



**PREDICCIÓN DEL PRECIO PONDERADO DE LA ENERGÍA EN EL MERCADO
MAYORISTA EN COLOMBIA MEDIANTE MODELOS DE DEEP LEARNING CON
UN HORIZONTE DE PREDICCIÓN DE 24 HORAS, UTILIZANDO DATOS
HISTÓRICOS DE XM RECOLECTADOS DURANTE 10 AÑOS**

Yeferson García Vásquez

Felipe Ruiz Zea

Monografía presentada para optar al título de Especialista en Analítica y Ciencia de Datos

Asesor

Gabriel Darío Uribe Guerra

Universidad de Antioquia

Facultad de Ingeniería

Especialización en Analítica y Ciencia de Datos

Medellín, Antioquia, Colombia

2024

Cita

(García Vásquez & Ruiz Zea, 2024)

Referencia

Garcia Vasquez, Y.F., & Ruiz Zea, F. (2024). *predicción del precio ponderado de la energía en el mercado mayorista en Colombia mediante modelos de deep learning con un horizonte de predicción de 24 horas, utilizando datos históricos de xm recolectados durante 10 años*]. Universidad de Antioquia, Medellín, Colombia.

**Estilo IEEE
(2020)**



Especialización en Análítica y Ciencia de Datos, Cohorte VII.

Centro de Investigación Ambientales y de Ingeniería (CIA).



Centro de Documentación Ingeniería (CENDOI)

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

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano: Julio Cesar Saldarriaga Molina

Jefe departamento: Diego José Luis Botia Valderrama

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

Tabla de contenido

Resumen	9
Abstract	10
1. Descripción del problema.....	11
1.1. Problema de negocio	11
1.2. Aproximación desde la analítica de datos.....	12
1.3. Origen de los datos	13
1.4. Métricas de desempeño.....	13
2. Objetivos.....	16
2.1. Objetivo general	16
2.2. Objetivos específicos	16
3. Datos.....	17
3.1. Datos originales.....	17
3.2. Datasets	19
3.3. Analítica descriptiva	20
3.3.1. Comportamiento de las variables.....	20
3.3.1. Distribuciones de las Variables Principales.....	21
3.3.2. Análisis de Autocorrelación y Autocorrelación Parcial del Precio de Bolsa.....	22
3.3.3. Importancia de Variables mediante Random Forest.....	23
3.3.4. Análisis de Correlación entre Variables	24
4. Proceso de Analítica.....	26
4.1. Pipeline principal.....	26
4.2. Preprocesamiento	29

4.3.	Modelos	29
4.3.1.	Descripción de las arquitecturas de los modelos implementados:	31
4.4.	Métricas	31
5.	Metodología	32
5.1.	Baseline	32
5.1.1.	Experimentos	33
5.2.	Validación	34
5.3.	Iteraciones y evolución	34
5.4.	Herramientas	35
6.	Resultados y discusión	36
6.1.	Métricas	37
6.1.1.	Experimento con todas las características:	37
6.1.2.	Experimento con la variable objetivo, precio de la bolsa nacional:.....	41
6.1.3.	Experimento con las Variables más Influyentes:.....	45
6.2.	Evaluación cualitativa.....	50
6.2.3	Desempeño General de los Modelos.....	50
6.2.3	Impacto de la Inclusión de Variables Adicionales.....	50
6.2.4	Limitaciones ante Volatilidad del Mercado.....	51
6.3.	Consideraciones de producción.....	51
6.3.1	Monitoreo del Desempeño del Modelo.....	51
6.3.2	Integración con Flujos de Datos en Tiempo Real	51
6.3.3	Infraestructura en la Nube y Escalabilidad.....	51
6.3.4	Automatización y Orquestación del Ciclo de Vida del Modelo	52
6.3.5	Gestión de Versiones y Reentrenamiento.....	52
6.3.6	Optimización de Recursos y Costos.....	52

6.3.7 Interoperabilidad y Accesibilidad	52
6.3.8 Manejo de la Volatilidad y Adaptabilidad.....	52
7. Conclusiones	54
8. Recomendaciones.....	56
References	57

Lista de tablas

Tabla 1. Fuentes de datasets.....	13
Tabla 2. Características de variables.....	18
Tabla 3. Precio de la energía por Franjas horarias.....	20
Tabla 4. Hiperparámetros fijos y variados.	30
Tabla 5. Descripción de capas de modelos implementados.	31
Tabla 6. Resultados para experimento 'Todos los features'.	37
Tabla 7. Resultados para experimento 'Solo precio'.	41
Tabla 8. Resultados para experimento 'Variables más influyentes'.	46

Lista de figuras

Figura 1. Comportamiento de las diferentes características	21
Figura 2. Distribución de las variables.....	22
Figura 3 Autocorrelación del precio de bolsa.....	23
Figura 4 Random Forest - Importancia de Variables.....	24
Figura 5 Matriz de correlación	25
Figura 6 Promedio precio Bolsa Nacional por 10 Años	26
Figura 7. Pipeline-Flujo de Trabajo.	28
Figura 8. Comportamiento de las métricas para las mejores iteraciones del experimento 'Todos los features'.	38
Figura 9 Comparación de la Estabilidad y Generalización durante el Entrenamiento para del experimento 'Todos los features'.	39
Figura 10. Desempeño de ventanas para experimento 'Todos los features'.	40
Figura 11. Comportamiento de las métricas para las mejores iteraciones del experimento 'Solo precio'.	42
Figura 12. Comparación de la Estabilidad y Generalización durante el Entrenamiento para del experimento 'Solo precio'.	43
Figura 13. Desempeño de ventanas para experimento 'Solo precio'.	45
Figura 14. Comportamiento de las métricas para las mejores iteraciones del experimento 'Variables más influyentes'.	47
Figura 15. Comparación de la Estabilidad y Generalización durante el Entrenamiento para del experimento 'Variables más influyentes'.	48
Figura 16. Desempeño de ventanas para experimento 'Variables más influyentes'	49

Siglas, acrónimos y abreviaturas

mm.	Milímetros
Kwh	Kilowatio Hora
COP	Pesos colombianos
UdeA	Universidad de Antioquia
MAE	Error Cuadrático Medio
RMSE	Raíz del Error Cuadrático Medio
ACF	Función de autocorrelación
STD	Desviación Estándar
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
DL	Deep Learning

Resumen

En Colombia, el precio de la energía ha mostrado un aumento irregular debido a diversos factores especulativos y operativos, lo que ha generado un alza significativa en los últimos años. Este comportamiento refleja la complejidad del mercado energético colombiano, influenciado por la interconexión de los sistemas de generación y la competencia en el mercado. Estas dinámicas hacen que, para los diferentes actores de la producción energética, sea complejo predecir la oferta con la que deben competir. Para enfrentar este desafío, se propone desarrollar un modelo predictivo basado en Deep Learning utilizando datos históricos del sistema energético de XM, considerando variables como el volumen útil de los embalses, la demanda energética y los precios históricos. Los objetivos del modelo incluyen anticipar las fluctuaciones en los precios de la energía y evaluar su desempeño con datos reales. La metodología empleará herramientas de ciencia de datos con para la limpieza, exploración y modelado de los datos, así como la implementación de modelos de aprendizaje profundo (DL), como LSTM, RNN y GRU, con el objetivo de predecir el precio de la energía eléctrica en un horizonte de 24 horas. Este enfoque busca mejorar la planificación y gestión del sistema energético colombiano, optimizando la toma de decisiones estratégicas y aumentando la resiliencia del sistema frente a las incertidumbres del mercado energético.

GitHub Repos:

- [Proyecto-Seminario](#)

Palabras clave: Energía, Hidroeléctrica, Series de tiempo, Deep Learning, Predicción, Python, XM.

Abstract

In Colombia, energy prices have experienced irregular increases due to various speculative and operational factors, leading to a significant rise in recent years. This behavior reflects the complexity of the Colombian energy market, influenced by the interconnection of generation systems and market competition. These dynamics make it challenging for energy producers to predict the supply they must compete against. To address this challenge, a predictive model based on Deep Learning is proposed, utilizing historical data from the XM energy system and considering variables such as reservoir storage levels, energy demand, and historical prices. The model aims to anticipate energy price fluctuations and evaluate its performance using real data. The methodology involves leveraging data science tools for data cleaning, exploration, and modeling, as well as implementing Deep Learning models, such as LSTM, RNN, and GRU, to predict energy prices within a 24-hour horizon. This approach seeks to enhance the planning and management of Colombia's energy system, optimizing strategic decision-making and increasing the system's resilience to market uncertainties.

Keywords: Energy, Hydroelectric, Time Series, Deep Learning, Prediction, Python, XM.

1. Descripción del problema

1.1. Problema de negocio

Desde un enfoque social, la energía es un pilar fundamental de la calidad de vida moderna. Las ciudades desarrolladas ofrecen un acceso continuo que permite un estilo de vida activo y productivo a cualquier hora del día, mientras que las zonas rurales o en desarrollo suelen carecer de un suministro constante y asequible. Estas disparidades en el acceso a la energía crean diferencias significativas que impactan aspectos esenciales de la vida cotidiana, como la refrigeración de alimentos y medicinas, la conectividad y la movilidad [1].

En el contexto colombiano, donde el 70% la matriz energética depende de hidroeléctricas [2], los precios de la energía son altamente volátiles debido a la incertidumbre inherente en el mercado. Esto hace fundamental anticipar las fluctuaciones de precios con un horizonte de 24 horas, un período clave que responde a la dinámica del negocio implementada por los actores y productores de energía en el país[3], [4]. La tarea de predicción se vuelve particularmente compleja frente a fenómenos climáticos extremos, como El Niño y La Niña [5], [6] así como por la interconexión inherente entre los sistemas meteorológico, hidrológico y energético. Además, factores como las interrupciones por fallas de mantenimiento, eventos imprevistos y la especulación en el mercado subrayan la necesidad de modelos predictivos confiables para optimizar la gestión de estos recursos.

Con base en las perspectivas expuestas anteriormente, el precio de la energía en Colombia ha aumentado hasta un 37% en 2023[7], impulsado por diversos factores especulativos y operacionales, especialmente debido a la alta dependencia de las centrales hidroeléctricas. Anticipar las variaciones en estos precios es complejo debido la demanda fluctuante y la competencia entre los diferentes actores del mercado.

Para abordar esta dificultad, se propone implementar un modelo predictivo basado en DL que integre datos históricos operacionales del Sistema Interconectado de Energía Mayorista de Colombia (XM)[8], como el volumen útil de los embalses, la capacidad de generación energética (hidráulica, solar y térmica), la demanda energética y los precios históricos de la bolsa de energía, Este enfoque permitiría a los actores del mercado reducir la incertidumbre y mejorar la toma de decisiones.

Para este estudio, la capacidad de los modelos de redes neuronales en predicciones de series temporales ha sido ampliamente respaldada en la literatura, como en [9], donde se implementaron modelos Long Short-Term Memory (LSTM), combinaciones de LSTM y Transformación Discreta de Ondas (DWT-LSTM), y Máquinas de Soporte Vectorial (SVM) para predicciones en Irán y Alemania, alcanzando eficiencias entre el 0.59% y el 4.2% en MAPE. Otro estudio ha planteado horizontes de predicción similares al propuesto en este trabajo de 24 horas, utilizando redes neuronales bayesianas para los mercados belga e italiano, con valores de MAPE de hasta un 8.2%, promediando cada mes por separado [10]. A nivel nacional, algunos autores han propuesto el uso de modelos de Deep Learning como LSTM, LSTM-CNN, DNN y GRU para la predicción de generación de energía fotovoltaica y de precios de energía. Este último se realizó con datos históricos entre 2013 y 2018, logrando resultados precisos en periodos cortos, con un MAPE de hasta 6.98% [11], [12].

1.2. Aproximación desde la analítica de datos

Este proyecto desarrolla modelos predictivos que permiten anticipar las fluctuaciones en los precios de la energía eléctrica en Colombia con un horizonte de 24 horas. Usando técnicas de analítica de datos y aprendizaje profundo, se analiza el comportamiento histórico del mercado energético. Los modelos incorporan variables operativas y climáticas que impactan los precios, tales como la demanda energética, la capacidad de generación y factores meteorológicos [13].

De esta manera los modelos predictivos permiten a las empresas del sector energético y a los reguladores mejorar la planificación y gestión de sus recursos. Al anticipar los precios de la energía con precisión, las generadoras pueden ajustar sus estrategias de producción, para alinear la oferta con las expectativas del mercado [14]. Lo que no solo contribuye a reducir los costos, sino también a maximizar la eficiencia en la utilización de las fuentes de energía disponibles.

Además, los comercializadores y distribuidores de energía tienen la oportunidad de utilizar estas predicciones para diseñar ofertas más competitivas y gestionar mejor el riesgo asociado a las variaciones de los precios en el mercado mayorista [15]. Esta anticipación facilita la implementación de estrategias efectivas que balancean la oferta y la demanda, promoviendo la sostenibilidad y la estabilidad del sistema energético.

Desde una perspectiva más amplia, los modelos predictivos pueden ser empleados para simular escenarios futuros bajo diferentes condiciones operativas o climáticas, como sequías o fenómenos de El Niño y La Niña [16]. Esto permite evaluar el impacto potencial de eventos extremos y desarrollar planes de contingencia que aumenten la resiliencia del sistema energético frente a las incertidumbres del mercado.

En resumen, la aplicación de la analítica de datos a través de estos modelos predictivos proporciona herramientas valiosas para la toma de decisiones estratégicas en el sector energético colombiano, promoviendo una gestión más eficiente y sostenible de los recursos energéticos.

1.3. Origen de los datos

Los datos utilizados en este proyecto provienen de fuentes de acceso público. La información de los datos del sistema energético interconectado se obtiene utilizando la base de datos oficial de XM, la cual es la entidad reguladora en el país para las tarifas eléctricas, a través de su API o su biblioteca disponible en Python. Ambos conjuntos de datos abarcan un período de 10 años, desde 2014 hasta 2024.

Además, los datos del **Índice Oceánico de El Niño (ONI)** se obtienen del Centro de Predicción Climática (Climate Prediction Center - CPC) de la **NOAA**. Al igual que los datos proporcionados por XM, los datos del ONI abarcan un período de 10 años, desde 2014 hasta 2024.

Tabla 1. Fuentes de datasets.

DataSet	Fuente	Tipo de Acceso
Energía/Precio, Nivel de embalses, vertimientos, consumos...	XM , Pydataxm	Público (Consumo API o descarga directa)
Índice Oceánico de El Niño (ONI)	Climate Prediction Center - ONI	Público (Descarga directa)

1.4. Métricas de desempeño

Para evaluar el desempeño del modelo analítico, se utilizan métricas estadísticas como la Raíz del Error Cuadrático Medio, el Error Medio Absoluto y el Error Porcentual Absoluto Medio, por sus siglas en inglés RMSE, MAE y MAPE respectivamente, así como posibles combinaciones de estas para comparar el precio estimado con el precio real. En cuanto a las métricas del negocio,

resulta relevante analizar los porcentajes de aumento en los precios durante periodos de sequía. Por ejemplo, es posible simular escenarios con el modelo entrenado, reduciendo la dependencia de la generación hidroeléctrica y compensándola con fuentes alternativas, como las plantas térmicas, para estabilizar el comportamiento del precio. A continuación, se presentan las formulaciones matemáticas correspondientes [17].

- **Mean Absolute Error (MAE):** Mide el promedio de los errores absolutos entre los valores reales y los predichos.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

- **Root Mean Squared Error (RMSE):** Mide la magnitud de los errores entre los valores reales y los predichos, penalizando más los errores grandes.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

- **Mean Absolute Percentage Error (MAPE):** Mide el error promedio absoluto en forma de porcentaje entre los valores reales y los predichos.

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Donde:

- y_i son los valores reales
- \hat{y}_i son los valores predichos
- n es el número de observaciones

Para predecir el precio de la energía con un horizonte de 24 horas utilizando modelos de machine learning y deep learning, es fundamental establecer métricas de rendimiento aceptables, específicamente en términos del MAPE (Mean Absolute Percentage Error o Error Porcentual Absoluto Medio).

El MAPE es una métrica ampliamente utilizada para evaluar la precisión de los modelos de predicción en series de tiempo, ya que proporciona el error promedio en términos porcentuales, lo que facilita su interpretación y comparación.

Valores aceptables de MAPE en la predicción de precios de energía:

- MAPE inferior al 10%: Indica una precisión excelente y es generalmente difícil de alcanzar en mercados volátiles como el energético[18].
- MAPE entre 10% y 20%: Considerado aceptable y práctico para aplicaciones operativas y de planificación[18].
- MAPE superior al 20%: Puede ser insuficiente para aplicaciones que requieren alta precisión, pero podría ser aceptable para análisis a largo plazo o tendencias generales[18].

2. Objetivos

2.1. Objetivo general

Desarrollar un modelo de series de tiempo basado en Deep Learning, para prever el impacto de las fluctuaciones de precipitación sobre la producción de energía, con un enfoque en la disponibilidad y la dinámica de precios. Esto incluye la recopilación, la limpieza, la exploración, la modelación y la visualización de conjuntos de datos públicos relacionados a la generación de energía hidroeléctrica como volumen útil del embalse, vertimientos, consumos y precio.

2.2. Objetivos específicos

- Limpiar y preprocesar los datos recopilados del sistema Interconectado de Energía Mayorista de Colombia (XM) para asegurar su calidad y consistencia, eliminando o imputando valores faltantes, removiendo datos anormales, y normalizando o estandarizando las variables.
- Explorar los datos mediante análisis descriptivo y visualización para identificar las relaciones entre las variables operativas. Esto incluirá el cálculo de correlaciones entre las principales variables y la identificación de posibles variables adicionales que influyan en los precios de la energía.
- Definir y construir un modelo predictivo basado en DL (LSTM, GRU u otro similar), ajustando sus parámetros clave, para predecir el precio de la energía. El modelo será evaluado dividiendo los datos en conjuntos de entrenamiento y prueba, utilizando métricas de evaluación como MAE, RMSE, MAPE, para medir la precisión y rendimiento del modelo.
- Comparar los resultados obtenidos con estudios previos o modelos alternativos, realizando una revisión de la literatura para contextualizar los hallazgos y evaluar si el modelo desarrollado mejora la precisión en comparación con enfoques tradicionales o previos.

3. Datos

3.1. Datos originales

Datos del Mercado Energético (API de XM): La base de datos utilizada para el entrenamiento del modelo fue obtenida a través de la API de XM, la cual proporciona información detallada sobre el Mercado de Energía Mayorista en Colombia. Esta API es de acceso público y no requiere autenticación, lo que facilita la extracción y manipulación de los datos necesarios para el análisis. La información fue recopilada en formato tabular y almacenada en archivos CSV.

Para interactuar con la API y extraer la información relevante, se empleó la librería `pydataxm` en Python. Esta librería proporciona una interfaz sencilla que actúa como un cliente, facilitando la comunicación con la API de XM. Al utilizar `pydataxm`, es posible enviar solicitudes y recibir datos de manera eficiente y flexible, sin la necesidad de implementar manualmente las llamadas a la API. De esta manera, la librería se encarga de gestionar las conexiones y la obtención de los datos, simplificando el proceso y permitiendo enfocarse en el análisis.

Los conjuntos de datos seleccionados abarcan registros horarios desde el año 2014 hasta el 2024, proporcionando un marco temporal extenso y detallado para el análisis [19].

Datos del Fenómeno de El Niño: Además de los datos energéticos, se incorporó información climatológica relacionada con el Fenómeno de El Niño, obtenida de los registros oficiales de la Administración Nacional Oceánica y Atmosférica (NOAA) de Estados Unidos. Esta información es fundamental para evaluar el impacto de las variaciones climáticas en el sector energético. Los datos del Fenómeno de El Niño provienen de la sección "Cold & Warm Episodes by Season" de la NOAA e incluyen el Índice Oceánico de El Niño (ONI), que mide las anomalías de la Temperatura Superficial del Mar (SST) en la región del Niño 3.4 (5°N-5°S, 120°-170°W). La NOAA actualiza estos datos automáticamente el primer jueves de cada mes, con valores de ONI expresados en °C que representan anomalías de la SST. Estos valores pueden ser ajustados hasta dos meses después de su publicación inicial debido al filtro de alta frecuencia aplicado a los datos ERSSTv5.

La serie temporal de datos climáticos se extiende desde 1950 hasta 2024, proporcionando un contexto histórico amplio, aunque para este estudio se delimita al periodo de 2014 a 2024. Los datos se presentan en agregados estacionales trimestrales, permitiendo el análisis de patrones climáticos y su posible influencia en el comportamiento del sector energético[20].

Datos categóricos: capturar variaciones temporales en el comportamiento del precio de la energía, **se agregaron las variables categóricas** FranjaHoraria y FinDeSemana. La variable FranjaHoraria segmenta los registros en tres franjas horarias (Valle, Llana o Punta), permitiendo analizar el impacto de la demanda en diferentes periodos del día. La variable FinDeSemana identifica si el registro corresponde a un fin de semana, cuando los precios suelen ser más bajos debido a la menor demanda por el cierre de industrias y oficinas, permitiendo la aplicación de la tarifa valle durante todo el día. Estas **variables** optimizan el análisis predictivo y la modelización de precios energéticos [21].

Característica de los datos:

El conjunto de datos comprende registros horarios que abarcan el periodo del 1 de agosto de 2014 al 1 de agosto de 2024, totalizando 87,672 registros con 10 características, detalladas en la Tabla 2. Características de variables. Con un tamaño aproximado de 5.0 MB, su manejo es viable en equipos de cómputo convencionales, sin necesidad de infraestructura de alto rendimiento.

Tabla 2. Características de variables

Variable	Descripción	Unidad de medida	formato
VoluUtilDiarEner	Aportes energéticos medios históricos, reflejando el promedio histórico de aportes energéticos al sistema.	(kWh)	Decimal
AporEnerMediHist	Capacidad útil diaria de energía, indicando la capacidad máxima de almacenamiento energético diaria.	(kWh)	Decimal
CompBolsNaciEner	Demanda real de energía en el sistema, medida de la energía consumida en tiempo real.	(kWh)	Decimal
CompBolsNaciEner	Componente de bolsa nacional de energía, representando valores económicos asociados al mercado energético.	(kWh)	Decimal
DemaReal	Demanda real de energía en el sistema, medida de la energía consumida en tiempo real.	(kWh)	Decimal
PrecBolsNaci	Precio de bolsa nacional, indicando el costo de la energía en el mercado mayorista.	(COP/kWh)	Decimal
Hidraulica	Generación hidroeléctrica, cantidad de energía producida por fuentes hidráulicas.	(kW)	Decimal
Termica	Generación termoeléctrica, cantidad de energía producida por fuentes térmicas.	(kW)	Decimal
FinDeSemana	Categóricos temporales	Valor adimensional	Entero
FranjaHoraria	Categóricos temporales	Valor adimensional	Entero

La granularidad horaria de los datos posibilita un análisis detallado de las fluctuaciones en el mercado energético, así como de su correlación con factores climáticos y económicos. Los datos han sido obtenidos de una fuente oficial y confiable, lo que garantiza su precisión y relevancia, proporcionando una base sólida para el análisis y la modelización en el marco de los objetivos del estudio.

3.2.Datasets

Dado que se quiere predecir una variable que cambia con el tiempo, es crucial mantener el orden temporal de los datos para preservar sus dependencias naturales. Por eso, se dividió el conjunto de datos de forma secuencial siguiendo el orden cronológico:

- **Conjunto de entrenamiento (80% inicial):** Utilizamos los primeros datos para entrenar el modelo. Dentro de este grupo, se reservó el 20% más reciente para validación interna, lo que ayudara a ajustar los hiperparámetros y previene el sobreajuste.
- **Conjunto de prueba (20% final):** Los datos más recientes se utilizan para probar el modelo final. Este conjunto no se usa durante el entrenamiento ni la validación interna, lo que permite evaluar cómo el modelo se desempeña con datos nuevos.

Esta metodología asegura que el modelo aprende de datos pasados y se evalúa con datos futuros, emulando condiciones reales de predicción y respetando la secuencia temporal de la serie.

Las variables categóricas FranjaHoraria y FinDeSemana se codificaron adecuadamente mediante la creación de columnas que reflejan estos valores en formato numérico, para ser interpretadas por los algoritmos de aprendizaje automático. La variable FranjaHoraria clasifica las observaciones en tres bloques horarios según el comportamiento del precio de la energíaTabla 3

Tabla 3. Precio de la energía por Franjas horarias

Franja	Horarios	Características
Franja Valle (0)	12:00 a.m. a 8:00 a.m.	Menor demanda y precios más bajos
Franja Llana (1)	- 8:00 a.m. a 10:00 a.m.	Precios moderados
	- 2:00 p.m. a 6:00 p.m.	
	- 10:00 p.m. a 12:00 a.m.	
Franja Punta (2)	- 10:00 a.m. a 2:00 p.m.	Demanda y precios alcanzan picos máximos
	- 6:00 p.m. a 10:00 p.m.	

La variable FinDeSemana identifica si una observación corresponde a un fin de semana (1) o día laboral (0), considerando que los precios tienden a ser más bajos durante los fines de semana debido a la menor demanda por el cierre de industrias y oficinas. Esta codificación permite capturar patrones temporales significativos que afectan la variabilidad del precio de la energía, mejorando la capacidad predictiva del modelo sin introducir sesgos o sobreajustes.

3.3. Analítica descriptiva

3.3.1. Comportamiento de las variables

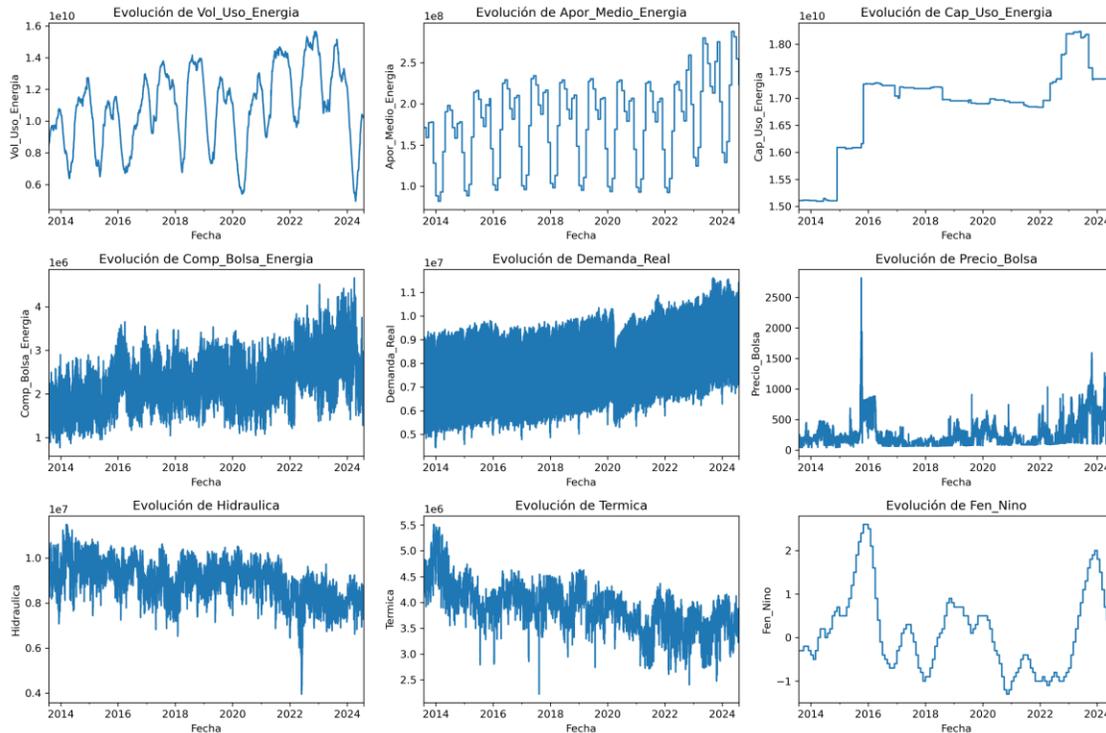
La Figura 1. Comportamiento de las diferentes características muestra, como primera parte del análisis exploratorio, el comportamiento temporal de las variables a lo largo del tiempo. El Volumen de Uso de Energía exhibe fluctuaciones estacionales, posiblemente asociadas a variaciones en la demanda relacionadas con las estaciones del año. Los Aportes Energéticos Medios Históricos presentan variaciones periódicas que reflejan cambios en la generación, influenciados por factores climáticos y operativos. Se observan incrementos puntuales en la Capacidad de Almacenamiento Energético, lo que sugiere mejoras en la infraestructura y capacidad del sistema eléctrico.

El Componente Económico de la Bolsa de Energía evidencia un crecimiento sostenido, alineado con el aumento continuo de la Demanda Real de Energía, asociado al desarrollo económico y al incremento del consumo energético. El Índice del Fenómeno de El Niño, que refleja sus ciclos históricos, indica una influencia directa en la dinámica de generación y precios, subrayando la importancia de incluir factores climáticos en el análisis del mercado energético.

El Precio de la Energía muestra alta volatilidad, con incrementos en períodos específicos, posiblemente vinculados a cambios en la oferta, condiciones externas o eventos imprevistos en el

mercado. La Generación Hidroeléctrica mantiene una tendencia estable con variabilidad moderada, mientras que la Generación Termoeléctrica exhibe mayor variabilidad en respuesta a condiciones operativas, precios de combustibles y fluctuaciones en la demanda.

Figura 1. Comportamiento de las diferentes características



3.3.1. Distribuciones de las Variables Principales

El estudio de las distribuciones estadísticas de las principales variables permite evaluar su comportamiento y variabilidad ver Figura 2:

Volumen Útil Diario de Energía: Presenta una distribución multimodal con picos que indican fluctuaciones en la disponibilidad de energía almacenada, reflejando diferentes niveles de reserva energética.

Aportes Energéticos Medios Históricos: Exhiben alta variabilidad, sugiriendo períodos con distintas condiciones de aporte energético.

Capacidad Útil de Energía: Muestra picos concentrados, posiblemente asociados a períodos de máxima capacidad operativa o inversiones en infraestructura.

Componente de Bolsa Nacional de Energía: Presenta una distribución aproximadamente normal, centrada en valores que indican estabilidad en la participación económica del mercado.

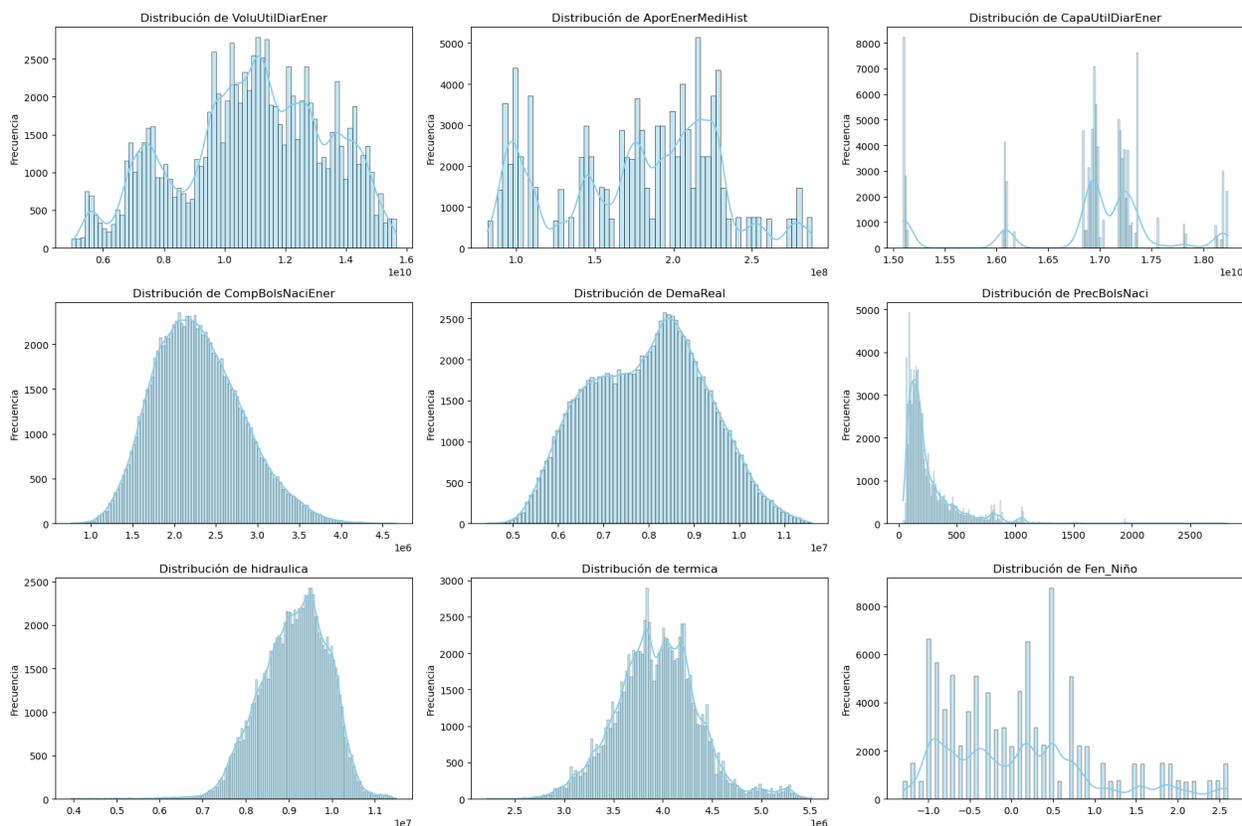
Demanda Real: Muestra una distribución sesgada a la izquierda, reflejando un consumo sostenido con picos de alta demanda en ciertos períodos.

Precio de Bolsa Nacional: Exhibe alta asimetría y curtosis, con predominio de valores bajos y algunos picos significativos, indicando volatilidad y posibles eventos extremos.

Generación Hidroeléctrica y Termoeléctrica: La hidroeléctrica muestra mayor estabilidad y una distribución más concentrada, mientras que la termoeléctrica refleja mayor variabilidad, asociada a su flexibilidad operativa.

Índice del Fenómeno de El Niño: Muestra alta variabilidad, con valores que sugieren diferentes intensidades del fenómeno a lo largo del período estudiado.

Figura 2. Distribución de las variables



3.3.2 Análisis de Autocorrelación y Autocorrelación Parcial del Precio de Bolsa

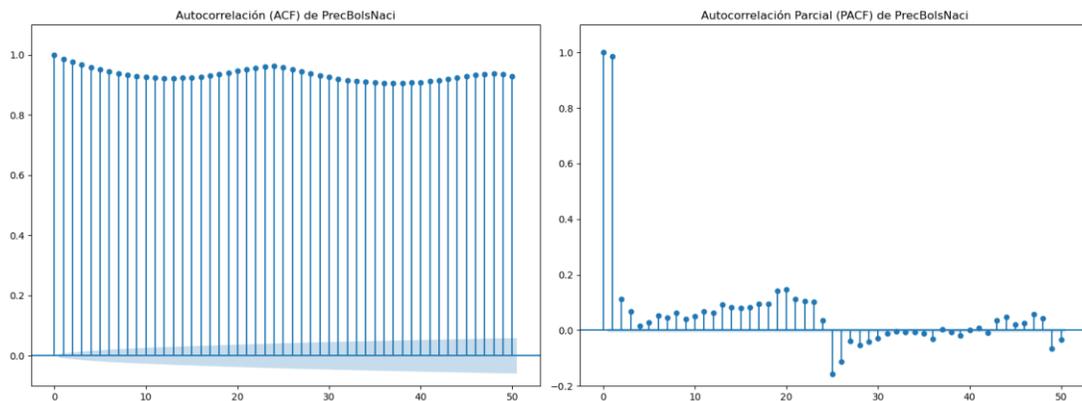
El análisis de las funciones de autocorrelación o por sus siglas en inglés ACF y autocorrelación parcial (PACF) del Precio de Bolsa proporciona información clave sobre su estructura temporal ver Figura 3:

ACF: Muestra alta persistencia a lo largo de varios rezagos, indicando una fuerte dependencia temporal y sugiriendo un comportamiento no estacionario con memoria prolongada.

Esto implica que los valores futuros están altamente correlacionados con valores pasados, lo cual puede ser indicativo de patrones estacionales o tendencias a largo plazo.

PACF: Presenta un valor significativo en el primer rezago, seguido de valores decrecientes en los rezagos sucesivos. Esto indica que la serie temporal puede ser modelada con una estructura autorregresiva de primer orden (AR(1)), aunque la persistencia en la ACF sugiere la necesidad de modelos más complejos, como ARIMA o SARIMA, para capturar adecuadamente la dinámica de la serie.

Figura 3 Autocorrelación del precio de bolsa



3.3.3 Importancia de Variables mediante Random Forest

La gráfica de importancia de variables obtenida mediante el modelo Random Forest, mostrada en la Para explorar estas posibles correlaciones, se puede complementar el análisis con la gráfica de correlación *Figura 5 Matriz de correlación*, que muestra las relaciones entre estas variables y el Precio de Bolsa Nacional. Al considerar conjuntamente estas variables, se resalta su importancia en la dinámica de precios, evitando segmentaciones que podrían obviar su influencia colectiva en el modelo predictivo.

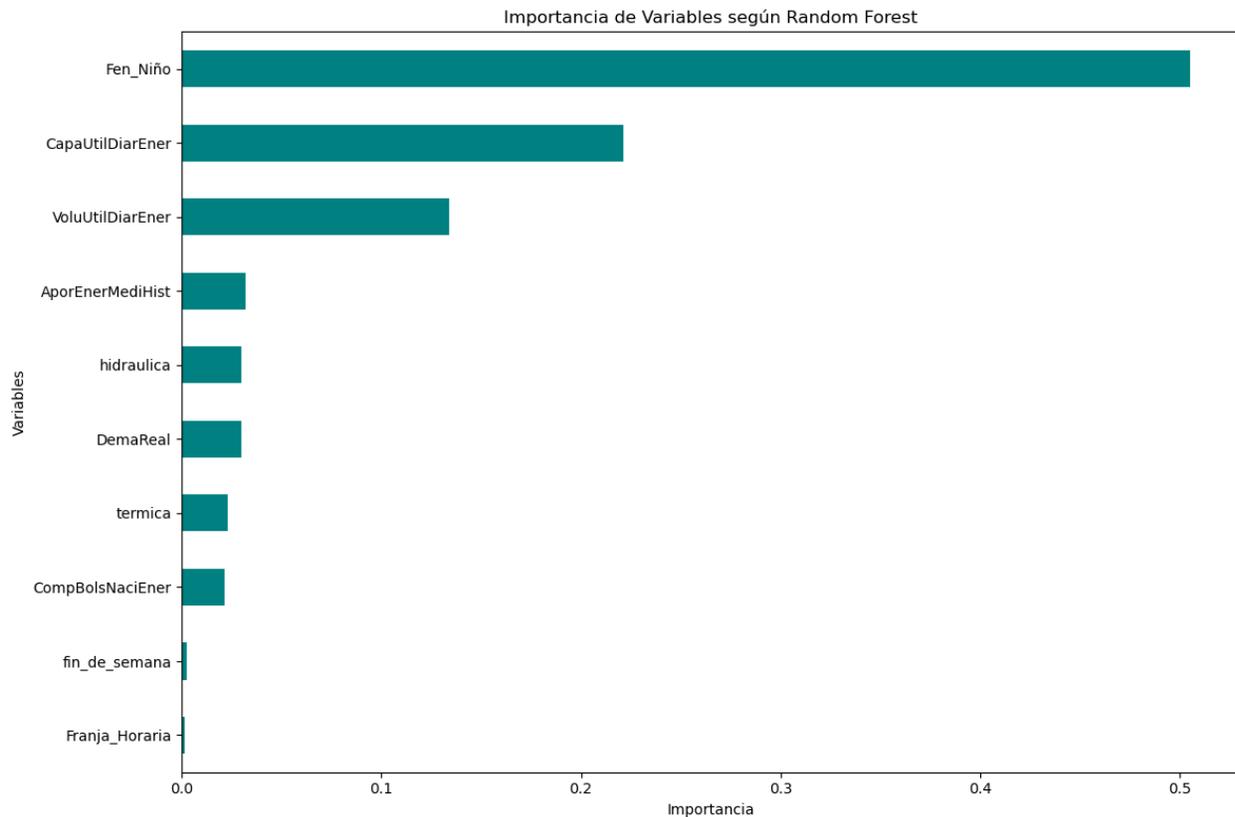
Figura 4. indica cuáles características tienen mayor influencia en la predicción del Precio de Bolsa Nacional. Se aprecia que el Índice del Fenómeno de El Niño es la variable más significativa, lo que sugiere que las variaciones climáticas asociadas a este fenómeno tienen un impacto considerable en la volatilidad de los precios energéticos.

Después de estas tres variables principales, se observa que las siguientes cinco variables— Demanda Real, Aportes Energéticos Históricos, Generación Hidroeléctrica, Generación Termoeléctrica y el Componente de Bolsa Nacional de Energía—tienen niveles de importancia muy similares. Esto sugiere que, aunque individualmente su influencia es menor que la de las tres primeras variables, colectivamente contribuyen de manera significativa al comportamiento de los

precios. La similitud en sus niveles de importancia podría indicar relaciones interdependientes o correlaciones entre ellas y el precio de la energía.

Para explorar estas posibles correlaciones, se puede complementar el análisis con la gráfica de correlación *Figura 5 Matriz de correlación*, que muestra las relaciones entre estas variables y el Precio de Bolsa Nacional. Al considerar conjuntamente estas variables, se resalta su importancia en la dinámica de precios, evitando segmentaciones que podrían obviar su influencia colectiva en el modelo predictivo.

Figura 4 Random Forest - Importancia de Variables



3.3.4 Análisis de Correlación entre Variables

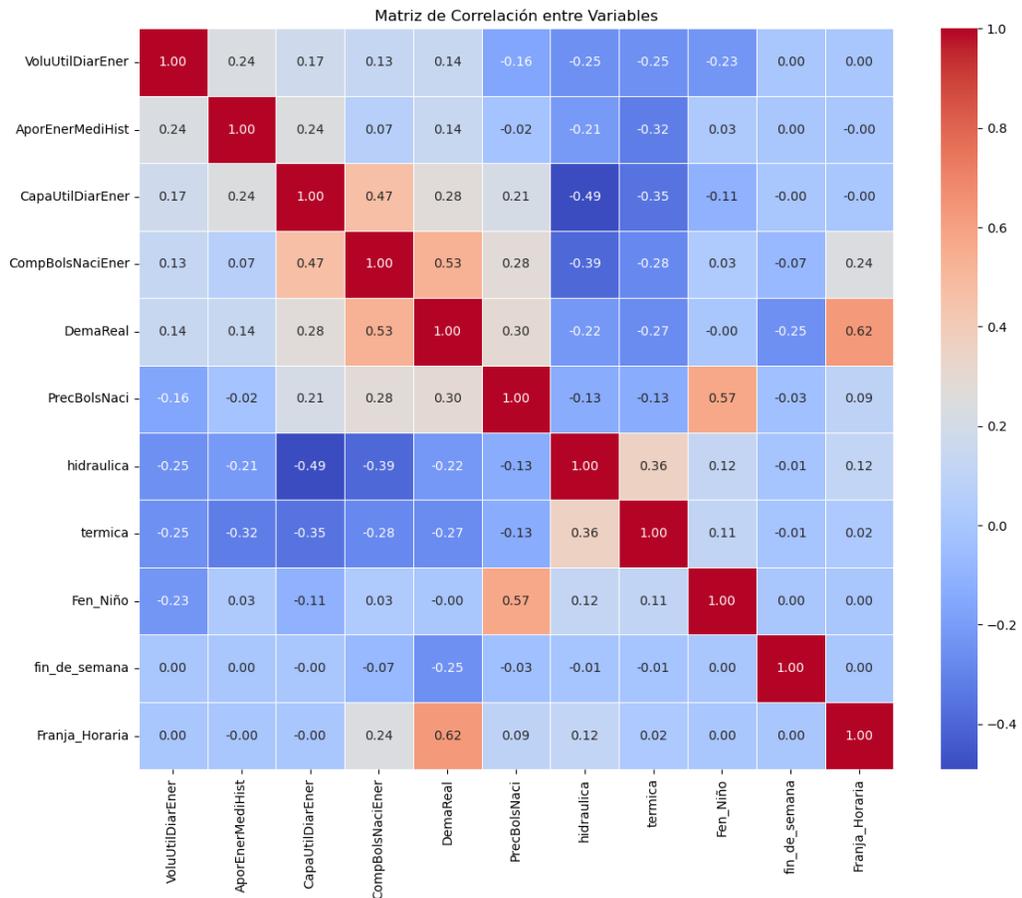
La matriz de correlación complementa el análisis anterior al proporcionar una visión cuantitativa de las relaciones lineales entre las variables *Figura 5*:

Correlaciones Positivas Significativas: Se observa que la variable con mayor correlación con el Precio de Bolsa Nacional es el Índice del Fenómeno de El Niño. Las siguientes tres variables—Demanda Real, Generación Termoeléctrica e Hidroeléctrica y Capacidad Útil Diaria de Energía—presentan correlaciones positivas con coeficientes entre 0.3 y 0.2. Aunque estas correlaciones indican una relación moderada y no muy significativa, sugieren que una mayor capacidad y demanda están asociadas con incrementos en los valores económicos del mercado. Sin

embargo, dado que estas correlaciones no son fuertes, es probable que otros factores también influyan de manera considerable en la volatilidad de los precios energéticos.

Correlaciones Negativas: Se observa una correlación negativa entre Generación Hidroeléctrica y Termoeléctrica, reflejando la dinámica de sustitución entre estas fuentes de generación. Cuando la generación hidroeléctrica disminuye, la termoeléctrica aumenta para suplir la demanda, y viceversa.

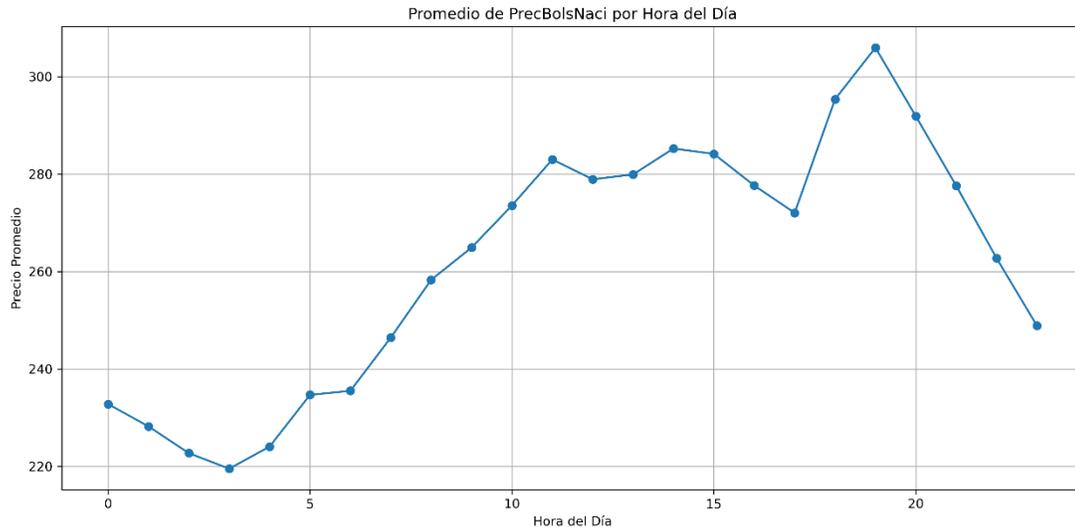
Figura 5 Matriz de correlación



Variación Promedio del Precio de Bolsa Nacional por Hora del Día

La gráfica de la variación promedio del **Precio de Bolsa Nacional** por hora del día revela patrones intradía significativos Figura 6. Los precios más bajos se registran durante las primeras horas de la madrugada, entre las 00:00 y las 05:00 horas, cuando la demanda de energía es mínima. A partir de las 05:00 horas, se observa un incremento gradual del precio, alcanzando su punto máximo alrededor de las 19:00 horas, reflejando el aumento en el consumo residencial, comercial e industrial

Figura 6 Promedio precio Bolsa Nacional por 10 Años



La identificación de estos patrones horarios es crucial para la planificación estratégica, ya que permite a generadores y comercializadores optimizar la programación de generación y gestionar recursos de manera eficiente. Además, facilita la predicción de fluctuaciones, ayudando a anticipar cambios en los precios horarios, maximizar la rentabilidad y minimizar riesgos. Por último, contribuye al equilibrio entre oferta y demanda, reduciendo el riesgo de sobrecostos, congestiones o déficits de energía, y asegurando la confiabilidad del suministro eléctrico.

4. Proceso de Analítica

4.1. Pipeline principal

A continuación, se detalla el flujo de trabajo implementado. El proceso inicia con la interacción directa con la API de Sinergox XM, utilizando la librería **pydataxm** para la extracción de datos. La mayoría de las variables se encuentran disponibles con una granularidad diaria; sin embargo, ciertas columnas simulan diferentes horas del día, como: *'CompBolsNaciEner'*, *'DemaReal'*, *'PrecBolsNaci'*, *'DispoReal'* (que incluye categorías como *'hidráulica'*, *'solar'* y *'térmica'*). Por otro lado, algunas variables presentan exclusivamente datos diarios, como: *'VoluUtilDiarEner'*, *'AporEnerMediHist'* y *'CapaUtilDiarEner'*.

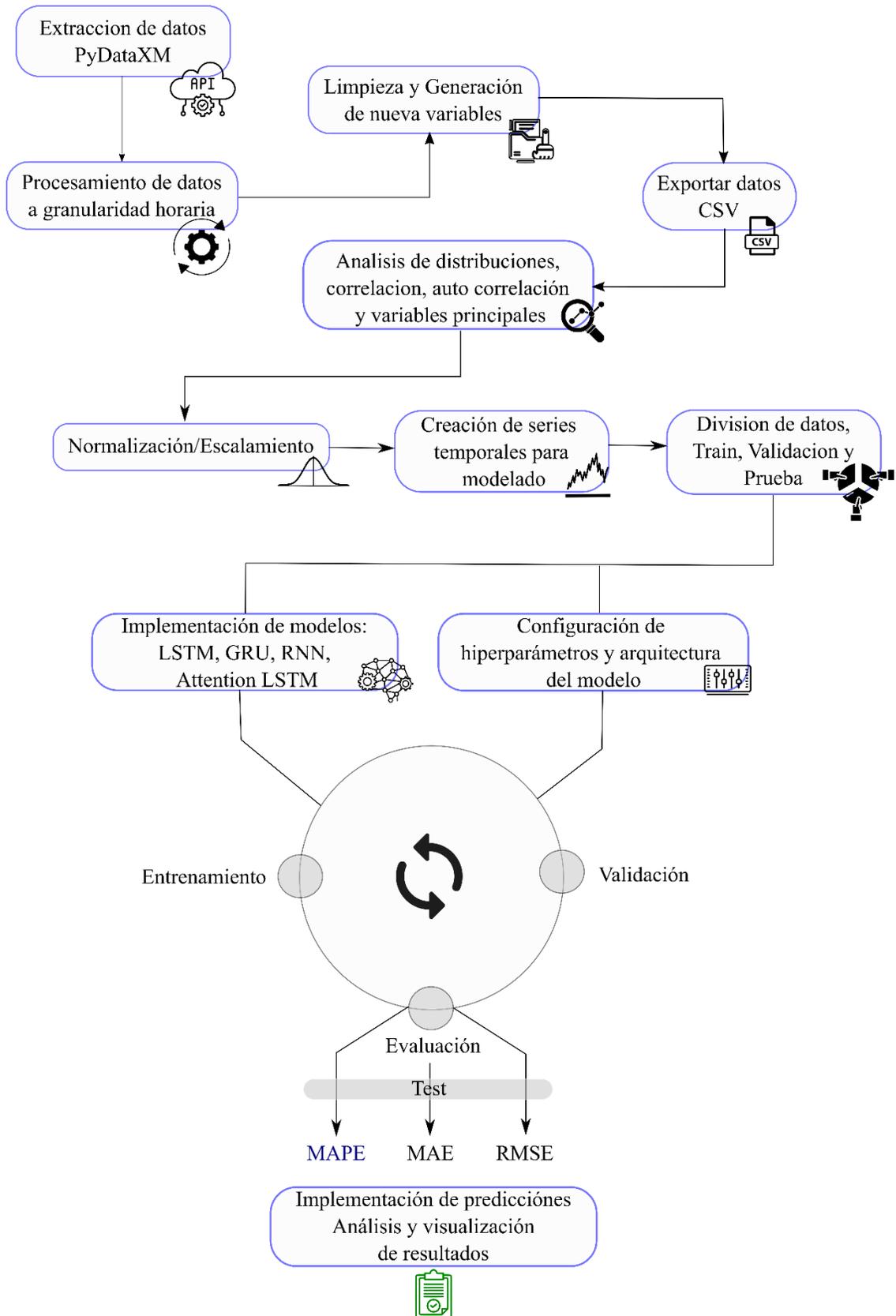
Los datos crudos obtenidos se almacenan en un archivo CSV con el propósito de mejorar la eficiencia en el manejo de datos y mitigar problemas asociados a la disponibilidad intermitente de la API. Este enfoque permite acelerar los tiempos de ejecución y reducir la dependencia de las consultas en tiempo real.

Posteriormente, la información es cargada y sometida a un proceso de transformación. Uno de los pasos clave en esta etapa es la segmentación de las variables diarias. Se replican los valores de las variables diarias para los 24 períodos del día, creando así un dataset con datos en resolución horaria. Al mismo tiempo, las columnas que contienen datos en formato por hora se transforman de columnas a filas, lo que expande el tamaño del dataset al incluir las 24 horas del día en las fechas correspondientes.

Además, la variable "Disponibilidad Real" se clasifica en tres grupos principales de generación energética: hidráulica, térmica y solar. Esto facilita el análisis y la construcción de modelos especializados para cada fuente de energía. Otras variables como 'FinDeSemana' y 'FranjaHoraria' son extraídos con análisis expuestos anteriormente. En paralelo, se realiza una exhaustiva limpieza de datos, buscando valores nulos, duplicados, y se efectúa un análisis exploratorio para detectar correlaciones entre variables. Este análisis también incluye la visualización de las distribuciones, la autocorrelación y la identificación de valores atípicos que puedan influir en la precisión de los modelos y el grado en que cada variable independiente influye en la varianza del precio de la energía en la bolsa nacional.

Con los datos procesados, se procede a la implementación de modelos de aprendizaje profundo. Se prueba una variedad de arquitecturas de redes neuronales y se exploran distintos hiperparámetros, tanto propios de los modelos como de la arquitectura de las capas, con el fin de identificar la combinación óptima para obtener predicciones precisas de la variable objetivo 'PrecBolsNaci'.

Figura 7. Pipeline-Flujo de Trabajo.



4.2. Preprocesamiento

Una vez descargados los datos crudos, se procedió a la limpieza y preparación de estos. Durante este proceso, se eliminó la columna correspondiente a los aportes de energía solar debido a su baja variabilidad y la limitada contribución a la generación total, lo cual tiene un efecto casi nulo en los modelos predictivos. Además, para asegurar la compatibilidad de las diferentes escalas de las variables y facilitar la convergencia de los modelos DL, se aplicó un escalamiento utilizando la técnica `MinMaxScaler`, normalizando todas las variables en un rango de $[0, 1]$.

Dado que muchas variables estaban disponibles solo a nivel diario, fue necesario realizar una aumentación de datos para transformar las entradas a nivel horario. Para ello, se hizo una réplica de los valores diarios en las 24 horas correspondientes de cada día, lo que asegura que los modelos puedan manejar datos con una resolución temporal adecuada para capturar la dinámica horaria del sistema energético.

Adicionalmente, se generan nuevas características como tendencias y lags de la variable objetivo 'PrecBolsNaci' para capturar patrones temporales en la serie, optimizando así la capacidad predictiva del modelo. Además, se realizan algunos análisis por bimestre, trimestre o semestre, con el fin de analizar si existe algún fenómeno estacionario que permita implementar otro tipo de modelos como ARIMA o SARIMAX[22], de lo cual no se obtuvo un resultado concluyente que representara estacionalidad en los datos, por lo cual no es incluido en este estudio.

4.3. Modelos

En este trabajo se implementan modelos de aprendizaje profundo con un enfoque en redes LSTM para abordar el desafío de la predicción de series temporales. Las redes LSTM son consideradas debido a su capacidad para capturar dependencias a largo plazo en datos secuenciales, lo que las hace especialmente eficaces en series temporales con patrones complejos y estacionales[23].

Se evalúan varias arquitecturas de LSTM, en donde se incorpora regularización mediante capas de dropout para reducir el riesgo de sobreajuste. Además, se optimiza la estructura del modelo y los hiperparámetros, tales como el número de unidades, la tasa de aprendizaje, el tamaño del batch y las ventanas de predicción, utilizando como técnica de ajuste un `GridSearch` manual.

En paralelo, se implementan otros modelos para evaluar su rendimiento frente al modelo LSTM convencional. Entre estos, se incluye el modelo LSTM RNN, lo cual combina la capacidad

de las LSTM con un enfoque recurrente para modelar relaciones temporales, lo cual podría representar rendimientos superiores[24].

Además, se implementa el modelo GRU, por sus siglas en ingles Gated Recurrent Unit, como alternativa a las LSTM. Las redes GRU, al contar con una arquitectura más simple, ofrecen un rendimiento similar, pero con tiempos de entrenamiento más rápidos, lo que resulta beneficioso dada la magnitud del conjunto de datos y la necesidad de hacer predicciones en tiempo real [25].

Por último, se integra un modelo Attention LSTM para evaluar el impacto de la atención en la mejora de las predicciones. Este modelo utiliza mecanismos de self-attention para ponderar la relevancia de diferentes partes de la secuencia temporal, combinando los beneficios de las LSTM con la capacidad de atención para capturar dependencias a largo plazo de manera más eficiente[26].

Cada uno de estos modelos fue ajustado y optimizado para lograr un mejor desempeño en la predicción del precio de la e, teniendo en cuenta tanto las características temporales como las complejidades del conjunto de datos.

A continuación, se describen las arquitecturas utilizadas y los parámetros variados:

Tabla 4. Hiperparámetros fijos y variados.

Hiperparámetro	Valores Variados	Valor Fijo
Número de neuronas (units)	20, 40, 60, 80, 100, 120	-
Tasa de aprendizaje (learning rate)	0.01, 0.001	-
Tamaño de batch (batch size)	25, 50, 100	-
Tamaño de la ventana (look_back)	24, 48, 72, 96	-
Horizonte de predicción (n_horizon)	-	24
Función de pérdida (loss)	-	'mse'
Optimizador	-	Adam
Épocas (epochs)	-	500
Escalado de datos	-	Min-Max Scaler
Early Stopping (Patience)	-	10 (patience)
Monitor de Early Stopping	-	'val_loss'

Como se aprecia en la tabla anterior (Tabla 4), se ha utilizado la función de pérdida **Error Cuadrático Medio (MSE)**, Es importante notar que esta es una función útil en series de tiempo, ya que penaliza fuertemente los errores grandes al elevarlos al cuadrado. Esto permite que el modelo ajuste mejor en situaciones donde los errores extremos son críticos para el rendimiento. Al ser una función diferenciable, facilita el uso de algoritmos de descenso de gradiente, lo que hace el

entrenamiento más eficiente y garantiza una convergencia estable. Además, el MSE es fácil de interpretar, al expresar el error en unidades relacionadas con la variable objetivo, lo que permite evaluar la precisión del modelo de manera intuitiva[27].

4.3.1. Descripción de las arquitecturas de los modelos implementados:

En la Tabla 5, se hace una breve descripción de las capas de los modelos implementados:

Tabla 5. Descripción de capas de modelos implementados.

Modelo	Capa de entrada	Capa(s) intermedia(s)	Capa de salida	Características adicionales
LSTM	Recibe la secuencia de datos temporales	1 capa LSTM con unidades ajustables, activación 'relu'	Capa Dense: predicción de 24 horas (número de unidades = horizonte)	Ninguna
LSTM RNN	Recibe la secuencia de datos temporales	1 capa LSTM con activación 'relu'	Capa Dense: predicción de 24 horas	Predicción inicializada con matriz de ceros; Capa Reshape para organizar la salida
GRU	Recibe la secuencia de datos temporales	1 capa GRU con activación 'relu'	Capa Dense: predicción de 24 horas	Ninguna
RNN	Recibe la secuencia de datos temporales	1 capa SimpleRNN con unidades ajustables, activación 'relu'	Capa Dense: predicción de 24 horas	Ninguna
Attention LSTM	Recibe la secuencia de datos temporales	1 capa LSTM con activación 'relu', Self-attention, otra capa LSTM	Capa Dense: predicción de 24 horas	Concatenación de salida de LSTM y atención; Dropout para regularización

4.4.Métricas

Se utilizaron diversas métricas para evaluar el desempeño del modelo en la predicción de la variable objetivo 'PrecBolsNaci'. Cada métrica aporta una perspectiva distinta sobre el error presentado por el modelo al realizar las predicciones. El **MAE** mide la magnitud media del error entre los valores predichos y los valores reales, sin considerar su dirección, proporcionando una interpretación directa de la desviación promedio en las unidades de la variable de salida. Por otro

lado, la **RMSE** otorga mayor peso a los errores grandes, siendo particularmente útil cuando son penalizados más severamente en el contexto del análisis. Además, se empleó el **MAPE** para evaluar el porcentaje de error relativo, lo que facilita la interpretación del error en términos porcentuales respecto a los valores reales [17][28]. Estas métricas en conjunto proporcionan una visión integral de la precisión y robustez del modelo, especialmente al tratarse de datos de series temporales relacionadas con la predicción de precios energéticos. Sus representaciones matemáticas se encuentran descritas en ‘Métricas de desempeño’:

5. Metodología

La metodología adoptada se centra en el diseño, implementación y validación de modelos de aprendizaje profundo para la predicción de series temporales en el contexto de precios energéticos. Se estructura en etapas que abarcan desde el preprocesamiento de los datos, la experimentación con configuraciones de entrada y arquitectura, hasta la validación de los resultados obtenidos.

5.1. Baseline

Con el dataset correctamente procesado, se plantea la implementación de varios modelos de DL, entre ellos se incluyen el **LSTM**, **LSTM RNN**, **GRU**, **RNN** y **Attention LSTM**, para esto se utiliza 3 conjuntos de datos; uno exclusivo para pruebas, otro para validación, y un último para prueba, dividido de acuerdo con lo mencionado en la sección Datasets.

Inicialmente, se configuran las secuencias de entrada utilizando varios tamaños de ventana (`look_back`), con un valor de 24 horas, manteniendo un horizonte de predicción fijo de 24 pasos (24 horas) en el futuro. Para capturar patrones y fluctuaciones en los datos, se establece un paso del 50% del tamaño de cada ventana entre secuencias consecutivas, estos parámetros se varían basados en [29] donde se muestra la sensibilidad de los modelos al overlapping o superposición de ventanas y su tamaño. Este enfoque de overlapping del 0.5 permite captar variaciones sutiles sin el costo computacional asociado con el desplazamiento de un solo paso, logrando un equilibrio entre la capacidad de detección de patrones y la eficiencia en el procesamiento.

Además de la ventana de 24 horas, los modelos también se evalúan con ventanas más amplias, como 48, 72 y 96 horas, con la finalidad de analizar cómo perturba el tamaño de la ventana el rendimiento y capacidad del modelo para aprender patrones a distintas escalas temporales. Estas

configuraciones se aplican a cada modelo para identificar la combinación óptima de tamaño de ventana y arquitectura en función del horizonte de predicción.

Este enfoque permite comprender el rendimiento de cada modelo en función del tamaño de la ventana y el paso entre secuencias, así como su capacidad para capturar patrones estacionales diarios de corto plazo, que son básicos para una predicción precisa en series temporales.

Con este enfoque de ventaneo y partición de datos, además de las distintas variaciones de parámetros mencionadas en la sección de ‘Modelos’, se hacen 3 experimentos para verificar la influencia de variables adicionales en la predicción, como se muestra a continuación.

5.1.1. Experimentos

Se llevaron a cabo tres experimentos, variando el conjunto de características de entrada en cada uno:

5.1.1.1. Todos los features: Incluye todas las variables del dataset, como 'PrecBolsNaci', 'hidraulica', 'termica', 'CapaUtilDiarEner', 'Fen_Niño', 'FinDeSemana', 'FranjaHoraria', 'DemaReal', 'VoluUtilDiarEner', 'AporEnerMediHist', 'CompBolsNaciEner'.

5.1.1.2. Solo precio: Utiliza únicamente la variable 'PrecBolsNaci', como única entrada/salida para el modelo, lo que permite explorar su comportamiento específico, lo cual puede brindar mayor sensibilidad al contexto histórico.

5.1.1.3. Variables más influyentes: Basado en el análisis de correlación y la relevancia de las variables (aplicando RandomForest), se seleccionan las características más relevantes: 'PrecBolsNaci', 'VoluUtilDiarEner', 'Fen_Niño', 'CapaUtilDiarEner'.

En cada experimento, se ajustan los hiperparámetros siguiendo el proceso descrito en la sección de ‘Modelos’. La elección final del modelo se apoyó en la métrica **MAPE** con el resultado más bajo obtenido en el conjunto de test, lo cual certifica una precisión en la predicción de la variable 'PrecBolsNaci'(Precio de la Bolsa Nacional). Los resultados de los modelos, configuraciones de ventana y conjuntos de características se exponen en la sección de Resultados y discusión, permitiendo una comparación detallada de cada arquitectura y configuración.

5.2. Validación

El proceso de validación de los modelos se diseñó de manera estructurada para obtener resultados consistentes y representativos. Esto incluyó la limpieza y normalización de los datos, la generación de secuencias temporales con ventanas de 24, 48, 72 y 96 horas, y un horizonte de predicción fijo de 24 horas. También se incorporaron variables derivadas como 'FinDeSemana' y 'FranjeHoraria' para captar patrones relevantes en los datos.

Los datos se dividieron en un 80% para entrenamiento (reservando un 20% adicional para validación dentro de los datos de entrenamiento) y un 20% para prueba. La optimización de hiperparámetros se llevó a cabo mediante búsqueda manual, ajustando parámetros como unidades, tamaño de lote, tasa de aprendizaje y ventanas de predicción.

El rendimiento de los modelos se evaluó mediante métricas como MAE, RMSE y MAPE, con el objetivo de analizar su precisión y su comportamiento en diferentes escenarios de predicción.

5.3. Iteraciones y evolución

A través de las iteraciones, el enfoque principal fue optimizar tanto el preprocesamiento como la selección de modelos y la afinación de sus parámetros, buscando siempre mejorar la precisión en el desempeño de la serie temporal.

5.3.1. Iteraciones de Preprocesamiento:

Inicialmente, se probaron distintas estrategias de preprocesamiento, como el escalado Min-Max, la estandarización y el uso de filtros como el `savgol_filter` de SciPy o suavizados (rolling) en variables como 'Demanda_Real' y 'Comp_Bolsa_Energia' que presentan un alto ruido o dispersión. También se añadieron variables temporales, como 'FinDeSemana' para identificar fines de semana y 'FranjaHoraria' para clasificar el día en zonas horarias asociadas con variaciones de precios. Adicionalmente, se exploraron segmentaciones basadas en meses, días festivos y días de la semana; sin embargo, estas últimas no mostraron representatividad en el modelo y fueron descartadas del análisis.

5.3.2. Iteraciones para Definir el Tamaño de Ventana y Horizonte de Predicción:

En la siguiente fase, se evaluaron diversas configuraciones de ventana temporal y horizontes de predicción para observar su impacto en el rendimiento. Partiendo de una ventana y horizonte de predicción 1 hora, hasta llegar a 24 horas que es el estándar del negocio, se probaron también ventanas de varios días en su equivalente en horas como 48, 72 y 96 permitiendo capturar comportamientos tanto de corto como de mediano plazo en el espectro de las horas. Asimismo, se

variaron los pasos entre las secuencias, lo que permitió encontrar un equilibrio entre la precisión con la captura del comportamiento temporal y la eficiencia computacional con una superposición entre ventanas del 50%.

5.3.3. Iteraciones de Selección de Modelos y Configuración de Capas:

A continuación, se evaluaron diferentes modelos (LSTM, LSTM RNN, GRU, RNN y Attention LSTM). En esta fase además de estudiar las distintas redes neuronales, se experimentó con el uso de capas de regularización, como dropout, para reducir el riesgo de sobreajuste. Estas iteraciones ayudaron a identificar las arquitecturas que mejor captaban las dependencias temporales en los datos, aumentando así el rendimiento en las métricas.

5.3.4. Iteraciones de Afinación de Hiperparámetros:

Finalmente, se llevaron a cabo iteraciones enfocadas en la sintonización de hiperparámetros clave, tales como el número de unidades o neuronas en las capas de los modelos, la tasa de aprendizaje, el tamaño del lote, las distintas ventanas y la selección de variables adicionales para la predicción. A través de una búsqueda manual (implementando bucles entre los distintos parámetros variados), se identificó la combinación óptima de parámetros que maximizaba la precisión en el conjunto de test. Esta fase permitió mejorar de manera considerable la capacidad de generalización del modelo.

Cada grupo de iteraciones aportó ajustes significativos al modelo, lo cual permitió alcanzar un resultado final que integra un preprocesamiento adecuado, una arquitectura de red óptima y parámetros calibrados específicamente para la predicción de series temporales en este contexto.

5.4. Herramientas

- **Google Colab:** Plataforma de desarrollo colaborativo basada en la nube, utilizada para la implementación y prueba de código en entornos compartidos.
- **Postman:** Herramienta para la interacción con APIs, específicamente utilizada para la conexión con el API de XM en este proyecto.
- **Anaconda:** Entorno de gestión de paquetes y desarrollo en Python, con versión 3.10.4 instalada para garantizar la compatibilidad de las bibliotecas utilizadas.
- **Visual Studio Code:** Entorno de desarrollo integrado (IDE) local para la creación y edición de código.

- **OneDrive:** Plataforma de almacenamiento en la nube que permite la edición colaborativa de documentos y archivos en tiempo real.
- **Mendeley:** Gestor de referencias utilizado para organizar y acceder a la bibliografía relevante del proyecto.
- **GitHub:** Repositorio de código fuente que facilita el control de versiones y la colaboración en el desarrollo del proyecto.
- **Kaggle:** Plataforma utilizada para el análisis de datos y el acceso a datasets relevantes para el estudio.

Entorno de Computación

Todos los modelos de DL fueron entrenados y evaluados en el siguiente entorno de hardware y software:

- **Hardware:**
 - **Procesador:** (Core i9)
 - **Memoria RAM:** (16 GB)
 - **Unidad de Almacenamiento:** (SSD de 1 TB)
 - **Tarjeta Gráfica (GPU):** (NVIDIA T600 Memoria de video dedicada 4096 MB GDDR6)
- **Software:**
 - **Sistema Operativo:** (Windows 11)
 - **Entorno de Desarrollo:** (VS Code)
 - **Librerías y Versiones:**
 - **Python:** (Python 3.10.4)
 - **TensorFlow:** (v2.8.0)

6. Resultados y discusión

A continuación, se detallan las métricas de desempeño de cada modelo en los conjuntos de prueba, como se ha mencionado anteriormente, se utilizan el MAE [en \$COP/kWh], MAPE [en %] y RMSE [en \$COP/kWh] como las métricas clave. Estas se calcularon con base en los valores reales; después de realizar las predicciones, se aplicó la inversa de la escala para obtener valores en las unidades originales.

6.1. Métricas

Para esta sección, se presentan los resultados detallados de los experimentos con distintos modelos y combinaciones de hiperparámetros mencionados en ‘Baseline’, con el objetivo de identificar el conjunto que mejora la precisión en el pronóstico del precio energético.

6.1.1. Experimento con todas las características:

Para este primer experimento, denominado ‘Todos los features’, los resultados son expuestos en la Tabla 6, donde se detalla el mejor desempeño de cada modelo en las diferentes configuraciones de ventana y parámetros.

Tabla 6. Resultados para experimento ‘Todos los features’.

Ventana	Modelo	Hiperparámetros			MAPE	MAE	RMSE
		units	l_r^*	b_s^*			
24	GRU	40	0.01	100	15.39	60.26	92.59
48	GRU	80	0.01	50	13.19	49.17	76.62
72	GRU	40	0.01	25	15.36	52.76	83.87
96	GRU	40	0.01	50	13.13	47.02	76.33
24	LSTM	40	0.001	100	14.43	53.89	85.94
48	LSTM	120	0.01	100	13.45	48.33	75.69
72	LSTM	40	0.01	50	15.82	61.31	96.08
96	LSTM	80	0.01	50	13.99	50.14	81.16
24	Attention LSTM	120	0.01	50	14.61	57.25	87.22
48	Attention LSTM	40	0.01	50	13.41	49.69	77.67
72	Attention LSTM	120	0.01	25	16.7	63.44	91.86
96	Attention LSTM	60	0.01	25	15.63	61.1	92.18
24	LSTM RNN	40	0.01	50	14.92	58.1	89.76
48	LSTM RNN	40	0.01	100	13.1	49.13	77.43
72	LSTM RNN	120	0.01	100	15.69	54.94	85.77
96	LSTM RNN	80	0.01	25	14.64	55.25	84.62
24	RNN	40	0.001	25	15.53	55.41	87.71
48	RNN	40	0.001	25	14.11	52.98	81.24
72	RNN	20	0.001	100	15.69	51.65	83.28
96	RNN	40	0.01	50	13.88	55.2	86.64
Promedio					14.63	54.35	84.68

* l_r : learning rate; b_s : batch size

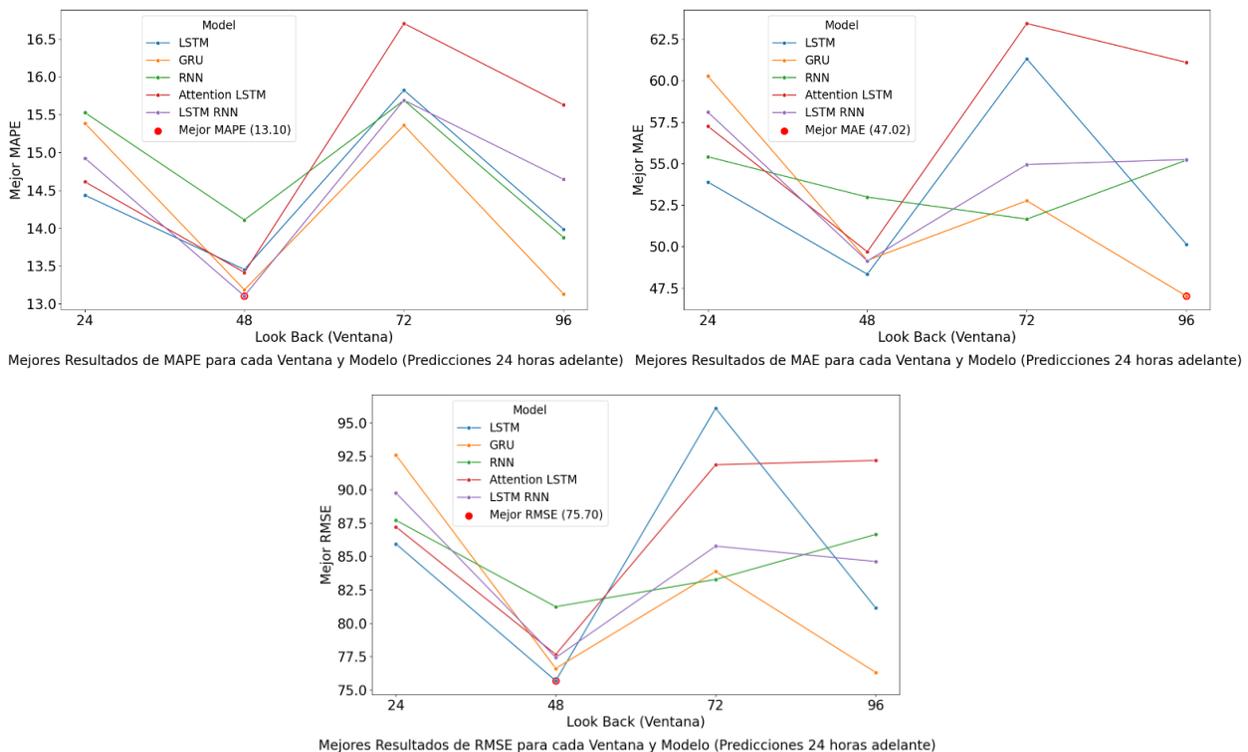
Los resultados obtenidos reflejan cómo cada modelo reacciona ante distintos tamaños de ventana. Se observa un menor MAPE al usar ventanas de 48 y 96 horas en varios modelos,

especialmente en los modelos GRU y LSTM, lo cual sugiere que estas arquitecturas son más efectivas para capturar patrones de corto o mediano plazo en la serie de tiempo.

En términos de desempeño específico, el modelo GRU destaca con las ventanas de 48 y 96 horas, logrando un MAPE tan bajo como 13.13%. Sin embargo, el menor valor de MAPE fue alcanzado por el modelo LSTM RNN con una ventana de 48 horas y una configuración de 40 unidades, una tasa de aprendizaje de 0.01 y un tamaño de lote de 100, logrando un MAPE de 13.1%.

Además de la tabla anterior, La Figura 8 muestra el MAPE, MAE y RMSE para cada modelo y tamaño de ventana, proporcionando una comparación visual del desempeño entre arquitecturas y configuraciones de ventana, y resaltando cómo el ajuste de la ventana afecta la precisión en la serie temporal.

Figura 8. Comportamiento de las métricas para las mejores iteraciones del experimento 'Todos los features'.

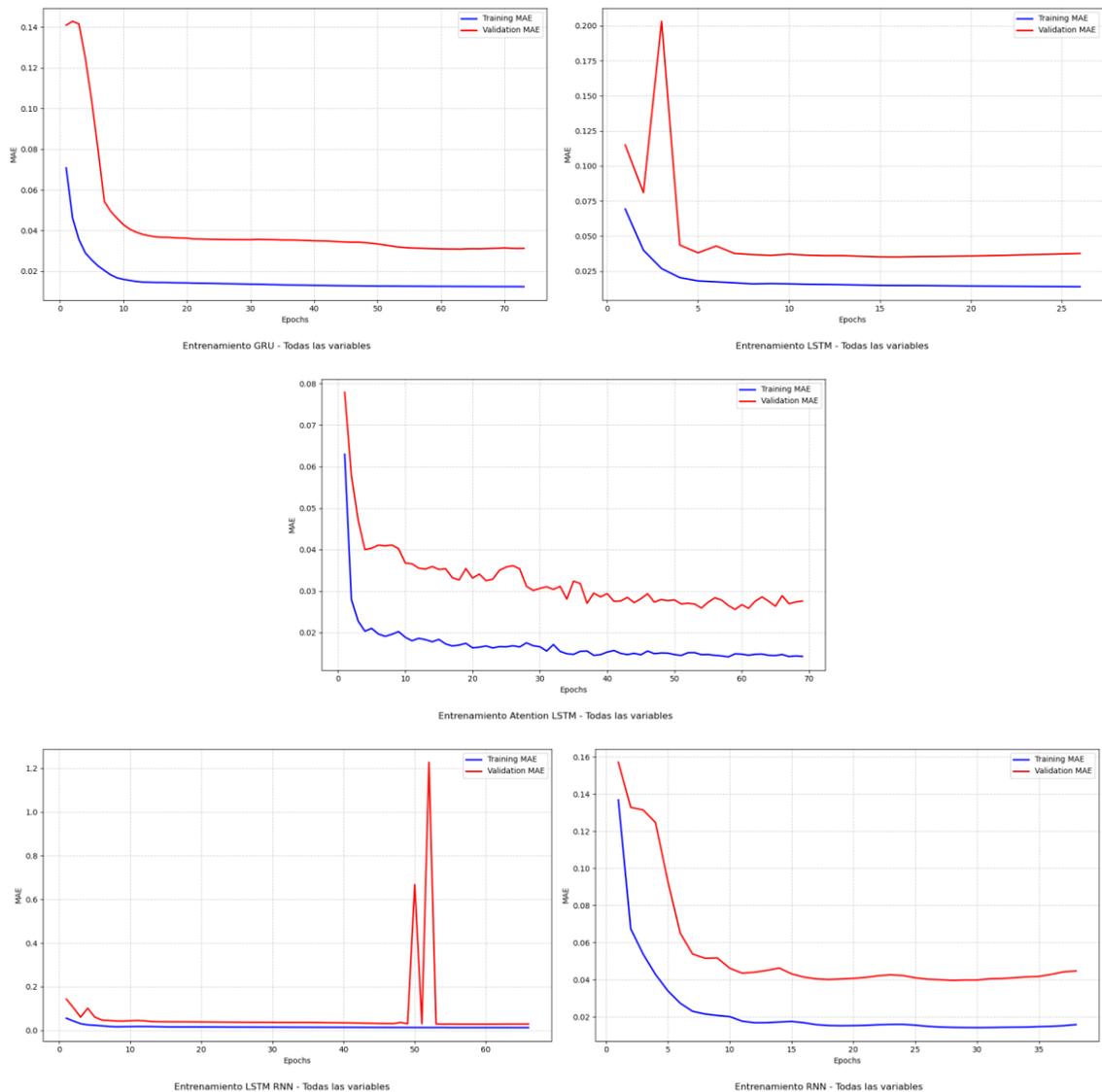


Aunque el modelo LSTM RNN con ventana de 48 horas destaca como el mejor en términos de MAPE, no presenta el mejor desempeño en cuanto a las métricas MAE y RMSE. En particular, el modelo GRU, con una ventana de 48 horas, logra un MAE 47.02 [COP/kWh], siendo este el menor del experimento, mientras que el modelo LSTM con una ventana de 96 horas obtiene un RMSE 75.62 [COP/kWh].

Si bien, la selección del modelo óptimo se fundamenta principalmente en el MAPE, es relevante destacar estos desempeños superiores en MAE y RMSE para otros modelos, ya que resaltan una precisión más estable. Estas diferencias subrayan cómo ciertas arquitecturas pueden tener ventajas en métricas específicas, lo cual debe tenerse en cuenta al evaluar la robustez y precisión del modelo en diferentes escenarios de predicción.

En la Figura 9, se ilustra como para este experimento con múltiples variables, el rendimiento y la estabilidad de cada modelo variaron significativamente basados en sus graficas de perdida para entrenamiento y validación.

Figura 9 Comparación de la Estabilidad y Generalización durante el Entrenamiento para del experimento 'Todos los features'.

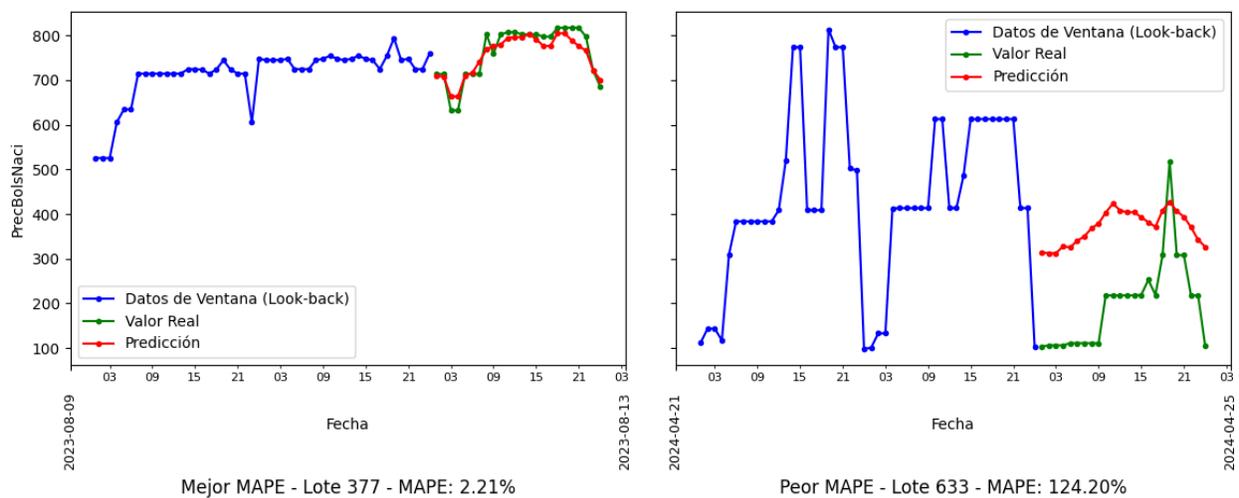


Los modelos GRU y LSTM mostraron curvas de entrenamiento y validación estables, aunque con una diferencia notable entre ambas métricas, lo que sugiere posibles limitaciones para captar completamente la complejidad de los datos. Sin embargo, ambos lograron una tendencia consistente durante el entrenamiento, lo que refleja una capacidad de estabilidad en su comportamiento, aunque no necesariamente de generalización.

Por otro lado, los modelos Attention LSTM y LSTM RNN enfrentaron mayores desafíos de estabilidad. Aunque el modelo LSTM RNN presentó un pico pronunciado en la época 50, las curvas de entrenamiento y validación se mantienen relativamente cercanas entre sí, sugiriendo una alineación en el aprendizaje a pesar de la inestabilidad temporal. En general, esta comparación resalta que los modelos GRU y LSTM muestran tendencias más estables en el entrenamiento frente a la complejidad de los datos multivariados, aunque todos enfrentan cierto grado de dificultad en su generalización.

En la Figura 10 se presentan dos ventanas de los datos de prueba, correspondientes a las mejores y peores predicciones según la métrica clave MAPE.

Figura 10. Desempeño de ventanas para experimento 'Todos los features'.



Comparación de las Predicciones: Mejor y Peor MAPE

El mejor caso muestra un MAPE de 2.21%, lo que evidencia la alta capacidad predictiva del modelo en esa ventana. En contraste, el peor caso muestra un MAPE de 124.2%, lo que indica que, en ciertas circunstancias, el modelo no logra capturar adecuadamente los patrones subyacentes de los datos, reflejando una pérdida significativa de precisión. No obstante, al observar con más

detalle, se puede notar que el modelo intenta seguir el comportamiento de la ventana anterior, en este caso, las 48 horas previas. Esto sugiere que cuando el mercado presenta comportamientos erráticos o se introduce más aleatoriedad, el modelo tiene dificultades para predecir con exactitud. Esto se hace aún más evidente cuando la peor predicción ocurre casi al final del conjunto de datos de prueba (que son 2 años en total), lo que implica que, en el largo plazo, la alta variabilidad de los precios dificulta la capacidad del modelo para hacer predicciones estables.

6.1.2. Experimento con la variable objetivo, precio de la bolsa nacional:

En este experimento, denominado 'Solo Precio', los resultados se presentan en la Tabla 7. Esta muestra cómo los modelos responden a las distintas configuraciones de ventana, donde se destaca que los mejores resultados en MAPE se obtuvieron con ventanas de 48 y 96 horas en los modelos GRU y LSTM RNN, mientras que la variabilidad en el desempeño, especialmente en las métricas MAE y RMSE, permite observar la adaptación de cada arquitectura al horizonte temporal seleccionado.

Tabla 7. Resultados para experimento 'Solo precio'.

Ventana	Modelo	Hiperparámetros			MAPE	MAE	RMSE
		units	l_r*	b_s*			
24	GRU	40	0.001	100	13.39	47.06	80.54
48	GRU	20	0.001	25	12.43	45.96	76.78
72	GRU	20	0.001	50	14.43	48.34	82.23
96	GRU	80	0.001	25	12.47	46.27	79.31
24	LSTM	120	0.001	50	13.66	47.88	79.87
48	LSTM	80	0.001	25	13.14	48.42	78.29
72	LSTM	100	0.01	25	15.29	50.66	83.15
96	LSTM	80	0.01	25	13.77	52.23	83.8
24	Attention LSTM	100	0.001	50	13.62	49.06	80.73
48	Attention LSTM	80	0.001	25	13.15	49.24	77.01
72	Attention LSTM	40	0.01	25	14.97	53.07	84.29
96	Attention LSTM	80	0.01	100	13.4	48.5	79.88
24	LSTM RNN	120	0.001	50	13.45	48.53	81.19
48	LSTM RNN	100	0.001	100	13.09	47.8	77.29
72	LSTM RNN	80	0.01	25	15.46	50.81	83.74
96	LSTM RNN	80	0.001	25	13.62	49.85	81.38
24	RNN	60	0.001	100	13.89	48.2	80.31
48	RNN	20	0.01	50	12.76	48.14	77.78
72	RNN	80	0.01	25	15.62	55.25	88.78
96	RNN	20	0.01	50	12.97	47.58	80.86

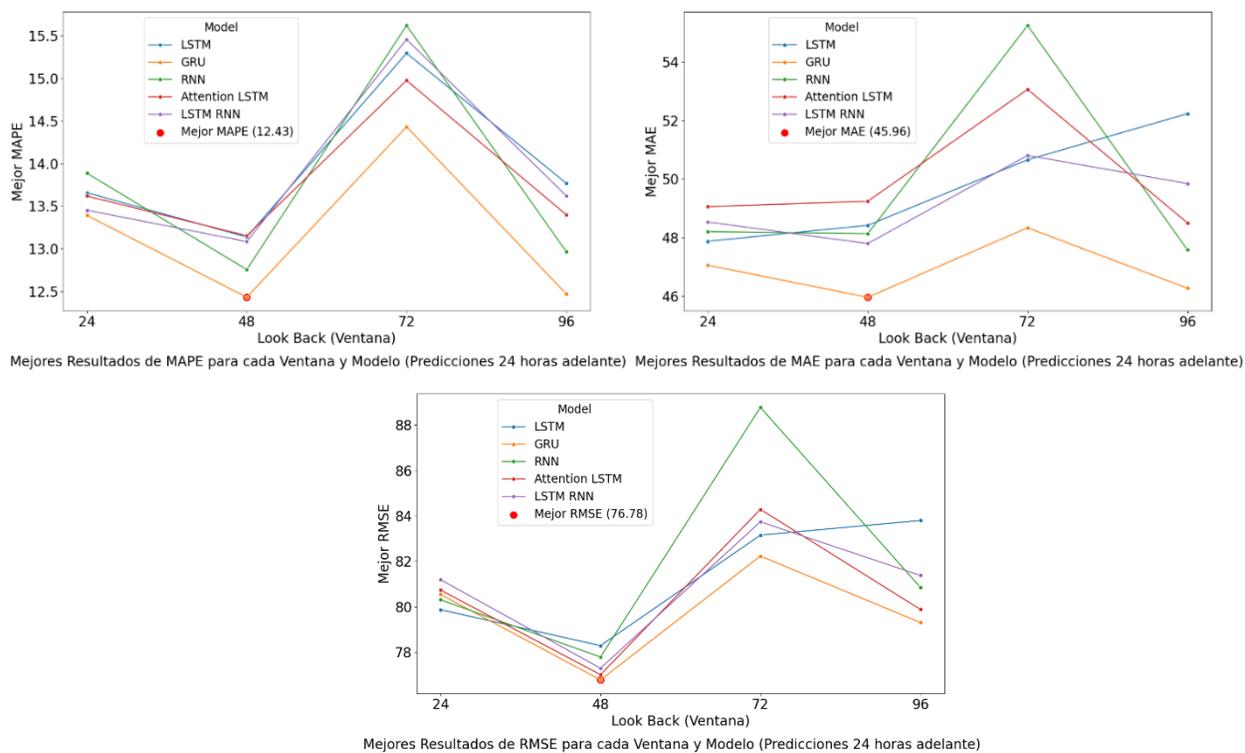
Promedio	13.73	49.14	80.86
----------	-------	-------	-------

* l_r : learning rate; b_s : batch size

En general, el modelo GRU con ventana de 48 horas y el modelo LSTM RNN con la misma configuración, destacan como las configuraciones más precisas, logrando un MAPE de 12.43% y 13.09%, respectivamente, lo cual muestra un rendimiento superior comparado con el experimento anterior donde el mejor resultado fue 13.1%. Esto sugiere que estos modelos capturan adecuadamente las tendencias y fluctuaciones del precio de la bolsa en horizontes de corto plazo, lo cual es relevante para aplicaciones donde la estabilidad de los precios tiene alta variabilidad.

La Figura 11 presenta los resultados de las métricas MAPE, MAE y RMSE para cada modelo.

Figura 11. Comportamiento de las métricas para las mejores iteraciones del experimento 'Solo precio'.



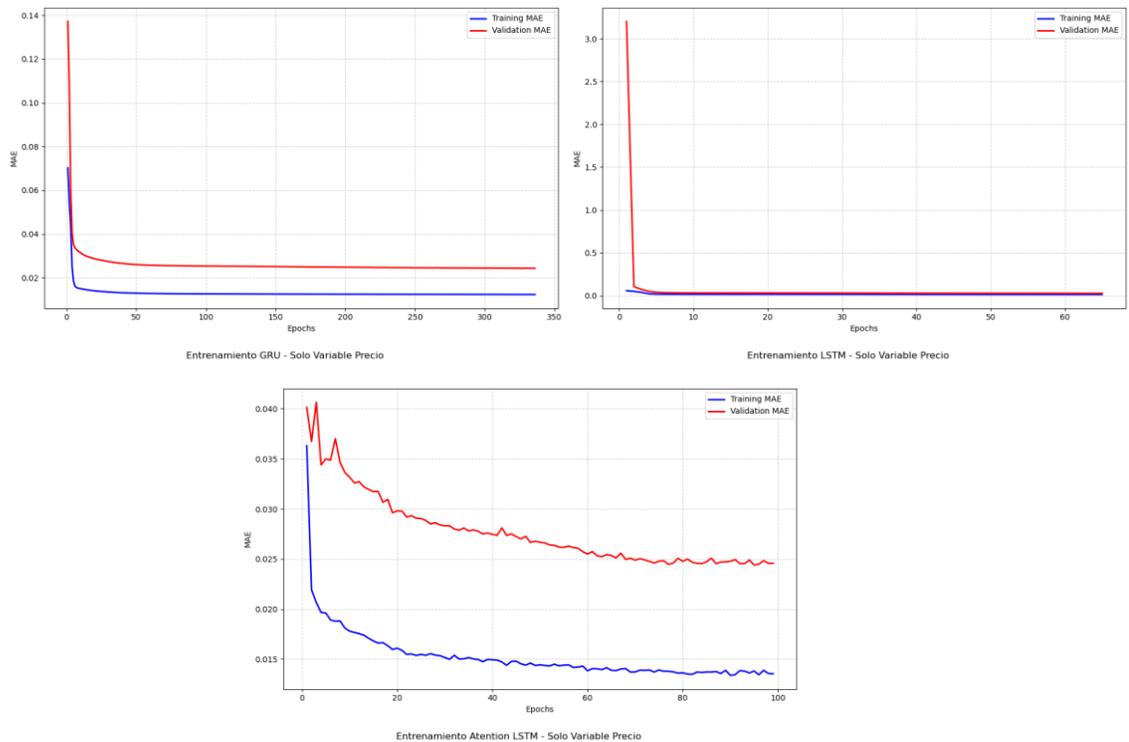
En la figura anterior se observa que, en términos de MAPE, el modelo GRU con una ventana de 48 horas alcanzó el mejor desempeño con un 12.43%, destacando su capacidad para capturar el comportamiento de la serie temporal en ese horizonte. También posee el mejor RMSE

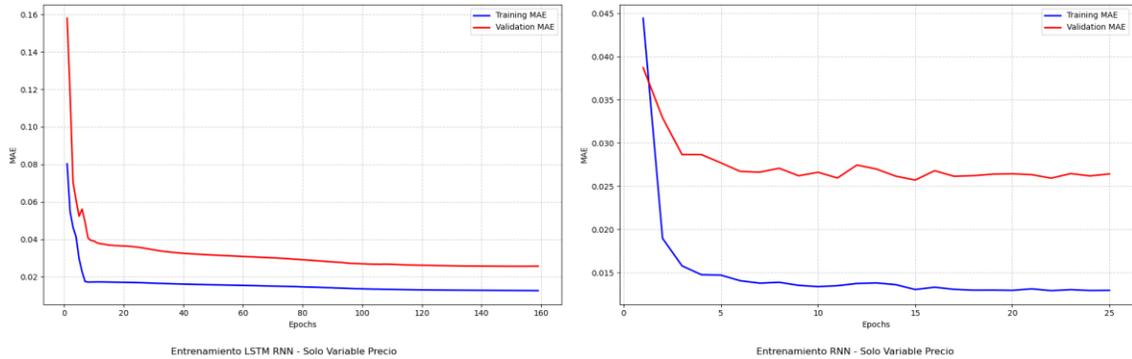
de 76.78 [COP/kWh], lo que indica una menor desviación en comparación con otros modelos, además de sobresalir con el mejor MAE, con un valor de 45.96 [COP/kWh].

Aunque el MAPE fue la métrica principal utilizada para seleccionar el modelo óptimo, los resultados adicionales en términos de MAE y RMSE también ofrecen información valiosa. Algunos modelos que no fueron elegidos debido a la prioridad otorgada al MAPE muestran diferencias mínimas en MAPE, pero presentan un mejor desempeño en cuanto a MAE y RMSE. Esto indica que, en aplicaciones donde sea esencial minimizar errores absolutos o cuadráticos y se requiera mayor estabilidad en las predicciones, podría ser beneficioso considerar alternativas en la configuración de la ventana y la arquitectura del modelo. Este ajuste podría proporcionar una ventaja significativa en contextos donde el enfoque no solo esté en la precisión porcentual, sino también en reducir los errores de manera más consistente.

Como se puede apreciar en la Figura 12, se presentan las gráficas de pérdida para los mejores modelos resaltados en la Tabla 7.

Figura 12. Comparación de la Estabilidad y Generalización durante el Entrenamiento para del experimento 'Solo precio'.



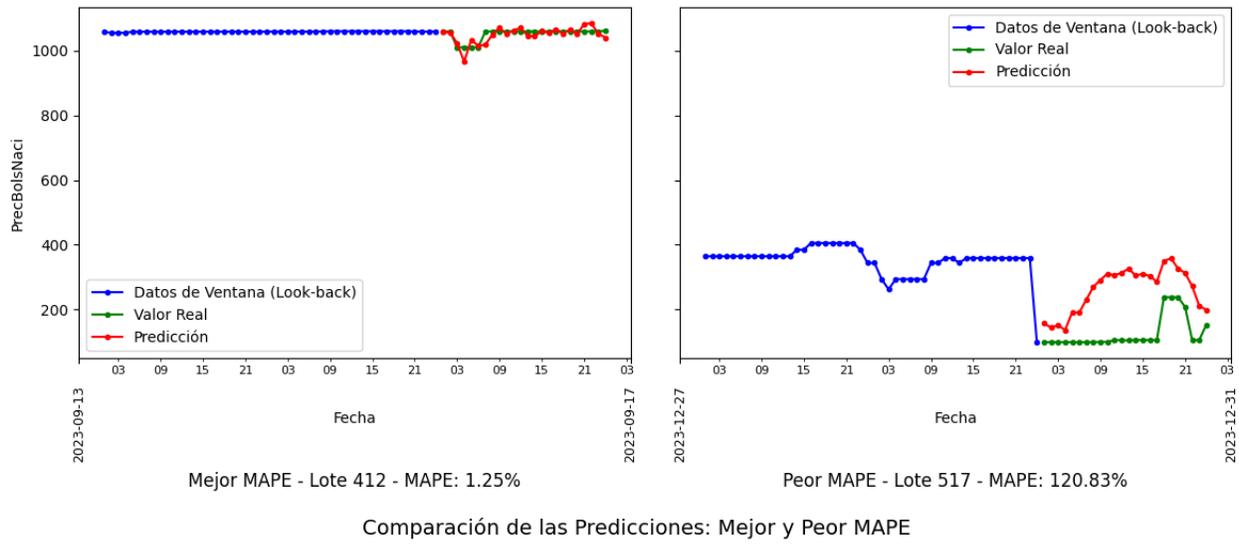


El modelo Attention LSTM muestra la convergencia más estable entre todos los modelos, con una disminución suave tanto en el MAE de entrenamiento como en el de validación a lo largo del tiempo. La curva de validación sigue de cerca la de entrenamiento con mínimas fluctuaciones, lo que sugiere que el modelo generaliza bien sin mostrar signos de sobreajuste. Esta estabilidad hace que el modelo Attention LSTM sea potencialmente la opción más confiable para capturar patrones complejos en este experimento.

En comparación, los modelos GRU, LSTM y RNN muestran mayores diferencias entre las curvas de entrenamiento y validación, lo que puede indicar problemas de generalización. Aunque el modelo LSTM también muestra buena estabilidad, los modelos GRU y RNN presentan mayor variabilidad en sus curvas de validación, lo que sugiere que podrían tener más dificultades para capturar los patrones subyacentes de los datos. En general, el modelo Attention LSTM destaca como el de mejor rendimiento para esta tarea de entrenamiento.

En la Figura 13 se muestran dos casos de predicción dentro del conjunto de prueba, el cual abarca desde el 1 de agosto de 2022 hasta el 1 de agosto de 2024.

Figura 13. Desempeño de ventanas para experimento 'Solo precio'.



En el lado izquierdo, el gráfico presenta la mejor predicción del modelo en términos de MAPE, con un valor de 1.25% para el Lote 412. En este caso, el modelo logra capturar el comportamiento del precio de la bolsa con notable fidelidad, mostrando un pronóstico muy cercano a los valores reales y reflejando una comprensión acertada de la dinámica subyacente en esa ventana.

En contraste, el gráfico del lado derecho ilustra la predicción con el mayor error (Lote 517), donde el MAPE alcanza el 120.83%. Este alto valor sugiere que el modelo no pudo ajustarse adecuadamente a las condiciones de esta ventana, reflejando una desconexión significativa con los valores reales. Se observa que la predicción intenta seguir el patrón previo, pero se ve afectada por posibles fluctuaciones abruptas y mayor volatilidad en el mercado.

Este análisis revela que, si bien el modelo es capaz de generar predicciones precisas en periodos más estables, encuentra mayores desafíos ante escenarios de cambio rápido y comportamiento errático, como aquellos observados al final de la serie temporal al igual que el experimento anterior.

6.1.3. Experimento con las Variables más Influyentes:

Para este experimento denominado 'Variables más influyentes', los resultados son expuestos en la Tabla 8:

Tabla 8. Resultados para experimento 'Variables más influyentes'.

Ventana	Modelo	Hiperparámetros			MAPE	MAE	RMSE
		units	l_r^*	b_s^*			
24	GRU	80	0.001	100	14.22	51.4	84.1
48	GRU	60	0.001	50	13.00	48.53	78.64
72	GRU	40	0.001	100	15.40	55.71	88.04
96	GRU	60	0.001	100	13.22	50.15	80.9
24	LSTM	80	0.001	100	14.37	53.1	83.62
48	LSTM	60	0.010	25	13.69	52.3	80.73
72	LSTM	20	0.001	25	15.36	53.43	85.16
96	LSTM	20	0.001	25	13.27	49.59	81.11
24	Attention LSTM	80	0.001	50	14.28	52.34	82.53
48	Attention LSTM	80	0.001	50	13.91	53.14	81.18
72	Attention LSTM	80	0.010	25	15.63	55.08	85.27
96	Attention LSTM	80	0.010	25	14.90	59.27	90.59
24	LSTM RNN	40	0.010	25	14.74	54.55	86.14
48	LSTM RNN	100	0.010	100	13.51	50.95	80.85
72	LSTM RNN	100	0.001	100	15.86	54.19	83.91
96	LSTM RNN	120	0.010	25	15.10	60.98	89.78
24	RNN	20	0.001	100	14.52	50.2	82.85
48	RNN	40	0.001	50	12.98	47.21	76.51
72	RNN	20	0.001	25	15.18	50.53	83.9
96	RNN	60	0.001	100	13.57	50.55	81.19
Promedio					14.34	52.66	83.35

* l_r : learning rate; b_s : batch size

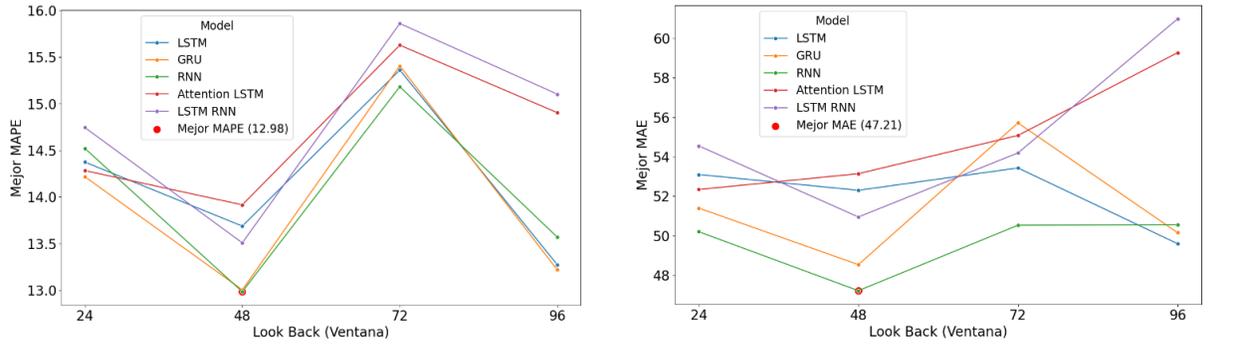
El modelo GRU, con un MAPE mínimo de 13% en la ventana de 48 horas y el LSTM con un MAPE de 13.27% en la ventana de 96 horas, presentan buenos resultados. Pero en general, el modelo RNN con un MAPE de 12.98% en la ventana de 48 horas, presenta el mejor resultado para este experimento.

La inclusión de las variables más influyentes no mejoró consistentemente el desempeño del modelo comparado con el uso exclusivo del precio. Esto sugiere que, para predicciones a corto plazo, un modelo basado en el precio puede ser igualmente efectivo o incluso superior.

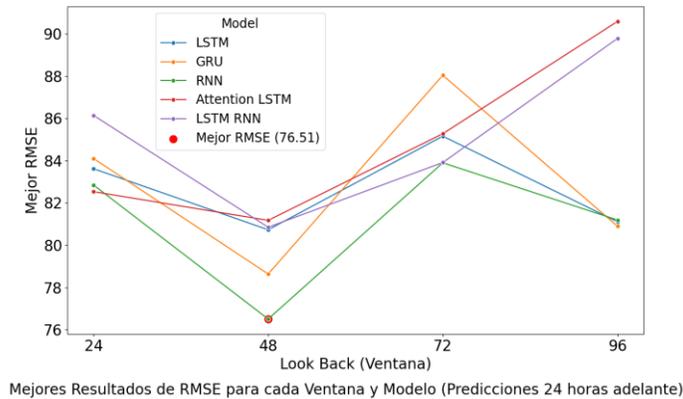
En la Figura 14 se presenta la distribución de las métricas, destacando que el modelo RNN con ventana de 48 horas logró el mejor desempeño en todas ellas. Asimismo, se observa una tendencia sólida en la que la ventana de 48 horas generalmente ofrece los mejores resultados,

mientras que la ventana de 72 horas tiende a mostrar el peor rendimiento, una tendencia que se confirma en los tres experimentos realizados.

Figura 14. Comportamiento de las métricas para las mejores iteraciones del experimento 'Variables más influyentes'.



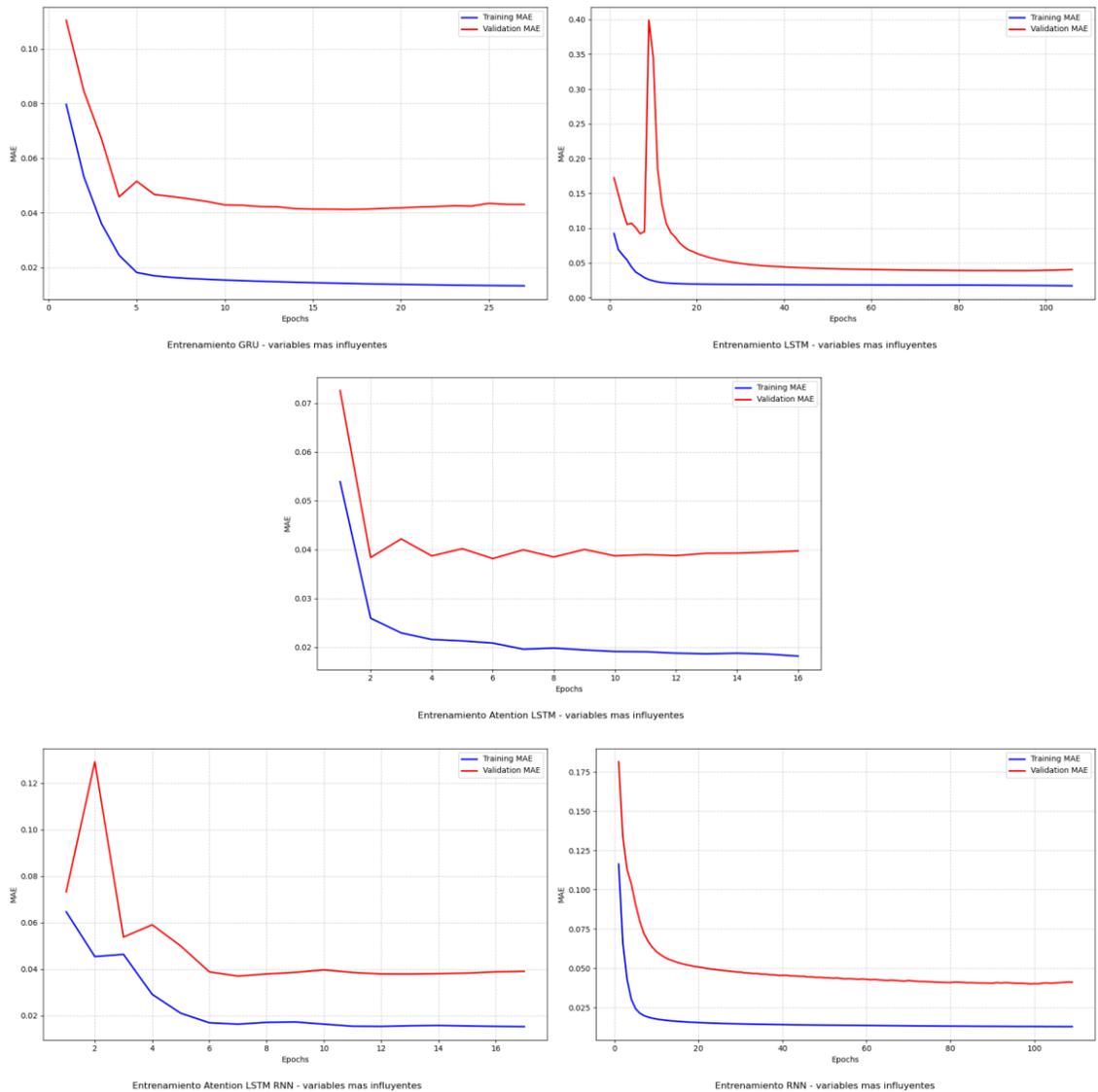
Mejores Resultados de MAPE para cada Ventana y Modelo (Predicciones 24 horas adelante) Mejores Resultados de MAE para cada Ventana y Modelo (Predicciones 24 horas adelante)



Mejores Resultados de RMSE para cada Ventana y Modelo (Predicciones 24 horas adelante)

Los resultados de entrenamiento y validación en los modelos GRU, LSTM, Attention LSTM, LSTM RNN y RNN muestran diferencias significativas en términos de estabilidad y capacidad de generalización como se muestra en la Figura 15.

Figura 15. Comparación de la Estabilidad y Generalización durante el Entrenamiento para del experimento 'Variables más influyentes'.



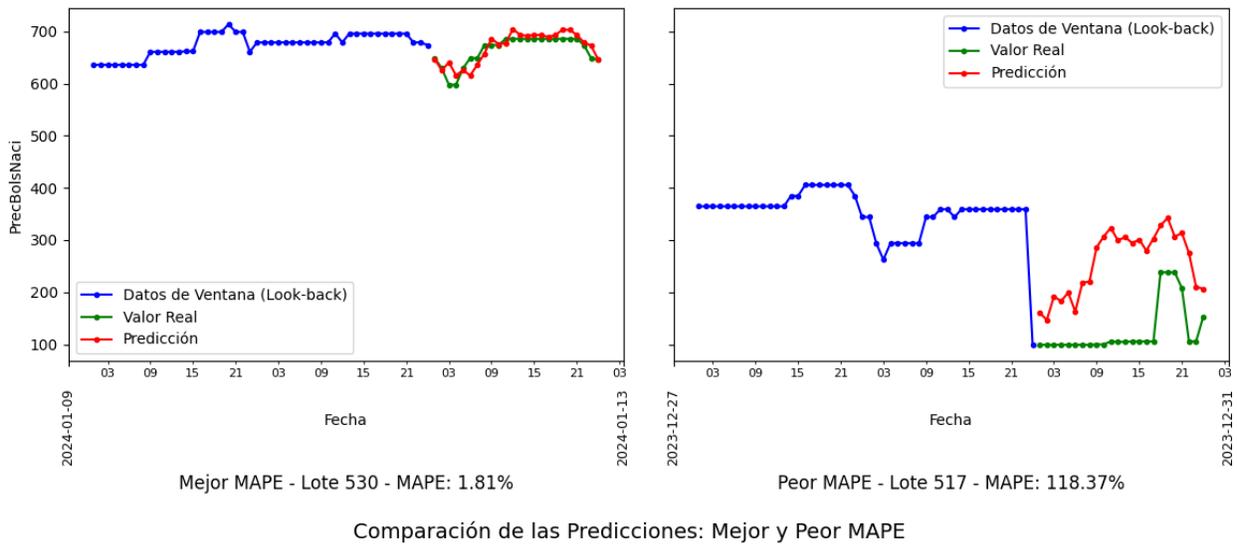
El modelo GRU presenta una curva de error de entrenamiento estable, pero con una discrepancia notable frente al error de validación, lo que indica problemas de generalización. Por otro lado, el modelo LSTM experimenta inestabilidad en las épocas intermedias, manifestada en un aumento considerable del error en validación, lo cual sugiere sobreajuste. En contraste, los modelos Attention LSTM y LSTM RNN presentan curvas de error más estables tanto en entrenamiento como en validación, y aunque muestran una mejor generalización que el modelo LSTM, aún existe una brecha en los errores de validación. Finalmente, el modelo RNN evidencia

una capacidad de generalización sólida, con una curva de error de validación que se mantiene cercana a la del entrenamiento, reflejando una alta estabilidad en el proceso de aprendizaje.

La baja diferencia entre los errores de entrenamiento y validación en el modelo RNN refuerza su capacidad de adaptarse a datos no vistos, destacándose como la opción más robusta y confiable para el problema en estudio.

En la Figura 16, se presentan dos casos de predicción dentro del conjunto de prueba.

Figura 16. Desempeño de ventanas para experimento 'Variables más influyentes'



El gráfico a la izquierda ilustra el caso con la mejor predicción del modelo, basado en el MAPE, con un valor de 1.81% para el Lote 530. Mientras que, el gráfico a la derecha muestra el caso con el mayor error (Lote 517), donde el MAPE alcanza el 118.37%. Este contraste reafirma el comportamiento observado previamente: cuando el modelo se enfrenta a ventanas de datos significativamente alejadas de los utilizados para el entrenamiento, especialmente en escenarios de cambios abruptos en el precio de la energía, su capacidad de predicción disminuye. No obstante, el modelo mantiene un buen desempeño en términos generales, logrando capturar patrones macro dentro de las ventanas de predicción. Como se ha concluido anteriormente, aunque el modelo muestra un desempeño robusto en condiciones de estabilidad, su efectividad se ve reducida cuando se enfrenta a comportamientos irregulares y rápidos cambios en los datos.

6.2. Evaluación cualitativa

En esta sección se presenta un análisis cualitativo de los resultados obtenidos en los tres experimentos realizados para predecir el precio de la energía. Se exploraron diferentes modelos de redes neuronales recurrentes y variantes, utilizando ventanas de entrada de 24, 48, 72 y 96 horas, y se evaluó el impacto de incluir distintas variables de entrada.

6.2.3 Desempeño General de los Modelos

Los resultados indican que la ventana de 48 horas consistentemente ofreció el mejor desempeño en términos de las métricas evaluadas (MAPE, MAE y RMSE). Esto sugiere que este tamaño de ventana es adecuado para capturar patrones relevantes en la serie temporal del precio de la energía.

Los modelos basados únicamente en la variable objetivo (precio de la bolsa nacional) lograron un desempeño ligeramente superior en comparación con aquellos que incluían todas las variables o las más influyentes. Esto indica que, para horizontes de predicción cortos, el historial del precio es la información más relevante, y la inclusión de variables adicionales no aporta mejoras significativas en las arquitecturas usadas.

6.2.3 Impacto de la Inclusión de Variables Adicionales

La adición de más variables al modelo no resultó en una mejora en el desempeño de las predicciones. Una posible explicación es que, en ventanas de tiempo tan cortas como 24 a 96 horas, las variables adicionales no presentan una correlación fuerte con la variable objetivo. Esto limita su capacidad para aportar información útil al modelo y puede introducir ruido que afecta negativamente la precisión de las predicciones.

6.2.3 Estabilidad y Generalización de los Modelos

Los modelos GRU y LSTM demostraron ser efectivos en capturar patrones a mediano plazo, especialmente con ventanas de 48 horas. Sin embargo, presentaron diferencias entre los errores de entrenamiento y validación, lo que sugiere limitaciones en su capacidad de generalización.

El modelo Attention LSTM mostró mayor estabilidad y una mejor generalización en algunos experimentos, lo que lo convierte en una alternativa prometedora. Sin embargo, su desempeño no superó significativamente al de los modelos más simples cuando se usó únicamente la variable objetivo.

6.2.4 Limitaciones ante Volatilidad del Mercado

Una observación común en los experimentos es la dificultad de los modelos para predecir con precisión en condiciones de alta volatilidad y cambios abruptos en el mercado, especialmente hacia el final del periodo de prueba. Esto indica que, aunque los modelos son efectivos en capturar tendencias generales, enfrentan desafíos para adaptarse a comportamientos cambiantes del mercado.

6.3. Consideraciones de producción

La implementación de los modelos de predicción del precio de la energía en un entorno de producción requiere una planificación detallada para garantizar su eficacia y sostenibilidad. A continuación, se presentan las principales consideraciones técnicas que deben abordarse para asegurar un despliegue exitoso.

6.3.1 Monitoreo del Desempeño del Modelo

Establecer mecanismos de monitoreo continuo es esencial para evaluar en tiempo real el rendimiento del modelo. Esto incluye el seguimiento de métricas clave como MAPE, MAE y RMSE, lo que permite detectar desviaciones en el desempeño que podrían indicar la necesidad de reentrenamiento o ajuste de hiperparámetros. La implementación de sistemas de alerta temprana facilita una respuesta oportuna ante posibles degradaciones en la precisión predictiva.

6.3.2 Integración con Flujos de Datos en Tiempo Real

La integración con sistemas de *streaming* de datos puede resultar fundamental para que el modelo reciba información actualizada de manera inmediata. Tecnologías como Apache Kafka o Apache Flink pueden utilizarse para gestionar grandes volúmenes de datos en tiempo real, asegurando baja latencia y alta disponibilidad. Una arquitectura orientada a eventos permite que el modelo procesa y analice datos al instante, mejorando la relevancia y actualidad de las predicciones.

6.3.3 Infraestructura en la Nube y Escalabilidad

El despliegue en plataformas de computación en la nube ofrece ventajas en términos de escalabilidad, flexibilidad y costo. Servicios como AWS SageMaker, Google Cloud AI Platform o Azure Machine Learning [30], [31], [32] proporcionan entornos optimizados para modelos de aprendizaje automático, facilitando el entrenamiento, despliegue y gestión. La utilización de contenedores Docker y orquestadores como Kubernetes podría permitir una escalabilidad

horizontal y una gestión eficiente de los recursos, adaptándose a las futuras demandas del mercado[33].

6.3.4 Automatización y Orquestación del Ciclo de Vida del Modelo

Adoptar prácticas de MLOps es clave para automatizar y gestionar eficientemente el ciclo de vida del modelo. Herramientas como MLflow, Kubeflow o TensorFlow Extended (TFX) [34] pueden emplearse para orquestar el entrenamiento, validación, despliegue y monitoreo, facilitando la integración y el despliegue continuos (CI/CD). Esto garantiza que las actualizaciones del modelo se implementen de manera estable y reproducible.

6.3.5 Gestión de Versiones y Reentrenamiento

El control de versiones de modelos y conjuntos de datos es esencial para mantener la trazabilidad y reproducibilidad. Implementar estrategias de versionamiento permite comparar diferentes iteraciones del modelo y seleccionar la más adecuada. Además, programar reentrenamientos periódicos asegura que el modelo incorpore nuevos datos y se adapte a posibles cambios en los patrones del mercado, manteniendo su relevancia y desempeño a lo largo del tiempo.

6.3.6 Optimización de Recursos y Costos

La eficiencia en el uso de recursos computacionales es fundamental para minimizar costos operativos. Esto implica dimensionar adecuadamente la infraestructura, aprovechar instancias de cómputo escalables y utilizar servicios sin servidor (*serverless*) cuando sea apropiado. El monitoreo continuo del consumo de recursos y el ajuste dinámico de la capacidad permiten optimizar el rendimiento económico sin comprometer la calidad del servicio.

6.3.7 Interoperabilidad y Accesibilidad

Exponer el modelo mediante APIs estandarizadas facilita su integración con otros sistemas y aplicaciones. El diseño de interfaces RESTful o gRPC, junto con una documentación clara y detallada, permite que desarrolladores y usuarios consuman las predicciones del modelo de manera sencilla y eficiente. Adoptar estándares abiertos promueve la interoperabilidad y evita dependencias tecnológicas.

6.3.8 Manejo de la Volatilidad y Adaptabilidad

Dado que los modelos enfrentan desafíos para predecir en escenarios de alta volatilidad, es pertinente incorporar técnicas que mejoren su robustez. Esto puede incluir la utilización de modelos híbridos que combinen enfoques estadísticos y de ML, la incorporación de variables

exógenas relevantes (como indicadores económicos o climáticos) y la implementación de técnicas de regularización y ajuste de hiperparámetros que mejoren la generalización del modelo.

7. Conclusiones

En conclusión, el presente estudio desarrolló modelos predictivos basados en aprendizaje profundo para anticipar el precio de la energía en Colombia, utilizando arquitecturas como LSTM, GRU, RNN y variantes de estas. Los experimentos realizados con diferentes conjuntos de variables y tamaños de ventana temporal (24, 48, 72 y 96 horas) revelaron que una ventana de 48 horas proporcionó consistentemente el mejor desempeño en las métricas evaluadas. El modelo más destacado, que empleó únicamente la variable del precio histórico de la bolsa nacional, alcanzó un MAPE (Error Porcentual Absoluto Medio) de 12,43%, un MAE (Error Absoluto Medio) de 45,96 COP/kWh y un RMSE (Raíz del Error Cuadrático Medio) de 76,78 COP/kWh. Estos resultados indican que el modelo tiene una capacidad adecuada para predecir tendencias generales del precio de la energía a corto plazo, como se menciona en ‘Métricas de desempeño’, donde un MAPE entre el 10%-20%, es aceptable y práctico para labores operativas.

Sin embargo, se observó que los modelos presentan dificultades para predecir con precisión durante períodos de alta volatilidad y cambios repentinos en el mercado, especialmente hacia el final del período de prueba. Esto sugiere que, aunque los modelos son efectivos para capturar patrones en condiciones de relativa estabilidad, su capacidad predictiva disminuye en escenarios de mayor incertidumbre. El análisis identificó que la variable más influyente en la predicción del precio de la energía es el propio precio histórico de la bolsa nacional. Aunque se consideraron otras variables como el volumen útil diario de energía, el índice del Fenómeno de El Niño y la capacidad útil diaria de energía, su inclusión en el modelo no mejoró significativamente el desempeño en horizontes de predicción cortos. Esto indica que, para predicciones a corto plazo, el precio histórico encapsula la información más relevante, y las variables adicionales podrían no aportar un valor predictivo sustancial en este contexto.

Los resultados obtenidos tienen implicaciones significativas para la industria energética en Colombia. La capacidad de predecir el precio de la energía con razonable precisión a corto plazo puede ayudar a las empresas y reguladores a mejorar la planificación de costos y la gestión financiera. Esto permite optimizar la operación de los recursos energéticos, ajustando la oferta según las expectativas de demanda y precios, desarrollando estrategias más eficientes para la gestión de la oferta y la demanda. Aunque los modelos presentan limitaciones en condiciones de alta volatilidad, su aplicación puede ser valiosa en períodos de estabilidad relativa del mercado,

contribuyendo a aumentar la resiliencia del sistema energético frente a las incertidumbres del mercado.

8. Recomendaciones

Para mejorar la precisión de los modelos predictivos en contextos volátiles, es crucial explorar enfoques avanzados que maximicen la integración de variables de gran relevancia, como indicadores económicos, climáticos y sociales, que influyen directamente en los precios de la energía. La implementación de modelos híbridos que combinen técnicas de ML y métodos estadísticos tradicionales resulta particularmente efectiva, ya que estos modelos híbridos pueden captar tanto patrones comunes como eventos o cambios inesperados en el mercado.

Además, para asegurar que el modelo se adapte a las condiciones cambiantes del mercado, es recomendable implementar mecanismos de adaptación continua. Esto se logra mediante reentrenamientos periódicos con datos actualizados y la aplicación de técnicas avanzadas de gestión de datos, como el aprendizaje en línea o incremental. Estas técnicas permiten que el modelo se ajuste de manera dinámica a nuevas realidades sin la necesidad de una reestructuración total. Asimismo, técnicas de regularización y ajustes de hiperparámetros más avanzados pueden optimizar la generalización y reducir el riesgo de sobreajuste, mejorando la robustez del modelo.

Un análisis profundo de los períodos de alta volatilidad es también esencial para identificar las limitaciones específicas del modelo y formular estrategias que mitiguen sus deficiencias. La colaboración con especialistas en mercados energéticos y en análisis financiero puede brindar perspectivas valiosas, ayudando a refinar las predicciones del modelo y a mejorar su aplicabilidad en escenarios complejos.

Este enfoque integral contribuirá significativamente a la precisión y adaptabilidad de los modelos predictivos en mercados energéticos volátiles.

References

- [1] B. A. Bridge, D. Adhikari, and M. Fontenla, “Electricity, income, and quality of life,” *Soc Sci J*, vol. 53, no. 1, pp. 33–39, 2016, doi: <https://doi.org/10.1016/j.soscij.2014.12.009>.
- [2] Minenergía, “Colombia apuesta a la aplicación de un estándar mundial de sostenibilidad con el apoyo de la Cooperación Económica y Desarrollo (SECO) de Suiza y la Asociación Internacional de Energía hidroeléctrica (IHA),” Minenergía. Accessed: Apr. 02, 2024. [Online]. Available: <https://www.minenergia.gov.co/es/sala-de-prensa/noticias-index/colombia-como-uno-de-los-líderes-latinoamericanos-en-energía-hidroeléctrica-le-apuesta-a-la-aplicación-de-un-estándar-mundial-de-sostenibilidad-con-el-apoyo-de-la-cooperación-económica-y-desar>
- [3] “Así funciona la bolsa de energía en Colombia.” Accessed: Nov. 11, 2024. [Online]. Available: <https://www.valoraanalitik.com/asi-funciona-bolsa-de-energia-colombia/>
- [4] “Preguntas frecuentes.” Accessed: Nov. 11, 2024. [Online]. Available: <https://www.xm.com.co/herramientas/preguntas-frecuentes>
- [5] J. T. Fasullo, B. L. Otto-Bliesner, and S. Stevenson, “ENSO’s Changing Influence on Temperature, Precipitation, and Wildfire in a Warming Climate,” *Geophys Res Lett*, vol. 45, no. 17, pp. 9216–9225, Sep. 2018, doi: 10.1029/2018GL079022.
- [6] “¿Qué es el fenómeno de El Niño y La Niña? | National Geographic.” Accessed: Nov. 11, 2024. [Online]. Available: <https://www.nationalgeographic.es/medio-ambiente/que-es-el-fenomeno-de-el-nino-y-la-nina>
- [7] “Precios de electricidad subieron 19,51% anual y la Costa Caribe es donde más crece.” Accessed: May 26, 2024. [Online]. Available: <https://www.larepublica.co/economia/precios-de-electricidad-subieron-19-51-anual-y-la-costa-caribe-es-donde-mas-crece-3566130/>
- [8] “XM Administradores del mercado eléctrico.” Accessed: Nov. 19, 2024. [Online]. Available: <https://www.xm.com.co/>
- [9] S. A. Nabavi, S. Mohammadi, N. H. Motlagh, S. Tarkoma, and P. Geyer, “Deep learning modeling in electricity load forecasting: Improved accuracy by combining DWT and LSTM,” *Energy Reports*, vol. 12, pp. 2873–2900, 2024, doi: <https://doi.org/10.1016/j.egy.2024.08.070>.

- [10] A. Brusaferrri, M. Matteucci, P. Portolani, and A. Vitali, “Bayesian deep learning based method for probabilistic forecast of day-ahead electricity prices,” *Appl Energy*, vol. 250, pp. 1158–1175, 2019, doi: <https://doi.org/10.1016/j.apenergy.2019.05.068>.
- [11] P. Asesor and S. Takahashi, “MODELO DE PREDICCIÓN DE GENERACIÓN FOTOVOLTAICA EN COLOMBIA BASADO EN REDES NEURONALES,” 2024.
- [12] W. Mauricio Villa Álvaro Jaramillo Duque, “PREDICCIÓN DEL PRECIO DE LA ENERGÍA ELÉCTRICA EN COLOMBIA CON HORIZONTE DE MUY CORTO PLAZO UTILIZANDO MODELOS DE DEEP.” [Online]. Available: www.udea.edu.co
- [13] “Ministerio de Minas y Energía.” Accessed: Nov. 10, 2024. [Online]. Available: <https://www.minenergia.gov.co/en/>
- [14] R. Weron, “Electricity price forecasting: A review of the state-of-the-art with a look into the future,” *Int J Forecast*, vol. 30, no. 4, pp. 1030–1081, Oct. 2014, doi: [10.1016/J.IJFORECAST.2014.08.008](https://doi.org/10.1016/J.IJFORECAST.2014.08.008).
- [15] J. Lago, F. De Ridder, and B. De Schutter, “Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms,” *Appl Energy*, vol. 221, pp. 386–405, Jul. 2018, doi: [10.1016/J.APENERGY.2018.02.069](https://doi.org/10.1016/J.APENERGY.2018.02.069).
- [16] “(PDF) Electricity Price Forecasting via Statistical and Deep Learning Approaches: the German case.” Accessed: Nov. 10, 2024. [Online]. Available: https://www.researchgate.net/publication/366090440_Electricity_Price_Forecasting_via_Statistical_and_Deep_Learning_Approaches_the_German_case
- [17] Minenergía, “Colombia apuesta a la aplicación de un estándar mundial de sostenibilidad con el apoyo de la Cooperación Económica y Desarrollo (SECO) de Suiza y la Asociación Internacional de Energía hidroeléctrica (IHA),” Minenergía. Accessed: Apr. 03, 2024. [Online]. Available: <https://www.minenergia.gov.co/es/sala-de-prensa/noticias-index/colombia-como-uno-de-los-líderes-latinoamericanos-en-energía-hidroeléctrica-le-apuesta-a-la-aplicación-de-un-estándar-mundial-de-sostenibilidad-con-el-apoyo-de-la-cooperación-económica-y-desar>
- [18] J. J. Montaña Moreno, A. Palmer Pol, A. Sesé Abad, and B. Cajal Blasco, “El índice R-MAPE como medida resistente del ajuste en la previsión,” *Psicothema*, vol. 25, no. 4, pp. 500–506, 2013, doi: [10.7334/psicothema2013.23](https://doi.org/10.7334/psicothema2013.23).

- [19] “GitHub - EquipoAnaliticaXM/API_XM.” Accessed: Nov. 10, 2024. [Online]. Available: https://github.com/EquipoAnaliticaXM/API_XM
- [20] “Climate Prediction Center - ONI.” Accessed: Nov. 10, 2024. [Online]. Available: https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php
- [21] “CIFRAS RELEVANTES DEL MERCADO DE ENERGÍA MAYORISTA (MEM) EN FEBRERO 2023: Transacciones Variación Mensual % Variación Anual % feb-23 ene-23 feb-22.”
- [22] “ARIMA vs. SARIMA vs. SARIMAX. Exploring Time Series Forecasting... | by sophiamsac | Medium.” Accessed: Nov. 10, 2024. [Online]. Available: <https://medium.com/@sophiamsac/arima-vs-sarima-vs-sarimax-03dd04fc7c66>
- [23] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, “Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network,” *IEEE Trans Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019, doi: 10.1109/TSG.2017.2753802.
- [24] S. M. Al-Selwi *et al.*, “RNN-LSTM: From applications to modeling techniques and beyond—Systematic review,” *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 5, p. 102068, 2024, doi: <https://doi.org/10.1016/j.jksuci.2024.102068>.
- [25] X. He, W. Zhao, Z. Gao, L. Zhang, Q. Zhang, and X. Li, “Short-term load forecasting by GRU neural network and DDPG algorithm for adaptive optimization of hyperparameters,” *Electric Power Systems Research*, vol. 238, p. 111119, 2025, doi: <https://doi.org/10.1016/j.epsr.2024.111119>.
- [26] A. Meng *et al.*, “Electricity price forecasting with high penetration of renewable energy using attention-based LSTM network trained by crisscross optimization,” *Energy*, vol. 254, p. 124212, 2022, doi: <https://doi.org/10.1016/j.energy.2022.124212>.
- [27] “MAE, MSE, RMSE, and F1 score in Time Series Forecasting | by Ottavio Calzone | Medium.” Accessed: Nov. 10, 2024. [Online]. Available: <https://medium.com/@ottaviocalzone/mae-mse-rmse-and-f1-score-in-time-series-forecasting-d04021ffa7ce>
- [28] “Las métricas de IPM Insights incluyen MAPE (Error porcentual absoluto medio).” Accessed: Nov. 10, 2024. [Online]. Available: https://docs.oracle.com/cloud/help/es/pbcs_common/PFUSU/insights_metrics_MAPE.htm#PFUSU-GUID-C33B0F01-83E9-468B-B96C-413A12882334

- [29] W. Mauricio Villa Álvaro Jaramillo Duque, “PREDICCIÓN DEL PRECIO DE LA ENERGÍA ELÉCTRICA EN COLOMBIA CON HORIZONTE DE MUY CORTO PLAZO UTILIZANDO MODELOS DE DEEP.”
- [30] “Azure Machine Learning - ML as a Service | Microsoft Azure.” Accessed: Nov. 11, 2024. [Online]. Available: <https://azure.microsoft.com/en-us/products/machine-learning/>
- [31] “AI & Machine Learning Products & Services | Google Cloud.” Accessed: Nov. 11, 2024. [Online]. Available: <https://cloud.google.com/products/ai>
- [32] “Machine Learning Service - Amazon SageMaker - AWS.” Accessed: Nov. 11, 2024. [Online]. Available: <https://aws.amazon.com/sagemaker/>
- [33] “Overview | Kubernetes.” Accessed: Nov. 11, 2024. [Online]. Available: <https://kubernetes.io/docs/concepts/overview/>
- [34] “MLOps: A deep dive into TFX, Kubeflow, ZenML, and MLflow | by Ransford Addai | Medium.” Accessed: Nov. 11, 2024. [Online]. Available: <https://medium.com/@ransford.addai/mlops-a-deep-dive-into-tfx-kubeflow-zenml-and-mlflow-847f57c47b03>