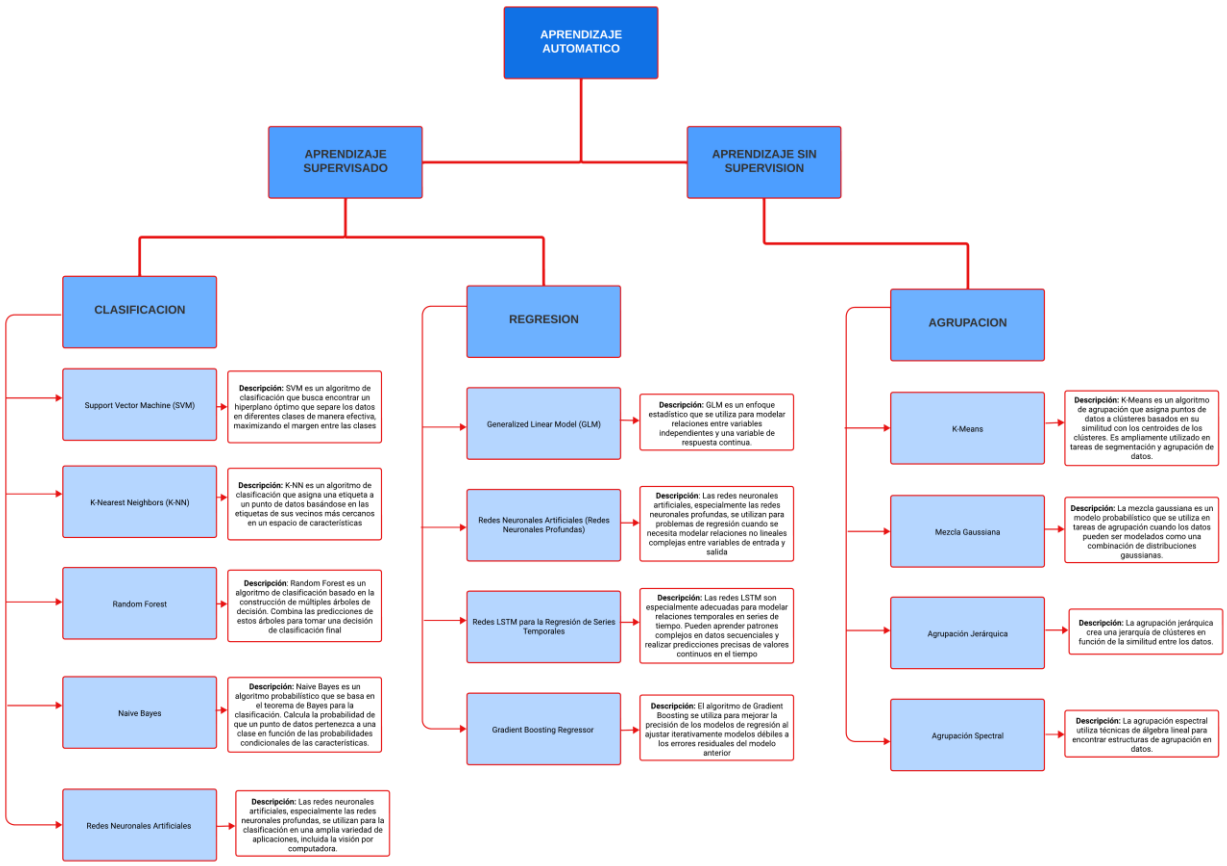


Figura 3-2. Clasificación de los algoritmos de Machine Learning (**Adaptado de** [10])

3.3 Redes Neuronales Artificiales (ANN)

Las Redes Neuronales Artificiales, también referidas como RNAs, constituyen un modelo de aprendizaje automático que toma inspiración en el funcionamiento del cerebro humano. Están configuradas por unidades de procesamiento denominadas neuronas artificiales, dispuestas en capas interconectadas. Cada neurona recibe múltiples entradas, ejecuta cálculos ponderados en base a estas entradas y genera una respuesta de salida. La capacidad de adaptación y aprendizaje de las Redes Neuronales Artificiales se logra mediante la automatización de la adaptación de los pesos de las conexiones a través de procedimientos de entrenamiento algorítmicos[11]. Esto les concede la facultad de modelar relaciones altamente complejas en los datos y resolver una diversidad de problemas, que abarcan desde el reconocimiento de patrones hasta el procesamiento de imágenes y el procesamiento del lenguaje natural.

Como se mencionó previamente, en cada nivel de una red neuronal, se encuentra neuronas interconectadas que desempeñan un papel crucial en el procesamiento de datos. La capa de entrada es la encargada de recibir los datos de entrada y posteriormente transmitirlos a través de la red, mientras que la capa de salida es responsable de generar la salida final de la red. Entre estas capas de entrada y salida, es común encontrar una o varias capas ocultas, las cuales tienen la responsabilidad de llevar a cabo el procesamiento y la transformación de la información. Para una representación visual y esquemática de la estructura y los componentes fundamentales de una red neuronal, se puede consultar la Figura 3-3.

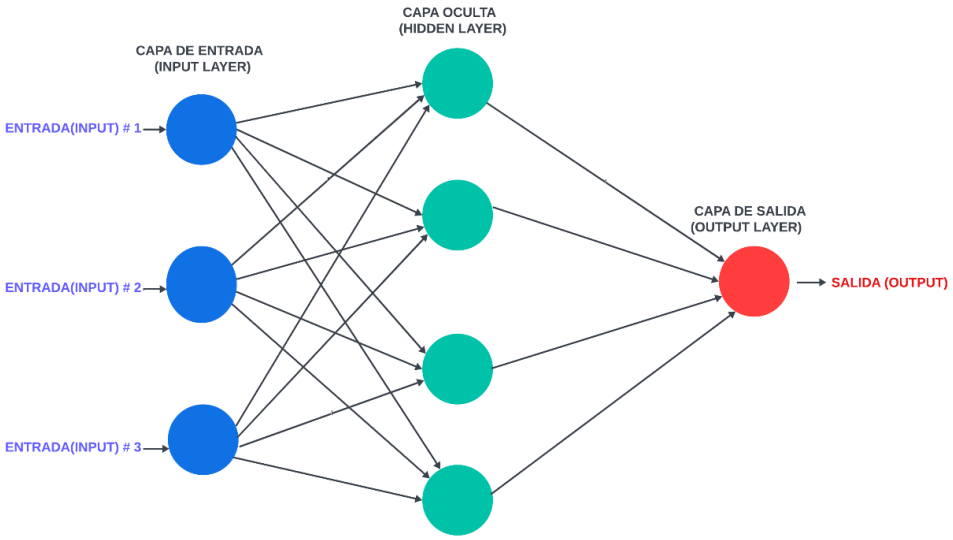


Figura 3-3. Estructura y componentes fundamentales de una red neuronal. **Fuente Propia.**

3.3.1 Entrenamiento y Ajuste de pesos

El proceso de entrenamiento de una red neuronal implica proporcionar datos de entrada y esperar una salida predefinida[11]. Este enfoque de entrenamiento, conocido como "entrenamiento supervisado", se basa en el concepto de supervisar y guiar activamente el aprendizaje de la red. Si la salida no coincide con la salida deseada, se ajustan los pesos de las conexiones en la red para mejorar la precisión. Este proceso iterativo se repite en múltiples ocasiones hasta que la red adquiere la capacidad necesaria para generar la salida deseada. Además, en el contexto de las redes neuronales, existen varios algoritmos de entrenamiento, cada uno diseñado con propósitos específicos.

Entre estos algoritmos, se encuentra modelos de redes neuronales especializados como las Redes Neuronales Recurrentes (RNN) y las Redes LSTM (Long Short-Term Memory). Las RNN están diseñadas para trabajar con datos secuenciales, permitiendo a la red mantener información sobre estados anteriores en su proceso de cálculo. Por su parte, las redes LSTM son una variante de las RNN que se enfocan en abordar problemas de memoria a largo plazo en secuencias de datos.

3.3.2 Funciones de Activación

Las funciones de activación desempeñan un papel crucial en el funcionamiento de las redes neuronales, ya que

determinan la salida de cada neurona en función de su entrada respectiva [11].

Estas funciones se clasifican en dos categorías principales: lineales y no lineales. Las funciones de activación lineales realizan una operación de ponderación simple, multiplicando la entrada por un peso y sumando un sesgo. Por otro lado, las funciones de activación no lineales son más complejas y permiten a la red modelar relaciones no lineales en los datos, lo que las hace especialmente adecuadas para tareas de aprendizaje más sofisticadas.

La elección de la función de activación es una decisión crítica que puede influir de manera significativa en la capacidad de la red neuronal para aprender y generalizar a partir de nuevos datos. Entre las funciones de activación ampliamente utilizadas se encuentra la función sigmoide, la función ReLU (Rectified Linear Unit) y la función tangente hiperbólica las cuales se observa en la Figura 3-4. Cada una de estas funciones tiene sus propias características y aplicaciones particulares en el contexto del aprendizaje automático y las redes neuronales.

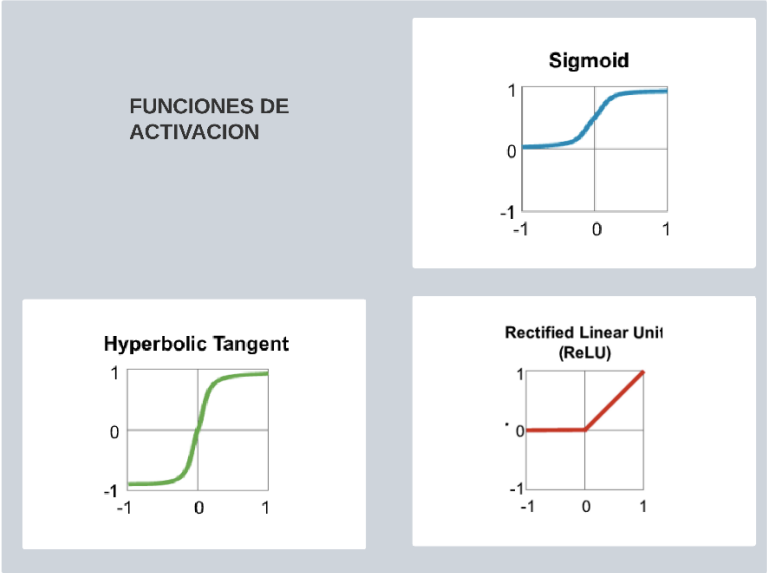


Figura 3-4. Funciones de Activación **Fuente Propia**

3.4 Redes Convolucionales (CNN)

Si bien las redes neuronales convolucionales (CNN) son reconocidas principalmente por su eficacia en el procesamiento de imágenes y videos debido a su habilidad para detectar patrones y características visuales en datos bidimensionales, también demostran ser aplicables de manera prometedora en el análisis de series temporales[12]. A pesar de que las series temporales son datos unidimensionales con propiedades temporales, las CNN han sido adaptadas con éxito para este contexto, considerando la dimensión temporal como una dimensión espacial. Esta adaptación implica que las capas convolucionales de una CNN pueden aprender filtros y características que son sensibles a patrones en datos unidimensionales, como tendencias, oscilaciones y eventos recurrentes. De esta manera, las CNN tienen la capacidad de identificar y capturar información relevante en series temporales, lo que las convierte en una herramienta efectiva para diversas aplicaciones, como la predicción de series temporales y el análisis de señales biomédicas, entre otros campos. En resumen, aunque las CNN son ampliamente reconocidas por su aplicación en imágenes, su versatilidad y capacidad de adaptación las hacen valiosas para analizar y extraer patrones de datos secuenciales en diversas disciplinas científicas.

3.5 Redes Neuronales Recurrentes (RNN)

Las redes neuronales recurrentes (RNN) son un tipo de arquitectura neuronal en la que las neuronas están conectadas de tal manera que la salida de una neurona se convierte en la entrada de otra neurona, creando una especie de bucle de retroalimentación [13]. En otras palabras, la información fluye a través de estas redes de una

manera circular, permitiendo que las conexiones anteriores influyan en las conexiones posteriores, lo que las hace adecuadas para tratar con secuencias y datos temporales. Esta propiedad conferirá a la RNN la capacidad de retener información previa, otorgándole así una "memoria" de los datos de entrada anteriores. Dicha característica se revela particularmente beneficiosa en aplicaciones que involucra secuencias de datos, como el procesamiento del lenguaje natural o la predicción meteorológica.

No obstante, es esencial resaltar que el proceso de entrenamiento de una red neuronal recurrente (RNN) es significativamente más complejo en comparación con el de una red neuronal feedforward, que se caracteriza por su flujo unidireccional y representa la forma más básica de las redes neuronales[14]. Esta complejidad deriva de la incorporación de conexiones retroalimentadas en la RNN. En este contexto, se hacen uso de algoritmos de aprendizaje, como el conocido algoritmo de retropropagación a través del tiempo (BPTT, por sus siglas en inglés), con el propósito de ajustar los pesos que gobiernan las interconexiones entre las neuronas. De esta manera, se busca potenciar la precisión de la salida generada por la red, lo cual reviste importancia en aplicaciones que requieren la captura de relaciones temporales y de secuencias en los datos de entrada[15].

En la Figura 3-5, se exhibe una representación gráfica de una neurona recurrente, acompañada de una neurona recurrente en el contexto temporal. La figura ilustra la estructura intrínseca de una neurona recurrente, resaltando su capacidad para mantener y propagar información a lo largo del tiempo, lo que la distingue de las neuronas convencionales de una red neuronal feedforward.

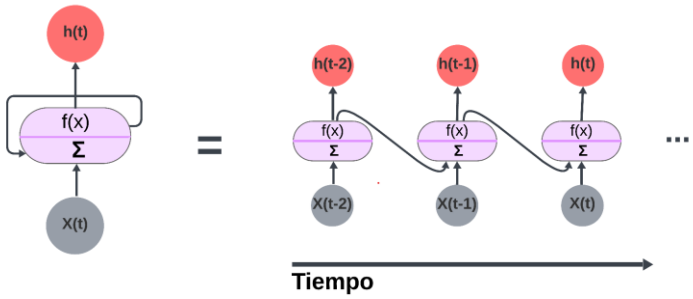


Figura 3-5. Representación de una neurona recurrente y neurona recurrente en el contexto temporal. **Fuente Propia**

3.6 Redes Nueronales Long Short-Term Memory (LSTM)

Las redes LSTM han surgido como una evolución de las redes neuronales recurrentes (RNN) y, además de su amplio rango de aplicaciones, abordan eficazmente los desafíos relacionados con las dependencias a largo plazo, gracias a su diseño basado en una unidad de almacenamiento única. Estas redes LSTM se han destacado especialmente en la predicción de series temporales financieras. En el campo del aprendizaje profundo, las redes neuronales se han consolidado como predictores populares debido a su capacidad para realizar aproximaciones no lineales y su capacidad de autoaprendizaje adaptativo[16].

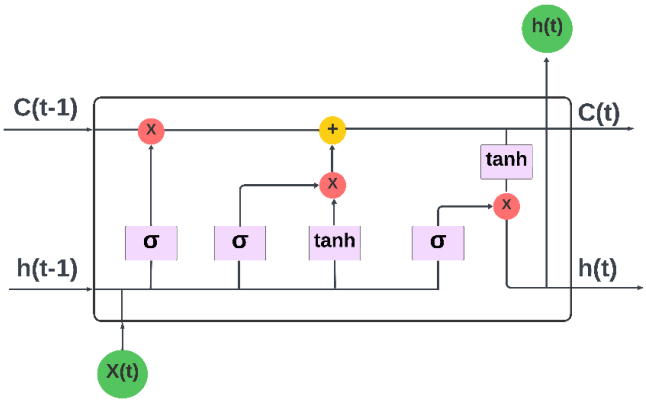


Figura 3-6. Estructura de una Celda LSTM. **Fuente Propia**

En contraste con una unidad de red recurrente convencional, la celda LSTM, representada en la Figura 3-6, presenta notables diferencias, siendo una de las más destacadas la incorporación de una celda de estado adicional. Esta celda de estado se crea como un elemento central en el funcionamiento de las redes LSTM y se asemeja a una cinta transportadora que tiene la capacidad de recibir adiciones de datos o expulsar información de la memoria de la red. Para gestionar este proceso de interacción con la memoria[17], se emplean tres compuertas distintas:

- **Gate de Olvido:** Este mecanismo posibilita la eliminación de componentes de la memoria.
- **Gate de Actualización:** Facilita la inclusión de nuevos elementos en la memoria.
- **Gate de Salida:** Se encarga de la generación del estado oculto actualizado.

Las compuertas opera de manera análoga a válvulas, permitiendo un flujo completo de información cuando están completamente abiertas y bloqueándolo por completo cuando están completamente cerradas. Cada compuerta consta de tres componentes esenciales: una red neuronal, que incluye funciones sigmoidales y otros elementos, junto con un multiplicador [17].

En este contexto, la variable (X(t)) representa la información de entrada en el tiempo (t), la variable (h(t-1)) hace referencia a la salida previamente calculada, la variable (h(t)) denota la nueva salida generada, y, por último C(t-1) que es la información (memoria) que es pasada por la anteiror etapa, por lo que C(t) es la información que pasa al etapa siguiente[17].

3.7 Modelo Autorrgresivo Integrado de Promedio Móvil (ARIMA)

El modelo ARIMA, que significa Autorregresión Integrada de Medias Móviles (por sus siglas en inglés), es una potente técnica estadística ampliamente utilizada en el análisis y predicción de datos temporales[18]. Este modelo se descompone en tres componentes clave: (p, d, q), que representan respectivamente Autorregresión, Integración y Medias Móviles.

- La componente de Autorregresión (p) en el modelo ARIMA se centra en la interdependencia entre el valor presente de la serie temporal y sus valores históricos. En términos más precisos, el modelo ARIMA considera la autocorrelación en la serie, es decir, cómo las observaciones previas influyen en la observación actual. La variable "p" en la especificación ARIMA denota el orden de la parte autorregresiva y corresponde al número de pasos temporales previos que se incorpora en el proceso de predicción.
- La componente de Integración (d) en el modelo ARIMA se relaciona con la operación de diferenciación aplicada a la serie temporal con el propósito de convertirla en una serie estacionaria. Una serie temporal estacionaria exhibe propiedades estadísticas invariables a lo largo del tiempo, lo que simplifica considerablemente su modelado y análisis. La variable "d" presente en la especificación ARIMA indica el grado de diferenciación necesario, es decir, cuántas veces debe aplicarse la operación de diferencia para lograr la estacionariedad deseada.
- Finalmente, la componente de Media Móvil (q) en el modelo ARIMA involucra la relación entre la observación actual de la serie temporal y los residuos de predicciones anteriores. La letra "q" en la notación ARIMA denota el orden de la parte de media móvil, indicando el número de residuos pasados que se incorpora en el proceso de predicción.

El modelo ARIMA se utiliza en una variedad de aplicaciones, como pronósticos financieros, predicción de demanda, análisis climático y muchos otros campos donde se requiere la modelización y predicción de datos temporales[18]. Su capacidad para capturar tendencias y patrones en datos secuenciales lo convierte en una herramienta valiosa en la investigación científica y el análisis de datos. Continuación de presenta la Ecuación 1 la cual representa una forma simplificada del modelo ARIMA.

$$(Y_t = Y_t - Y_{t-d}) \quad Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_P Y_{t-p} - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} + e_t \qquad (3\text{–}1)$$

Donde:

- $\boldsymbol{Y_t}$**:** Esta es la observación en el tiempo t de la serie temporal que estás analizando. Es el valor actual que deseas predecir o modelar.
- $\boldsymbol{c}$**:** Este es el término es una constante
- $\boldsymbol{\phi_1, \phi_2, \ldots \phi_p}$**:** Son los coeficientes de autorregresión de la parte Autoregresiva del modelo, con "p" representando el orden de esta parte.
- $\boldsymbol{Y_{t-1}}$**:** Estos son los valores anteriores de la serie temporal en los retrasos t-1, t-2, etc.
- $\boldsymbol{e_t}$**:** Este término representa el error o residuo en la predicción. Es la diferencia entre el valor observado real Y_t y el valor predicho por el modelo.
- $\boldsymbol{\theta_1, \theta_1, \ldots \theta_q}$**:** Son los coeficientes de la parte de media móvil (Moving Average) del modelo, con "q" representando el orden de esta parte.

En resumen, la fórmula ARIMA encapsula el proceso de modelado y predicción de series temporales, considerando factores como la dependencia en el tiempo, las transformaciones de diferenciación y el impacto de errores previos. La determinación adecuada de los valores de los los componentes (p, d, q), son pasos cruciales en la configuración de un modelo ARIMA. Estos pasos son fundamentales para lograr un modelo efectivo que capture las complejas dinámicas de una serie temporal.

3.8 Varibles Exógenas

La teoría de las variables exógenas en modelos de predicción se fundamenta en la premisa de que elementos externos poseen la capacidad de ejercer una influencia considerable en los resultados de un modelo predictivo. Estas variables exógenas, originadas fuera del ámbito del proceso analizado, se caracterizan por ser independientes, si bien se presume que pueden ejercer un impacto significativo en las predicciones generadas por el modelo. En resumen, esta teoría argumenta que la inclusión de variables exógenas en el modelo posibilita la consideración de factores adicionales que, en última instancia, contribuyen a mejorar la precisión y la validez del modelo predictivo.

En lo concerniente a la selección de variables exógenas, es un paso crucial en este proceso. Requiere la meticulosa elección de variables que se perciban pertinentes al fenómeno en estudio y que se haya identificado que ejercen un efecto significativo en las predicciones resultantes del modelo.

3.8.1 Variables exógenas en RNN LSTM

Dentro del marco de los modelos RNN LSTM el concepto de variables exógenas se refiere a las características o entradas adicionales que se insertan en el modelo en conjunción con la serie temporal primordial. Estas variables exógenas representan datos independientes de la serie temporal y se presume que poseen la capacidad de ejercer influencia en la predicción o el comportamiento de la serie.

La inclusión estratégica de variables exógenas tiene como propósito enriquecer y ampliar la capacidad predictiva de estos modelos, al considerar factores adicionales que pueden desempeñar un papel determinante en la dinámica de la serie temporal.

3.8.2 ARIMAX (Modelo Autorregresivo integrado de promedio con variable exógena)

El modelo ARIMAX, es una extensión del modelo ARIMA, destinado a mejorar la predicción de series temporales al incorporar variables exógenas. Las variables exógenas, ajenas a la serie principal, se considera factores relevantes para entender su comportamiento. Por ejemplo, al pronosticar la demanda de energía eléctrica, las variables exógenas pueden comprender datos climáticos como temperatura y humedad, factores que influyen en la demanda de energía.

El valor clave del ARIMAX radica en su capacidad para considerar factores externos que impactan en la serie temporal, permitiendo una captura más precisa de patrones complejos y tendencias en los datos. Esto resulta especialmente útil cuando las fluctuaciones en la serie no pueden ser completamente explicadas por relaciones internas. En resumen, el ARIMAX es una herramienta valiosa para mejorar la precisión en la predicción de series temporales al integrar factores exógenos relevantes.

3.9 Ensamblador

Un ensamblador se define como una técnica que integra múltiples modelos predictivos individuales en un modelo único, con mayor robustez y precisión. Su principal objetivo es elevar la calidad de las predicciones al aprovechar la diversidad de enfoques provenientes de varios modelos base.

- **Modelos Base:** En el contexto de un ensamblador, se emplea diversos modelos predictivos individuales conocidos como "modelos base". Estos modelos pueden pertenecer a diferentes categorías, como regresión, árboles de decisión, redes neuronales LSTM, ARIMA, ARIMAX entre otros.
- **Combinación de Predicciones:** El ensamblador combina las predicciones individuales de los modelos base para generar una predicción conjunta. Esta unión se efectúa a través de diversos procedimientos, como promedios, ponderaciones o mediante el empleo de algoritmos avanzados como Bagging.
- **Mejora de la Precisión:** La idea fundamental en el ensamblaje es que, al fusionar varios modelos que exhiben fortalezas y debilidades distintas, se incrementa la precisión global de las predicciones. Los errores inherentes a los modelos base pueden compensarse mutuamente, engendrando un modelo ensamblado que es más resistente y fiable.

En resumen, el ensamblaje de modelos se refiere a la combinación de varios modelos predictivos con el fin de mejorar la precisión de las predicciones en contraste con la utilización de un solo modelo. Esta técnica resulta particularmente beneficiosa cuando se aborda la predicción en contextos que involucra conjuntos de datos complicados o con ruido. En la sección de metodología, se proporcionará una explicación detallada sobre el tipo de ensamblador empleado.

4 METODOLOGÍA

Lo que importa verdaderamente en la vida no son los objetivos que nos marcamos, sino los caminos que seguimos para lograrlo.

- Peter Bamm-

En la presente sección, se expone los diversos procedimientos llevados a cabo en el transcurso de la investigación. Con el objetivo de facilitar una comprensión más clara, se incluirá un diagrama de procesos, representado en la Figura 4-1, que describe la metodología empleada en el estudio. Esta metodología comprende una serie de etapas, a saber: el tratamiento de datos, el raspado web, la construcción de modelos de predicción y la fase del ensamblador. Se proporciona información más detallada de cada una de las fases mencionadas en los apartados subsecuentes, con el fin de brindar una exposición más exhaustiva y técnica de los procedimientos llevados a cabo.

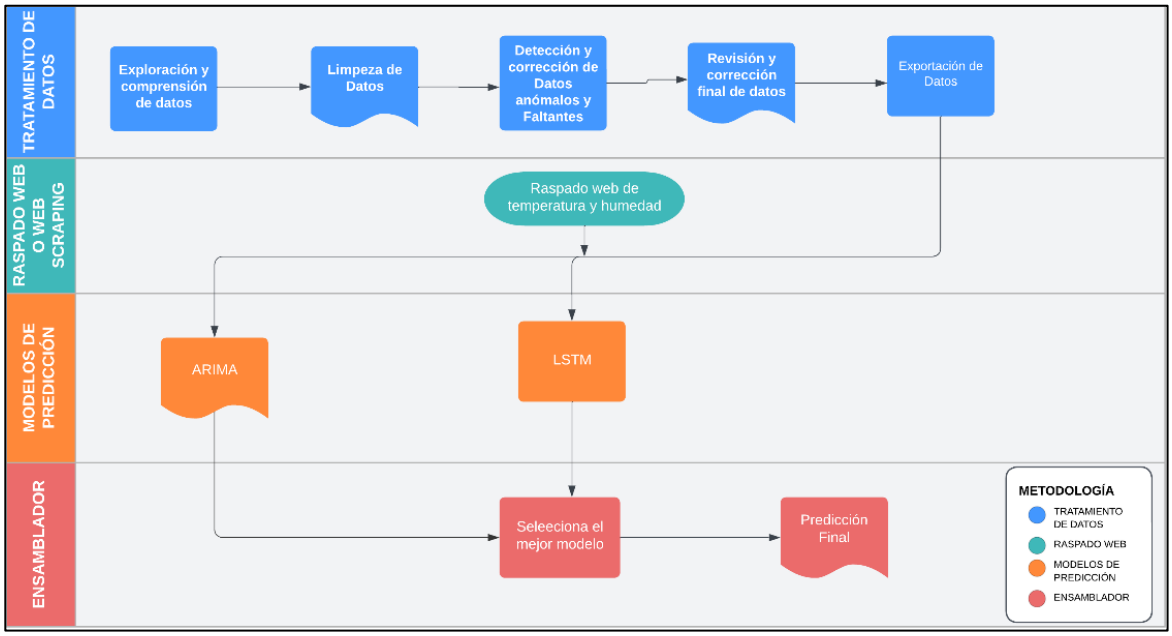


Figura 4-1. Esquema de procesos de Metodología. **Fuente Propia**

4.1 Tratamiento de Datos

Para llevar a cabo el procesamiento de datos, se optó por la utilización del lenguaje de programación Python, reconocido por su extensa colección de bibliotecas y su posición preeminente como uno de los lenguajes de programación más ampliamente adoptados en la industria contemporánea. La elección de Python se basa en su versatilidad inherente y su capacidad para simplificar considerablemente las tareas asociadas con el procesamiento y análisis de datos. A través de la explotación de diversas bibliotecas disponibles, se logró abordar eficazmente actividades fundamentales tales como la limpieza, transformación y preparación de los datos. Esto, a su vez, ha posibilitado el desarrollo de procesos de tratamiento de datos más eficientes y efectivos. Cabe destacar que las bases de datos analizadas de los centros de transformación se encuentran en discretización horaria, por lo que el tratamiento de datos se realiza en la misma discretización.

4.1.1 Exploración y comprensión de datos

Antes de iniciar el tratamiento de datos, es crucial realizar una exploración detallada para comprender a fondo la naturaleza de la información. Esta fase implica revisar la estructura del conjunto de datos y detectar posibles valores atípicos o datos faltantes. Para ilustrar este proceso, en las Figuras 4-2 y 4-3, que representan el caso de la Energía dada en [Wh] en los centros de transformación, se pueden identificar claramente valores anómalos, los cuales se encuentran encerrados en una circunferencia de color rojo.

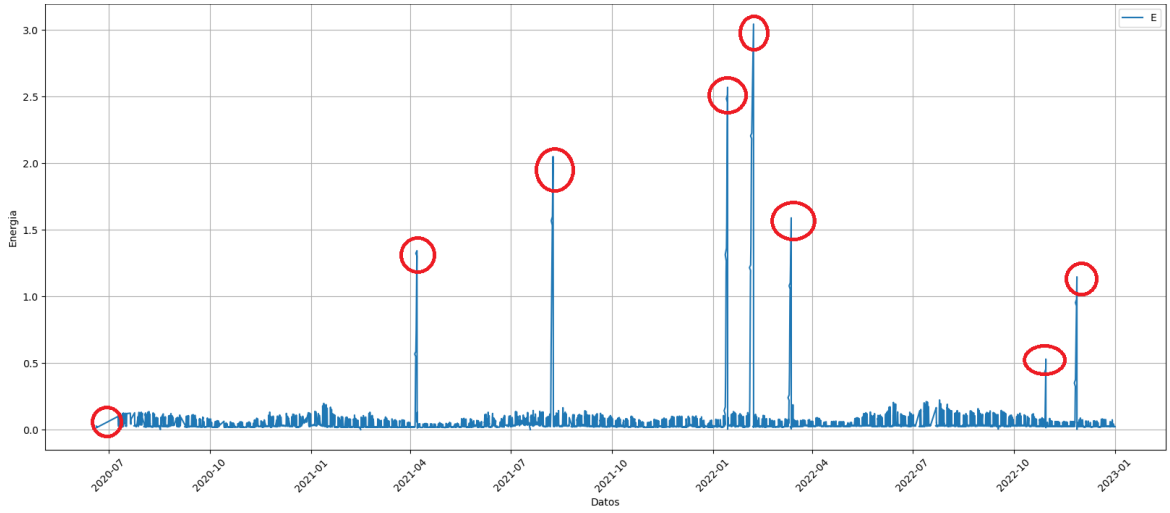


Figura 4-2. Datos de Energía en el centro de transformación A, **Fuente Propia**

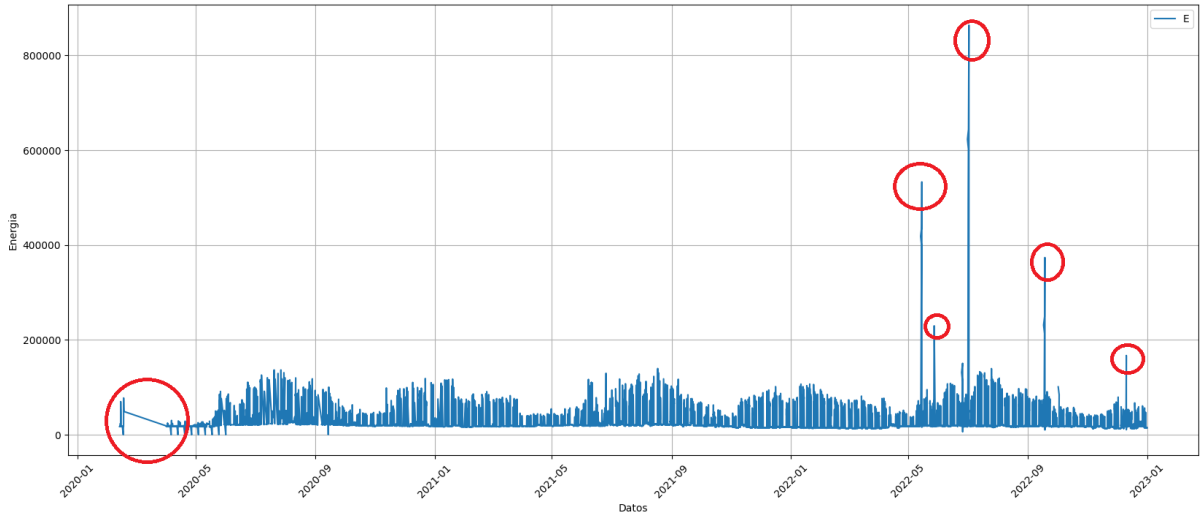


Figura 4-3. Datos de Energía en el centro de transformación B, **Fuente Propia**

Por otro lado, en la Figura 4-4 se representa el caso de la Temperatura en los centros de transformación, se observa una menor presencia de valores anómalos en principio. Sin embargo, también se marca con una circunferencia de color rojo aquellos valores que se considera anómalos. Estas representaciones gráficas son de gran utilidad para visualizar y comprender rápidamente la distribución de datos y destacar puntos atípicos que pueden requerir una atención especial en el posterior tratamiento de la información.

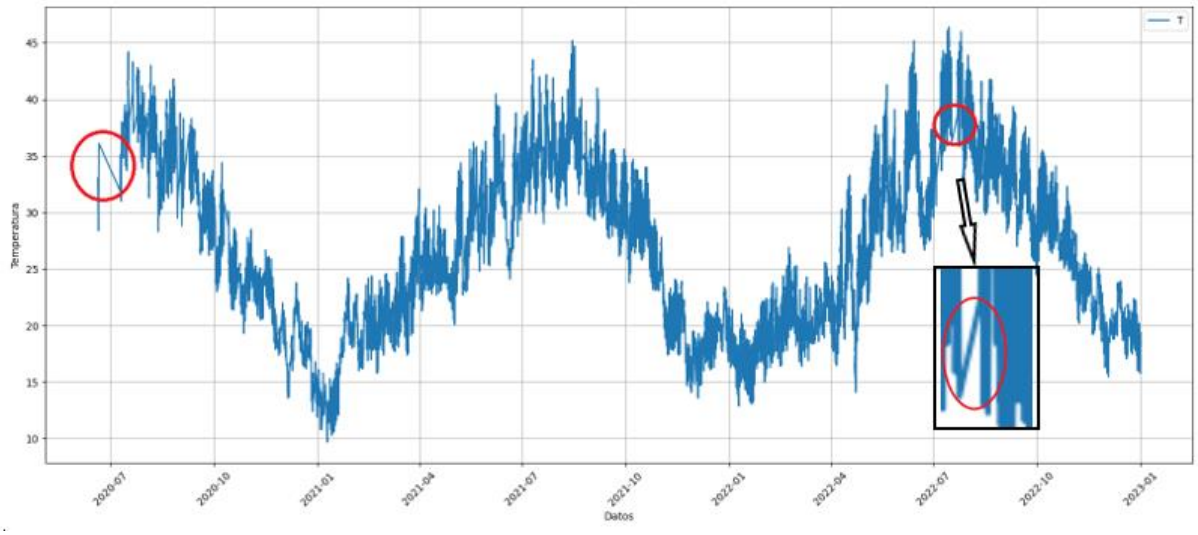


Figura 4-4. Datos de Temperatura en el centro de transformación A, **Fuente Propia**

4.1.2 Limpieza de Datos

En esta etapa, se abordan los datos faltantes y los valores atípicos. Los datos faltantes pueden ser tratados mediante métodos estadísticos, eliminación de filas o columnas con datos faltantes o mediante técnicas avanzadas de completación. Los valores atípicos pueden ser corregidos o eliminados, según el contexto y el impacto en el análisis.

En ambos casos, tanto para los datos de energía que se visualiza en la Figura 4-5 y 4-6 como para la temperatura que se visualiza en la Figura 4-7 y 4-8 en los centros de transformación, se aplicó un método de reindexación. Este proceso implicó la creación de un índice con las fechas de inicio y final, junto con la hora correspondiente, con el propósito de reorganizar la base de datos y detectar los datos faltantes. Además, en el caso de los datos de energía, se llevó a cabo un proceso de eliminación de valores atípicos o outliers que podrían afectar la calidad de los resultados. Estas acciones de reindexación y limpieza contribuyeron significativamente a mejorar la consistencia y precisión de los datos, preparándolos para el análisis y la posterior toma de decisiones informadas.

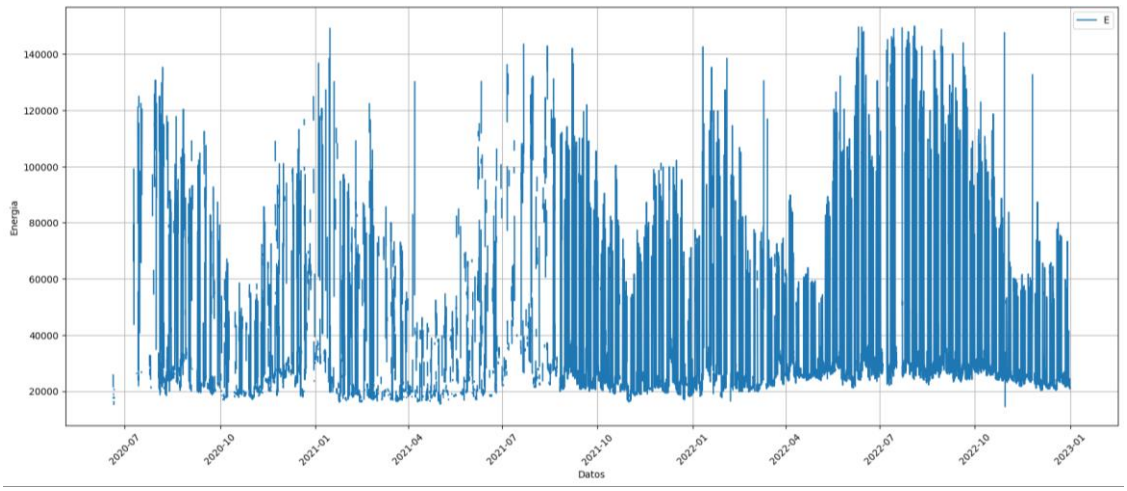


Figura 4-5. Datos de Energía Completos Centro de transformación A (Reindexación y Tratamiento de Outliers), **Fuente Propia**

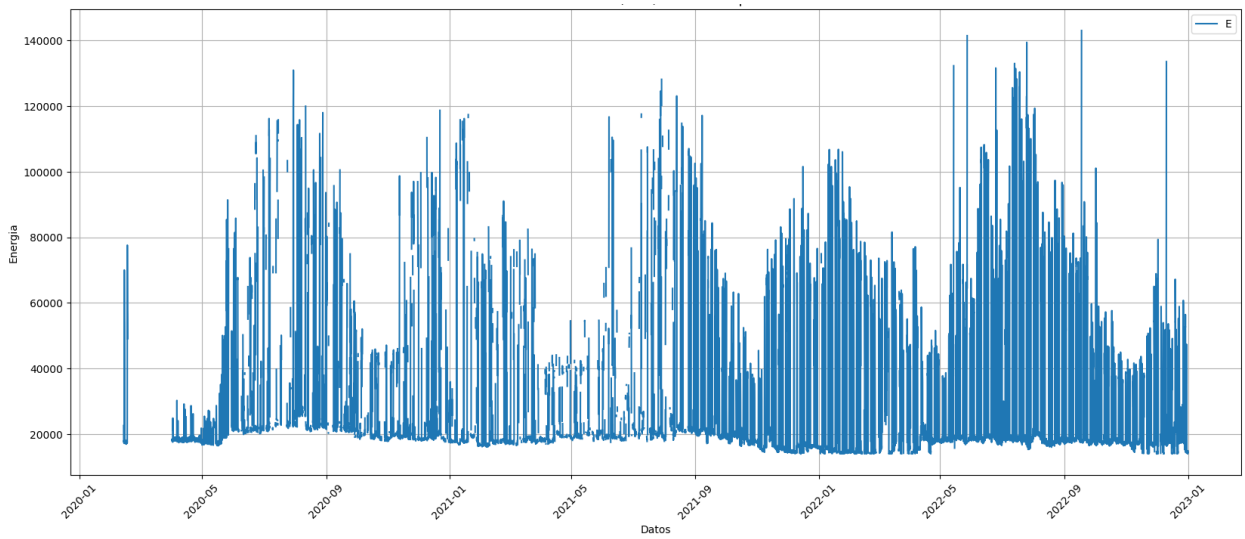


Figura 4-6. Datos de Energía Completos Centro de transformación B (Reindexación y Tratamiento de Outliers), **Fuente Propia**

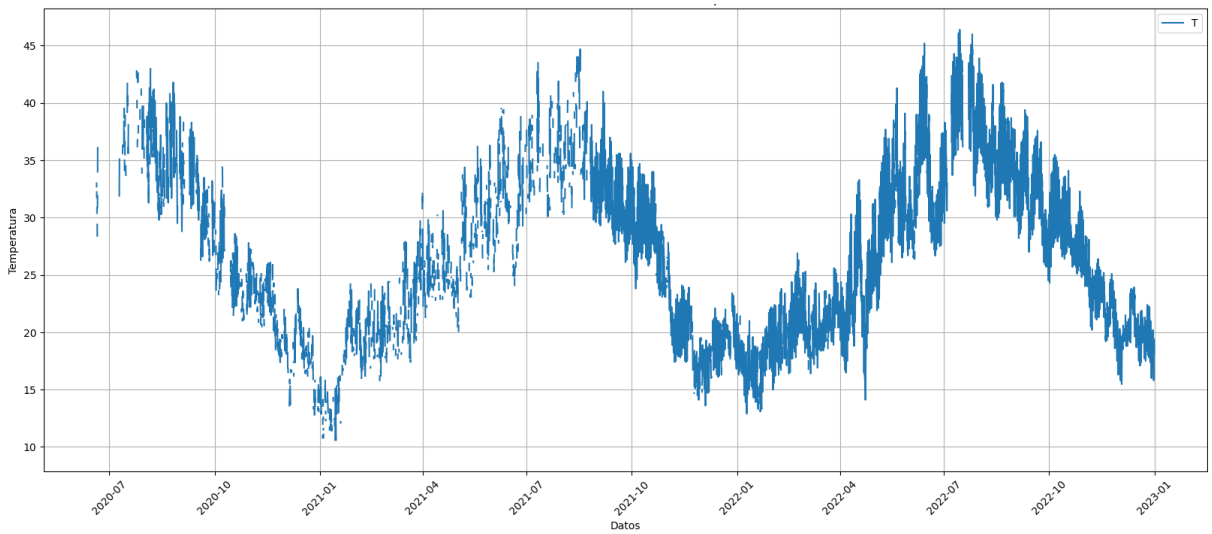


Figura 4-7. Datos de Temperatura Completos Centro de transformación A (Reindexación), **Fuente Propia**

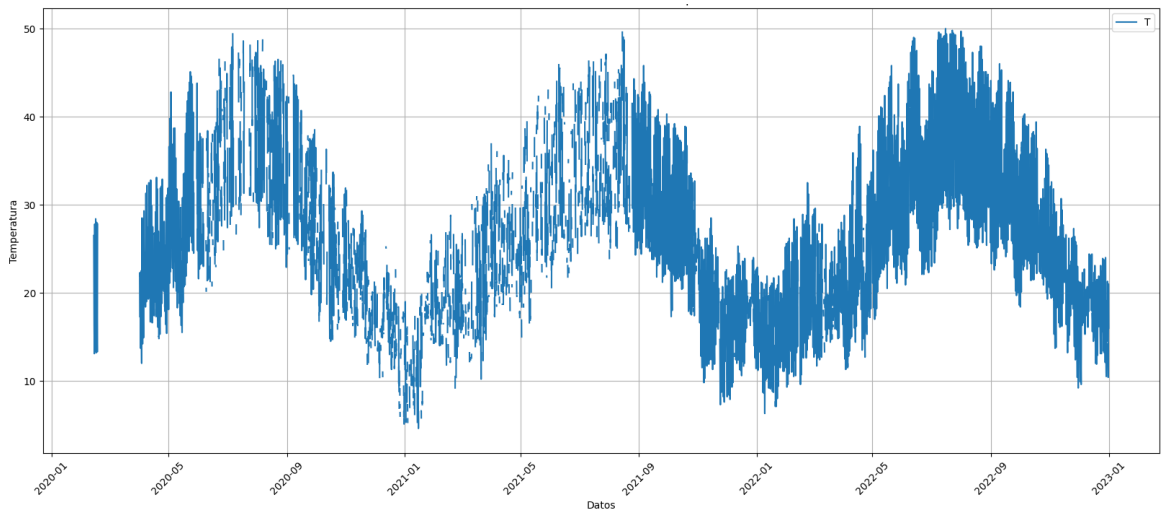


Figura 4-8. Datos de Temperatura Completos Centro de transformación B (Reindexación), **Fuente Propia**

Se presenta dos tablas que brinda una visión clara del estado inicial de la base de datos antes de ser tratada. En esta Tabla 4-1 y 4-2, se proporciona la información total de datos presentes en la base de datos, junto con el porcentaje de datos faltantes en cada una de las variables. Esta visualización es de gran utilidad para comprender la magnitud y relevancia de los datos faltantes, lo que permitirá enfocar adecuadamente el proceso de tratamiento y garantizar la integridad de los resultados.

Tabla 4–1 Información de la Base de Datos Centro de transformación A, **Fuente Propia**

Datos Totales	Datos Faltantes de Energía [%]	Datos Faltante de Temperatura [%]
22205 (100%)	30.61 %	28.95 %

Tabla 4–2 información de la Base de Datos Centro de transformación B, **Fuente Propia**

Datos Totales	Datos Faltantes de Energía [%]	Datos Faltante de Temperatura [%]
25255 (100%)	30.59 %	28.68 %

4.1.3 Detección y corrección de Datos anómalos y Faltantes

En el proceso de tratamiento de datos, la detección y corrección de datos anómalos y faltantes juega un papel crítico. Durante esta fase, se lleva a cabo un análisis exhaustivo para identificar valores erróneos, como repeticiones de datos en horas consecutivas, valores nulos y datos faltantes. Para abordar esta tarea, se utilizó el lenguaje de programación Python tal y como se ha estado realizado en las secciones anteriores del tratamiento de datos, brindando la flexibilidad necesaria para realizar una comprobación minuciosa de los datos.

En la Figura 4-9, se muestra el código utilizado para realizar la verificación de datos nulos y faltantes en la base de datos de los centros de transformación. Este código es ejecutado como parte del proceso de análisis de datos, con el propósito de identificar cualquier ausencia de información en los registros de los centros de transformación. La comprobación de datos nulos y faltantes es un paso crucial para asegurar que los datos estén completos y listos para su posterior tratamiento y análisis, lo que garantiza la integridad de los resultados obtenidos.

```
#Cuenta cuantos valores de 0 se tiene en la Energia_Wh
ceros_p = df['Energia_Wh'] == 0
#Cuenta cuantos valores de 0 se tiene en la Temperatura
ceros_t = df['Temp'] == 0

# conocer cuantos valores 0 existen
num_ceros_p = ceros_p.sum()
num_ceros_t = ceros_t.sum()
print('Existen',num_ceros_p,'de valores de Energia_Wh')
print('Existen',num_ceros_t,'de valores de Temperatura')


# Cuantos valores NaN existen
num_nan = df['Temp'].isna().sum().sum()
print('Hay', num_nan, 'valores NaN en el DataFrame.')
num_nan_E = df['Energia_Wh'].isna().sum().sum()
print('Hay', num_nan_E, 'valores NaN en el DataFrame.')
```

Figura 4-9. Código en Python para la verificación de valores nulo y faltantes en base de datos CT A, B **Fuente Propia**

Para mejorar la calidad de nuestra base de datos, implementamos un proceso de filtrado riguroso después de organizar los datos por día y hora. El primer paso en este proceso implicó la aplicación de una media móvil con una ventana de tamaño 15 para analizar la columna de Energía_Wh. Este enfoque se basó en un bucle 'for' que recorrió la columna de Energía_Wh, lo cual se replicó en todos los filtros posteriores.

Este primer filtro se centró en identificar y corregir los valores incorrectos detectados mediante la ventana móvil. Se elige una ventana de tamaño impar para asegurar de que el valor en cuestión siempre ocupara la posición central de la ventana. Esto garantizó una evaluación justa de los valores circundantes. Para sustituir los valores incorrectos, se calcula la media de la ventana móvil, lo que preservó la tendencia general de la serie temporal y proporcionó una aproximación más precisa de los valores corregidos.

La elección de una ventana de tamaño relativamente grande se justificó debido a la presencia de varios valores nulos (NaN) consecutivos en los datos. Una ventana más grande permitió capturar datos de energía en algunos casos donde existían varios valores NaN presentes, lo que a su vez permitió calcular la media con dichos valores, mejorando la precisión de la corrección. Este enfoque aseguró que nuestra base de datos estuviera libre de valores erróneos y proporcionó una base sólida para análisis posteriores.

Luego del primer filtro, se implementa una comparación entre cada dato y su predecesor para asegurar que no hubiera valores repetidos en un intervalo menor a 2 horas. Posteriormente, se aplicó un segundo filtro, nuevamente una media móvil con una ventana de tamaño 5, siguiendo el mismo principio que el primer filtro. Dado que algunos datos se corrigieron en la primera etapa de filtrado, el segundo filtro pudo corregir aún más valores. En este caso, si un dato no difería de su predecesor, se reemplazaba por un valor NaN.

Finalmente, se ejecuta un tercer y último filtro manteniendo el mismo enfoque que el segundo filtro. Con la diferencia que este último filtro se integró en un ciclo 'while' que se mantuvo activo hasta que no se detectaron más valores incorrectos en los datos. Este enfoque aseguró que los datos mantengan coherencia con la tendencia general de la serie y que los valores faltantes se completen utilizando la media de los valores adyacentes.

Es importante resaltar que en este estudio se adopta la decisión de no eliminar ningún dato, incluso en casos donde no existieron valores, es decir, donde no había registros de datos. Esta estrategia se basó en la premisa de que la ausencia de datos también es una información valiosa y puede proporcionar insights[1] relevantes en ciertos contextos. Mantener todos los registros, incluso aquellos con valores faltantes, contribuye a mantener la integridad y la completitud de la serie temporal. En este caso, los faltantes son reemplazados por valores que respetan la distribución de la serie, es decir, se sustituyen por el valor más probable que pudiese haber estado en su lugar.

La decisión de no eliminar datos, incluso cuando no había información presente, se tomó con el propósito de maximizar la utilización de todos los datos disponibles y evitar la pérdida potencial de información relevante. Además, esto refleja una aproximación más realista a la gestión de datos en situaciones prácticas donde la falta de datos es común.

Para validar la efectividad del modelo de filtrado empleado en este estudio, se propone considerar la utilización de datos provenientes de centros de transformación con una calidad conocida y una alta precisión. Esto implica la extracción de un conjunto de datos específico, lo que permitiría realizar una evaluación exhaustiva de cómo el modelo de filtrado aborda la ausencia de datos y su capacidad para reconstruir información faltante en comparación con datos de alta calidad.

A través de esta comparativa, se podrá valorar con mayor precisión la eficacia del modelo de filtrado al tratar con datos que carecen de existencia y se podrá determinar su impacto en la mejora de la calidad de los datos. Este enfoque de validación proporcionará una evaluación sólida y objetiva de la capacidad del modelo para lidiar con la ausencia de datos, lo cual es esencial para garantizar la fiabilidad de los resultados y conclusiones basados en los datos corregidos.

Esto se ilustra de manera gráfica en la Figura 4-10, que proporcionará una representación visual que facilitará la comprensión del proceso de tratamiento de datos mediante el uso de la ventana móvil.

[1] Insights se refiere a conocimientos o entendimientos significativos que se obtienen al reconocer que la falta de datos también puede ser una fuente valiosa de información en ciertos contextos de análisis.

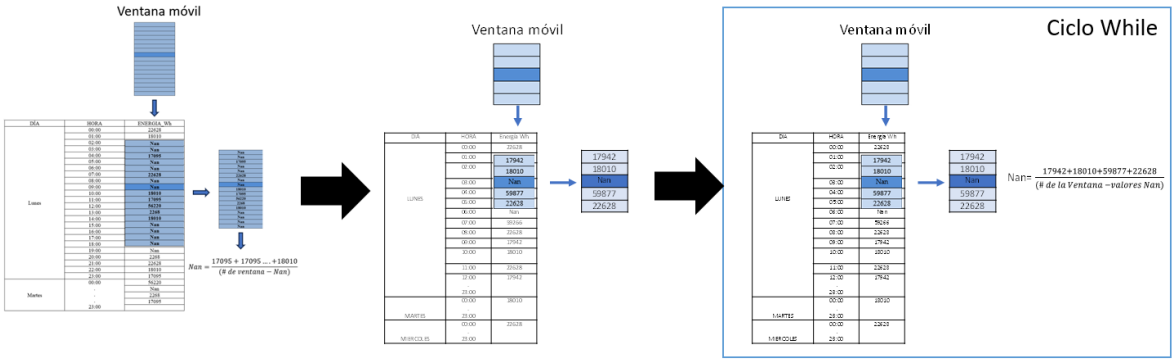


Figura 4-10. Proceso de tratamiento de datos con la ventana móvil. **Fuente Propia**

4.1.4 Revisión y corrección final de datos

En este apartado, se realiza una revisión de los datos previamente tratados mediante una media móvil diaria, con el propósito de identificar los outliers que se encontraran fuera de 2.3 sigma. La desviación estándar de 2.3 sigma representa una distancia de aproximadamente 2.3 veces la desviación estándar de la media, y se eligió un umbral de 2.3 de manera empírica después de llevar a cabo diversas pruebas y considerar la naturaleza de los datos.

La desviación estándar es una medida que indica cuánto se dispersan los datos alrededor de la media. Al utilizar 2.3 sigma, se busca detectar valores inusuales que se encuentren relativamente lejos de la media, lo que permite identificar outliers con mayor precisión.

Una vez identificados los outliers, se procedió a reemplazarlos por la media de los 5 vecinos superiores e inferiores. Para realizar este cálculo, se utilizó una ventana de tamaño 11 en la media móvil diaria, lo que permitió suavizar los datos y obtener una estimación más confiable de los valores a remplazar.

La elección de 2.3 sigma y la ventana de tamaño 11 para la media móvil se basaron en consideraciones estadísticas y en la necesidad de obtener resultados fiables en el proceso de tratamiento de datos. Al aplicar este enfoque, se logró mejorar la calidad de los datos y prepararlos para un análisis más preciso y confiable.

La Figura 4-11,4-12 representa como ejemplo de revisión que se realizó en ambos centros de transformación los datos de Energía del centro de transformación A, donde se destacan visualmente los umbrales superior e inferior de la media diaria, respetando el valor de 2.3 sigma asignado. Además, se marcan con puntos de color rojo aquellos valores que se encuentran fuera de estos umbrales, identificándolos como outliers. Esta representación proporciona una visualización más clara y concisa de lo explicado anteriormente, facilitando la identificación y comprensión de los datos atípicos.

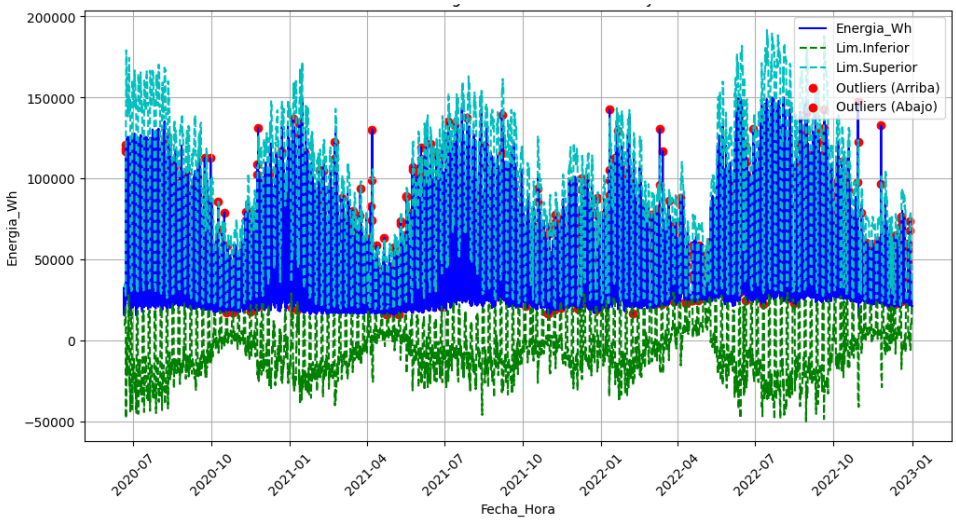


Figura 4-11. Revisión final de datos tratados energía en el CT A e identificación de Outliers **Fuente Propia**

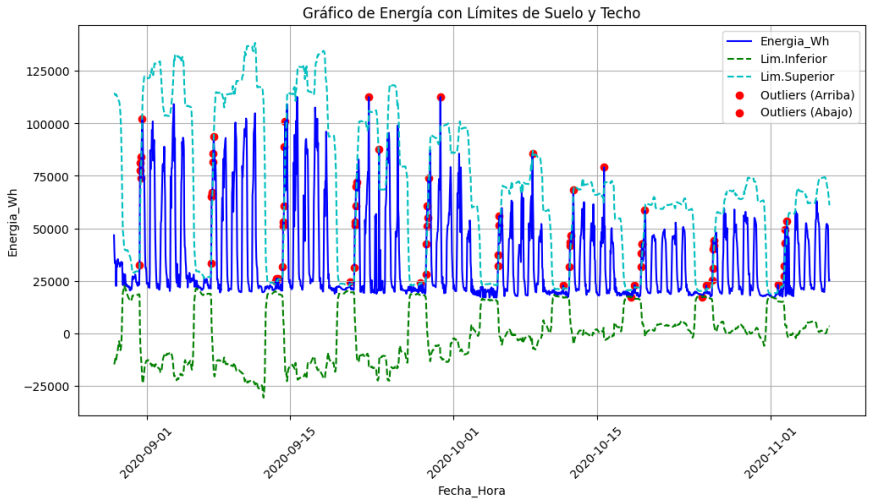


Figura 4-12. Revisión final de un grupo de datos tratados energía en el CT A e identificación de Outliers **Fuente Propia**

La Figura 4-13 y 4-14 muestran los datos ya corregidos del centro de transformación, donde se puede apreciar que todos los valores se encuentran dentro del umbral de desviación estándar establecido. Esta corrección ha permitido preparar los datos para su exportación, garantizando que estén libres de outliers y listos para su posterior análisis o uso.

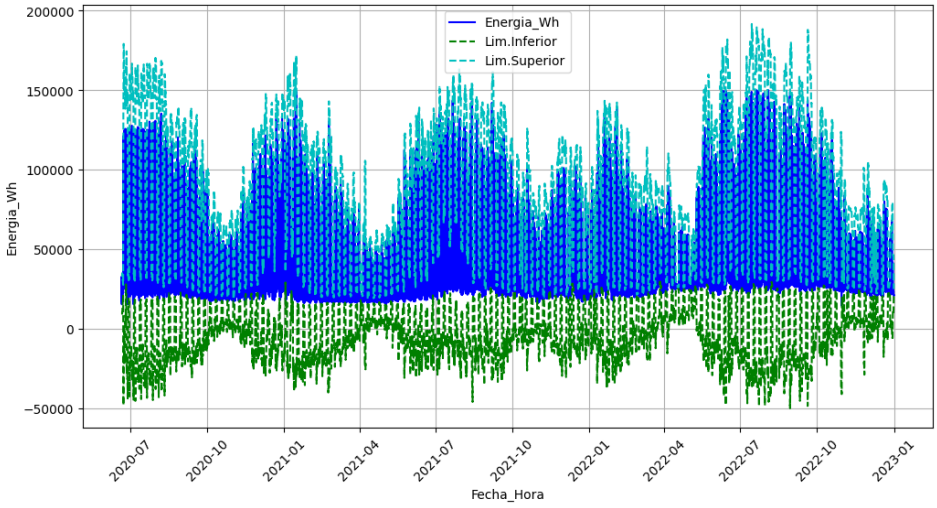


Figura 4-13. Revisión y verificación de Datos Corregidos energía CT A **Fuente Propia**

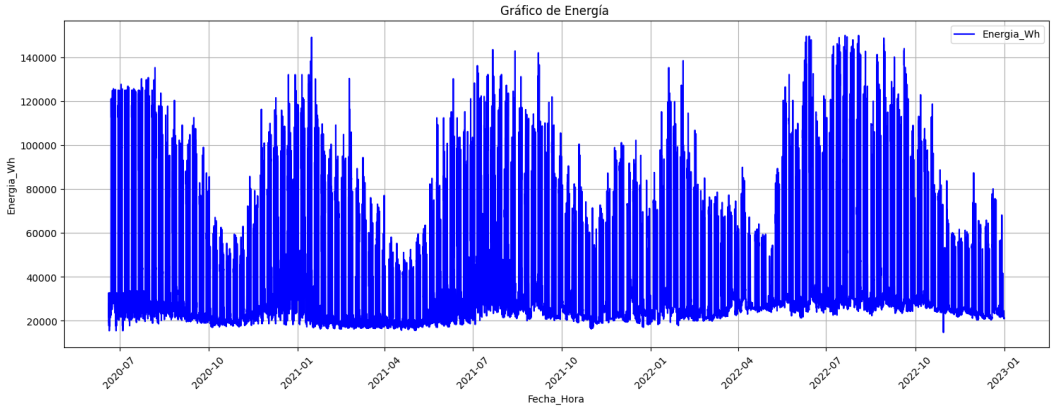


Figura 4-14. Datos Corregidos energía CT A **Fuente Propia**

4.1.5 Exportación de Datos

Una vez realizada la corrección de la base de datos, esta se encuentra preparada para su utilización y es el momento de proceder con la exportación de los datos. Esta fase es esencial para permitir que los datos tratados sean accesibles en el formato adecuado para su análisis y posterior aplicación en los modelos predictivos. Al exportar la base de datos, se asegura que los datos estén limpios y coherentes, lo que proporciona una sólida base para realizar pronósticos precisos y valiosos en el campo de la predicción de demanda eléctrica. Esta exportación optimiza la eficiencia y la viabilidad de los análisis posteriores, permitiendo una implementación efectiva de los modelos de predicción y contribuyendo a una toma de decisiones informada en el ámbito de la gestión energética y el suministro eléctrico.

4.2 Raspado Web o Web Scraping

En esta sección, se llevó a cabo la extracción de datos de la página web "https://www.tutiempo.net/registros/lemg/" a través de la implementación de un script personalizado. Este proceso revistió un alto grado de utilidad, dado que posibilita la obtención de información precisa relacionada con las variables de temperatura y humedad, todo ello sin requerir una posterior depuración de los datos.

El empleo de esta metodología permitió la adquisición de registros concernientes a la temperatura y la humedad correspondientes a momentos específicos en horas y fechas predefinidas. Este logro resulta de suma importancia en la obtención de variables exógenas de elevada calidad destinadas a su incorporación en los modelos de predicción. La disponibilidad de datos precisos y actualizados conlleva una mejora sustancial en la confiabilidad y precisión de los modelos predictivos, enriqueciendo así la capacidad de tomar decisiones informadas y eficaces en el contexto de la gestión de recursos y la planificación.

La utilización del web scraping se ha revelado como una herramienta valiosa para enriquecer el análisis y la elaboración de modelos predictivos con información relevante y de confianza. La obtención de datos directamente de la fuente de origen contribuye a mantener la integridad y la precisión de la información, lo que fortalece los resultados de los análisis y pronósticos realizados.

La Figura 4-15 exhibe una captura de pantalla de la página web https://www.tutiempo.net/registros/lemg/ en la que se realizó la extracción de datos. En esta captura, se resaltan con círculos de color rojo las variables extraídas de la mencionada página. Este procedimiento se denomina Web Scraping y se emplea para recopilar información de la web de manera automatizada.

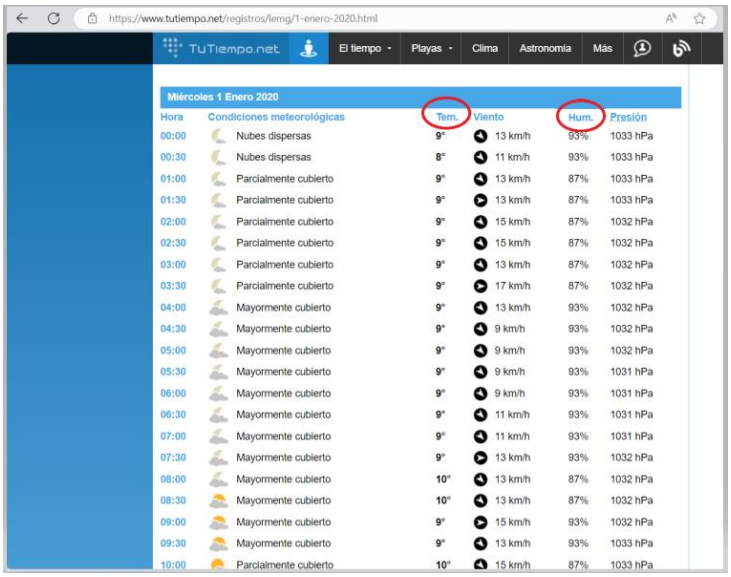


Figura 4-15. Captura de pantalla de página Web de extracción de datos.

4.3 Estacionariedad

La estacionariedad en una serie temporal representa una característica esencial en el análisis de datos secuenciales. Esta propiedad se define por la constancia en el tiempo de las propiedades estadísticas fundamentales de la serie. En términos más precisos, una serie temporal se considera estacionaria cuando su media, varianza y autocorrelación permanecen inalteradas a lo largo de su evolución temporal.

El procedimiento empleado para determinar la presencia de estacionariedad en una serie temporal se basa en la aplicación de la Prueba ADF (Augmented Dickey-Fuller). Esta prueba estadística es utilizada con el propósito de evaluar la estacionariedad en una serie cronológica específica lo cual se ve en la Tabla 4-3. Su metodología se fundamenta en la comparación entre la serie de tiempo original y una versión diferenciada de la misma. La premisa subyacente es que, en el caso de que la serie original sea no estacionaria, la aplicación de una operación de diferenciación, es decir, la obtención de la diferencia entre observaciones consecutivas resultará en una serie que exhiba estacionariedad.

Tabla 4–3 Prueba ADF a serie temporal

Prueba ADF datos de Energía			
Estadística ADF	-16.7438215673183		
Valor P	1.3513625374568317e-29		
Valores Críticos	**1%**= -3.4306451689197694	**5%**= -2.861670454988149	**10%**= -2.5668394375040435

La Tabla 4-3 presenta el resultado de una prueba de raíz unitaria conocida como la prueba de Dickey-Fuller Aumentada (ADF, por sus siglas en inglés). Esta prueba se utiliza comúnmente para determinar si una serie temporal es estacionaria o no. Aquí está la explicación de los valores que obtuviste:

Estadística ADF: El valor de la estadística ADF es -16.7438215673183 en tu prueba. Esta estadística es una medida que indica cuán lejos está la serie temporal de ser no estacionaria. En general, cuanto más negativo sea este valor, mayor evidencia hay en contra de la hipótesis nula de que la serie temporal no es estacionaria. En tu caso, el valor altamente negativo de la estadística ADF (-16.74) sugiere fuertemente que la serie temporal es estacionaria.

Valor p: El valor p asociado a la estadística ADF es 1.3513625374568317e-29, lo que es esencialmente cero. El valor p es una medida de la probabilidad de obtener una estadística ADF al menos tan extrema como la observada si la serie no fuera estacionaria (hipótesis nula). Un valor de p cercano a cero indica una fuerte evidencia en contra de la hipótesis nula de no estacionariedad. En otras palabras, prácticamente no hay evidencia que respalde la idea de que los datos no son estacionarios.

Valores críticos: Los valores críticos son umbrales que se utilizan para comparar con la estadística ADF. En tu resultado, se proporcionan valores críticos para tres niveles de significancia: 1%, 5% y 10%. La estadística ADF (-16.74) es mucho más negativa que los valores críticos en todos los niveles de significancia, lo que refuerza la conclusión de que la serie es estacionaria.

En resumen, los resultados de la prueba ADF indican que los datos de la serie temporal son altamente estacionarios.

4.4 Modelos de Predicción

En esta sección, se presenta en detalle la metodología aplicada en el desarrollo de los modelos utilizados en el estudio. Esto permite una comprensión más completa de la ejecución de dicho desarrollo. Los dos modelos que se aborda en este estudio son el Modelo de Redes Neuronales Recurrentes (RNN) LSTM y el Modelo Autorregresivo Integrado de Media Móvil (ARIMA). Cabe destacar que se proporciona descripciones detalladas de ambos modelos en secciones posteriores para una comprensión más profunda.

4.4.1 RNN LSTM

En esta sección, se detalla el proceso de diseño, configuración y entrenamiento de los modelos LSTM utilizados en el estudio. Se explora varias configuraciones de modelos con variaciones en el número de capas y otros hiperparámetros con el objetivo de seleccionar el mejor modelo de predicción.

4.4.1.1 Preprocesamiento de datos:

Para este propósito, se emplea una utilidad llamada MinMaxScaler (), integrada en la librería Scikit-learn (sklearn) de Python. Scikit-learn es un conjunto de herramientas en Python destinado a tareas de análisis y extracción de información de datos. MinMaxScaler () es una función proporcionada por esta librería que ajusta las características de tus datos para que estén contenidas en un intervalo específico, generalmente entre 0 y 1. Esto garantiza que todas las características compartan una misma escala, previniendo posibles sesgos hacia una característica con mayor magnitud en el modelo resultante. La fórmula utilizada por MinMaxScaler () para llevar a cabo la normalización es la siguiente:

$$X_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{4–1}$$

Donde:

- x es el valor original de la característica
- X_{norm} es el valor normalizado
- $\min(x)$ es el valor minimo de la caracteristica en el conjunto de datos
- $\max(x)$ es el valor máximo de la caracteristica en el conjunto de datos

- **Segmentación de datos**

La segmentación de datos resulta fundamental en el ámbito del aprendizaje automático, en especial al abordar el análisis de series temporales. En este contexto de estudio, se adopta una distribución tripartita de los datos: aproximadamente un 70-80% para entrenamiento, un 10-20% para validación y un 10% para prueba tal y como se muestra en la Tabla 4-4. Estos porcentajes son comúnmente empleados para garantizar una evaluación precisa y un ajuste efectivo de los modelos LSTM.

De manera más detallada, el conjunto de entrenamiento habilita al modelo para internalizar patrones y relaciones inherentes a los datos. La etapa de validación, desligada del entrenamiento, facilita la calibración de hiperparámetros y la mitigación del sobreajuste. Por su parte, el conjunto de prueba sirve para calibrar la eficacia del modelo en contextos inexplorados.

Como inicio del proceso, se crea una función que transforma los datos en un DataFrame en conjuntos de entrada y etiquetas. Estos conjuntos son ideales para entrenar modelos de aprendizaje automático, especialmente aquellos diseñados para manejar secuencias, como las Redes Neuronales Recurrentes (RNN). Esta función simplifica la tarea de preparar los datos en formatos adecuados para estos tipos de modelos.

Tabla 4–4 Segmentación de Datos de centros de transformación

	Centro de Transformación-A	**Centro de Transformación-B**
Entrenamiento	80% = 17764 dts	80% = 20204 dts
Validación	10% = 2220.5 dts	10% = 2525.5 dts
Test	10% = 2220.5 dts	10% = 2525.5 dts

4.4.1.2 Selección de Arquitectura de Red y Entrenamiento

En este apartado, se presenta un análisis de diversas arquitecturas de redes neuronales LSTM, que son objeto de estudio en el proyecto. La Tabla 4-5 recopila estos modelos junto con sus respectivas arquitecturas. Se observa que, en algunas de estas arquitecturas, existen variaciones en el número y tipo de capas. La primera capa, que se refiere a las entradas, establece el historial necesario de datos para la predicción, siendo comúnmente un historial de 2 semanas en la mayoría de los modelos.

En cuanto a la capa densa, el último valor corresponde a la salida de la red neuronal LSTM y se relaciona con el horizonte de predicciones. En la mayoría de los modelos, se ha configurado el horizonte de predicciones en 36 unidades, lo que equivale a anticipar los eventos futuros en un período de 36 horas. No obstante, se ha implementado un modelo adicional con un horizonte de 24 unidades con fines de referencia y comparación. Es importante destacar que el enfoque principal del estudio se centra en optimizar el horizonte de predicciones a 36 unidades, debido a su mayor grado de anticipación y relevancia en el contexto de la aplicación.

En algunos de estos modelos, se incorporan variables exógenas, mientras que en otros no. La presencia o ausencia de variables exógenas en cada modelo se detalla en la columna correspondiente de la tabla.

Es esencial enfatizar que la inclusión de variables exógenas en una red LSTM se realiza mediante la incorporación de estas variables como características adicionales en los datos de entrada de la red. Esta operación habilita a la red para considerar información externa o contextual que puede influir en las predicciones y el desempeño general del modelo. Importante destacar que la cantidad de datos correspondientes a las variables exógenas es igual a la cantidad de datos en las series temporales, lo cual garantiza un tratamiento imparcial y equitativo de ambas fuentes de información.

Para una representación visual de esta incorporación de variables exógenas en una arquitectura de red LSTM, se puede consultar la Figura 4-16, que ilustra claramente cómo estas variables adicionales se integran en el proceso de modelado de la red.

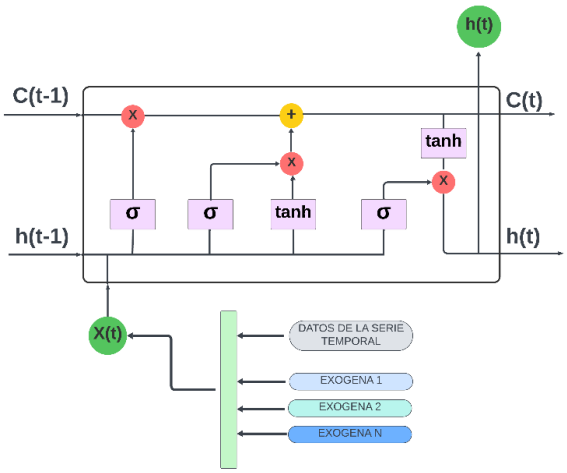
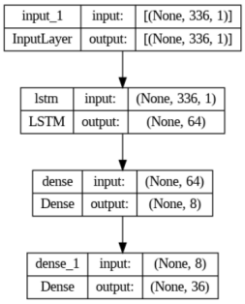
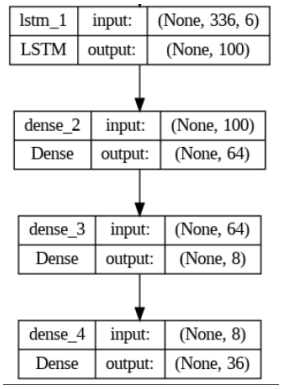
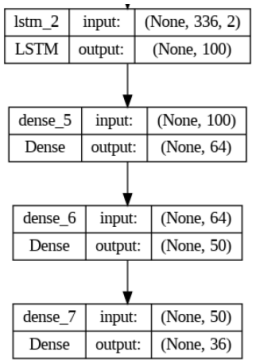
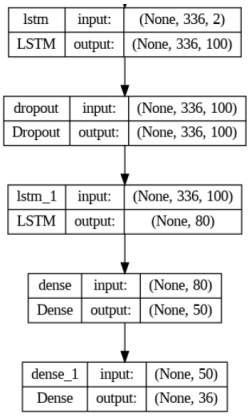
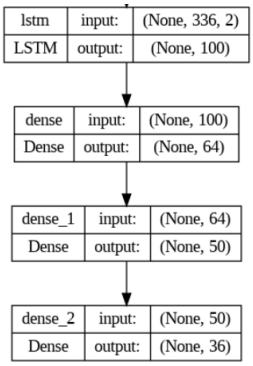
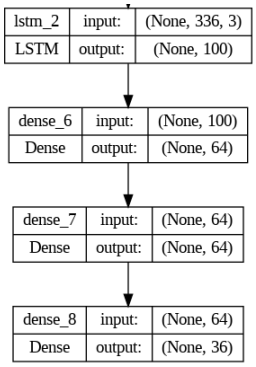
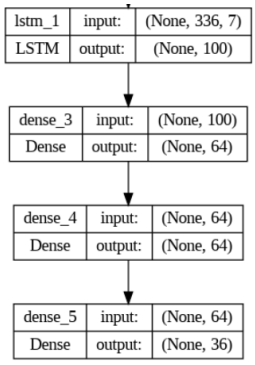
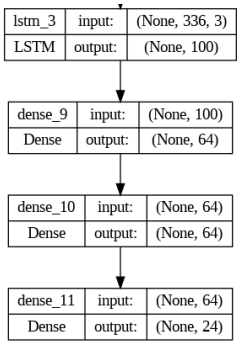


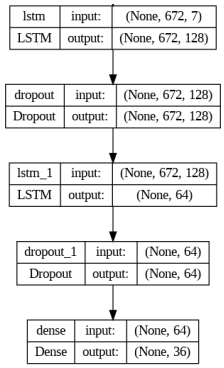
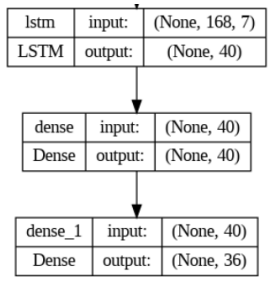
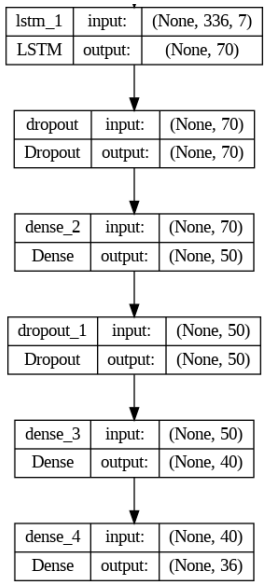
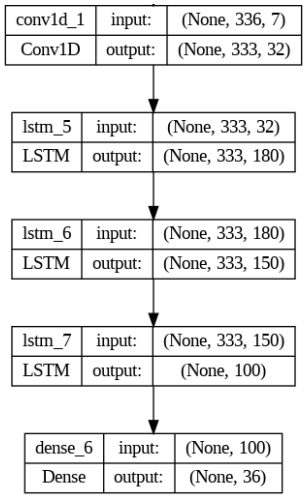
Figura 4-16. Arquitectura LSTM con variables exógenas **Fuente Propia**

En la Figura 4-16, se ve de manera clara cómo las variables adicionales se incorporan al proceso de modelado en la red LSTM. Estas variables se suman a los datos de la serie temporal y se utilizan como entradas extra en la red LSTM. Esto significa que la red toma en cuenta información adicional, como datos externos, al hacer sus predicciones. La Figura 4-16 te muestra cómo estas variables extras se integran de manera natural en el flujo de datos de la red LSTM, lo que ayuda a entender cómo influyen en el proceso de modelado.

Tabla 4–5 Descripción y caracteristicas de Modelos LSTM

Modelo	Descripción	Variable Exógena	Arquitectura	Épocas
Modelo 1 (model000)	Arquitectura LSTM	Ninguna		26
Modelo 2 (model001)	Arquitectura LSTM	Fecha, Temperatura		26
Modelo 3 (model002)	Arquitectura LSTM	Temperatura		26
Modelo 4 (model003)	Arquitectura LSTM	Temperatura		26

Modelo 5 (model004)	Arquitectura LSTM	Humedad		26
Modelo 6 (model006)	Arquitectura LSTM	Temperatura, Humedad		26
Modelo 7 (model005)	Arquitectura LSTM	Tiempo, Temperatura, Humedad.		26
Modelo 8 (model007)	Arquitectura LSTM	Tiempo, Temperatura, Humedad		26

Modelo 9 (model005-2)	Arquitectura LSTM	Tiempo, Temperatura, Humedad		30
Modelo 10 (model005-3)	Arquitectura LSTM	Tiempo, Temperatura, Humedad		10
Modelo 11 (model005-4)	Arquitectura LSTM	Tiempo, Temperatura, Humedad.		25
Modelo 12 (model005-5)	Arquitectura LSTM	Tiempo, Temperatura, Humedad.		25

Una vez presentada la Tabla 4-5, se procede a llevar a cabo una validación cruzada en series temporales utilizando la técnica conocida como TimeSeriesSplit. Esta estrategia de validación implica la división de la serie temporal en múltiples segmentos para evaluar el rendimiento de los modelos de manera más exhaustiva. En este estudio, se ha optado por dividir la serie temporal en tres bloques para asegurar que cada segmento contenga suficientes datos para entrenar y evaluar los modelos de manera efectiva.

Esta división en tres segmentos o también llamados pliegues permite una evaluación detallada del rendimiento de los modelos en diferentes secciones de la serie temporal. Cada bloque se utiliza en turnos como conjunto de validación, mientras que los otros dos se utilizan como conjuntos de entrenamiento, se puede tener una idea de lo mencionado en la Tabla 4-6. Este enfoque brinda una visión más completa de cómo nuestros modelos se desempeñan en diversas partes de los datos, lo que es esencial para comprender su capacidad de generalización. Los resultados de esta validación se presentan en detalle en el capítulo de resultados.

Es importante destacar que la metodología aplicada es coherente para ambos centros de transformación. Esto implica que las características y la arquitectura de los modelos no se modifican; la única variación radica en los datos específicos del centro de transformación en cuestión que se utilizan para el entrenamiento. Cabe resaltar que un modelo entrenado con datos de un centro de transformación no será igualmente preciso al aplicarlo a una base de datos de otro centro de transformación, ya que los datos pueden comportarse de manera diferente en cada caso. Por esta razón, se procede a entrenar un modelo con los datos apropiados según el enfoque requerido para cada centro de transformación específico.

Tabla 4–6. Representación de validación cruzada en serie temporal

Segmento o Pliegues	Conjunto de Entrenamiento	Conjunto de Validación
1	Período 1 a 5	Período 6
2	Período 1 a 6	Período 7
3	Período 1 a 7	Período 8

La Tabla 4-6 Representa una validación cruzada de series temporales con 3 pliegues en una serie de tiempo hipotética.

4.4.2 ARIMA

El modelo ARIMA está compuesto por tres componentes fundamentales: el componente autoregresivo (p), el componente de integración (d) y el componente de media móvil (q). La estimación de los parámetros (p, d, q) asociados a estos componentes implica un proceso matemático de considerable extensión. En este estudio, se llevó a cabo una serie de procedimientos con el propósito de determinar los valores adecuados para estos parámetros.

Inicialmente, se realizó un enfoque en el parámetro "p," el cual corresponde al componente autoregresivo de nuestro modelo. Este parámetro es fundamental ya que cuantifica la cantidad mínima de observaciones previas que deben considerarse al realizar las predicciones. Para determinar el valor óptimo de "p," se adopta la estrategia de discretizar la serie temporal en intervalos de 4 horas. Esta elección se fundamenta en la conveniencia de trabajar con múltiplos de 24 horas, lo que facilita la división equitativa del día.

La discretización en intervalos de 4 horas tuvo un impacto significativo en la mejora de la visualización de la serie temporal, lo que contribuyó a nuestro proceso analítico. Al realizar un análisis gráfico detallado, se identifica un patrón recurrente que se manifestaba cada 42 observaciones, equivalente a una semana en el contexto de la discretización de 4 horas. Este hallazgo es esclarecedor, ya que indicó que para comprender de manera adecuada el comportamiento de la serie y realizar pronósticos precisos para 36 observaciones futuras, no sería suficiente considerar únicamente unas pocas horas o días de datos previos.

En consecuencia, se determina que el mínimo número de observaciones previas a considerar debería corresponder a una semana completa, lo que se traduce en 168 observaciones en la discretización horaria que se

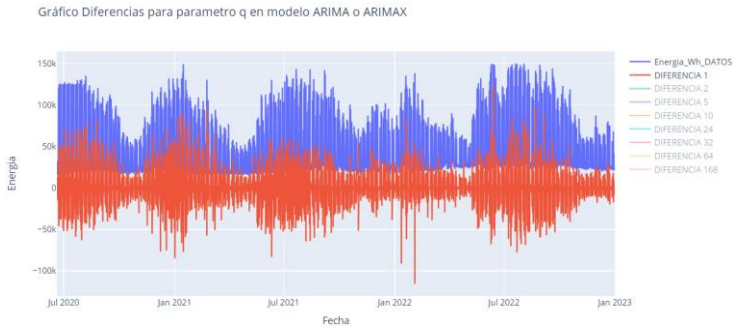
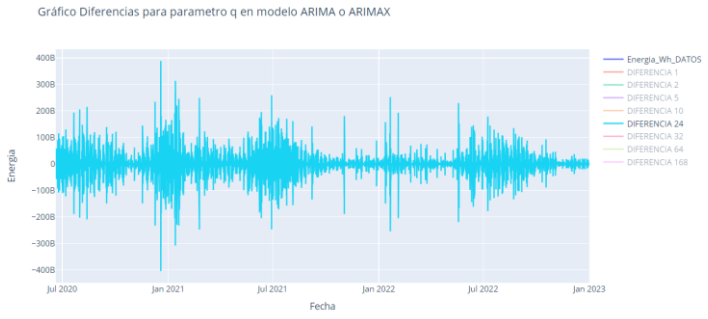
emplea en este estudio. Cabe destacar que esta elección de "p" resulta en una demanda computacional considerable debido a la inclusión de un gran número de valores previos en el modelo. Por lo tanto, de manera pragmática, se opta por reducir este valor empíricamente, equilibrando así la necesidad de precisión con las restricciones computacionales.

En relación al parámetro "d", que está vinculado a la diferenciación de la serie temporal, se llevaron a cabo diversas iteraciones de diferenciación. Se nota que la primera diferenciación lograba la estacionariedad de los datos. Por lo tanto, se opta por un valor inicial de "d" igual a 1. Posteriormente, se llevaron a cabo varias pruebas de modelos para determinar si una diferencia de 1 produce resultados satisfactorios o superiores en comparación con otros valores de "d". Este enfoque se respalda con evidencia documentada en la Tabla 4-7, que muestra las distintas diferenciaciones aplicadas a la serie con el fin de ilustrar lo mencionado anteriormente de manera gráfica.

Finalmente, la determinación del parámetro "q" se llevó a cabo mediante un enfoque empírico que implicó la evaluación de múltiples valores hasta alcanzar una selección óptima para el componente de media móvil (MA). Este proceso de selección de parámetros se presenta de manera detallada en la Tabla 4- 8.

Es importante destacar que el proceso de determinación de los parámetros (p, d, q) no se limita exclusivamente a un modelo ARIMA, ya que, durante las diversas pruebas realizadas, se consideraron varios modelos que incluían variables exógenas. Esto llevó a la identificación de dos tipos de modelos: ARIMA y ARIMAX, los cuales se registraron en la Tabla 4-8, específicamente en la columna de "Modelos".

Tabla 4–7. Diferencias de la serie temporal para parámetro "d" ARIMA, ARIMAX

Numero de Diferencias	Gráfica de Diferencias vs Datos de serie Temporal (Energía_Wh)
1	
2	
24	

168	

Tabla 4–8. Arquitectura de Modelo ARIMA, ARIMAX

Nombres de Modelos	Coeficiente p: Orden de la parte autoregresiva (AR).	Coeficiente d: Orden de la integración (diferenciación).	Coeficiente q: Orden de la parte de media móvil (MA).	Variable Exógena	Modelos
ARIMA_1	12	0	0	Sin exógena	ARIMA
		0	0	Temperatura	ARIMAX
		0	1	Sin exógena	ARIMA
		0	1	Temperatura	ARIMAX
		0	2	Sin Exógena	ARIMA
		0	2	Temperatura	ARIMAX
		0	2	Temperatura y Humedad	ARIMAX
		1	0	Sin exógena	ARIMA
		1	0	Temperatura	ARIMAX
		1	1	Sin exógena	ARIMA
		1	1	Temperatura	ARIMAX
ARIMA_2	24	0	0	Sin exógena	ARIMA
		0	0	Temperatura	ARIMAX
		0	0	Temperatura y Humedad	ARIMAX
		0	1	Temperatura	ARIMAX
		0	2	Sin Exógena	ARIMA
		0	2	Temperatura	ARIMAX
		0	2	Temperatura y Humedad	ARIMAX
		1	0	Temperatura	ARIMAX
		1	0	Temperatura y Humedad	ARIMAX
		1	1	Sin exógena	ARIMA
		1	1	Temperatura	ARIMAX
		1	1	Temperatura y Humedad	ARIMAX
		1	2	Temperatura	ARIMAX

		1	2	Temperatura y Humedad	ARIMAX
		2	0	Sin exógena	ARIMA
		2	0	Temperatura	ARIMAX
		2	0	Temperatura y Humedad	ARIMAX
ARIMA_3	42	0	0	Sin exógena	ARIMA
		0	0	Temperatura	ARIMAX
		0	0	Temperatura y Humedad	ARIMAX
		0	2	Sin exógena	ARIMA
ARIMA_4	48	1	1	Temperatura	ARIMAX
ARIMA_5	72	0	0	Sin exógena	ARIMA
		0	0	Temperatura	ARIMAX
		0	2	Sin Exógena	ARIMA

En última instancia, la elección adecuada entre los modelos ARIMA o ARIMAX se basa en una evaluación exhaustiva de la serie temporal utilizando métricas de rendimiento clave, como el Error Cuadrático Medio (MSE), el Error Absoluto Medio (MAE) y la Raíz del Error Cuadrático Medio (RMSE). Estas métricas desempeñan un papel fundamental en la medición de la calidad de las predicciones generadas por los modelos.

Los valores específicos de cada una de estas métricas se detallan y analizan en profundidad en la sección de resultados. A través de esta evaluación rigurosa, se identificará el modelo que mejor se ajusta a los datos observados y que logra las predicciones más precisas.

Es importante destacar que, en el caso del modelo ARIMAX, además de los parámetros tradicionales de ARIMA (p, d, q), también se considera la influencia de las variables exógenas (denotadas como "X"). Estas variables exógenas pueden representar factores externos que impactan en la serie temporal. La elección adecuada de estas variables es esencial para mejorar la capacidad predictiva del modelo ARIMAX.

Esta selección del mejor modelo no solo permitirá tomar una decisión informada sobre si utilizar un modelo ARIMA o ARIMAX, sino que también proporcionará los valores óptimos para los parámetros ARIMA (p, d, q) y las variables exógenas (X). Estos parámetros y variables son críticos para capturar de manera efectiva las características fundamentales de la serie temporal en análisis, lo que garantiza que el modelo sea una herramienta adecuada para la toma de decisiones y pronósticos precisos.

4.5 Ensamblador

En esta sección, se desarrolla un metamodelo con la capacidad de combinar las cualidades sobresalientes de los modelos previamente entrenados. Este metamodelo se construye mediante el uso de la biblioteca sklearn.linear_model y, en particular, empleando la clase LinearRegression.

La elección de LinearRegression como componente del metamodelo se basa en su capacidad para capturar y combinar las características más relevantes de los modelos previos. Esta clase realiza una regresión ponderada de las predicciones de los modelos base y ajusta los coeficientes de manera óptima para minimizar el error cuadrático medio. En esencia, busca la combinación lineal de las predicciones individuales que mejor se ajuste a los datos. La figura 4-17 y 4-18. Representa gráficamente el concepto de agrupar y sintetizar las características de mayor relevancia presentes en los modelos, con la finalidad de forjar un metamodelo como resultado final

Se diseña dos ensambladores de modelos utilizando la clase LinearRegression, previamente descrita. El primer ensamblador se fundamenta en la implementación de modelos de redes neuronales recurrentes (RNN) LSTM, aprovechando así sus características sobresalientes. Esta estrategia ha conducido a resultados altamente

prometedores, haciendo uso de los modelos especificados en la Tabla 4-5.

Por otro lado, el segundo ensamblador se caracteriza por combinar el mejor modelo RNN LSTM con el mejor modelo ARIMA o ARIMAX, en función del rendimiento de los modelos que están siendo sometidos a evaluación, como se detalla en la Tabla 4-7. Esta combinación ha generado igualmente resultados alentadores. Los análisis adicionales, se exponen detalladamente en la sección de resultados.

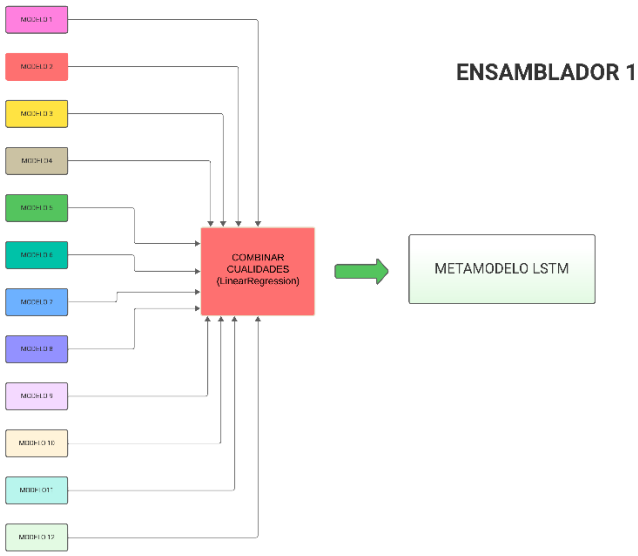


Figura 4-17. Ensamblador 1(Ensamblador LSTM) **Fuente Propia**

La Figura 4-17 ilustra el Ensamblador 1, el cual consta de los 12 modelos de RNN LSTM seleccionados para este estudio, con el propósito de identificar y combinar sus características destacadas para la creación de un metamodelo denominado "METAMODELO LSTM".

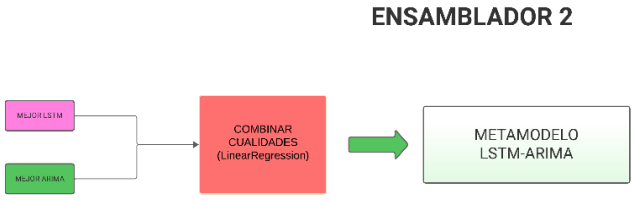


Figura 4-18. Ensamblador 2 (Ensamblador LST-ARIMA, ARIMAX) **Fuente Propia**

En la Figura 4-18 se presenta el Ensamblador 2, cuyo objetivo es similar al del Ensamblador 1. Ambos ensambladores buscan identificar y combinar las características sobresalientes de los modelos para crear un metamodelo. Sin embargo, la distinción radica en que el Ensamblador 2 utiliza el mejor modelo de LSTM en conjunto con el mejor modelo ARIMA. Esta combinación da lugar a la creación del metamodelo denominado "METAMODELO LSTM-ARIMA-ARIMAX".

5 RESULTADOS Y DISCUSIONES

La verdadera sabiduría está en reconocer la propia ignorancia.

- Sócrates-

En esta sección, se expone los resultados, análisis y comparaciones de los modelos entrenados con el objetivo de identificar el modelo óptimo. En primer lugar, se presentan individualmente cada uno de los modelos junto con sus respectivos resultados. Posteriormente, se llevará a cabo un análisis y comparación detallados que serviren de base para la selección del modelo de predicción más adecuado en el contexto de este estudio.

5.1 Modelos

En la siguiente sección, se detallan los resultados de los modelos de predicción utilizados en este estudio con el objetivo de proporcionar un análisis exhaustivo de su rendimiento y eficacia.

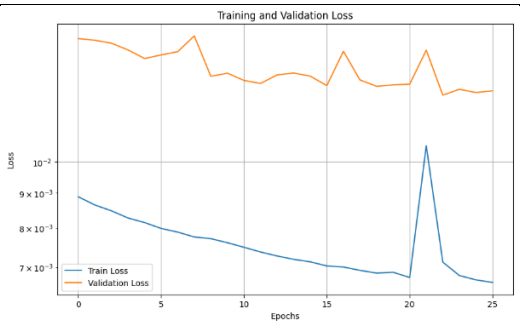
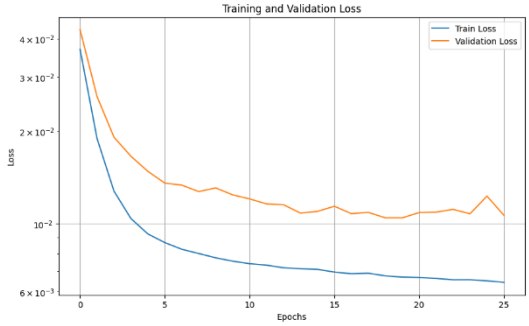
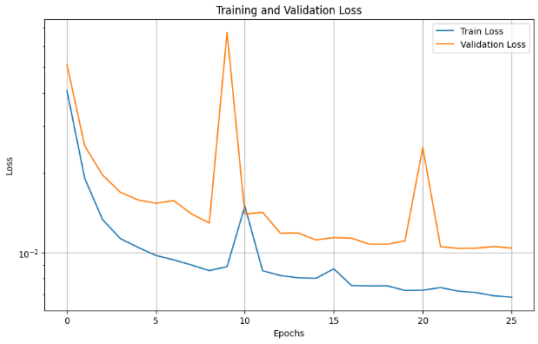
5.1.1 RNN LSTM

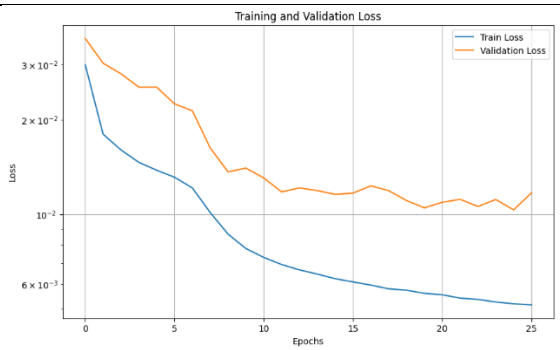
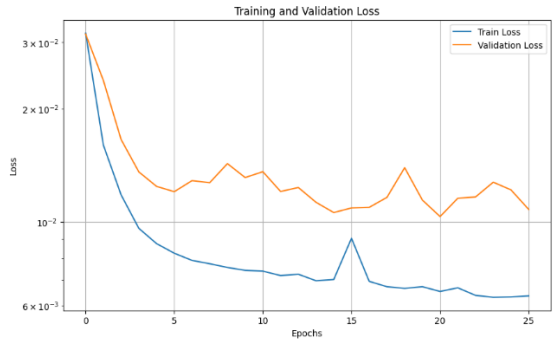
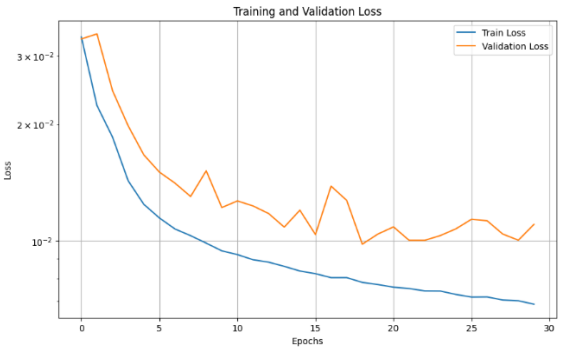
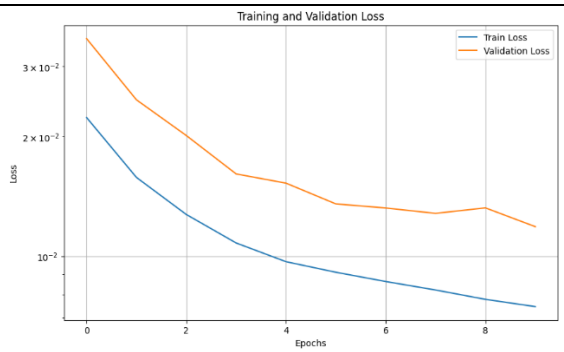
Al discutir los aspectos relevantes en el capítulo 3, se presenta los resultados derivados de la aplicación de diversos modelos de Redes Neuronales Recurrentes LSTM. Estos resultados se exhiben en la Tabla 5-1. Los valores presentes en esta tabla se mantienen dentro del contexto de normalización previamente mencionado. El propósito de esta sección es proporcionar una comprensión más profunda de la relevancia y las implicaciones de estos resultados dentro del contexto de este estudio.

Es importante destacar que, dado que los resultados muestran similitudes en términos de métricas de validación, pérdidas y curvas de pérdida, se ha optado por presentar los datos de un centro de transformación en particular.

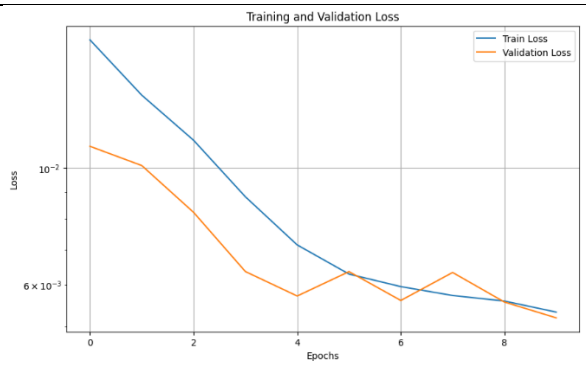
Tabla 5–1. Resultados modelos de RNN LSTM

Modelo	Loss Train	Loss Val	Loss RMSE(Val)	Grafica de Curvas de perdidas, entrenamiento y validación
Modelo 1 (model000)	0.0080	0.0126	0.1122	

Modelo 2 (model001)	0.0066	0.0127	0.1128	
Modelo 3 (model002)	0.0064	0.0106	0.1032	
Modelo 4 (model003)	0.0063	0.0102	0.1008	
Modelo 5 (model004)	0.0068	0.0104	0.1021	
Modelo 6 (model006)	0.0065	0.0107	0.1034	

Modelo 7 (model005)	0.0055	0.0109	0.1045	
Modelo 8 (model007)	0.0065	0.0092	0.1016	
Modelo 9 (model005-2)	0.0072	0.100	0.0959	
Modelo 10 (model005-3)	0.0075	0.0119	0.1089	
Modelo 11 (model005-4)	0.0090	0.0107	0.1033	

Modelo 12 (model005-5)	0.0049	0.0088	0.0940



En la Tabla 5-1 se observa que el Modelo 12 exhibe pérdidas más bajas tanto en el conjunto de entrenamiento como en el de validación. A primera vista, se selecciona el Modelo 12 como el mejor, ya que muestra resultados superiores en términos de pérdidas, específicamente con valores de 0.0049 en entrenamiento y 0.0940 para las pérdidas RMSE en validación, el RMSE (Root Mean Square Error o Error Cuadrático Medio de la Raíz) es una métrica que mide la discrepancia entre las predicciones y los valores reales, lo que facilita la interpretación.

Adicionalmente, se incluye gráficas que representan las curvas de pérdidas tanto en el conjunto de entrenamiento como en el de validación de los modelos. Estas curvas desempeñan un papel crucial, ya que proporcionan una visualización dinámica de cómo el rendimiento del modelo evoluciona durante el proceso de entrenamiento. La Curva de Pérdida de Entrenamiento muestra cómo la pérdida del modelo disminuye gradualmente a medida que se ajusta a los datos de entrenamiento, lo que refleja su capacidad de aprendizaje. La Curva de Pérdida de Validación es esencial para evaluar si el modelo generaliza de manera efectiva a nuevos datos, ya que indica si la pérdida en el conjunto de validación permanece baja o comienza a aumentar, lo que podría señalar un problema de sobreajuste.

Estas curvas de pérdidas son fundamentales para tomar decisiones informadas sobre la elección del mejor modelo, ya que permite evaluar de manera dinámica cómo se desempeña el modelo a lo largo del tiempo y si está aprendiendo y generalizando de manera efectiva.

No obstante, se reconoce que la evaluación basada únicamente en la inspección visual no es suficiente para validar de manera sólida los modelos, especialmente en el contexto de series temporales. Por lo tanto, se opta por aplicar la validación cruzada, una técnica ampliamente reconocida para validar modelos en el ámbito de las series temporales de manera precisa y confiable.

La validación cruzada en series temporales implica dividir la serie temporal en múltiples segmentos secuenciales en el caso de este estudio de lo hizo en 3, asegurando que la información futura no influya en la información pasada. Esto simula cómo se usaría el modelo en situaciones reales, donde solo se dispone de datos históricos para hacer predicciones futuras. Luego, el modelo se entrena y evalúa en múltiples iteraciones, utilizando conjuntos de datos de entrenamiento que preceden cronológicamente a los conjuntos de datos de validación.

En este contexto específico, se utilizan las métricas MSE (Error Cuadrático Medio) y MAE (Error Absoluto Medio) para evaluar el rendimiento del modelo en cada iteración de la validación cruzada en series temporales. El MSE mide el promedio de los errores al cuadrado entre las predicciones y los valores reales, siendo útil cuando se buscan minimizar discrepancias grandes. El MAE, por su parte, toma el promedio de los valores absolutos de los errores, lo que lo hace menos sensible a valores extremos y proporciona una medida directa de la discrepancia promedio entre las predicciones y los datos reales

Los resultados de esta validación cruzada se presentan de manera gráfica mediante un diagrama de barras, lo que facilitará la comprensión de la evaluación realizada en la Figura 5-1

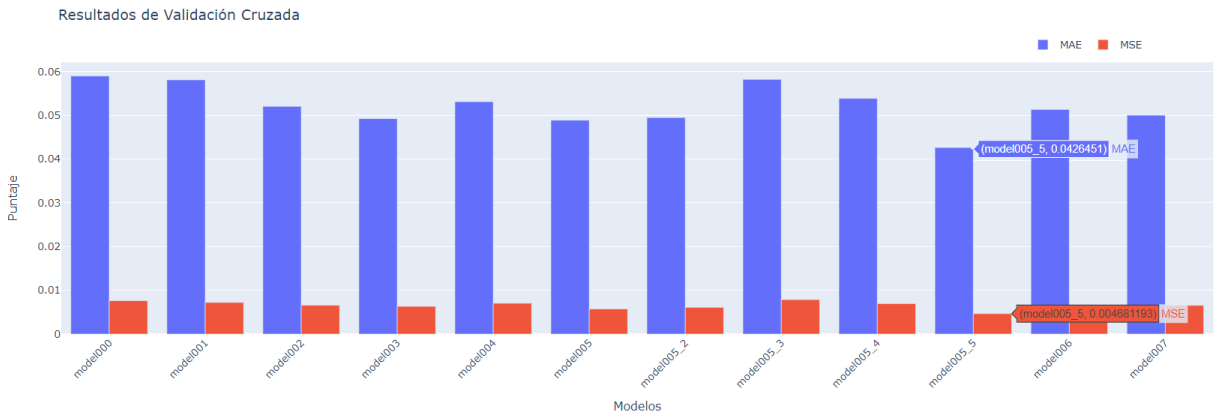


Figura 5-1. Resultados Validación Cruzada Modelos LSTM

Como resultado de este proceso de evaluación, se ha determinado que el model005_5, designado como Modelo 12, se destaca como el mejor en términos de su desempeño general en las métricas evaluadas. Este modelo logra un MAE (Error Absoluto Medio) de 0.0426451 la cual es representada con la barra de color azul en la Figura 5-1 y un MSE (Error Cuadrático Medio) de 0.004681, como se representa en la barra de color rojo en el diagrama de barras adjunto previamente. Este rendimiento sobresaliente se traduce en su habilidad para realizar predicciones significativamente más precisas en comparación con los otros modelos analizados.

Se muestra la predicción realizada por el Modelo 12 (model005_5) para el conjunto de prueba en la Figura 5-2 En esta representación visual, se puede apreciar que el modelo logra una predicción efectiva al mantener la tendencia de la serie temporal.

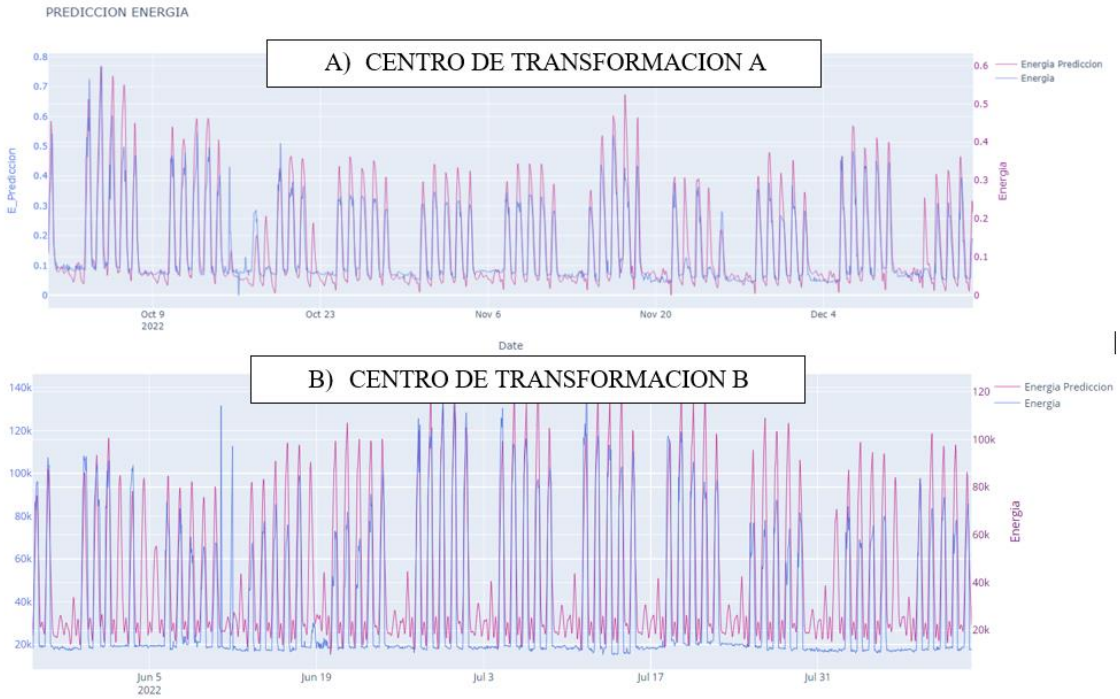


Figura 5-2. A) y B) Predicción de Energía del grupo de prueba (test)

La Figura 5-2 presenta dos gráficos: el primero, representado por la línea azul, muestra la serie temporal de la variable de energía en el grupo de prueba. El segundo, representado por la línea roja, ilustra las predicciones generadas por el Modelo 12. Al examinar estos gráficos, es evidente que el modelo ha logrado realizar predicciones efectivas que siguen de cerca la tendencia de la serie temporal de la variable de energía.

Mediante la Figura 5-3 muestra una similitud con la figura anterior, con la diferencia de que se ha agregado una línea verde que representa las predicciones para los 36 valores del último grupo de prueba, es decir, se trata de valores futuros del conjunto de prueba. Esto proporciona una visión anticipada de cómo el modelo se desempeña

en la predicción de datos aún no observados.

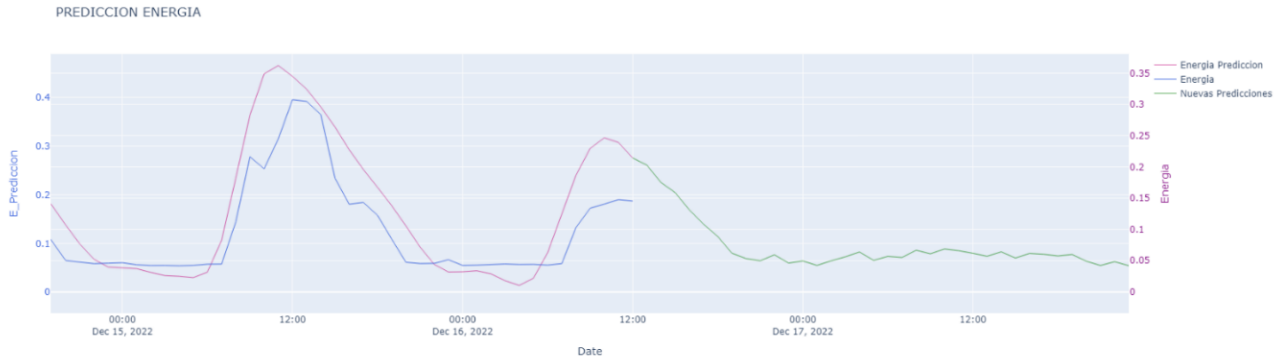


Figura 5-3. Predicción de energía del grupo prueba más 36h de predicción

La arquitectura del Modelo 12 implementa una combinación de técnicas para el procesamiento de series temporales. Comienza con una capa convolucional que utiliza 32 filtros y la función de activación ReLU (Rectified Linear Unit) para detectar patrones locales en los datos de entrada. La función ReLU es ampliamente utilizada en redes neuronales debido a su capacidad para introducir no linealidad al mantener los valores positivos intactos y convertir los valores negativos en cero, lo que ayuda al modelo a capturar relaciones no lineales en los datos. Luego, se aplican tres capas LSTM en secuencia con diferentes cantidades de unidades y tasas de abandono (dropout) para capturar tanto dependencias a corto como a largo plazo en la secuencia temporal. La elección de estos valores se basa en la experimentación y ajuste fino, buscando un equilibrio entre la capacidad del modelo para capturar patrones complejos y su capacidad para evitar el sobreajuste. Además, la capa de salida utiliza una función de activación lineal, que produce predicciones directas sin aplicar transformaciones adicionales a las salidas de las capas anteriores, lo que es apropiado para una tarea de regresión como la predicción de series temporales. Estos valores específicos se determinan mediante pruebas empíricas para lograr un rendimiento óptimo en la tarea de predicción de series temporales, permitiendo un aprendizaje efectivo y una generalización adecuada. La arquitectura expuesta se representa en la Figura 5-4

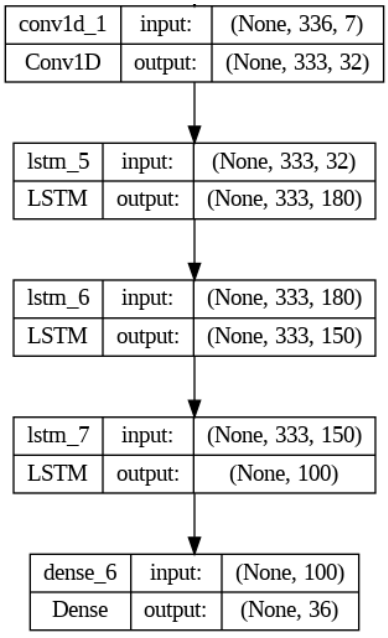


Figura 5-4. Arquitectura Modelo 12 (model005_5)

Para un analasis mas objetivo del desempeño de los modelos, se llevó a cabo una comparación entre los Modelos de redes neuronales LSTM. El propósito de esta comparación es evaluar el desempeño del modelo seleccionado en relación con sus contrapartes, así como identificar datos relevantes y sutiles que puedan derivarse de este análisis comparativo.

El enfoque principal de esta comparación se orienta hacia la evaluación de la precisión de las predicciones generadas por los modelos seleccionados en relación con los datos reales. La evaluación se basa en la medición de los errores expresados en términos de la diferencia con los valores reales.

En las Tablas 5-2, 5-3 y 5-4 se presentan las comparaciones de los modelos utilizando diversas métricas estadísticas en diferentes intervalos de discretización temporal. Como se ha mencionado previamente, el horizonte de predicciones se establece en 36 horas, manteniendo una resolución temporal discreta a nivel horario. No obstante, con el propósito de realizar un análisis más objetivo del desempeño de los modelos, se ha procedido a agrupar tanto los datos observados como las predicciones en intervalos de tiempo diarios y semanales. Es importante destacar que esta agrupación no implica que las predicciones se hayan efectuado en conformidad con dicha discretización, sino que se realiza una agrupación para facilitar la evaluación y comparación de los resultados en un marco de tiempo diario y semanal.

Tabla 5–2 Comparaciones Modelos LSTM Discretización Horaria

Modelos	RMSE Horario	Rango Estadístico	Varianza	Desviación Estándar	Coeficiente de Variación
Modelo 1	8395.041989	76378.324	7.048558e+07	8395.569183	49.159819
Modelo 2	10244.694116	90700.250	1.048986e+08	10242.002967	30.643208
Modelo 3	6914.173787	77374.382	4.773365e+07	6908.953992	22.081091
Modelo 4	7920.026059	87659.742	6.273139e+07	7920.314825	46.301368
Modelo 5	6993.235589	69777.679	4.892829e+07	6994.875911	119.163790
Modelo 6	6316.440698	69226.734	3.890322e+07	6237.244909	6.190429
Modelo 7	7354.353941	85683.695	5.276875e+07	7264.210473	6.260784
Modelo 8	26314.288937	162542.949	6.926098e+08	26317.481487	58.048287
Modelo 9	26134.51185	149199.726	6.821993e+08	26118.945288	23.339076
Modelo 10	7984.697805	87659.742	6.373444e+07	7983.385144	33.937329
Modelo 11	9078.335818	99152.117	8.238182e+07	9076.443280	32.334876
Modelo 12	6142.535174	75586.978	3.325420e+07	5766.645540	2.72009

Tabla 5–3. Comparaciones Modelos LSTM Discretización Diaria

Modelos	RMSE Diario	Rango Estadístico	Varianza	Desviación Estándar	Coeficiente de Variación
Modelo 1	12388.62168	28328.9823	4.601310e+07	6783.2954	0.6525
Modelo 2	11592.7281	23610.3134	3.720348e+07	6099.4655	0.6171
Modelo 3	12462.8461	30413.5433	5.781193e+07	7603.4155	0.7670
Modelo 4	12905.37522	29984.4447	5.944058e+07	7709.7719	0.7422
Modelo 5	12733.0758	31013.6212	5.959199e+07	7719.5848	0.7594

Modelo 6	12239.8013	30902.6338	6.572821e+07	8107.2938	0.8796
Modelo 7	13715.17054	28986.1504	5.905758e+07	7684.8931	0.6744
Modelo 8	12538.4090	29726.1077	6.258303e+07	7910.9435	0.8097
Modelo 9	11253.9851	25386.1002	4.503763e+07	6711.0078	0.7397
Modelo 10	12918.5949	29984.4447	5.842714e+07	7643.7646	0.7313
Modelo 11	12717.7438	27873.9015	5.182457e+07	7198.9285	0.6845
Modelo 12	11257.1897	31508.7525	6.182449e+07	7862.8552	0.9700

Tabla 5–4. Comparaciones Modelos LSTM Discretización Semanal

Modelos	RMSE Semanal	Rango Estadístico	Varianza	Desviación Estándar	Coeficiente de Variación
Modelo 1	11297.6390	20544.4386	2.529148e+07	5029.064011	0.491620
Modelo 2	10492.8928	14135.6254	1.431630e+07	3783.688847	0.384005
Modelo 3	11145.3883	20633.9974	2.648387e+07	5146.248438	0.514256
Modelo 4	11308.4124	18660.5756	2.119761e+07	4604.086522	0.441783
Modelo 5	11089.3991	20039.5607	2.528091e+07	5028.012532	0.502821
Modelo 6	10470.2312	20016.8442	2.505678e+07	5005.674557	0.537138
Modelo 7	12139.6299	19034.8224	2.612299e+07	5111.065490	0.459688
Modelo 8	11032.8982	17034.6231	2.054748e+07	4532.933348	0.446545
Modelo 9	9670.98941	15735.9853	1.594440e+07	3993.043328	0.449158
Modelo 10	11308.4124	18660.5756	2.119761e+07	4604.086522	0.441783
Modelo 11	11156.7980	16877.5591	2.164379e+07	4652.289181	0.454454
Modelo 12	9340.3633	21528.2363	2.935034e+07	5417.595522	0.696166

Para facilitar la comprensión de las tablas comparativas, se presentan las figuras 5-5 a 5-9, que ofrecen una visión más detallada de diversos aspectos del análisis. La figura 5-5 muestra los valores del Root Mean Square Error (RMSE) para los modelos LSTM en tres configuraciones de discretización: horaria, diaria y semanal. Es evidente que el Modelo 12 exhibe el RMSE más bajo en las tres discretizaciones, lo que indica su superioridad en términos de precisión en las predicciones en estos contextos temporales.

En la figura 5-6, se representa el rango estadístico de los modelos, que refleja la amplitud de los errores en las predicciones, un rango de error más amplio sugiere una mayor variabilidad en la calidad de las predicciones, incluyendo la posible presencia de valores atípicos o errores significativos en ciertos casos. Notablemente, el Modelo 6 supera al Modelo 12 en discretización horaria, mientras que el Modelo 2 supera al Modelo 12 en discretización diaria y semanal. Esto subraya que, aunque sea inferior en el rango estadistico posee un RMSE más bajo en el Modelo 12 lo que denota una mejor precisión promedio

La figura 5-7 muestra la variabilidad a través de la varianza, destacando que el Modelo 12 exhibe un rendimiento superior en el contexto de la discretización horaria. Sin embargo, es importante observar que el Modelo 2 sobresale en la discretización diaria y semanal. La menor varianza en la discretización horaria dentro del Modelo 12 indica una mayor consistencia en las predicciones a nivel de hora en este enfoque particular. No obstante, es fundamental reconocer que esta dinámica puede cambiar al modificar la escala de tiempo a diaria o semanal, debido a las diferencias en la distribución de datos y las características inherentes a cada modelo en dichos casos.

La elección de la discretización adecuada dependerá de los objetivos específicos de su aplicación y de la escala temporal relevante.

En la Figura 5-8, se presenta el coeficiente de variación (CV), donde se destaca que el Modelo 12 exhibe un CV notablemente menor en el contexto de la discretización horaria. Este hallazgo sugiere que, a nivel horario, las predicciones de dicho modelo mantienen una mayor consistencia en relación a su valor promedio en comparación con los otros modelos. En otras palabras, las predicciones del Modelo 12 tienden a mantenerse más cercanas a su valor medio en la escala horaria, en contraste con los otros modelos. Sin embargo, es importante resaltar que en la discretización diaria y semanal, el Modelo 2 recupera su superioridad en comparación con el Modelo 12. Esto se debe a una reducción significativa en el CV al modificar la escala de tiempo, efecto que está vinculado a la influencia de dicha escala en la variabilidad de los datos. La transición a intervalos temporales más extensos, como los diarios y semanales, suele ir acompañada de una disminución en la variabilidad relativa de los datos, lo cual se refleja en la reducción del CV en todos los modelos. No obstante, es crucial comprender que esta reducción no necesariamente implica una mejora en la precisión de los modelos, sino que se traduce en una suavización de la variabilidad intradiaria en la discretización diaria y semanal.

Finalmente, en la figura 5-9 se explora la desviación estándar, donde el Modelo 12 presenta una menor dispersión en la discretización horaria, indicando que sus predicciones a nivel horario tienden a estar más cerca de su valor promedio en comparación con otros modelos en la misma discretización. No obstante, al cambiar la discretización a diaria y semanal, la variabilidad aumenta, lo que lleva al predominio del Modelo 2 en estas configuraciones temporales más amplias.

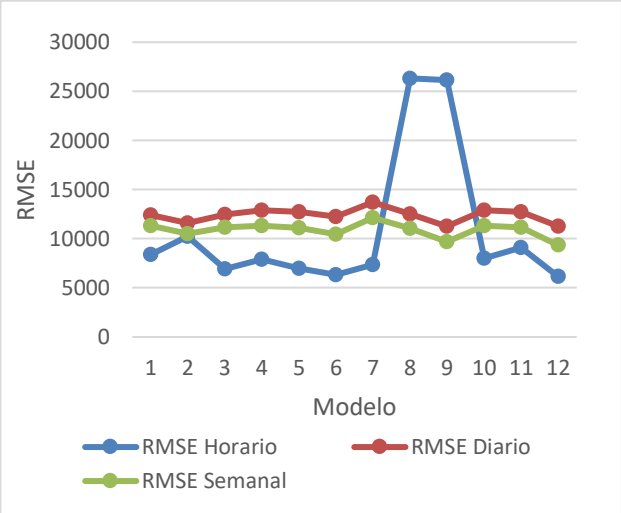


Figura 5-5. RMSE

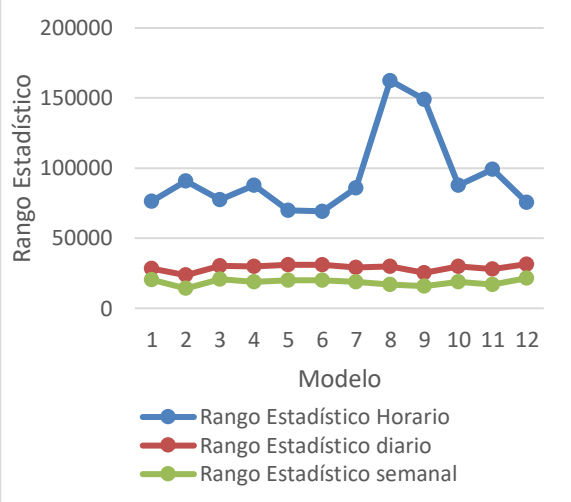


Figura 5-6. Rango Estadístico

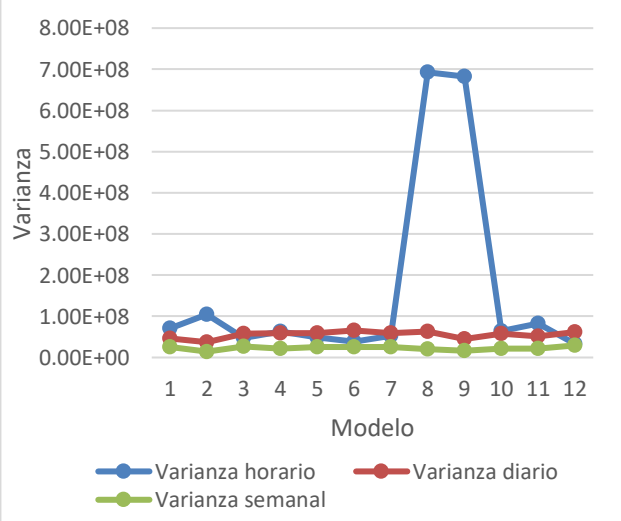


Figura 5-7. Varianza

Figura 5-8.Coeficiente de Variación

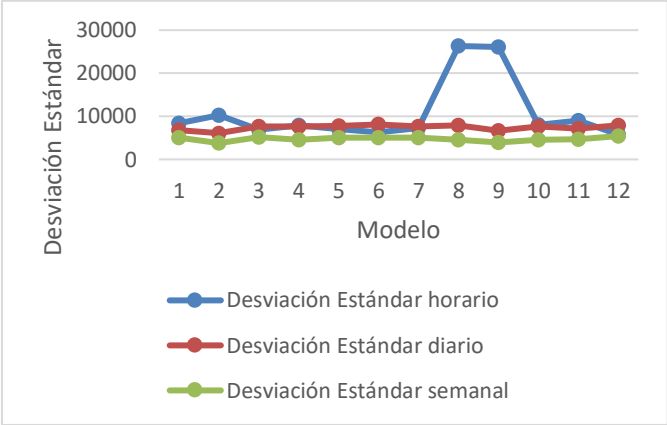


Figura 5-9. Desviación Estándar

La elección del mejor modelo se basa en el análisis de las tablas y gráficos presentados previamente. Tras la evaluación, se concluye que el Modelo 12 exhibe un desempeño superior en la discretización horaria. Dado que se trabaja con una serie de tiempo en dicha discretización, se selecciona el Modelo 12 como la opción preferente.

Por otro lado, el Modelo 2 muestra resultados prometedores en situaciones de discretización diaria y semanal en varios casos, sugiriendo que podría ser una elección adecuada para este tipo de escenarios. En resumen, la elección del modelo adecuado depende de la escala de tiempo en la que se esté trabajando y de las especificidades de los datos y las necesidades del análisis.

Por último, se muestra la Tabla 5-5 que contiene información relevante sobre los valores máximos y mínimos de error en discretizacion horaria, junto con las fechas correspondientes. Esto puede ayudar a entender la naturaleza del error, ya que es posible que en esas fechas se trate de días festivos o haya actividades inusuales que influyan en las predicciones

Tabla 5–5. Valores Máximos y Mínimos de Error en Discretización Horaria

Modelos	Fecha y Hora Max	Valor Máximo Horario	Fecha y Hora Min	Valor Minimo Horario
Modelo 1	**2022-12-02 07:00:00**	38393.23828	**2022-12-15 04:00:00**	2.4609375
Modelo 2	**2022-10-17 17:00:00**	49248.4687	**2022-11-01 19:00:00**	0.078125
Modelo 3	**2022-10-29 10:00:00**	42732.5820	**2022-11-12 19:00:00**	1.52734375
Modelo 4	**2022-10-29 10:00:00**	49973.8476	**2022-12-31 19:00:00**	0.0234375
Modelo 5	**2022-10-18 17:00:00**	36413.6406	**2022-10-23 11:00:00**	1.8781
Modelo 6	**2022-10-29 10:00:00**	36184.6367	**2022-12-04 00:00:00**	1.1210

Modelo 7	**2022-10-29 10:00:00**	43044.6093	**2022-10-26 08:00:00**	0.3672
Modelo 8	**2022-10-18 17:00:00**	91340.5117	**2022-12-03 05:00:00**	2.66015
Modelo 9	**2022-10-18 17:00:00**	93796.9765	**2022-12-05 06:00:00**	22.7656
Modelo 10	**2022-10-29 10:00:00**	49973.8476	**2022-11-27 02:00:00**	1.8242
Modelo 11	**2022-10-29 10:00:00**	49973.8476	**2022-11-27 02:00:00**	1.8242
Modelo 12	**2022-10-18 17:00:00**	49009.0800	**2022-11-20 18:00:00**	0.23437

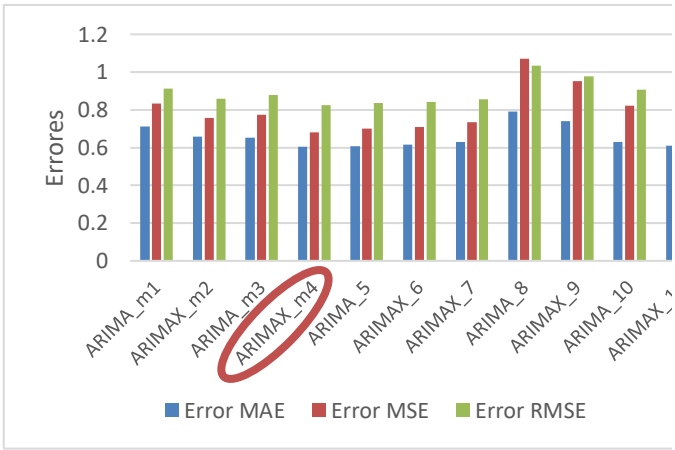
En la fecha "2022-10-29", se destaca la presencia de errores significativos en varios modelos, a pesar de que no se trata de un día festivo ni se relaciona con factores climáticos excepcionales. Esto sugiere la posibilidad de que los datos históricos utilizados por estos modelos muestren un comportamiento inusual o inesperado en comparación con datos anteriores. En consecuencia, los modelos pueden haber asumido que las predicciones para ese día serían similares a los días anteriores, lo que resulta en un error significativo en las predicciones.

5.1.2 ARIMA/ARIMAX

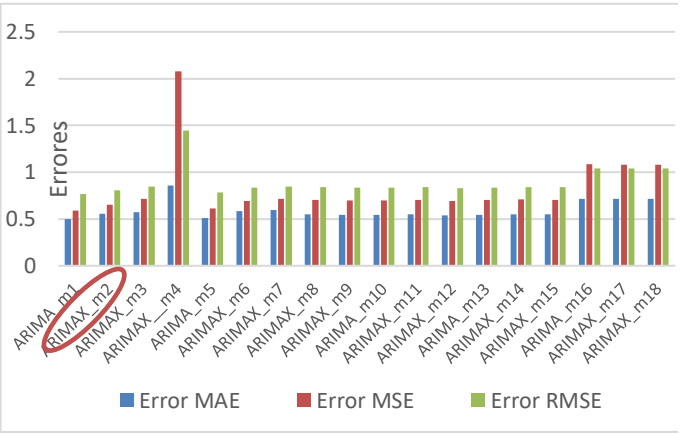
Luego de la exposición en el tercer capítulo, se procede a exhibir los resultados derivados de los diferentes modelos ARIMA y ARIMAX. Estos hallazgos son presentados de manera tabular en la denominada Tabla 5-6, seguida por una detallada explicación y análisis de los resultados. El propósito es ofrecer una comprensión más profunda de su relevancia y repercusiones dentro del marco de este estudio.

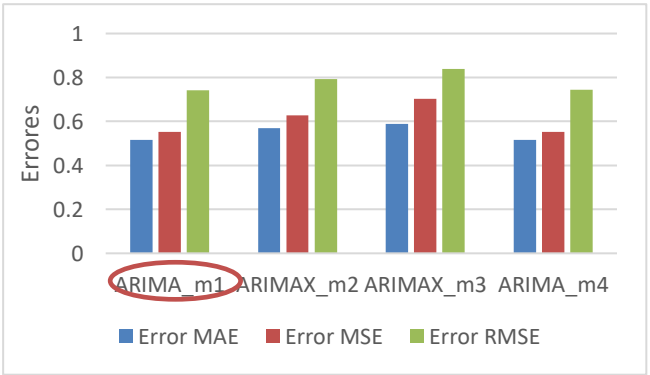
Tabla 5–6. Resultados de modelo ARIMA, ARIMAX

Modelos	Coeficiente p, d, q	Variable Exógena	Error MAE	Error MSE	Error RMSE	Grafica de barras Errores
ARIMA_m1	(12,0,0)	Sin exógena	0.71065	0.83384	0.91314	
ARIMAX_m2	(12,0,0)	Temperatura	0.65742	0.75583	0.85780	
ARIMA_m3	(12,0,1)	Sin exógena	0.65221	0.77428	0.87993	
ARIMAX_m4	(12,0,1)	Temperatura	0.60504	0.68071	0.82505	
ARIMA_5	(12,0,2)	Sin exógena	0.60744	0.70054	0.8369	
ARIMAX_6	(12,0,2)	Temperatura	0.61635	0.70870	0.8418	
ARIMAX_7	(12,0,2)	Temperatura y Humedad	0.62883	0.73450	0.8570	
ARIMA_8	(12,1,0)	Sin exógena	0.78991	1.06937	1.03410	

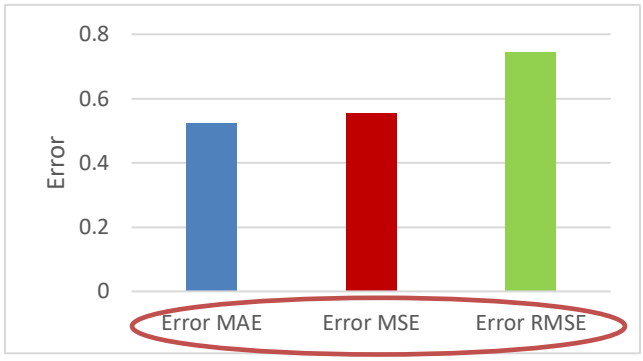


ARIMAX_9	(12,1,0)	Temperatura	0.73898	0.95299	0.97621
ARIMA_10	(12,1,1)	Sin exógena	0.63000	0.82072	0.90593
ARIMAX_11	(12,1,1)	Temperatura	0.61053	0.77178	0.87851
ARIMA_m1	(24,0,0)	Sin exógena	0.49910	0.58929	0.76765
ARIMAX_m2	(24,0,0)	Temperatura	0.55824	0.65234	0.80767
ARIMAX_m3	(24,0,0)	Temperatura y Humedad	0.57304	0.71287	0.8443
ARIMAX__m4	(24,0,1)	Temperatura	0.85902	2.08025	1.44231
ARIMA_m5	(24,0,2)	Sin exógena	0.50774	0.61177	0.78215
ARIMAX_m6	(24,0,2)	Temperatura	0.58135	0.69506	0.8337
ARIMAX_m7	(24,0,2)	Temperatura y Humedad	0.59324	0.71599	0.8461
ARIMAX_m8	(24,1,0)	Temperatura	0.55186	0.70338	0.8386
ARIMAX_m9	(24,1,0)	Temperatura y Humedad	0.54264	0.69520	0.8337
ARIMA_m10	(24,1,1)	Sin exógena	0.54202	0.69716	0.8349
ARIMAX_m11	(24,1,1)	Temperatura	0.55188	0.70389	0.8389
ARIMAX_m12	(24,1,1)	Temperatura y Humedad	0.53855	0.69070	0.8310
ARIMA_m13	(24,1,2)	Sin exógena	0.54205	0.70136	0.8374
ARIMAX_m14	(24,1,2)	Temperatura	0.55209	0.70694	0.8408
ARIMAX_m15	(24,1,2)	Temperatura y Humedad	0.54729	0.70261	0.8382
ARIMA_m16	(24,2,0)	Sin exógena	0.71307	1.08283	1.04059
ARIMAX_m17	(24,2,0)	Temperatura	0.71561	1.08271	1.0405
ARIMAX_m18	(24,2,0)	Temperatura y Humedad	0.71598	1.08208	1.04023
ARIMA_m1	(42,0,0)	Sin exógena	0.51713	0.55202	0.7429
ARIMAX_m2	(42,0,0)	Temperatura	0.56904	0.62807	0.7925
ARIMAX_m3	(42,0,0)	Temperatura y Humedad	0.58857	0.70260	0.8382
ARIMA_m4	(42,0,2)	Sin exógena	0.51640	0.55248	0.74329





ARIMAX_m1	(48_1_10)	Temperatura	0.56957	0.65129	0.8070
ARIMA_m1	(72,0,0)	Sin exógena	0.52461	0.55587	0.7455
ARIMAX_m2	(72,0,0)	Temperatura	0.59982	0.66914	0.8180



La Tabla 5-6 presenta los resultados de varios modelos que se desarrolla a lo largo del proceso. Esta tabla refleja el enfoque empírico seguido para el ajuste y prueba de distintos valores de los coeficientes p, d y q. Se observaron buenos resultados con los coeficientes (24, 0, 0). La elección de estos coeficientes se sustenta en las métricas estadísticas utilizadas para validar las predicciones en comparación con el conjunto de datos original, las cuales se detallan posteriormente. En la Tabla 5-7, se destacan algunos de los modelos más sobresalientes para facilitar su identificación.

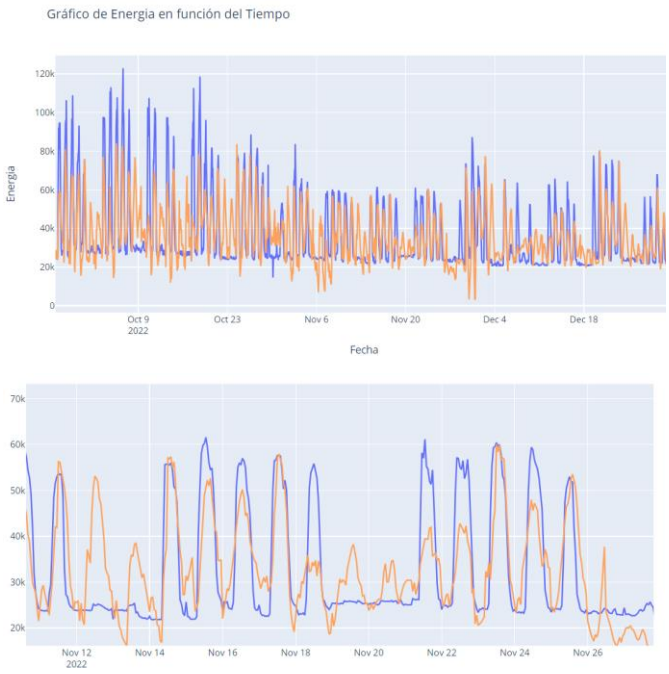
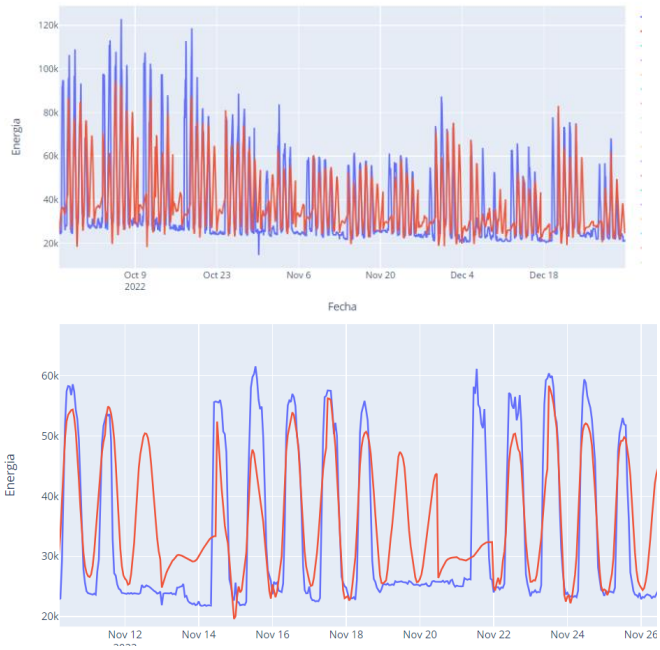
En la validación de predicciones, es esencial utilizar métricas adecuadas para evaluar la precisión del modelo. En este sentido, se emplearon las métricas más comúnmente utilizadas y reconocidas, como el Error Cuadrático Medio (MSE) que es una medida importante que cuantifica la diferencia entre los valores originales y los valores predichos, permitiendo evaluar el rendimiento del modelo en términos de precisión. Un MSE menor indica una mejor capacidad de predicción.

El Error Absoluto Medio (MAE), que mide la magnitud promedio de los errores de predicción. El MAE proporciona una medida de la precisión promedio de las predicciones sin considerar su dirección, lo que lo hace valioso en la evaluación del modelo. Otra métrica relevante empleada es la Raíz del Error Cuadrático Medio (RMSE) la cual ya es explicada previamente en las redes LSTM.

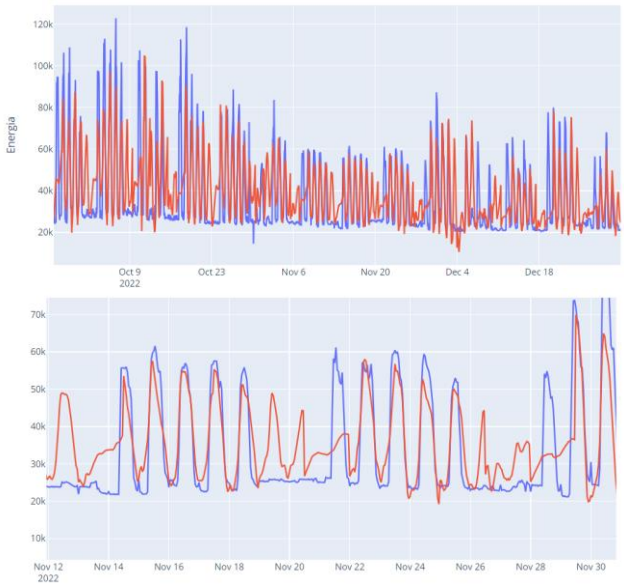
Por último, se exhiben en la Tabla 5-6 representaciones gráficas en forma de barras de las distintas métricas bajo consideración. Esta presentación gráfica posibilita una identificación expedita de la métrica óptima, se encierra en la gráfica y se resalta el valor asociado a la métrica más apropiada en el conjunto de datos analizados.

En la Tabla 5-7, presenta una síntesis de los modelos sobresalientes que son elegidos, acompañados de sus respectivas representaciones gráficas de las predicciones correspondientes. Este análisis suplementario robustece el proceso de toma de decisiones en lo que respecta a la selección del modelo ARIMA/ARIMAX más idóneo, fundamentándose en un enfoque metodológico y científico sólido.

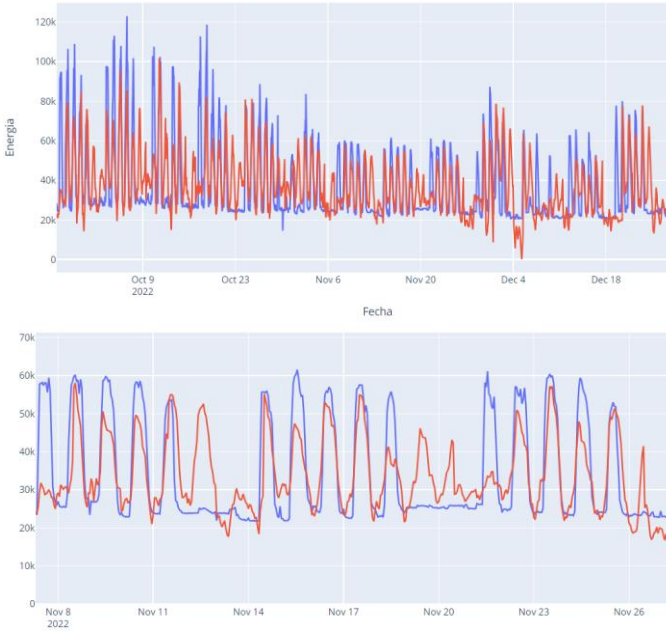
Tabla 5–7. Resumen de los mejores modelos ARIMA, ARIMAX

Modelos	Coeficiente p, d, q	Variable Exógena	Error MAE	Error MSE	Error RMSE	Grafica de predicción
ARIMAX	(12,0,1)	Temperatura	0.60504	0.68071	0.82505	
ARIMA	(24,0,0)	Sin exógena	0.49910	0.58929	0.76765	

ARIMA	(42,0,2)	Sin exógena	0.51640	0.55248	0.74329



ARIMAX	(48,1,10)	Temperatura	0.56957	0.65129	0.8070



El modelo ARIMA (24, 0, 0) demuestra ser el más eficiente en términos de rendimiento y uso de recursos computacionales. A pesar de que existen modelos con coeficientes superiores al elegido que demandan más recursos computacionales, el ARIMA (24, 0, 0) sigue siendo la elección preferida frente a otros modelos. A pesar de presentar inferioridad en la métrica de RMSE frente al ARIMA (42, 0, 0), la diferencia no es lo suficientemente significativa como para justificar el aumento en el uso de recursos computacionales.

Como se menciona en la discusión de los resultados de los modelos "RNN LSTM", el estudio se enfoca exclusivamente en un centro de transformación (Centro de Transformación A), del cual se muestran la totalidad de los resultados. Esto se debe a la aplicación de una metodología uniforme para ambos centros de transformación A y B, aunque los datos de cada centro de transformación pueden comportarse de manera ligeramente diferente, lo que puede justificar una posible variación en los coeficientes. Los resultados del mejor modelo aplicado a ambos centros de transformación se encuentran en la Figura 5-10. Esta representación tiene como objetivo ilustrar cómo se comporta el modelo en ambos conjuntos de datos y resaltar cualquier diferencia significativa en su rendimiento.

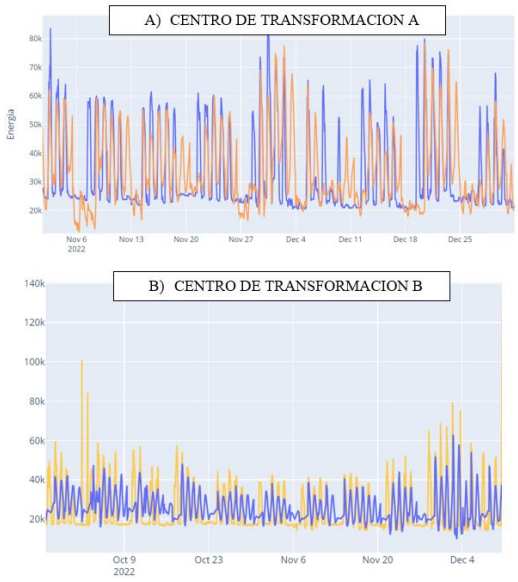


Figura 5-10. Predicción de modelo ARIMA (24,0,0) en ambos centros de transformación.

Con el objetivo de realizar una evaluación objetiva del rendimiento de los modelos, se lleva a cabo una comparación entre los Modelos ARIMA y ARIMAX que se encuentra en la Tabla5-7. Esta comparación se enfoca en analizar y contrastar el rendimiento del modelo seleccionado con sus contrapartes, además de identificar datos significativos que puedan surgir de este análisis comparativo.

Este enfoque de comparación se alinea con el utilizado en la evaluación de los modelos RNN LSTM. Las Tablas 5-8, 5-9 y 5-10 presentan las comparativas entre los modelos en términos de diversas métricas estadísticas, así como en relación con la discretización temporal.

Tabla 5–8. Comparacion Modelos ARIMA, ARIMAX. Discretización Horaria

Modelos	RMSE Horario	Rango Estadístico Horario	Varianza Horario	Desviación Estándar Horario	Coeficiente de Variación Horario
ARIMAX (12,0,1)	13173.4179	110393.309739	1.736239e+08	13176.6440	141.757193
ARIMA (24,0,0)	12114.718809	122248.730886	1.428740e+08	11952.9927	5.999459
ARIMA (42,0,2)	12175.964241	115882.688053	1.427328e+08	11947.0836	5.049338
ARIMAX (48,0,10)	13182.190461	119503.492725	1.734872e+08	11947.0836	21.466616

Tabla 5–9. Comparacion Modelos ARIMA, ARIMAX. Discretización Diaria

Modelos	RMSE Diario	Rango Estadístico Diario	Varianza Diario	Desviación Estándar Diario	Coeficiente de Variación Diario
ARIMAX (12,0,1)	12079.05159	32038.968755	4.050846e+07	6364.625966	0.618417
ARIMA (24,0,0)	13230.7932	32038.968755	2.687587e+07	5184.194234	0.425381
ARIMA (42,0,2)	13699.5711	27033.777255	3.016143e+07	5491.942233	0.437042
ARIMAX (48,0,10)	13267.7605	40443.272943	5.973125e+07	7728.599666	0.714271

Tabla 5–10. Comparación Modelos ARIMA, ARIMAX Discretización Semanal

Modelos	RMSE Semanal	Rango Estadístico Semanal	Varianza Semanal	Desviación Estándar Semanal	Coeficiente de Variación Semanal
ARIMAX (12,0,1)	11533.8259	11638.764686	9.613569e+06	3100.575524	0.278115
ARIMA (24,0,0)	13398.5874	9834.357534	1.087658e+07	3297.966184	0.253215
ARIMA (42,0,2)	13786.7428	9217.366604	1.200903e+07	3465.404643	0.258903
ARIMAX (48,0,10)	12281.7858	10910.149478	1.728155e+07	4157.108798	0.357612

Con el fin de simplificar la comprensión de los datos contenidos en las tablas previas, se generan las Figuras 5-11 a 5-15. Estas figuras facilitan la interpretación de la información presentada en las tablas, permitiéndonos identificar de manera más efectiva el modelo óptimo en las diversas métricas estadísticas y configuraciones de discretización temporal.

En la Figura 5-11, se presenta el valor del RMSE de los modelos seleccionados para su comparación. Se destaca que, en la discretización horaria, el Modelo ARIMA (24, 0, 0) exhibe un rendimiento superior en comparación con los demás modelos. Por otro lado, en las discretizaciones diaria y semanal, el Modelo ARIMAX (12, 0, 1) muestra una ventaja sobre los demás.

La Figura 5-12 ofrece una visión del rango estadístico, que refleja la amplitud de los errores en las predicciones. En la discretización horaria, el Modelo ARIMAX (12, 0, 1) se posiciona como el líder en este aspecto, mientras que, en las discretizaciones diaria y semanal, el Modelo ARIMA (42, 0, 2) demuestra su superioridad. A pesar de que el Modelo ARIMA (24, 0, 0) se ubica por debajo de ellos en diversas discretizaciones temporales, presenta una precisión promedio notablemente mejor.

En la Figura 5-13, gracias a la representación de la varianza, se observa que el Modelo ARIMA (24, 0, 0) sobresale en la discretización horaria y diaria, lo que indica que los datos no se dispersan ampliamente alrededor de la media. En la discretización semanal, el Modelo ARIMAX (12, 0, 1) lidera en este aspecto

La Figura 5-14 destaca el coeficiente de variación, donde el Modelo ARIMA (42, 0, 2) arroja los mejores resultados en la discretización horaria, lo que sugiere que las predicciones mantienen una mayor consistencia en relación a su valor promedio.

Por último, en la Figura 5-15 se presenta la desviación estándar. En la discretización diaria, el Modelo ARIMA (24, 0, 0) muestra una menor dispersión, mientras que en la discretización horaria, los modelos ARIMA (24, 0, 0), ARIMA (42, 0, 0) y ARIMAX (48, 0, 10) exhiben valores de desviación muy similares entre ellos. Por otro lado, el Modelo ARIMAX (12, 0, 1) se destaca en la discretización semanal.

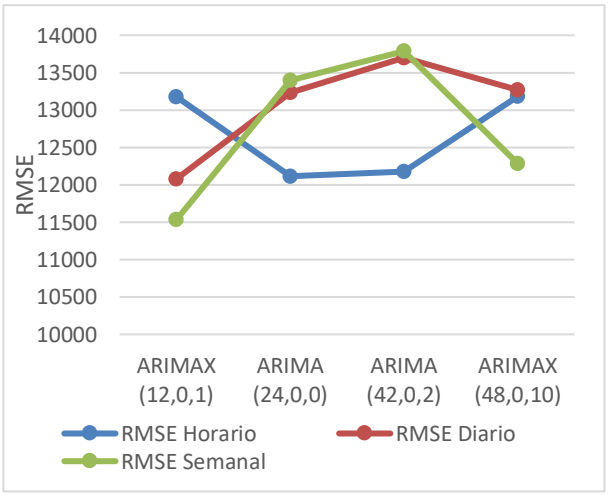


Figura 5-11. RMSE

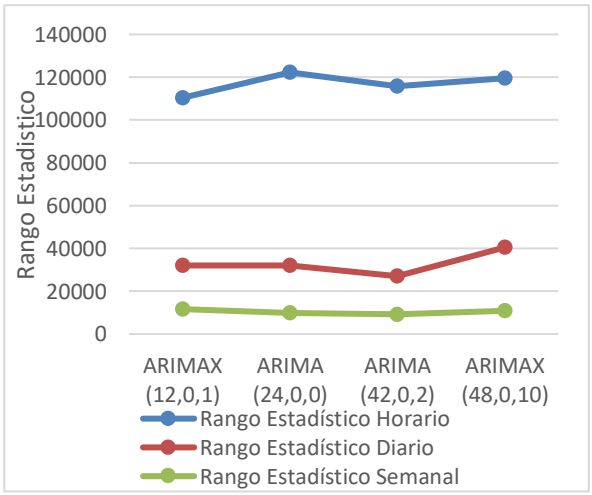


Figura 5-12. Rango Estadístico

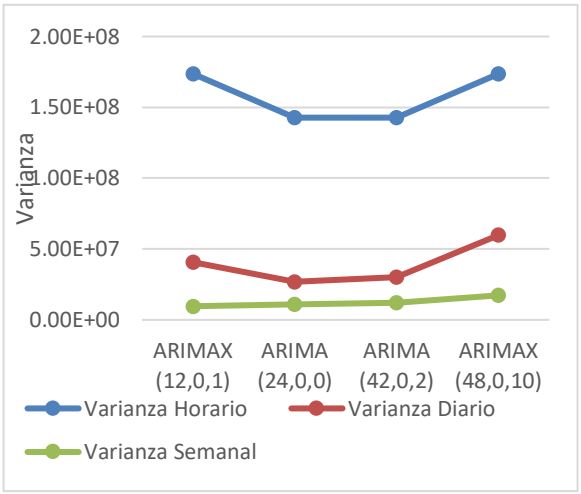


Figura 5-13.Varianza

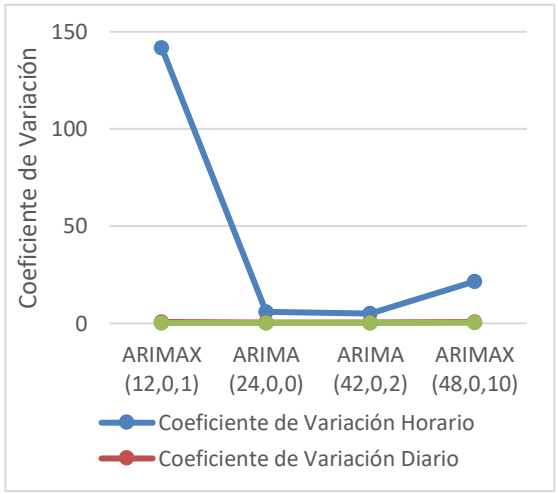


Figura 5-14. Coeficiente de Variación

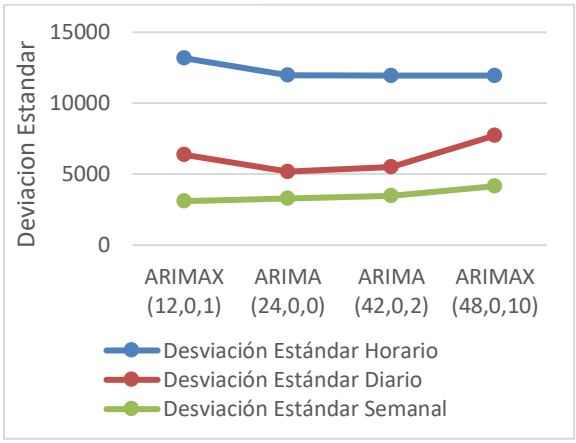


Figura 5-15. Desviación Estándar

Tras un análisis detallado que considera las diversas métricas presentadas anteriormente y criterios específicos, la elección del Modelo ARIMA (24, 0, 0) se justifica por su desempeño sobresaliente. En la discretización horaria, este modelo presenta un RMSE inferior, una varianza más baja y una desviación estándar que se asemeja estrechamente a varios otros modelos. Este conjunto de métricas destaca la prioridad dada a la precisión y consistencia en las predicciones, lo que respalda la elección de este modelo.

En la Tabla 5-11, se presentan datos relevantes que muestran los valores máximos y mínimos de error en la discretización horaria, siguiendo un enfoque similar al presentado para los modelos LSTM. Se destaca un patrón interesante: la fecha "2022-10-17 17:00:00" se repite en todos los modelos seleccionados, y en cada caso, registra el error máximo. Al igual que en las redes LSTM, no se ha identificado ninguna razón relacionada con días festivos o factores climáticos que explique este comportamiento. Por lo tanto, se mantiene el mismo criterio de que factores inusuales en los datos históricos pueden ser responsables de estos errores notables en las predicciones.

Tabla 5–11. Valores Máximos y Mínimos de Error en Discretización Horaria

Modelos	Fecha y Hora Max	Valor Máximo Horario	Fecha y Hora Min	Valor Minimo Horario
ARIMAX (12,0,1)	2022-10-17 17:00:00	69475.7808	2022-11-20 02:00:00	6.0727

ARIMA (24,0,0)	2022-10-17 17:00:00	79543.8647	2022-11-17 20:00:00	1.5931
ARIMA (42,0,2)	2022-10-17 17:00:00	73713.4715	2022-12-22 02:00:00	0.1269
ARIMAX (48,0,10)	2022-10-17 17:00:00	75465.0025	2022-12-23 04:00:00	1.7676

Tras la realización del análisis detallado expuesto previamente, se identifican los modelos óptimos en las categorías de ARIMA/ARIMAX y LSTM. Se lleva a cabo una comparación directa entre el modelo más destacado en la categoría LSTM, conocido como el "Modelo 12", y el mejor modelo perteneciente a la categoría ARIMA/ARIMAX, que corresponde al ARIMA (24, 0, 0). Los resultados de esta comparación se presentan en la Tabla 5-12. El propósito fundamental de este ejercicio es determinar cuál de estos dos enfoques sobresale como el modelo superior en términos de rendimiento. En definitiva, se busca identificar el modelo que sobresale como la elección superior entre los modelos ARIMA/ARIMAX y LSTM en este contexto específico.

Cabe mencionar que esta comparativa se realiza en el contexto de la discretización horaria ya que la serie temporal con la que se trabaja se encuentra originalmente en discretización horaria, y solo se aplicaron discretizaciones diarias y semanales con fines de análisis y comparación de modelos. Por lo tanto, el enfoque principal de esta comparación se centra en la discretización horaria.

Tabla 5–12. Comparación Modelo12 (LSTM) y ARIMA (24,0,0)

Modelos	RMSE Horario	Rango Estadístico Horario	Varianza Horario	Desviación Estándar Horario	Coeficiente de Variación Horario
Modelo 12	11257.1897	31508.7525	6.182449e+07	7862.8552	0.9700
ARIMA (24,0,0)	12114.718809	122248.730886	1.428740e+08	11952.9927	5.999459

El Modelo 12 destaca claramente frente al ARIMA (24, 0, 0) debido a su rendimiento superior en todas las métricas evaluadas. Esto indica una capacidad de predicción más sólida y una menor variabilidad en las predicciones, lo que respalda de manera concluyente la elección de este modelo como la mejor opción para abordar el problema de predicción de energía.

5.2 Ensamblador

En esta sección se presentan los resultados de dos ensambladores, denominados: ensamblador 1 y 2. El ensamblador 1 consiste en un conjunto de redes LSTM, y el ensamblador 2 combina el mejor modelo de LSTM con el mejor modelo entre ARIMA y ARIMAX.

Ensamblador 1: Compuesto por los modelos LSTM presentados en la Tabla 4-5, se lleva a cabo una minuciosa comparación de las métricas estadísticas detalladas en la Tabla 5-13. El propósito es identificar el ensamblador que exhiba un rendimiento sobresaliente y, en última instancia, seleccionar al mejor ensamblador de entre los evaluados. Además, en la Figura 5-16 se presenta las predicciones de este ensamblador y los datos originales para una visualización más clara de su desempeño.

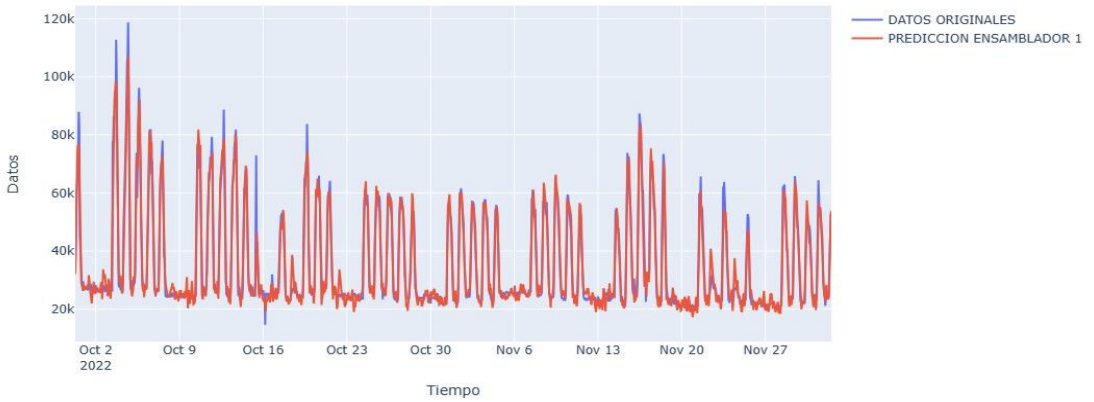


Figura 5-16. Prediccion Ensamblador 1

Ensamblador 2: Después de haber identificado previamente los modelos de mayor rendimiento tanto en la categoría de modelos LSTM como en la de modelos ARIMA y ARIMAX, el ensamblador 2 se compone de dos componentes clave: el "Modelo 12" de LSTM y el modelo "ARIMA (24, 0, 0)". Al igual que en el caso del ensamblador 1, los resultados de este ensamblador se encuentran detallados en la Tabla 5-13, y la capacidad predictiva de este ensamblador se representa en la Figura 5-17. Esto nos permitirá evaluar su eficacia y desempeño en el contexto de nuestras predicciones.

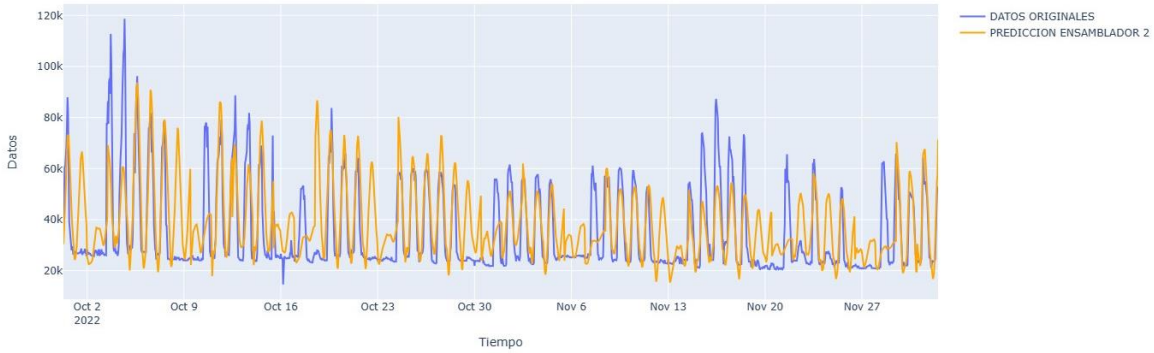


Figura 5-17. Prediccion Ensamblador 2

Se observa que el Ensamblador 1 muestra un rendimiento visualmente superior en comparación al Ensamblador 2. En un esfuerzo por mejorar el desempeño del Ensamblador 2, se ha tomado la decisión de sustituir el modelo de regresión lineal, que ha demostrado ser efectivo en el Ensamblador 1, por el RandomForestRegressor. Este nuevo modelo, denominado Ensamblador 3, se basa en árboles de decisión y es capaz de capturar relaciones complejas en los datos. El Ensamblador 3 mejora las predicciones visulamente en comparación al Ensamblador 2, las prediccion del Ensamblador 3 se presentan en la Figura 5-18, y las métricas estadísticas al igual que los anteriores ensambladores se encuentran en la Tabla 5-13

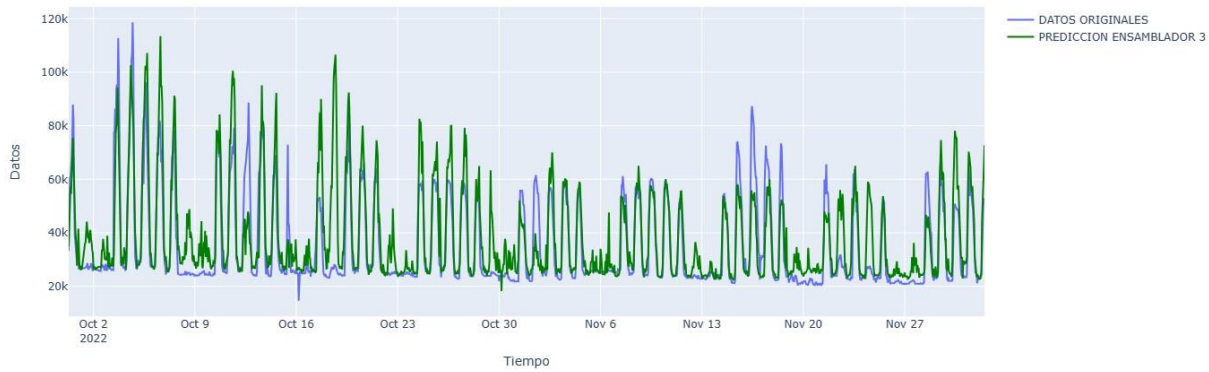


Figura 5-18. Prediccion Ensamblador 3

En la Tabla 5-13 se exhiben las métricas estadísticas que sirven como criterio de comparación entre los ensambladores, con el propósito de seleccionar el mejor entre ellos. Estas métricas proporcionan una evaluación objetiva y cuantitativa del rendimiento de los ensambladores

Tabla 5–13. Metricas Estadísticas Ensambladores

Modelos	RMSE Horario	Rango Estadístico Horario	Varianza Horario	Desviación Estándar Horario	Coeficiente de Variación Horario
Ensamblador 1	3435.199761	51717.6406	1.180836e+07	3436.329152	804.473381
Ensamblador 2	13380.652159	121084.4286	1.686180e+08	12985.298577	4.000697
Ensamblador 3	10869.352453	124074.200	1.075257e+08	10369.458474	3.171817

Los resultados evidencian que el Ensamblador 1 sobresale en la mayoría de las métricas de evaluación en comparación con los Ensambladores 2 y 3, consolidándose como el modelo líder. Sin embargo, en cuanto al coeficiente de variación, el Ensamblador 3 demuestra un rendimiento superior con respecto a los Ensambladores 1 y 2.

Con el fin de obtener una evaluación más completa del desempeño de estos ensambladores, se presenta la Tabla 5-14, que exhibe los valores máximos y mínimos de los errores asociados a cada modelo. Se destaca que el error máximo se asocia al Ensamblador 2, mientras que el mínimo de error se atribuye al Ensamblador 1. Sin embargo, al comparar el Ensamblador 1 con el Ensamblador 3, el Ensamblador 1 exhibe el valor máximo de error en relación con el Ensamblador 3.

Es importante comprender que los valores extremos máximos y mínimos no son suficientes para determinar la superioridad de un modelo, ya que pueden deberse a situaciones excepcionales o valores atípicos. Por lo tanto, para una evaluación más sólida, es necesario considerar una variedad de métricas estadísticas previamente mencionadas.

Tabla 5–14. Valores de Error Minimos y Máximos de los Ensambladores

Modelos	Fecha y Hora Max	Valor Máximo Horario	Fecha y Hora Min	Valor Minimo Horario
Ensamblador 1	2022-10-15 10:00:00	35218.5546	2022-10-03 09:00:00	0.06054
Ensamblador 2	2022-10-04 17:00:00	62311.0092	2022-11-18 05:00:00	16.62440
Ensamblador 3	2022-10-18 17:00:00	78873.3	2022-11-16 05:00:00	2.89500

6 CONCLUSIONES Y FUTURAS LINEAS DE INVESTIGACIÓN

Hay alguien tan inteligente que aprende de la experiencia de los demás.

- Voltaire-

En el transcurso de este trabajo, se desvela descubrimientos fundamentales que aportan una perspectiva valiosa al ámbito de la predicción de la demanda en instalaciones de transformación energética. La metodología empleada ha propiciado una comprensión profunda de los procesos inherentes, respaldada por datos exhaustivos y técnicas amparadas por la literatura científica. La interpretación de los resultados se ha enriquecido a través de vinculaciones con el marco teórico proporcionado por la revisión bibliográfica, lo cual ha contribuido de manera significativa a la apreciación de las implicaciones derivadas de la investigación.

Es imperativo destacar que las conclusiones presentadas no solo consolidan el corpus de conocimiento vigente en la disciplina, sino que también proporcionan directrices y perspectivas que orientan futuras investigaciones y aplicaciones prácticas. Su estudio ha identificado áreas de investigación promisorias y ha subrayado la imperante relevancia de la precisión en la predicción de la demanda en centros de transformación, lo cual posee el potencial de incidir de manera substancial en la eficiencia operativa y la toma de decisiones en dicho contexto. En última instancia, estas conclusiones se erigen como un recurso valioso para aquellos involucrados en el análisis y pronóstico de la demanda en entornos de transformación, y sientan los cimientos para futuros avances en este campo en constante desarrollo.

6.1 Conclusiones

En el análisis de los resultados derivados de los modelos ARIMA y ARIMAX, se discierne descubrimientos de singular interés en el contexto del caso de estudio y la naturaleza de los datos bajo consideración. Estos resultados indican que el modelo óptimo es, en efecto, el que excluye variables exógenas, es decir, un modelo ARIMA. Se ha observado que la especificación de los coeficientes p, d, q de dicho modelo no siempre resulta factible debido a múltiples restricciones. A pesar de la presencia de estacionalidad en los datos, la periodicidad es considerablemente amplia, lo que conlleva a una demanda sustancial de recursos computacionales, instando a ajustes iterativos de estos parámetros.

Además, se ha demostrado que la diferenciación aplicada a los modelos evaluados no ha conllevado mejoras sustanciales en términos de la precisión de las predicciones, lo cual sugiere que el modelo más apropiado es aquel que posee un valor de d igual a cero. A lo largo de diversas pruebas, se ha corroborado que la diferenciación ejerce un impacto notable en el costo computacional, lo que refuerza la elección del modelo ARIMA (24,0,0) como el más adecuado en estas circunstancias.

Por otro lado, en el análisis de los resultados del modelo de redes neuronales LSTM, se ha observado un mejor rendimiento cuando se incorpora variables exógenas como la temperatura y la humedad. Esto indica que, para este tipo de modelos, las variables exógenas enriquecen la capacidad predictiva y arrojan resultados altamente favorables en comparación con el modelo ARIMA. El modelo más sobresaliente resulta ser el Modelo 12, que combina múltiples capas LSTM con una capa convolucional, lo que contribuyó significativamente a la precisión de la predicción. La combinación de capas de diferentes tipos ha demostrado influir positivamente en la capacidad predictiva del modelo.

Se desarrollaron dos modelos de ensamblador. Dado que se ha identificado dos modelos sobresalientes de ARIMA y LSTM, el “ARIMA (24,0,0)” y el “Modelo 12” de las LSTM respectivamente, se procede a realizar un ensamblaje de los modelos con el objetivo de aprovechar las fortalezas inherentes a cada uno de ellos. En una primera instancia, el ensamblador, denominado "ensamblador 2," basado en LinearRegression, mostró

predicciones prometedoras. No obstante, al sustituir el algoritmo LinearRegression por RandomForestRegressor, se lograron mejoras significativas en el desempeño predictivo. El Ensamblador 3 demostró un rendimiento excepcional en combinación con el algoritmo RandomForestRegressor con respecto al Ensamblador 2.

Por otro lado, el Ensamblador 1 sobresale por su habilidad para combinar múltiples modelos previamente entrenados de la Red Neuronal Recurrente (RNN) LSTM, con el propósito de discernir las características más destacadas y elevar la calidad de las predicciones. Esta aproximación ha resultado altamente eficaz, superando significativamente en términos de capacidad predictiva al Ensamblador 2 y 3, sin embargo, el Ensamblador 1 tiene un mayor consumo de recursos computacionales que puede ser interpretado por un limitante.

6.2 Futuras lineas de investigación

La optimización computacional de modelos ARIMA en este estudio, ha identificado que el modelo ARIMA puede presentar desafíos computacionales significativos. Por lo tanto, surge una línea de investigación prometedora en la búsqueda de técnicas de optimización que mejoren la eficiencia computacional de los modelos ARIMA. Estas técnicas podrían incluir la exploración de algoritmos de optimización, estrategias de muestreo eficiente y paralelización de cálculos para acelerar el ajuste de modelos ARIMA.

El progreso en la elaboración de ensambladores avanzados va más allá de restringirse a conjuntos modelados de manera específica en los ensambladores tradicionales. Se abre el camino a la investigación y desarrollo de ensambladores más sofisticados que hacen uso de técnicas de selección dinámica de modelos o que ajustan automáticamente la composición de los ensambladores en función de la calidad intrínseca de los modelos base. Este enfoque puede potencialmente dar lugar a ensambladores más eficientes y adaptables, especialmente en el tratamiento de datos que poseen una naturaleza particular, como es el caso de los datos de energía.

La investigación de nuevas fuentes de datos contempla la viabilidad de incorporar fuentes de datos adicionales que podrían ser empleadas como variables exógenas. Estas fuentes adicionales abarcan datos de índole económica, variables climáticas, datos geoespaciales, así como otros factores pertinentes, con el propósito de optimizar la precisión en las predicciones de demanda y evaluar el proceso de incorporación de tales variables en los modelos. No obstante, esta expansión en la fuente de datos requiere de estudios adicionales de carácter multidisciplinario.

Las aplicaciones prácticas derivadas de esta categoría de investigaciones conllevan a una optimización en la administración de los recursos energéticos, incrementando su eficiencia. Los resultados alcanzados mediante tales aplicaciones tienen la capacidad de estimular la iniciación de nuevos estudios y estrategias, tales como la reducción del consumo energético, la imperante necesidad de transiciones desde modelos de generación energética vigentes hacia aquellos que ostentan un carácter más sostenible y, preferentemente, de naturaleza renovable. Estos desarrollos también poseen implicaciones significativas en una economía circular, así como impactos económicos tanto a nivel nacional como global.

REFERENCIAS

[1] I. Dyner and E. R. Larsen, "From planning to strategy in the electricity industry," *Energy Policy*, vol. 29, pp. 1145–1154, 2001.

[2] A. G. Kagiannas, D. T. Askounis, and J. Psarras, "Power generation planning: a survey from monopoly to competition," doi: 10.1016/j.ijepes.2003.11.003.

[3] A. M. Foley, B. P. Ó. Gallachóir, J. Hur, R. Baldick, and E. J. Mckeogh, "A strategic review of electricity systems models," doi: 10.1016/j.energy.2010.03.057.

[4] O. Trull Dominguez, "Predicción a corto plazo de la demanda horaria de energía eléctrica en España mediante modelos optimizados de Holt-Winters múltiple-estacionales Presentada por."

[5] "La Energía en España," Accessed: Oct. 11, 2023. [Online]. Available: www.miteco.gob.es.

[6] F. J. C. Diranzo, "INCIDENCIA DE LA CLIMATOLOGÍA EN EL CONSUMO DE GAS Y ELECTRICIDAD EN ESPAÑA Enric Valor i Micó** Hipòlit Torró i Enguix* Vicente Caselles Miralles**."

[7] S. Liu, *Zhi-Hua Zhou: Machine Learning*. 2021.

[8] D. P. Kroese, Z. I. Botev, T. Taimre, and R. Vaisman, "Data Science and Machine Learning: Mathematical and Statistical Methods," *Data Sci. Mach. Learn. Math. Stat. Methods*, pp. 1–515, Jan. 2019, doi: 10.1201/9780367816971/DATA-SCIENCE-MACHINE-LEARNING-DIRK-KROESE-ZDRAVKO-BOTEV-THOMAS-TAIMRE-RADISLAV-VAISMAN.

[9] "Data Clustering Algorithms - k-means clustering algorithm." https://sites.google.com/site/dataclusteringalgorithms/k-means-clustering-algorithm (accessed Oct. 11, 2023).

[10] M. Cabezón, "Implementación de redes neuronales recurrentes en Python . Miguel Cabezón Manchado Trabajo de fin de máster en Ingeniería Matemática," p. 43, 2018, [Online]. Available: https://eprints.ucm.es/49444/1/2018-MIGUEL CABEZON Memoria.pdf.

[11] X. B. Olabe, "REDES NEURONALES ARTIFICIALES Y SUS APLICACIONES."

[12] A. Moreno, "Autor: Álvaro Artola Moreno Tutor: José Antonio Pérez Carrasco," *Univ. Sevilla*, p. 80, 2019.

[13] "Qué son las redes neuronales y sus aplicaciones | OpenWebinars." https://openwebinars.net/blog/que-son-las-redes-neuronales-y-sus-aplicaciones/ (accessed Oct. 11, 2023).

[14] A. : Pau, A. Granell, and S. T. Porras, "Grado en Estadística."

[15] D. De Trabajo and C. Arana, "UNIVERSIDAD DEL CEMA Buenos Aires Argentina Serie," 2021, Accessed: Oct. 11, 2023. [Online]. Available: www.cema.edu.ar/publicaciones/doc_trabajo.html.

[16] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/NECO.1997.9.8.1735.

[17] U. Jorge, T. Lozano, D. Fabian, and H. Cofre, "Prediccí on para el mercado de acciones con Redes Neuronales LSTM," 2020.

[18] R. de Arce Ramón Mahía Dpto Economía Aplicada, "MODELOS ARIMA Mayo 2001."