



Análisis de Series Temporales y Predicción del Precio del kWh usando una serie de XM

Autores: Sebastian Castaño, Katherine Maldonado, Jhonatan Jurado

Fecha: febrero 7 de 2025

El Mercado Energético Colombiano

Contexto Empresarial

XM S.A. E.S.P.: Operación del SIN y administración del MEM en Colombia.



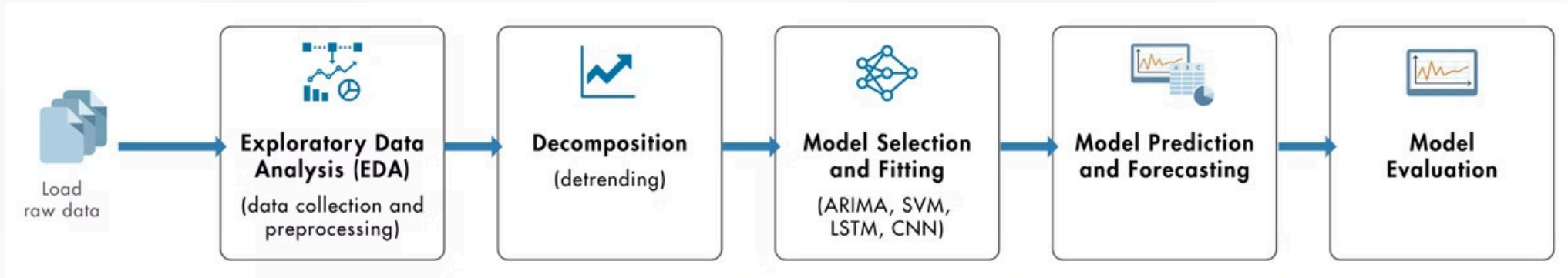
Factores Clave en la Determinación del Precio

- Oferta: Recursos de generación, disponibilidad de recursos.
- Demanda: Actividad económica, exportación de energía.
- Externos: Precio combustibles, políticas, geopolítica.

Objetivo



Comprender las fluctuaciones de la serie temporal y desarrollar un modelo predictivo del precio del kWh.



Metodología: Análisis de Series Temporales

1

EDA

Preprocesamiento y visualización de datos.

2

Descomposición

Separación en tendencia, estacionalidad, ciclos.

3

Modelo LSTM y XGBoots

Selección, ajuste y entrenamiento del modelo predictivo.

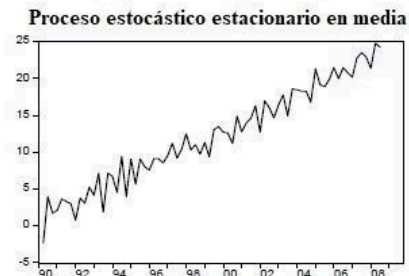
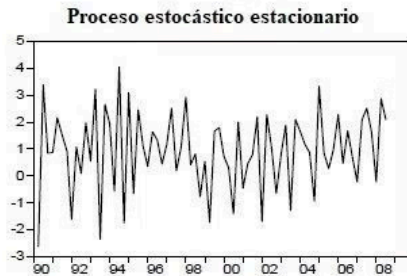
4

Evaluación y Pronósticos (Forecasting)

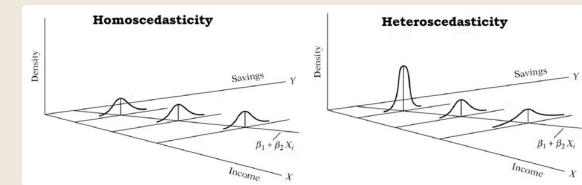
Validación con métricas y técnicas de validación cruzada.

Análisis Exploratorio

Identificar la naturaleza de la variable para el método de procesamiento y selección del modelo.



- **Heteroscedasticidad:** Varianza variable con el tiempo.



- **Grado de estocasticidad**
- **Estacionalidad:** Patrones que varían en el tiempo.
- **Autocorrelaciones**

Herramientas

Uso de diferentes gráficas como distribuciones de la serie, presentación de la serie en diferentes pasos de tiempo, diagrama de bigotes, etc.

Descomposición y Estacionalidad



Tendencia

Evolución general a largo plazo.



Estacionalidad

Patrones repetitivos en el tiempo.



Ruido

Variabilidad aleatoria no explicada.

Selección del Modelo: Redes LSTM y XGBoots

RNNs vs. LSTMs

RNNs: Limitaciones con datos secuenciales.

LSTMs: Memoria a largo plazo, modelo no lineal.

¿Por qué LSTM?

- Recuerda información a largo plazo.
- Robusto frente a ruido y tendencias.
- Maneja relaciones complejas.

¿Por qué XGBoots?

- Maneja grandes volúmenes de datos.
- Poco sensible ante variaciones.
- Altamente eficiente y puede entrenarse rápidamente.
- Menos costoso computacionalmente.

Funcionamiento de una Celda LSTM

1 Compuerta de Entrada

Actualiza la información en la memoria.

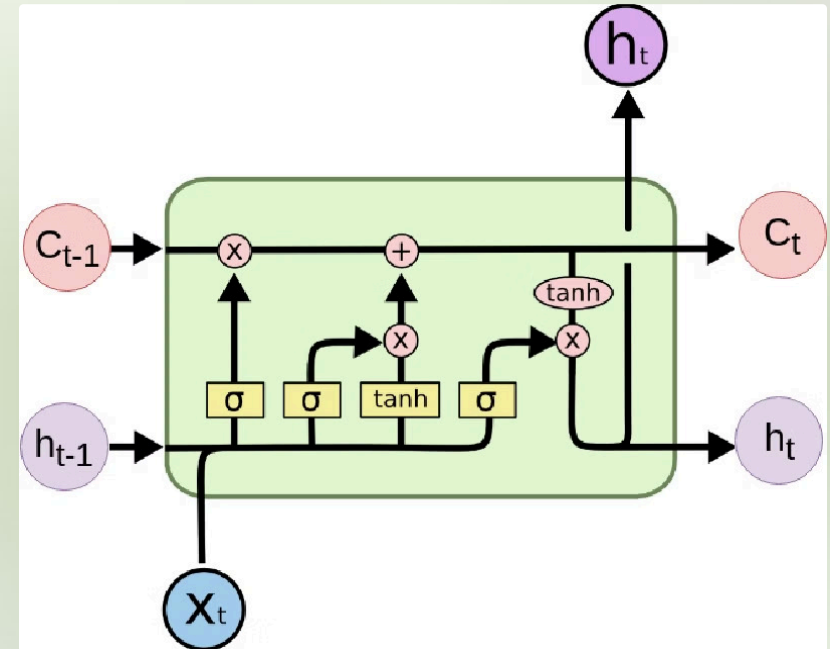
2 Compuerta de Olvido

Elimina información obsoleta.

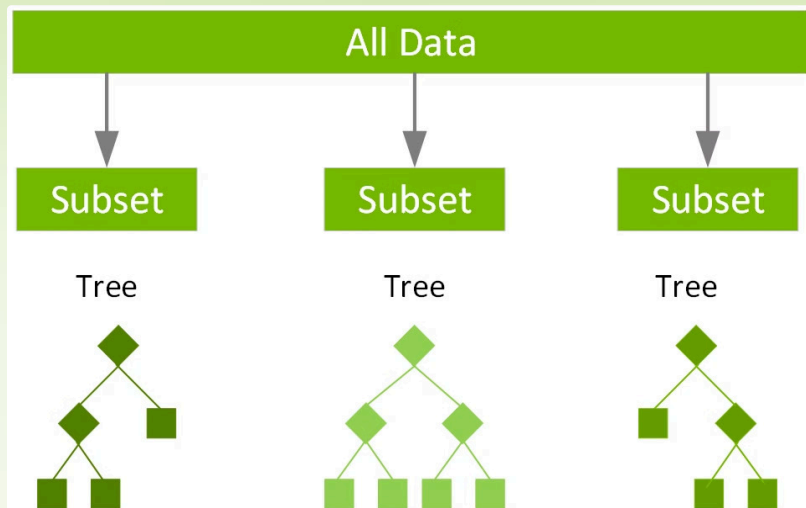
3 Compuerta de Salida

Determina la salida en cada momento.

Tres compuertas clave regulan el flujo de información dentro de la celda de memoria.



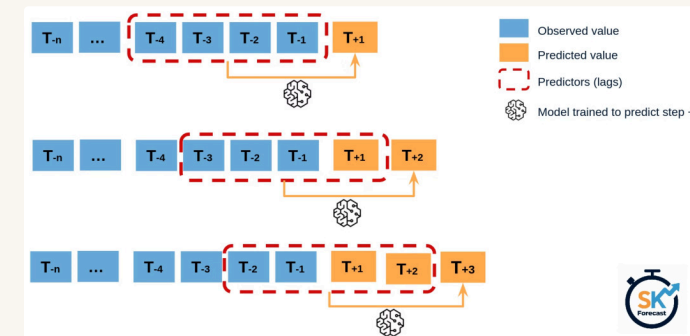
Procesamiento y predicción de la Serie



1

Transformación y estratificación

Serie en secuencias de entrada y salida.



2

Entrenamiento

Capas LSTM retienen información relevante.

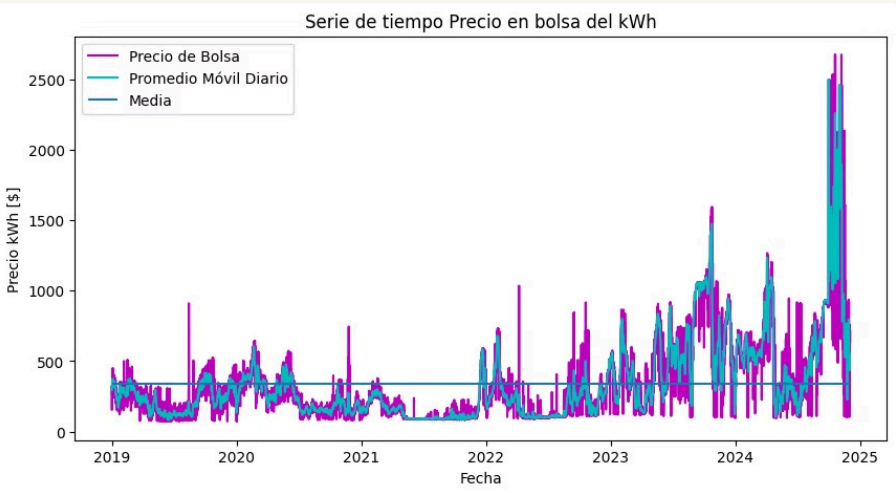
3

Predicción y pronósticos

Basada en patrones aprendidos.

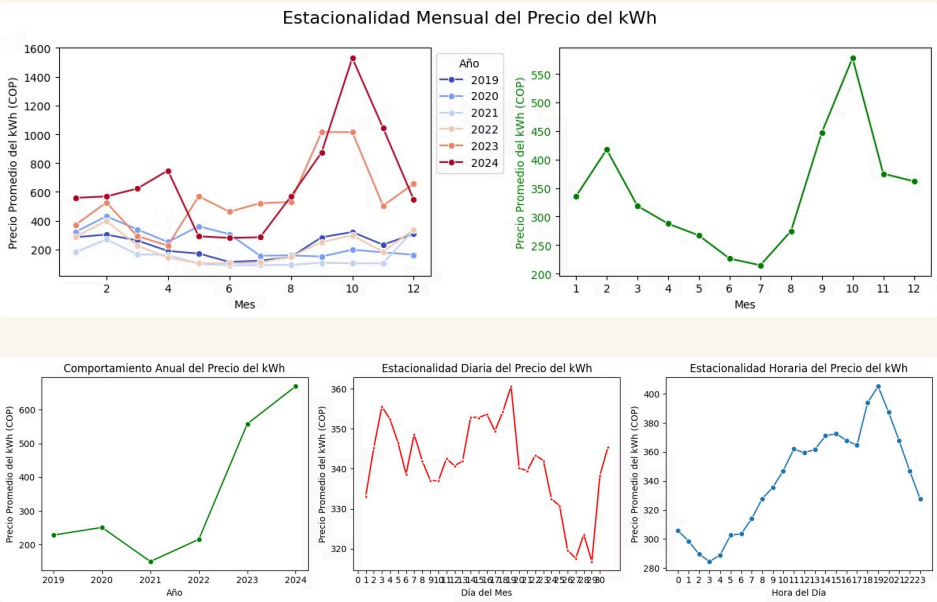
Resultados

1



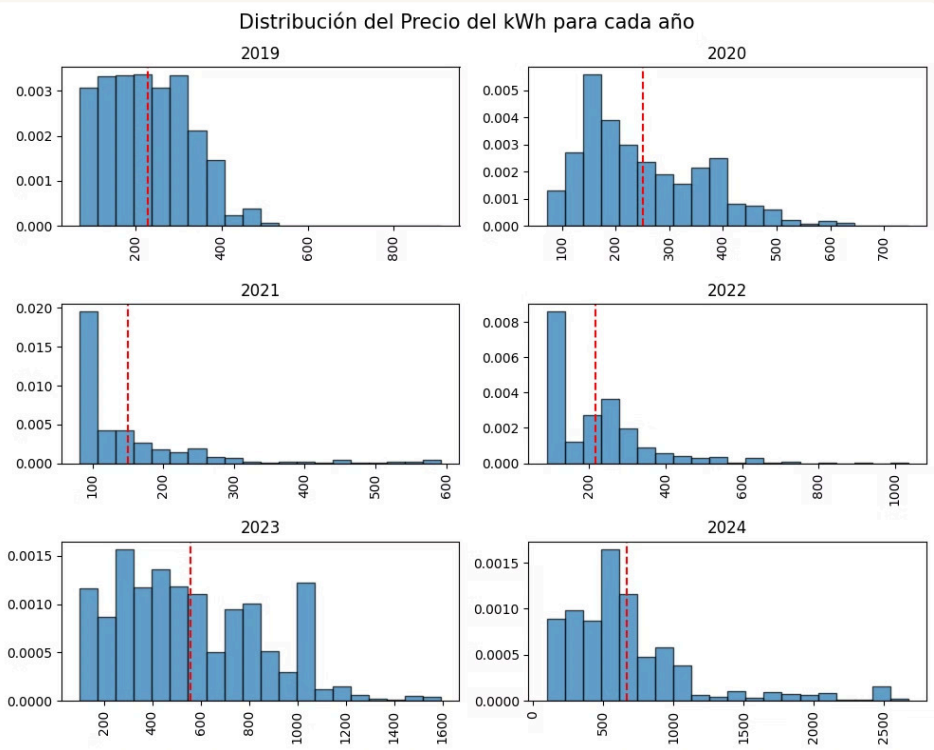
Serie de tiempo del precio del Kwh en la bolsa por horas y por promedio diario

2



Comportamiento del precio del kWh en intervalos de tiempo mensual, anual, diaria y horaria

3



Distribuciones del precio del kWh por año

4

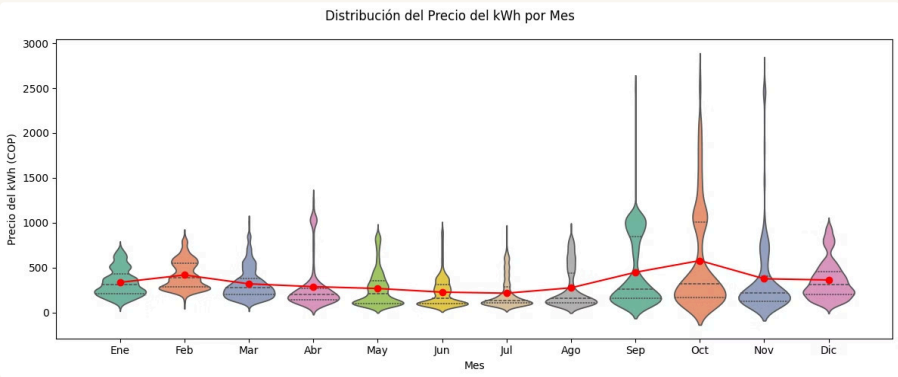
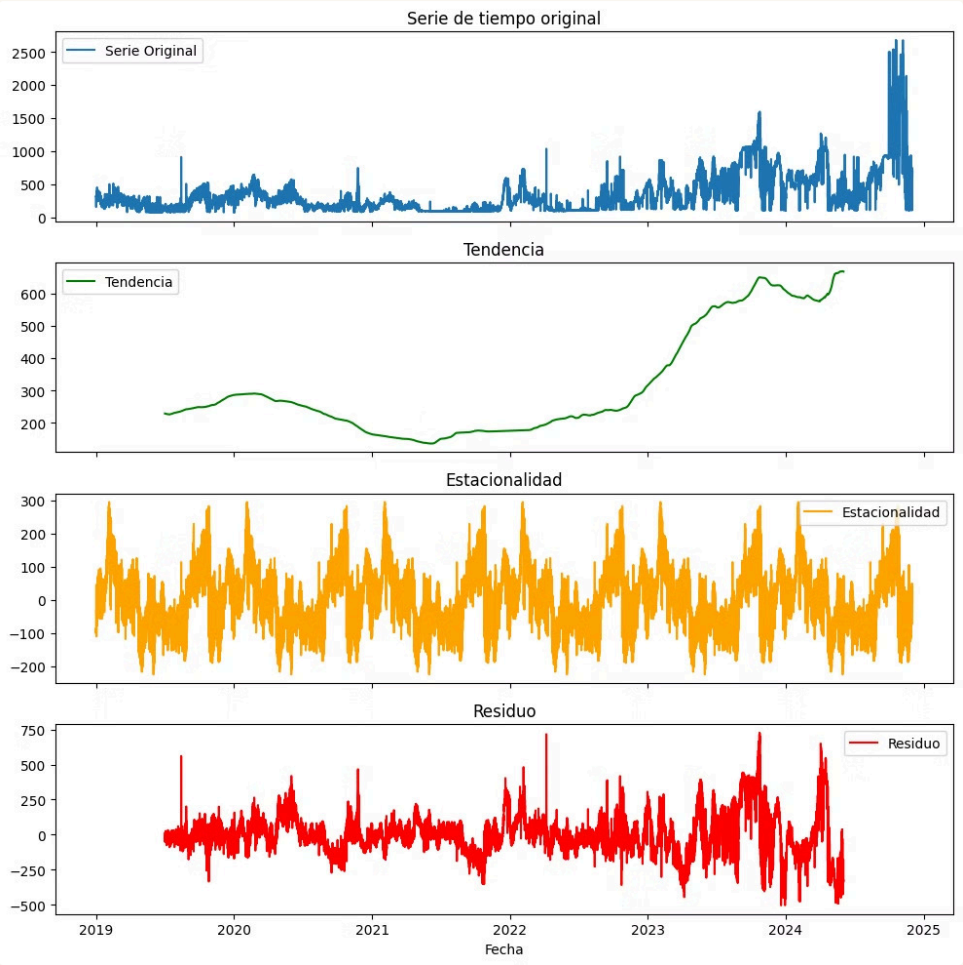


Diagrama de violin del precio del kWh por mes.

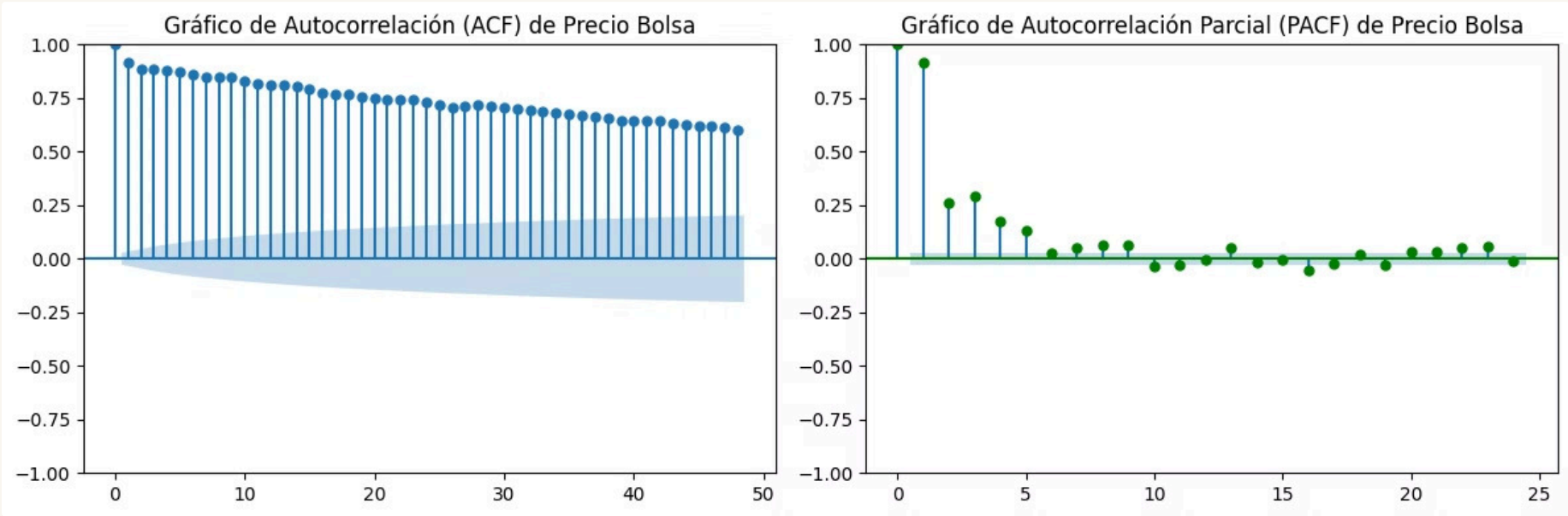
Resultados

1



Descomposición de la serie temporal por año. Tendencia, estacionalidad y residuo.

2

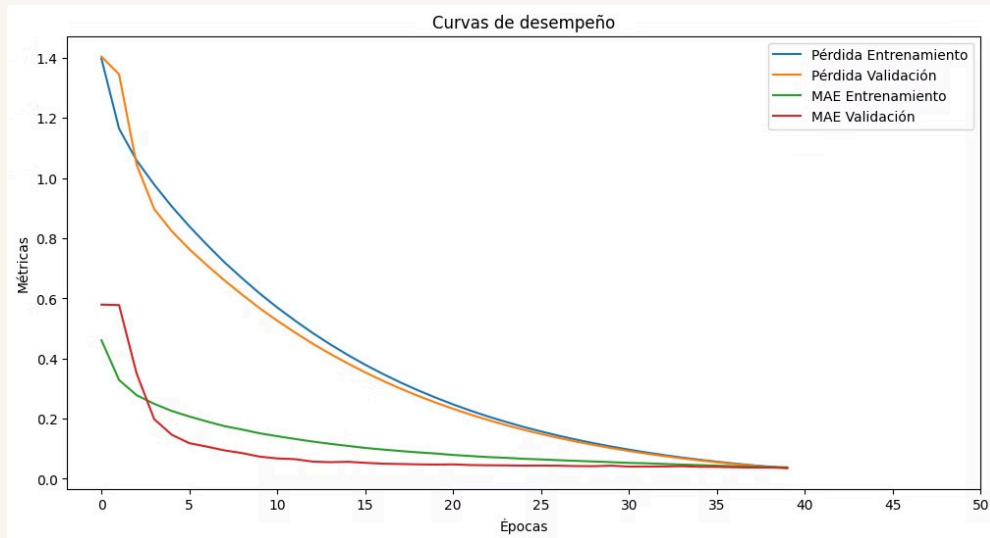


Gráficos de autocorrelación del precio del kWh.

Resultados

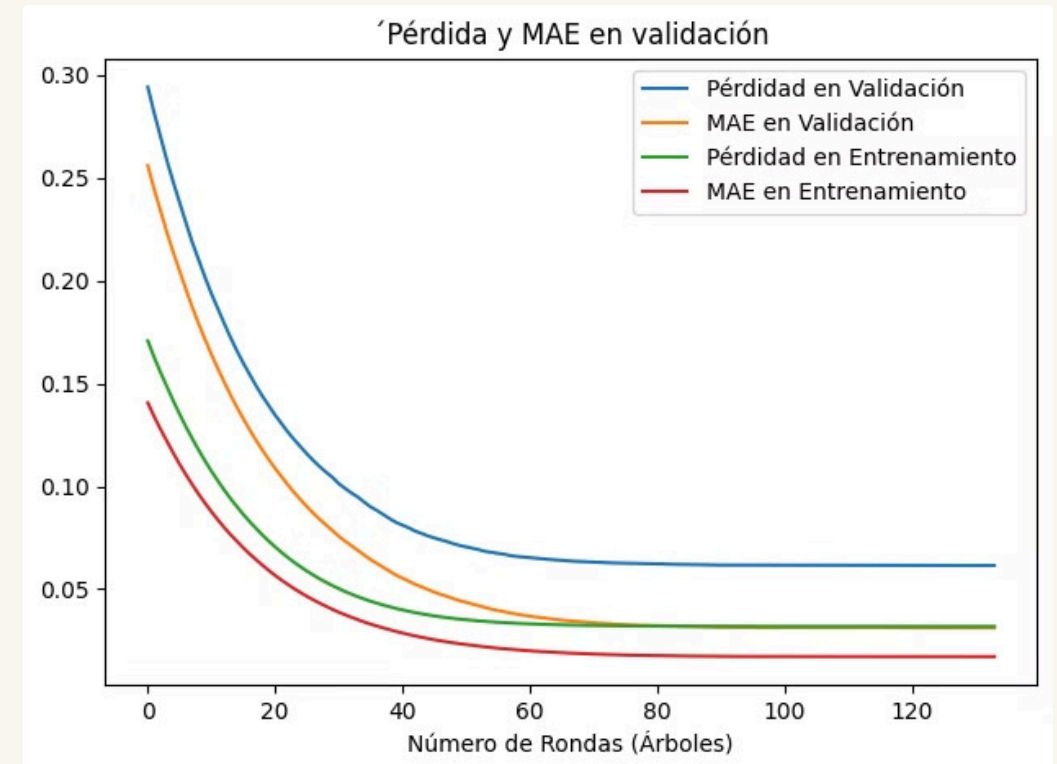
LSTM

- Error cuadrático medio (MSE): 0.004231
- Error absoluto medio (MAE): 0.036300
- Coeficiente de determinación (R^2): 0.876802



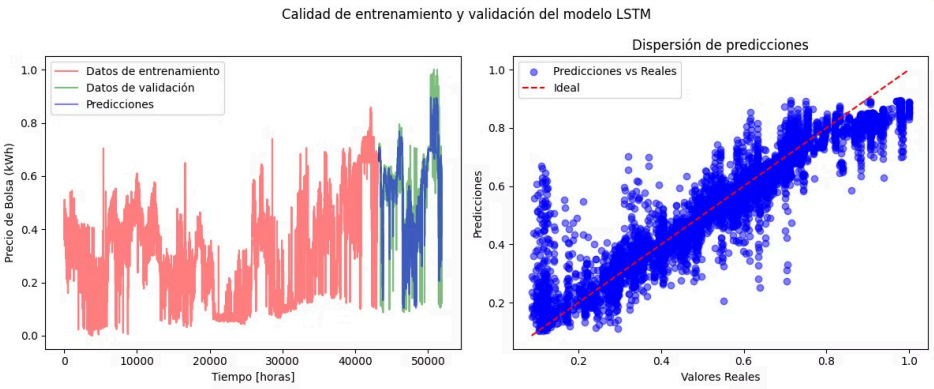
XGBoots

- Error cuadrático medio (MSE): 0.003774
- Error absoluto medio (MAE): 0.031135
- Coeficiente de determinación (R^2): 0.890109

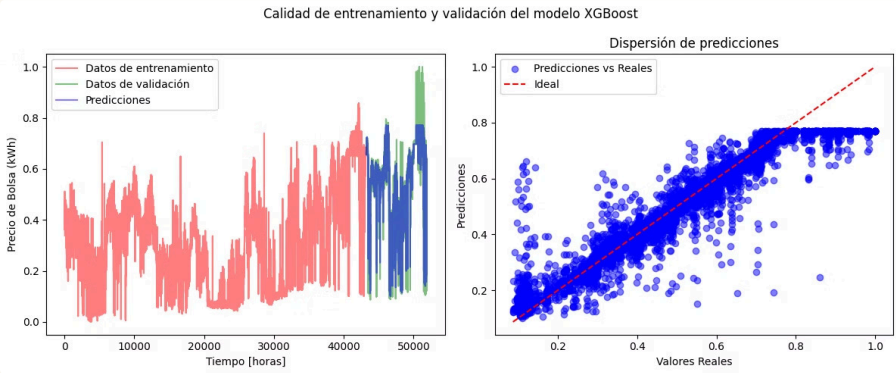


Resultados

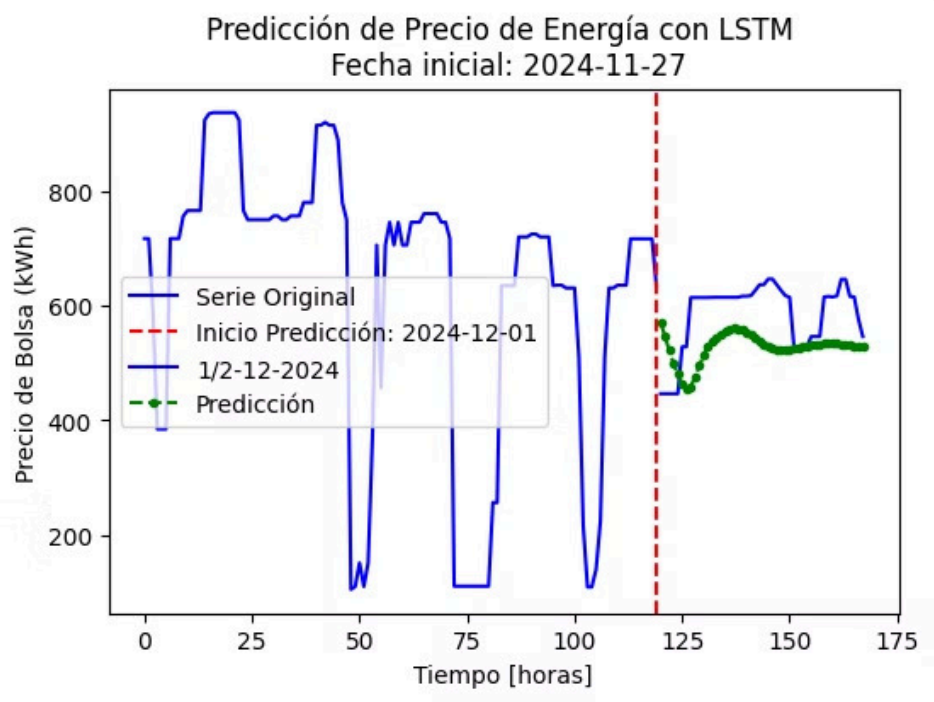
1



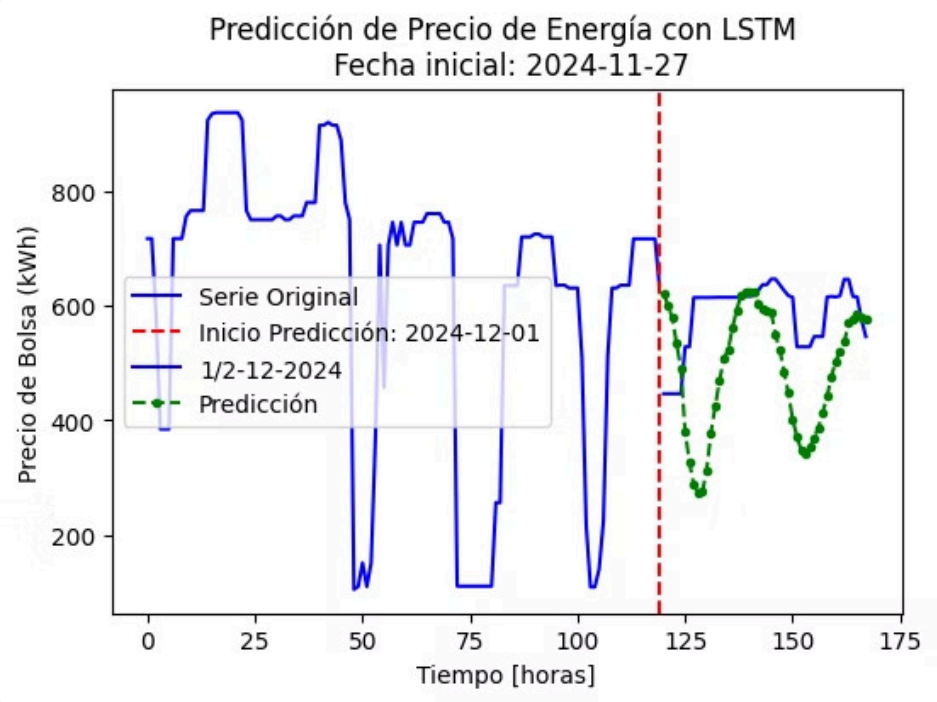
2



3



4



Conclusiones

Naturaleza de la serie

- Estocástica, no estacional.
- Gran variación del precio en los últimos años, no se encuentra una dependencia.

Datos

- El precio real es ponderado respecto a la energía transada por horas.

Predicción

- Poca autocorrelación en la serie
- Alto grado de volatilidad y aleatoriedad.
- Estimaciones a muy corto plazo y con muy baja precisión.
- El precio en bolsa depende de variables externas.



Mejoras

- Dataset más grande.
- Consideración de otras variables externas.

Referencias

- MathWorks. (s.f.). Análisis de series temporales. Recuperado de <https://la.mathworks.com/discovery/time-series-analysis.html>
- Yu, Y., Zhu, Y., Li, S., & Wan, D. (2014). Time series outlier detection based on sliding window prediction. Mathematical problems in Engineering, 2014(1), 879736.
- Ciencia de Datos. (s.f.). Forecasting time series with LightGBM. Recuperado de <https://cienciadedatos.net/documentos/py58-forecasting-time-series-with-lightgbm.html>
- Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. Frontiers in neurorobotics, 7, 21.
- Karim, F., Majumdar, S., Darabi, H., & Harford, S. (2019). Multivariate LSTM-FCNs for time series classification. Neural networks, 116, 237-245.
- MathWorks. (s.f.). Pronóstico de series de tiempo mediante deep learning. Recuperado de <https://la.mathworks.com/help/deeplearning/ug/time-series-forecasting-using-deep-learning.html>
- XM S.A. E.S.P. (s.f.). Generación Real del SIN. Recuperado de <https://sinergox.xm.com.co/oferta/Paginas/Informes/GeneracionSIN.aspx>
- OpenAI. (2025). ChatGPT-4 [Modelo de lenguaje de inteligencia artificial]. OpenAI. Recuperado de <https://openai.com>