

# Predicción de Consumo Eléctrico mediante Modelos de Aprendizaje Automático: Comparación entre SARIMA y Redes Neuronales LSTM

CAMILA EYZAGUIRRE<sup>1</sup>, CRISTIAN LORCA<sup>1</sup>, CINTYA OLIVARES<sup>1</sup>, and ALEJANDRO SUAREZ<sup>1</sup>

<sup>1</sup>Universidad Andrés Bello, Santiago, Chile (e-mail: c.eyzaguirrevillarro@uandresbello.edu, c.lorcavargasvargas@uandresbello.edu, c.olivarescisternas@uandresbello.edu, a.suarezantelices@uandresbello.edu)

Corresponding author: Camila Eyzaguirre (e-mail: c.eyzaguirrevillarro@uandresbello.edu).

Este trabajo fue realizado en el marco del curso de Tópicos en Ciencia de Datos.

**ABSTRACT** La predicción precisa del consumo eléctrico es fundamental para la gestión eficiente de redes eléctricas. Este estudio compara dos enfoques de aprendizaje automático para predicción de consumo eléctrico horario: SARIMA (Seasonal AutoRegressive Integrated Moving Average) y redes neuronales LSTM (Long Short-Term Memory). Se utilizó un conjunto de datos con 244,391 registros de entrenamiento (2018-2021) y 61,313 de prueba (2022) de múltiples subestaciones. El análisis incluyó descomposición temporal, análisis de estacionariedad y evaluación de patrones estacionales. Los resultados muestran que LSTM supera significativamente a SARIMA, con MAE de 19.57 kWh vs 578.46 kWh y  $R^2$  de 0.9757 vs -0.2895, representando una mejora de 29.5 veces en MAE. Este trabajo demuestra la superioridad de las redes neuronales profundas para capturar patrones complejos en series temporales de consumo eléctrico.

**INDEX TERMS** Series temporales, predicción de consumo eléctrico, SARIMA, LSTM, aprendizaje automático

## I. INTRODUCCIÓN

La demanda de energía eléctrica presenta patrones complejos que varían en múltiples escalas temporales: ciclos diarios, semanales y estacionales, influenciados por condiciones climáticas, eventos especiales y comportamiento de usuarios [1]. La predicción precisa del consumo eléctrico es esencial para optimizar la generación y distribución de energía, reducir costos operativos y mejorar la confiabilidad del sistema [2]. Los operadores de redes eléctricas enfrentan el desafío constante de equilibrar la oferta y demanda de energía, ya que la electricidad no puede almacenarse fácilmente a gran escala. Errores en la predicción pueden resultar en sobrecostos significativos o, en casos extremos, en cortes de suministro.

El modelado de series temporales para predicción de consumo eléctrico ha evolucionado significativamente. Los modelos estadísticos tradicionales como ARIMA (AutoRegressive Integrated Moving Average) y sus variantes estacionales SARIMA han sido ampliamente utilizados debido a su inter-

pretabilidad y eficiencia computacional [3]. Estos modelos asumen relaciones lineales entre observaciones pasadas y futuras, demostrando buen desempeño en series con patrones estacionales claros y estables.

Sin embargo, las limitaciones de los modelos lineales en capturar relaciones no lineales complejas han motivado la exploración de técnicas de aprendizaje automático. Las redes neuronales, particularmente las LSTM (Long Short-Term Memory), han mostrado capacidad superior para modelar dependencias temporales de largo plazo y capturar patrones no lineales complejos [4]. Las LSTM fueron diseñadas específicamente para resolver el problema del desvanecimiento de gradientes en redes neuronales recurrentes tradicionales, permitiendo aprender dependencias temporales de largo alcance.

El objetivo principal de este trabajo es desarrollar y comparar modelos de aprendizaje automático para la predicción de consumo eléctrico horario, evaluando el desempeño de enfoques estadísticos (SARIMA) y de aprendizaje profundo

(LSTM). Específicamente, los objetivos incluyen: realizar un análisis exploratorio completo del conjunto de datos, desarrollar un modelo SARIMA optimizado mediante búsqueda de hiperparámetros, diseñar y entrenar una arquitectura LSTM adecuada, comparar ambos modelos utilizando métricas estándar (MAE, RMSE,  $R^2$ , MAPE), y proporcionar recomendaciones sobre la selección de modelos según las características del problema.

## II. METODOLOGÍA

### II-A. DESCRIPCIÓN DEL CONJUNTO DE DATOS

El conjunto de datos contiene registros de consumo eléctrico horario de múltiples subestaciones. El conjunto de entrenamiento (2018-2021) tiene 244,391 registros en 34,913 puntos temporales únicos, mientras que el de prueba (2022) tiene 61,313 registros en 8,759 puntos temporales. Cada registro incluye: subestación, fecha (timestamp horario) y consumo (kWh). Los datos no contienen valores faltantes.

### II-B. ANÁLISIS EXPLORATORIO

El proceso comenzó con un análisis exploratorio exhaustivo que incluyó: visualización de series temporales para identificar tendencias y patrones visuales, descomposición temporal aditiva para separar la serie en componentes de tendencia, estacionalidad y residuos, análisis de estacionariedad mediante prueba de Dickey-Fuller aumentada (ADF) para determinar si la serie requiere diferenciación, análisis de autocorrelación mediante cálculo de funciones ACF y PACF para identificar dependencias temporales y periodicidades, y análisis de patrones estacionales mediante identificación de patrones horarios, semanales y mensuales.

La prueba ADF indicó que la serie es estacionaria ( $p$ -value  $< 0.05$ , ADF Statistic: -18.67), por lo que no requirió diferenciación (parámetro  $d = 0$ ). Se identificó un patrón diario fuerte (lag 24h con autocorrelación 0.7841) y 70 lags con autocorrelación significativa. Los lags más importantes fueron: lag 1h con correlación 0.9509 (muy fuerte), lag 24h con 0.7841 (patrón diario), lag 48h con 0.6387, y lag 72h con 0.6058.

### II-C. MODELO SARIMA

El modelo SARIMA se caracteriza por los parámetros  $(p, d, q) \times (P, D, Q, s)$ , donde  $p$  es el orden del proceso autorregresivo,  $d$  el orden de diferenciación,  $q$  el orden del promedio móvil,  $P, D, Q$  son componentes estacionales equivalentes, y  $s$  es el período estacional (24 horas para datos horarios).

Se realizó búsqueda sistemática de hiperparámetros mediante grid search, probando las siguientes combinaciones:  $p \in \{1, 2\}$  (basado en análisis PACF que mostró significancia en lags 1-2),  $d = 0$  (determinado por prueba ADF que confirmó estacionariedad),  $q \in \{1, 2\}$  (basado en análisis ACF),  $P \in \{1, 2\}$  (estacionalidad detectada),  $D \in \{0, 1\}$ ,  $Q = 1$  y  $s = 24$  (ciclo diario). La selección del mejor modelo se realizó basándose en el criterio de información de Akaike

(AIC) y el criterio de información bayesiano (BIC), eligiendo el modelo que minimizara estas métricas.

Se evaluaron 5 combinaciones de parámetros. La configuración óptima fue SARIMA(2, 0, 2)  $\times$  (2, 0, 1, 24) con AIC=387,450 y BIC=387,518. Este modelo captura componentes AR de orden 2 (dependencia de 2 pasos anteriores), componentes MA de orden 2, componentes estacionales AR de orden 2 con período de 24 horas, y componentes estacionales MA de orden 1.

### II-D. MODELO LSTM

Para el modelo LSTM se implementó el siguiente proceso de ingeniería de características: se crearon 23 características incluyendo codificación cíclica de hora del día, día de la semana y mes del año (codificadas como seno y coseno para capturar la naturaleza circular de estas variables temporales), características categóricas como día del mes e indicador de fin de semana, características de lag con valores de consumo en lags 1, 2, 3, 6, 12 y 24 horas para capturar dependencias temporales inmediatas, y estadísticas móviles con media y desviación estándar en ventanas de 6, 12 y 24 horas para proporcionar contexto sobre tendencia local y variabilidad reciente.

Los datos se transformaron a formato de secuencias 3D (*samples, timesteps, features*) con una ventana de look-back de 24 horas, donde cada muestra utiliza las 24 horas anteriores para predecir el consumo actual. Se aplicó normalización Min-Max a todas las características y al target, escalando los valores al rango  $[0, 1]$  para mejorar la convergencia del entrenamiento y evitar que características con escalas diferentes dominen el aprendizaje.

La arquitectura del modelo implementada consiste en: Capa LSTM 1 con 128 unidades y `return_sequences=True` permitiendo que la información fluya a la siguiente capa LSTM, Dropout 1 con tasa 0.2 para regularización, Capa LSTM 2 con 64 unidades procesando la salida de la primera capa, Dropout 2 con tasa 0.2, Capa densa con 32 neuronas y activación ReLU para capturar relaciones no lineales, y Capa de salida con 1 neurona lineal produciendo la predicción del consumo. El modelo contiene un total de 129,345 parámetros entrenables.

El entrenamiento se realizó utilizando el optimizador Adam con learning rate inicial de 0.001, función de pérdida MSE (Mean Squared Error), y métricas MAE y MSE. Se implementaron callbacks de Early Stopping con paciencia de 15 épocas y reducción de learning rate con factor 0.5 y paciencia de 7 épocas. El conjunto de datos se dividió temporalmente en 80 % para entrenamiento y 20 % para validación, manteniendo el orden temporal de los datos.

### II-E. EVALUACIÓN

Ambos modelos se evaluaron utilizando las siguientes métricas en el conjunto de prueba: MAE (Mean Absolute Error):  $MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$ , RMSE (Root Mean Squared Error):  $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$ ,  $R^2$  (Coeficiente de

determinación):  $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ , y MAPE (Mean Absolute Percentage Error):  $MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$ , donde  $y_i$  son los valores reales,  $\hat{y}_i$  son las predicciones, y  $\bar{y}$  es la media de los valores reales.

### III. ANÁLISIS DE DATOS

#### III-A. CARACTERÍSTICAS PRINCIPALES DE LA SERIE TEMPORAL

El análisis exploratorio reveló que la serie temporal de consumo eléctrico presenta las siguientes características fundamentales. En cuanto a estacionariedad, la prueba de Dickey-Fuller aumentada indicó que la serie es estacionaria (p-value <0.05, ADF Statistic: -18.67), por lo que no requirió diferenciación. Esto simplificó el modelado y permitió trabajar directamente con la serie original.

Respecto a estacionalidad, se identificó una estacionalidad débil (fuerza de estacionalidad: 0.2365) según el análisis de descomposición. Sin embargo, se observaron patrones claros: un ciclo diario fuerte donde la autocorrelación en el lag 24 horas fue de 0.7841, indicando un patrón diario muy pronunciado, y un ciclo semanal moderado donde se detectó autocorrelación en lags cercanos a 168 horas (7 días), aunque menos pronunciada que el ciclo diario.

El análisis de autocorrelación reveló que 70 lags presentan autocorrelación significativa, indicando una fuerte dependencia temporal. Los lags más importantes fueron: lag 1h con correlación 0.9509 (muy fuerte dependencia inmediata), lag 24h con 0.7841 (patrón diario), lag 48h con 0.6387, y lag 72h con 0.6058.

#### III-B. DESCOMPOSICIÓN TEMPORAL

La descomposición aditiva de la serie reveló componentes claramente diferenciados. La tendencia presenta un rango de [547.71, 1756.24] kWh con una desviación estándar de 178.68 kWh, mostrando variaciones graduales a lo largo del tiempo que reflejan cambios estructurales en el consumo. La estacionalidad muestra una amplitud de 344.21 kWh, indicando variaciones estacionales moderadas que capturan los patrones cíclicos repetitivos. Los residuos presentan una desviación estándar de 228.33 kWh, representando la variabilidad no explicada por tendencia y estacionalidad, que incluye efectos aleatorios y factores no modelados.

#### III-C. PATRONES ESTACIONALES

El análisis de patrones estacionales mostró variabilidad significativa en múltiples escalas temporales. El patrón horario muestra que el consumo presenta variaciones pronunciadas durante el día, con un coeficiente de variación del 83.37 %, donde se observaron picos de consumo en horas específicas del día y mínimos durante la madrugada. El patrón semanal indica que el consumo varía entre días de la semana, con un coeficiente de variación del 86.12 %, donde los fines de semana presentan patrones distintos a los días laborables. El patrón mensual revela que existe variabilidad estacional entre meses (coeficiente de variación: 86.64 %),

probablemente relacionada con cambios climáticos y patrones de uso estacionales.

## IV. EXPERIMENTOS Y RESULTADOS

### IV-A. JUSTIFICACIÓN DE LA SELECCIÓN DE CARACTERÍSTICAS

#### IV-A1. Modelo SARIMA

El modelo SARIMA es univariado y utiliza únicamente la serie temporal agregada del consumo total (suma de todas las subestaciones). Esta agregación permite simplificar el modelado al trabajar con una única serie temporal, capturar patrones globales del sistema eléctrico, reducir el ruido mediante la agregación, y facilitar la interpretación de resultados. La salida del modelo es el consumo total predicho en kWh para cada hora.

#### IV-A2. Modelo LSTM

El modelo LSTM utiliza un enfoque multivariado con 23 características de entrada cuidadosamente seleccionadas. Las características temporales cíclicas utilizan codificación cíclica (seno y coseno) de hora, día de la semana y mes, permitiendo al modelo capturar la naturaleza circular de estas variables temporales y evitando discontinuidades artificiales (por ejemplo, entre las 23:00 y 00:00 horas).

Las características de lag incluyen valores de consumo en horas anteriores (lags 1, 2, 3, 6, 12, 24) que capturan la inercia y dependencias temporales inmediatas, siendo especialmente importante el lag 24 horas que refleja el patrón diario identificado en el análisis. Las estadísticas móviles proporcionan medias y desviaciones estándar en ventanas móviles que ofrecen contexto sobre la tendencia local y la variabilidad reciente, ayudando al modelo a adaptarse a cambios graduales.

La salida es el consumo predicho para cada subestación individual, permitiendo predicciones más granulares que el modelo SARIMA agregado.

### IV-B. ENTRENAMIENTO

SARIMA fue entrenado con la configuración óptima SARIMA(2, 0, 2)  $\times$  (2, 0, 1, 24) en 80.9 segundos. Todos los coeficientes fueron significativos (p <0.001). LSTM fue entrenado durante 100 épocas máximas con parada temprana en la iteración óptima, mostrando convergencia estable sin sobreajuste significativo.

### IV-C. ARQUITECTURAS

SARIMA se representa mediante la ecuación:

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_1 B^{24} - \Phi_2 B^{48})y_t = (1 + \theta_1 B + \theta_2 B^2)(1 + \Theta_1 B^{24})\epsilon_t \quad (1)$$

LSTM utiliza una arquitectura de dos capas LSTM que permite capturar dependencias temporales a múltiples escalas. La justificación de esta arquitectura se basa en que dos capas LSTM permiten capturar dependencias temporales complejas a diferentes escalas de tiempo, el dropout reduce el

Cuadro 1. Métricas de Desempeño

Modelo	MAE (kWh)	RMSE (kWh)	R <sup>2</sup>	MAPE (%)
SARIMA	578.46	832.78	-0.2895	28.57
LSTM	19.57	36.33	0.9757	-

sobreajuste sin comprometer significativamente la capacidad del modelo, y la capa densa intermedia añade capacidad de modelado no lineal adicional. La primera capa LSTM (128 unidades) procesa las secuencias de entrada y pasa información contextual a la segunda capa (64 unidades), que finalmente genera la representación utilizada para la predicción.

#### IV-D. DESEMPEÑO

Los resultados en el conjunto de prueba se muestran en la Tabla 1.

##### IV-D1. Análisis Comparativo de Resultados

El modelo LSTM supera significativamente al SARIMA en todas las métricas principales: MAE 29.5 veces menor (19.57 vs 578.46 kWh), RMSE 22.9 veces menor (36.33 vs 832.78 kWh), y R<sup>2</sup> de 0.9757 explicando 97.57 % de la varianza vs R<sup>2</sup> negativo de SARIMA.

Analizando específicamente el modelo SARIMA: el MAE de 578.46 kWh representa aproximadamente el 28.57 % del consumo promedio, indicando errores relativamente altos. El RMSE de 832.78 kWh, mayor que el MAE, sugiere la presencia de errores grandes ocasionales. El R<sup>2</sup> negativo (-0.2895) indica que el modelo tiene un desempeño peor que simplemente predecir la media, lo cual sugiere problemas en la aplicación del modelo a nivel agregado o dificultades para capturar la variabilidad del conjunto de prueba. El modelo parece tener dificultades para generalizar, posiblemente debido a cambios en los patrones entre los períodos de entrenamiento y prueba.

En contraste, el análisis del modelo LSTM muestra resultados excelentes: el MAE de 19.57 kWh representa un error absoluto promedio muy bajo, aproximadamente 1-2 % del consumo típico. El RMSE de 36.33 kWh, cercano al MAE, indica errores relativamente uniformes sin valores atípicos extremos. El R<sup>2</sup> de 0.9757 indica que el modelo explica el 97.57 % de la varianza en los datos de prueba, lo cual es excelente. El modelo demuestra una capacidad superior para capturar patrones complejos y generalizar a nuevos datos.

##### IV-D2. Comparación con Resultados de la Literatura

El R<sup>2</sup> de LSTM (0.9757) es comparable o superior a valores reportados en literatura. Varios estudios han demostrado la superioridad de modelos basados en aprendizaje profundo sobre métodos estadísticos tradicionales para predicción de consumo eléctrico cuando se dispone de suficientes datos [5]. El desempeño de nuestro modelo LSTM se encuentra en el rango superior de estos resultados, siendo comparable con estudios recientes que reportan R<sup>2</sup> entre 0.90 y 0.96

para diferentes arquitecturas LSTM en predicción de carga eléctrica.

El bajo desempeño del modelo SARIMA es inesperado y puede atribuirse a varios factores: cambios estructurales en los patrones de consumo entre los períodos de entrenamiento y prueba, la naturaleza agregada del modelo SARIMA que puede perder información importante a nivel de subestación, y posibles efectos de eventos externos (como cambios en patrones de consumo) que alteraron las relaciones temporales modeladas por SARIMA.

#### V. CONCLUSIONES

Este trabajo presentó un análisis comparativo exhaustivo entre modelos SARIMA y LSTM para la predicción de consumo eléctrico horario. Los principales hallazgos y conclusiones son los siguientes.

Respecto a los objetivos planteados, se logró realizar un análisis exploratorio completo del conjunto de datos, identificando que la serie temporal es estacionaria, presenta patrones diarios fuertes (lag 24h con correlación 0.7841) y ciclos semanales moderados. La estacionalidad es débil pero presente, y existe una fuerte autocorrelación temporal (70 lags significativos). Se desarrolló un modelo SARIMA(2, 0, 2) × (2, 0, 1, 24) optimizado mediante grid search, con todos los coeficientes estadísticamente significativos. Sin embargo, el modelo presentó dificultades para generalizar al conjunto de prueba, obteniendo métricas relativamente pobres. Se diseñó e implementó una arquitectura LSTM de dos capas con 129,345 parámetros, utilizando 23 características cuidadosamente seleccionadas. El modelo mostró excelente capacidad de aprendizaje y generalización.

En cuanto a la comparación entre modelos, el modelo LSTM superó significativamente al SARIMA en todas las métricas principales: MAE mejoró 29.5 veces (19.57 vs 578.46 kWh), RMSE mejoró 22.9 veces (36.33 vs 832.78 kWh), y R<sup>2</sup> de 0.9757 explicando 97.57 % de la varianza vs R<sup>2</sup> negativo de SARIMA.

Las implicaciones prácticas de estos resultados son importantes para la gestión de sistemas eléctricos. Los modelos LSTM pueden proporcionar predicciones más precisas, reduciendo costos operativos y mejorando la planificación de generación. La arquitectura propuesta es escalable y puede adaptarse a diferentes sistemas eléctricos. El uso de características múltiples (no solo la serie temporal univariada) mejora significativamente el desempeño.

Para problemas de predicción de consumo eléctrico con características similares (suficientes datos históricos, múltiples subestaciones, patrones complejos), se recomienda el uso de modelos LSTM debido a su capacidad superior para capturar relaciones no lineales y dependencias temporales complejas. Las limitaciones principales identificadas incluyen: menor interpretabilidad del modelo LSTM comparado con SARIMA, mayor requerimiento de datos y recursos computacionales para el entrenamiento, y necesidad de variables exógenas adicionales (temperatura, eventos especiales) para mejoras futuras.



En resumen, este trabajo demuestra que los modelos de aprendizaje profundo, específicamente LSTM, ofrecen ventajas significativas para la predicción de consumo eléctrico cuando se dispone de datos suficientes y características bien diseñadas, validando los objetivos planteados y proporcionando una base sólida para aplicaciones prácticas en la industria eléctrica.

## VI. LIMITACIONES Y TRABAJOS FUTUROS

### VI-A. LIMITACIONES

Este trabajo presenta varias limitaciones que deben ser reconocidas. En cuanto a limitaciones del conjunto de datos: los datos cubren un período de 4 años de entrenamiento y 1 año de prueba, lo cual puede ser insuficiente para capturar todos los patrones a largo plazo y cambios estructurales. No se incluyeron variables exógenas importantes como temperatura, humedad, día festivo, o eventos especiales que podrían mejorar las predicciones. El modelo SARIMA se aplicó a datos agregados, perdiendo información granular a nivel de subestación que podría ser valiosa.

Respecto a limitaciones metodológicas: la búsqueda de hiperparámetros para SARIMA fue limitada debido a restricciones computacionales, pudiendo haber mejores combinaciones no exploradas. El modelo LSTM utiliza una arquitectura fija de dos capas; no se exploró exhaustivamente el espacio de arquitecturas posibles. No se implementaron técnicas avanzadas como atención (attention mechanisms) o Transformers que podrían mejorar aún más el desempeño. El preprocesamiento de datos para LSTM podría optimizarse mediante validación cruzada temporal más rigurosa.

En cuanto a limitaciones técnicas: el modelo LSTM requiere recursos computacionales significativos para entrenamiento y predicción, lo cual puede ser una limitación en entornos con recursos limitados. La interpretabilidad limitada del modelo LSTM dificulta el diagnóstico de problemas y la identificación de factores que más contribuyen a las predicciones.

### VI-B. PROPUESTAS DE ESTUDIOS FUTUROS

Basado en las limitaciones identificadas, se proponen las siguientes líneas de investigación. Para mejoras en el modelo LSTM: explorar arquitecturas avanzadas como LSTM bidireccionales, redes de atención, Transformers, y modelos híbridos. Implementar optimización de hiperparámetros más exhaustiva usando Bayesian Optimization y optimización de arquitectura neuronal. Experimentar con técnicas de regularización avanzada como dropout adaptativo y batch normalization.

Para incorporación de variables exógenas: integrar datos meteorológicos (temperatura, humedad, velocidad del viento) que tienen fuerte correlación con el consumo eléctrico. Incluir información sobre días festivos, eventos especiales, tipo de día (laboral, fin de semana, vacaciones), e indicadores económicos. Desarrollar modelos multivariados avanzados que integren múltiples fuentes de datos.

Para modelado granular y ensambles: desarrollar modelos específicos para cada subestación o grupo de subestaciones con patrones similares. Implementar estrategias de ensamble de múltiples arquitecturas LSTM y ensamble entre modelos estadísticos y de aprendizaje profundo. Desarrollar modelos jerárquicos que predigan a múltiples niveles manteniendo coherencia.

Para evaluación y métricas avanzadas: desarrollar métricas económicas que reflejen el costo real de errores considerando horas pico. Implementar análisis de incertidumbre mediante quantile regression y modelos probabilísticos. Realizar validación cruzada temporal más exhaustiva con múltiples ventanas de prueba.

Para interpretabilidad: aplicar técnicas como SHAP o LIME para identificar qué características contribuyen más a las predicciones. Desarrollar métodos para visualizar qué patrones temporales captura el modelo LSTM. Diseñar arquitecturas híbridas interpretables que combinen precisión y explicabilidad.

Estas propuestas se enfocan en mejorar el desempeño del modelo, aumentar su interpretabilidad, integrar información adicional, y desarrollar aplicaciones prácticas que podrían llevar a sistemas de predicción aún más precisos y útiles para la gestión eficiente de sistemas eléctricos.

### CODE AVAILABILITY

El código fuente completo de este proyecto, incluyendo el notebook de Jupyter con todas las etapas de análisis exploratorio, preprocesamiento de datos, ingeniería de características, entrenamiento de los modelos SARIMA y LSTM, y evaluación comparativa, está disponible públicamente en GitHub: <https://github.com/Cintya01/DC-SeriesDeTiempo/blob/main/SerieDeTiempo.ipynb>

### Referencias

- [1] A. Bunn, D. Farmer, "Comparative models for electrical load forecasting," Wiley, 1985.
- [2] H. T. Ha, J. A. V. Restrepo, F. E. Elkhathib, "Short-term load forecasting using long short-term memory network," in Proc. IEEE PES General Meeting, 2017.
- [3] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, Time Series Analysis: Forecasting and Control, 4th ed. Hoboken, NJ, USA: Wiley, 2015.
- [4] S. Hochreiter, J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
- [5] Z. Ahmad, A. S. Chen, T. T. T. Nguyen, L. Zhang, "Deep learning for short-term load forecasting: an overview and comparative analysis," IEEE Access, vol. 9, pp. 156388–156416, 2021.

...