



**Universidad
de Navarra**

DATAI
INSTITUTO DE CIENCIA DE LOS
DATOS E INTELIGENCIA ARTIFICIAL

DESARROLLO DE UN MODELO DE REDES NEURONALES PROFUNDAS PARA PREDECIR EL CONSUMO ENERGETICO DE UN EDIFICIO

Trabajo Fin de Máster

Máster Universitario en Big Data Science

Curso académico 2022-2023

AUTORES: Ana Cristina Montenegro Celada

Carlos Eduardo Vidales Bolaños

TUTOR ACADÉMICO: Montserrat-Ana Miranda Galcerán

TUTOR DE EMPRESA: Montserrat-Ana Miranda Galcerán

Madrid, 7 de julio de 2023

RESUMEN

El presente trabajo de fin de máster se centra en la construcción de un modelo de red neuronal recurrente (RNN, por sus siglas en inglés) para la predicción del consumo energético de un edificio no residencial. La necesidad de desarrollar modelos precisos y confiables para estimar el consumo de energía en edificios es fundamental para promover la eficiencia energética y tomar decisiones informadas en la gestión de los recursos.

El enfoque propuesto se basa en el uso de una RNN, una arquitectura de red neuronal especialmente diseñada para modelar secuencias de datos. La RNN tiene la capacidad de aprender patrones secuenciales y capturar dependencias a largo plazo en los datos de entrada, lo cual es especialmente relevante para la predicción del consumo energético en un contexto temporal.

Este proyecto demuestra el potencial de las redes neuronales recurrentes, en particular las LSTM, para la predicción del consumo energético en edificios no residenciales. Los resultados respaldan la utilidad de esta metodología en la gestión y optimización de los recursos energéticos, lo que puede contribuir significativamente a la eficiencia energética y la toma de decisiones informadas en el sector.



CONTENIDO

INTRODUCCION	2
1. OBJETIVOS	4
1.1 OBJETIVO GENERAL.....	4
1.2 OBJETIVOS ESPECIFICOS	4
1.3 HERRAMIENTAS UTILIZADAS.....	4
1.4 ESTADO DEL ARTE	5
1.5 HERRAMIENTAS UTILIZADAS.....	19
1.6 DIAGRAMA DE FLUJO	20
2. ANALISIS Y ESTRUCTURA DE LA BASE DE DATOS	20
2.1. PROCESAMIENTO DE LA BASE DE DATOS:	23
2.2. ANALISIS ESTADISTICO DE LOS DATOS.....	24
2.3. EXPLICACION DEL PROCESO DE LIMPIEZA	25
2.4. DEPENDENCIA ENTRE LAS VARIABLES	27
2.5. FEATURING ENGINEERING	28
2.6. COMPONENTE DE SERIE-TEMPORAL:	29
2.7. PREPARACION DE LOS DATOS PARA LOS MODELOS	32
2.8. ARQUITECTURA DE LOS MODELOS	32
3. RESULTADOS	38
4. DISCUSIÓN	44
5. CONCLUSIONES Y TRABAJO FUTURO	44
Bibliografía.....	47

INTRODUCCION

La gestión del consumo energético en edificios no residenciales representa un desafío significativo para la eficiencia energética y la sostenibilidad para Europa (1). Estos edificios, que incluyen oficinas, centros comerciales y hospitales, consumen grandes cantidades de energía eléctrica y térmica, y el control del consumo energético es fundamental para reducir costos y emisiones de gases de efecto invernadero. Sin embargo, la gestión de la energía en edificios no residenciales es complicada debido a la complejidad de los sistemas de climatización, iluminación y otros equipos, y la falta de información detallada sobre los patrones de consumo energético. Por lo tanto, existe una necesidad apremiante de desarrollar modelos de predicción precisos acerca del consumo energético en edificios no residenciales que permitan una mejor gestión y control de la energía.

En la actualidad, la gestión de la energía es un tema crucial en la Unión Europea (UE) debido a la creciente preocupación por el cambio climático. La predicción del consumo energético de edificios no residenciales se ha convertido en una herramienta clave para mejorar la eficiencia energética en el sector. La implementación de modelos de predicción que puede ayudar a los administradores de edificios a optimizar el consumo energético y reducir los costos de energía (2).

En los últimos años, se ha avanzado significativamente en el desarrollo de técnicas y metodologías para la predicción del consumo energético en edificios no residenciales. Además, se han implementado sistemas de monitorización y control de la energía, como los sistemas de gestión de la energía y los sistemas de automatización de edificios, para la captura de datos y la gestión del consumo energético en tiempo real. Estos sistemas permiten obtener información detallada sobre el consumo energético y las condiciones ambientales en los edificios, lo que puede mejorar la precisión de los modelos de predicción.

Sin embargo, según el autor Zhang, Jiang y Yu (3), aún existen desafíos y limitaciones en la predicción del consumo energético en edificios no residenciales, como la complejidad y la heterogeneidad de los edificios, la variabilidad en los patrones de consumo energético, la falta de datos históricos, la incertidumbre en las condiciones climáticas futuras, la ocupación y el funcionamiento de electrodomésticos.

Por lo tanto, En esta sección se explorará los avances y desafíos en la predicción del consumo energético de edificios no residenciales en la UE, y presentar algunas de las técnicas y metodologías más utilizadas en la actualidad. Además, se discutirán los desafíos y limitaciones actuales en la predicción del consumo energético, y se propondrán posibles soluciones y áreas de investigación futura para mejorar la eficiencia energética en edificios no residenciales.

1. OBJETIVOS

1.1 OBJETIVO GENERAL

Desarrollar un modelo de predicción del consumo energético para un edificio no residencial que permita predecir su consumo de energía y mejorar su eficiencia energética.

1.2 OBJETIVOS ESPECIFICOS

- Identificar y analizar los factores que influyen en el consumo energético del edificio en cuestión, como la iluminación, el clima, la ocupación, entre otros.
- Recopilar y analizar datos históricos del consumo de energía del edificio para entrenar el modelo de predicción y validar su precisión.
- Desarrollar un modelo de predicción en redes neuronales Long-Short-Term Memory (LSTM).
- Optimizar el modelo de predicción mediante el ajuste de los parámetros, la selección de características relevantes y la mejora de la precisión de la predicción en el conjunto de datos de prueba.

1.3 HERRAMIENTAS UTILIZADAS

En el proyecto, se emplearon diversas herramientas y librerías en Python para el análisis y modelado de datos. Entre ellas se encuentran pandas, matplotlib, seaborn, numpy, keras, scikit-learn, scipy y tensorflow.

Pandas fue utilizado para el procesamiento y manipulación de datos tabulares, ofreciendo funciones para filtrado, transformación y limpieza de datos. Matplotlib y seaborn se

utilizaron para la visualización de datos, creando gráficos informativos y visualmente atractivos.

En cuanto al modelado de datos, keras y tensorflow se utilizaron para el desarrollo de redes neuronales y aprendizaje automático, permitiendo la construcción y entrenamiento de modelos eficientes. Además, scikit-learn proporcionó herramientas para el preprocesamiento de datos, imputación de valores faltantes y evaluación de modelos. Por otro lado, scipy ofreció funcionalidades estadísticas, como pruebas y cálculos de correlación. El código relacionado con el documento se subió a un repositorio de Github.

1.4 ESTADO DEL ARTE

Existen varios modelos de predicción del consumo energético para edificios no residenciales, cada uno con sus propias características y distintos enfoques realizados durante la última década. Entre los métodos de predicción de energía existentes el enfoque con mayor aplicación han sido las redes neuronales – Artificial Neural Networks (ANNs), Support vector Machines (SVM), fuzzy logic, grey models techniques and *lineal regressions*.

1.4.1 REDES NEURONALES

Se introduce un modelo sencillo como la Redes Neuronales, es un modelo simplificado que emula el modo en que el cerebro humano procesa la información, donde las unidades de procesamiento se organizan en capas, existen tres partes:

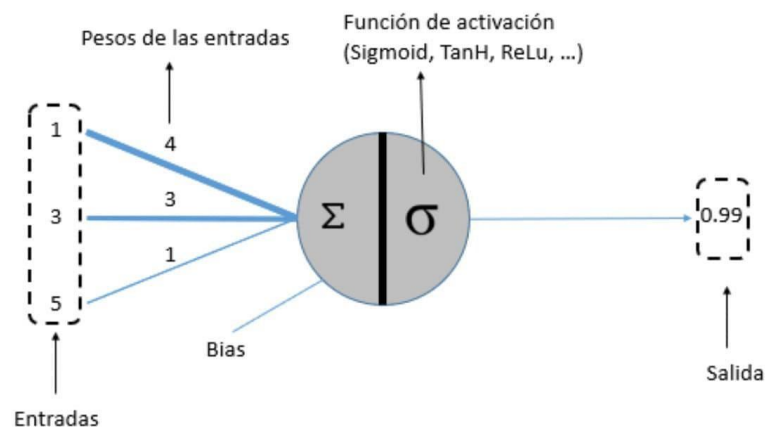
- **Capa de entrada (*Input layer*):** unidades que representa los campos de entrada
- **Capas ocultas y capa de salida (*Hidden layer and output layer*):** una o varias unidades que representan el campo o los campos de destino.

Las capas mencionadas se conectan con fuerzas de conexión o ponderaciones. Los datos de entrada se presentan en la primera capa, y los valores se propagan desde cada

neurona hasta cada neurona en la capa siguiente, teniendo como resultado la capa final.
(4)

En el caso de las neuronas artificiales, la suma de las entradas multiplicadas por sus pesos asociados determina el “impulso nervioso” que recibe la neurona. Este valor, se procesa en el interior de la célula mediante una función de activación que devuelve un valor que se envía como salida de la neurona, como se muestra la en la siguiente figura (5):

Figura No. 1
Red Neuronal Simple.



Pérez, J. (2022). Red neuronal [Captura de pantalla]. Xeridia.
<https://www.xeridia.com/blog/redes-neuronales-artificiales-que-son-y-como-se-entrenan-parte-i>

La figura anterior ilustra el funcionamiento de una neurona en una red neuronal simple, esta neurona recibe datos de entradas y las procesa ponderándolas con pesos específicos, ya sea que se asignan los valores o valores aleatorios. Luego, las entradas ponderadas se activan mediante una función de activación (este se define según la naturaleza del problema a solucionarse), produciendo así una salida.

1.4.2 MODELO DE NEURONA

La representación Algebraica de una red neuronal simple para el ejemplo anterior se describe mediante la siguiente ecuación:

$$z_j[t+1] = g \left(\sum_{i=1}^n W_{ji} z_i[t] + W_j \right)$$

En esta ecuación, se identifican cinco elementos básicos para j-enésima neurona de una red (6):

- Un conjunto de n señales de entrada $z_i[t]$, $i = 1, \dots, n$, que suministran a la neurona los datos del entorno
- Un conjunto de sinapsis, caracterizada por un peso propio W_{ji} , $i = 1, \dots, n$, donde el peso W_{ji} está asociado con la sinapsis que conecta la unidad i-enésima con la neurona j-enésima
- Un sesgo W_j cuya presencia aumenta la capacidad de procesamiento de la neurona y modifica la entrada a la neurona, según sea su valor positivo o negativo
- Un sumador o integrador que suma las señales de entrada, ponderadas con sus respectivos pesos, y el sesgo.
- Una función de activación g que suele limitar la amplitud de la salida de la neurona.

Utilizando la notación algebraica definida, podemos describir la operación de una neurona mediante la ecuación que determina su activación en el instante $t+1$

1.4.3 REDES NEURONALES RECURRENTES

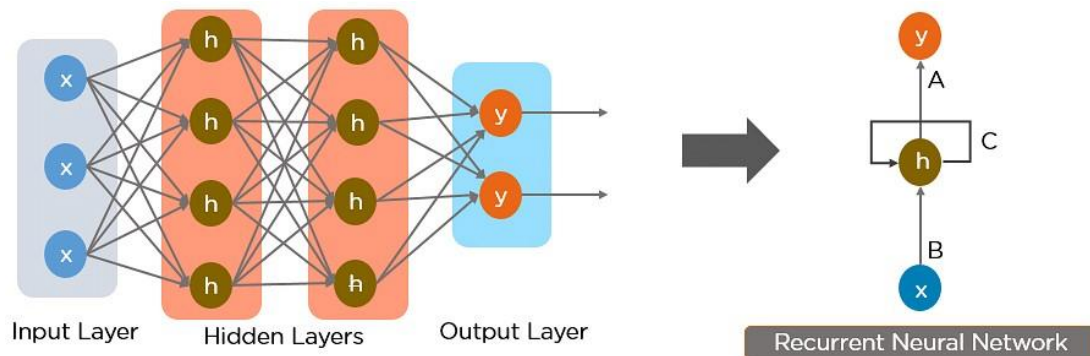
Dentro de la clasificación de las redes neuronales se encuentran las redes neuronales recurrentes (RNN, por sus siglas en inglés), las cuales se diseñan para procesar secuencias o colecciones de entradas. Estas redes aprovechan la naturaleza secuencial de los objetos de entrada y generan una salida secuencial correspondiente (7). En la figura No 2 se muestra la idea principal de una red neuronal recurrente.

Se observa el cambio en una red neuronal recurrente, los nodos en diferentes capas se comprimen para formar una sola capa de redes neuronales recurrentes. La capa de

entrada está representada por “x”, la capa oculta por “h” y la capa de salida por “y” que contiene las variables objetivo. Los parámetros A, B y C son utilizados en la red para mejorar la salida del modelo, cada uno siendo una matriz de pesos.

Figura No. 2

Idea principal de una red neuronal recurrente.



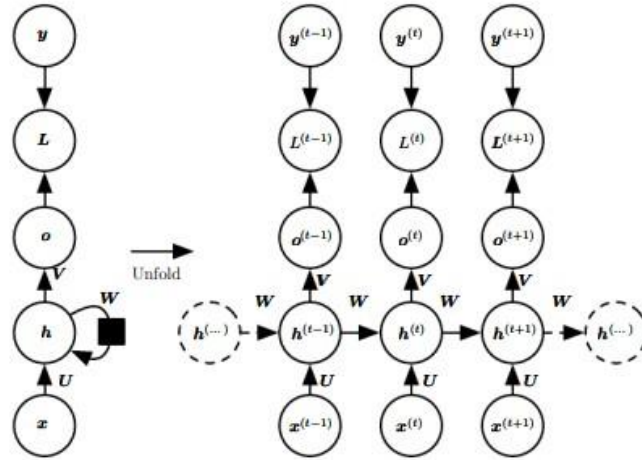
Fuente: Biswal, A. (10 de abril de 2023). Simple Recurrent Neural Network [Grafico Ilustrativo]. Recuperado de <https://www.simplilearn.com/tutorials/deep-learning-tutorial/rnn>

1.4.4 MODELO DE UN RED NEURONAL RECURRENTE

Usualmente las redes neuronales recurrentes profundas tiene un comportamiento sequence-vector. Como consecuencia de esta estructura, la salida de la red será un único elemento que se calculará en función del resto (8), como se muestra en la siguiente figura.

Figura No. 3

Topología de una Red Neuronal Recurrente



Fuente: Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press.
<http://www.deeplearningbook.org> (2016)

El modelo de una red neuronal recurrente se expresa algebraicamente de la siguiente manera:

$$a(t) = s + Wh_{(t-1)} + Ux(t) \quad (2.2.1) \quad h^{(t)}$$

$$= \tanh(a^{(t)}) \quad (2.2.2)$$

$$o^{(t)} = c + Vh^{(t)} \quad (2.2.3)$$

$$y^{(t)} = \text{softmax}(o^{(t)}) \quad (2.2.4)$$

Estas ecuaciones, se definen los siguientes elementos:

$a^{(t)}$: es el vector de activación en el instante de tiempo t en la capa oculta de la red neuronal. Se calcula el sesgo s , a esto se le suma el producto de la matriz de pesos W con el vector de estado anterior $h^{(t-1)}$, y el producto de matriz de pesos U con el vector de entrada $x^{(t)}$ en el tiempo t.

$h^{(t)}$: Es el vector de estado en el instante de tiempo t en la capa oculta. Se calcula aplicando la función de activación tangente hiperbólica (\tanh) al vector $a^{(t)}$, la función \tanh comprime los valores de $a^{(t)}$ entre -1 y 1, lo que permite que la capa oculta capture relaciones no lineales en los datos.

$\mathbf{o}^{(t)}$: Es el vector de salida en el instante de tiempo t . Se calcula sumando el sesgo c y el producto de la matriz de pesos V con el vector de estado $h^{(t)}$. Esto determina cómo la capa oculta influye en la capa de salida

$\mathbf{y}^{(t)}$: Es el vector de salida predicha en el instante de tiempo t . Se calcula aplicando la función softmax al vector $\mathbf{o}^{(t)}$, la función convierte los valores en $\mathbf{o}^{(t)}$ en probabilidades, lo que representa la distribución de probabilidades de las diferentes calases de salidas.
(9)

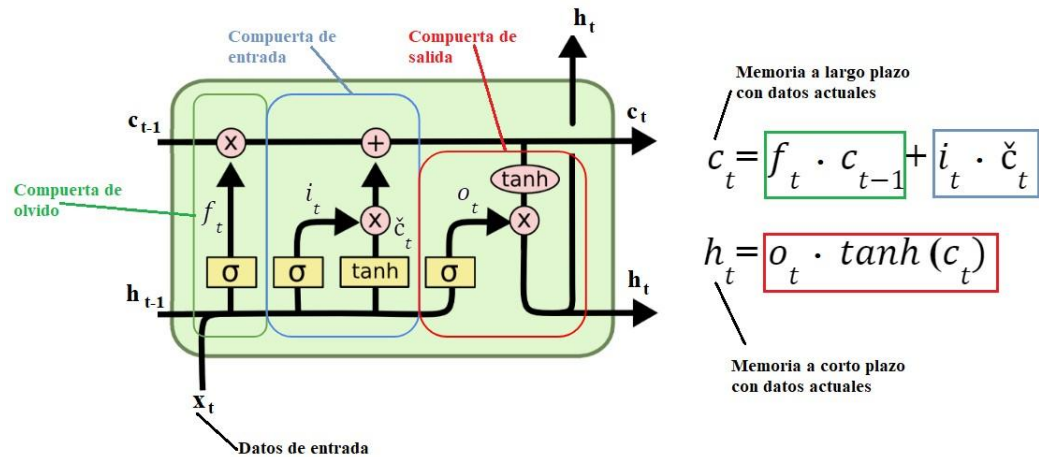
1.4.5 ARQUITECTURAS EN LAS REDES NEURONALES RECURRENTES

Existen distintas arquitecturas compatibles para redes neuronales recurrentes, en este caso solo hablaremos de LSTM, Bidireccionales y Attention.

- LSTM (Long Shor-Term Memory) es una arquitectura de red neuronal recurrente que se utiliza para modelar y capturar dependencias al largo plazo en secuencia de datos. Su estructura compone de una celda de memoria que puede almacenar información durante períodos prolongados, y consta de tres compuertas: la compuerta de entrada, la compuerta de olvido y la compuerta de salida, como se muestra en la figura No. 4. Estas compuertas permiten controlar el flujo de información hacia y desde la celda de memoria, facilitando así la retención de información relevante y el olvido de información no relevante. A diferencia de las redes neuronales recurrentes tradicionales, la LSTM aborda el problema del desvanecimiento y la explosión del gradiente mediante el uso de celdas de memoria y mecanismos de puerta. (10)

Figura No. 4

Arquitectura de LSTM



Fuente: Data Science PE. (Fecha de publicación). Series Temporales - OP 3: LSTM [Mensaje de blog]. <https://datasciencepe.substack.com/p/series-temporales-op-3-lstm> Descripción de la función de cada compuerta:

- Compuerta de Olvido: Esta compuerta determina cuánta información almacenada en la memoria a largo plazo debe olvidarse (11).

La información del estado oculto anterior y la entrada actual pasan a través de la función sigmoide (σ). La fórmula correspondiente en la siguiente

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f)$$

Donde h_{t-1} es el estado oculto de la celda anterior, x_t representa la entrada de la celda actual, W_f es la matriz de pesos y b_f es el vector de sesgo que se suma a la multiplicación de la matriz de pesos. (12)

- Compuerta de Entrada: Esta compuerta cuantifica la importancia de la información nueva para actualizar la memoria a largo plazo. (11) Primero, se aplica la función sigmoide (σ) al estado oculto anterior y a la entrada actual. Esto determina qué valores se deben de actualizar, transformando valores de 0 a 1. Un valor 0 significa que no es importante, 1 significa que es importante (Formula 2.3.1). Luego, el estado oculto y la entrada actual se pasan a través de la función \tanh , que los comprime entre -1 y 1 para

regular la red. Finalmente, la salida de tanh se multiplica por la salida sigmoide. La salida sigmoidea determinará qué información es importante y debe ser guardada por la salida tanh. La fórmula correspondiente es la siguiente:

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (2.3.1)$$

$$\tilde{c}_t = \tanh(W_c * [h_{t-1}, x_t] + b_c)$$

El estado de la celda anterior cambia según el grado de decisión para actualizar. Luego se suma \tilde{c}_t que es el nuevo estado de la celda, que cambia según el grado de decisión a actualizar. (12) Esta información permite determinar la memoria a largo plazo con datos actuales, formula presentada en la figura No. 4

- Compuerta de Salida: Esta compuerta controla qué información de la memoria a largo plazo (datos actuales) se hace visibles. El estado oculto anterior y la entrada actual se pasan a través de la función sigmoide (12):

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

Esto permite calcular la memoria a corto plazo con los datos actuales, como se muestra en la figura No. 4.

- LSTM Bidireccional: La arquitectura LSTM bidireccional se utiliza para procesar secuencias de datos en ambas direcciones: desde el pasado hacia el futuro y desde el futuro hacia el pasado. Esta arquitectura combina dos capas LSTM, una que recorre la secuencia en orden ascendente y otra que la recorre en orden descendente. Al hacerlo, la red neuronal puede capturar tanto la información contextual pasada como la futura, lo que permite un mejor modelado de las dependencias en la secuencia de datos. (13)
- Attention LSTM: La arquitectura Attention LSTM utiliza un mecanismo de atención para enfocarse en partes específicas de la información de entrada. Este enfoque

mejora la capacidad del modelo para capturar y centrarse en las partes relevantes de una secuencia de entrada. En lugar de tratar todas las partes de la secuencia por igual, el mecanismo de atención permite al modelo asignar diferentes pesos o niveles de importancia a diferentes partes de la secuencia, lo que mejora la capacidad del modelo para procesar de manera más efectiva secuencias largas y complejas. (14)

1.4.6 METRICAS DE EVALUACION EN LOS MODELOS

En esta sección se abordarán las métricas de evaluación en los modelos, es esencial ya que se busca medir su rendimiento y determinar su eficacia en la resolución de tareas específicas, además de elegir que modelo es el recomendable para el problema.

- Error Absoluto Medio (MAE, siglas en inglés): mide la magnitud promedio de los errores en un conjunto de predicciones, sin considerar su dirección. Es el promedio sobre la muestra de prueba de las diferencias absolutas entre la predicción y la observación real donde todas las diferencias individuales tiene el mismo peso.

$$MAE(\hat{y}, y) = \frac{1}{n} \sum_{k=1}^n (\hat{y}_k - y_k)$$

- Error Cuadrático Medio (MSE, siglas en inglés): medida de la calidad de un estimador, mide el promedio de los cuadrados de los errores es decir la diferencia cuadrada promedio entre los valores estimados y el valor real (15).

$$(\hat{y}, y) = \frac{1}{l - \tau - H} \sum_{k=\tau-H}^{lMSE} (\hat{y}_k - y_k)^2$$

- Raíz del Error Cuadrático Medio (RMSE, sigla en inglés): representa la raíz cuadrada del segundo momento muestral de las diferencias entre los valores predichos y los valores observados. Frecuentemente utilizadas para medir las diferencias entre valores estimados y el valor real (16)

$$RMSE(\hat{y}, y) = \sqrt{\frac{1}{l - \tau - H} \sum_{k=\tau-H}^l (\hat{y}_k - y_k)^2}$$

- Error cuadrático medio normalizado (NRMSE, por sus siglas en inglés) es una medida que facilita la comparación entre conjuntos de datos que facilita la comparación entre conjuntos de datos o modelos de diferentes escalas. Aunque no hay método consistente de normalización en la literatura las opciones comunes son el medio o el rango (definido como el valor máximo menos el valor mínimo) de los datos medios (17)

$$NRMSE(\hat{y}, y) = \sqrt{\frac{1}{l - \tau - H} \sum_{i=\tau-H}^l \left(\frac{y_k - \hat{y}_k}{y_{max}} \right)^2}$$

- Coeficiente de Pearson: es una medida de la dependencia lineal entre dos variables. (18) $E[(y - \mu_y)(\hat{y} - \mu_{\hat{y}})]$

$$\rho_{y, \hat{y}} = \frac{E[(y - \mu_y)(\hat{y} - \mu_{\hat{y}})]}{\sigma_y \sigma_{\hat{y}}}$$

1.4.7 PROYECTOS EN LA ACTUALIDAD Y SUS METODOLOGIAS

Este proyecto de investigación, llevado a cabo por Zhang, Jiang y Yu (2015), analiza el consumo de energía en edificios mediante la aplicación de técnicas de pronóstico, métodos de ingeniería y métodos estadísticos. El objetivo principal del proyecto es crear un modelo de predicción del consumo de energía eléctrica en edificios, que proporcione información útil para la gestión y optimización del consumo energético.

Con el objetivo de lograr esto, el modelo emplea diversas variables, como datos meteorológicos (temperatura, humedad) y variables de carga eléctrica del edificio (potencia activa, potencia reactiva). Sin embargo, no todas estas variables son relevantes para la predicción del consumo eléctrico, se emplea el análisis de componentes principales (PCA) para seleccionar las variables más significativas y reducir la dimensionalidad del conjunto de datos. El modelo de predicción del consumo de energía eléctrica en edificios utiliza redes neuronales artificiales como base. Se eligió este enfoque debido a que las redes neuronales artificiales han demostrado ser una técnica de aprendizaje automático altamente eficaz en la predicción de series de tiempo y la resolución de problemas complejos. Además, este tipo de modelo es capaz de manejar múltiples entradas, lo que lo hace adecuado para la predicción del consumo de energía eléctrica en edificios, donde hay múltiples factores que pueden afectar el consumo. Posteriormente, se emplea una red neuronal artificial (ANN) para realizar la predicción del consumo eléctrico. "La ANN se optimiza utilizando algoritmos genéticos para encontrar los valores óptimos de los parámetros de la red neuronal, tales como el número de neuronas, la tasa de aprendizaje y el número de épocas de entrenamiento.

El modelo propuesto por Zhang, Jiang y Yu se evaluó en un edificio de oficinas en España, y los resultados demuestran una mejora significativa en la precisión de la predicción del consumo eléctrico en comparación con otros modelos de referencia. Además, el modelo es capaz de identificar patrones en los datos, lo que proporciona a los administradores de edificios la capacidad de tomar decisiones informadas para la gestión y optimización del consumo energético. Además del artículo anterior, existe un trabajo, en el que se compartes el desarrollo e implementación de un sistema de monitoreo y predicción de consumo energético en edificios de oficinas. En este trabajo, se describe el desarrollo de un sistema de monitoreo y predicción del consumo energético en un edificio de oficinas mediante la utilización de un sistema de adquisición de datos, se realiza el procesamiento de los datos adquiridos utilizando técnicas de regresión y se emplean redes neuronales artificiales para la predicción del consumo

energético. El estudio proporciona una descripción detallada de los diversos componentes del sistema, junto con las técnicas empleadas para el procesamiento y análisis de los datos, la construcción y entrenamiento de las redes neuronales y la implementación del sistema en un edificio de oficinas. El enfoque central del estudio radica en el desarrollo de un modelo de predicción de la demanda energética en un edificio no residencial utilizando técnicas de inteligencia artificial, específicamente redes neuronales (ANN) y regresión lineal múltiple (MLR) (3).

En otros trabajos, se ha implementado un enfoque híbrido para la predicción de la demanda de electricidad a corto plazo utilizando una combinación de modelos estadísticos y de aprendizaje automático, incluyendo un modelo de regresión lineal múltiple (MLR) (3). Durante el proceso de modelado de la relación entre la demanda de electricidad y las variables climáticas, se recopilan y analizan los datos históricos de ambas variables. Posteriormente, se emplea un modelo de regresión lineal múltiple para establecer la relación entre la demanda de la electricidad y las variables climáticas. En este caso, la variable dependiente es la demanda de electricidad, mientras que las variables independientes son las variables climáticas. Al aplicar este modelo, se pretende determinar la influencia que las variables climáticas ejercen sobre la demanda de electricidad del edificio. Después de aplicar el modelo de regresión lineal múltiple (MLR) para modelar la relación entre la demanda de electricidad y las variables climáticas, los autores aplicaron un modelo de redes neuronales artificiales (ANN) para modelar la relación entre la demanda de electricidad y la hora del día, el día de la semana y la temperatura exterior. En otras palabras, el modelo de redes neuronales artificiales (ANN) se empleó para modelar las relaciones no lineales y complejas entre la demanda de electricidad y los factores que no se puede representar de manera adecuada mediante un modelo lineal. Posteriormente, se combinarán los resultados de ambos modelos para obtener una predicción híbrida de la demanda de electricidad.

Un tercer trabajo que merece ser mencionado (19) habla acerca del estudio comparativo que se desarrolla para la predicción del corto plazo del consumo energético se utilizaron

cinco técnicas de pronóstico diferentes, incluyendo el modelo ARIMA, el modelo SARIMA, el modelo de redes neuronales artificiales (ANN), el modelo de regresión de vectores de soporte (SVR) y el modelo de bosques aleatorios (RF). Se recopilieron datos de consumo de energía eléctrica de un edificio inteligente real en intervalos de 15 minutos durante un período de 2 años, y se utilizaron para entrenar y validar los modelos de pronóstico.

En general, el estudio concluyó que los modelos de ANN, SVR y RF son más adecuados para la predicción de corto plazo del consumo de energía eléctrica en edificios inteligentes. Sin embargo, también se destacó la necesidad de considerar otros factores, como la precisión de los sensores de medición de energía y las condiciones ambientales, para lograr una precisión aún mayor en la predicción del consumo de energía eléctrica en edificios inteligentes.

Todos estos autores concuerdan (19; 20; 3) en características de los datos para los modelos de predicción del consumo energético de edificios no residenciales incluyen:

- Características del edificio: es importante conocer las características físicas y operativas del edificio, como su tamaño, orientación, ubicación geográfica, materiales de construcción, sistemas de climatización e iluminación, para poder modelar su comportamiento energética, donde las características de los edificios que son importantes para la construcción de modelos de predicción, como el tamaño del edificio, la orientación, la ubicación, la zona climática, la envolvente del edificio, y los sistemas de calefacción, ventilación y aire acondicionado (20)
- Datos históricos de consumo: los datos de consumo energético históricos son esenciales para entrenar y validar los modelos de predicción. Estos datos deben ser precisos y representativos de las condiciones de uso y del clima en los que se ha producido el consumo, recogidos de sensores que están instalados dentro del

edificio, llevando un registro en intervalos de tiempo. En cuanto a las variables atributo a tomar en cuenta (21) para construir modelos son:

- Temperatura ambiente (°C) ○ Radiación Solar (W/m) ○ Velocidad del viento (m/s) ○ Dirección del viento
 - Presión ambiental (hPa) ○ Humedad Relativa del ambiente (%) ○ Intensidad de la lluvia (mm/Hr)
 - Temperatura media dentro del edificio (°C) ○ Humedad relativa media dentro del edificio (%) ○ Nivel medio de CO₂ dentro del edificio (ppm) ○ Ocupación total (No. De personas)
 - Tipo de día (Semana, entre semana y/o fiestas nacionales)
- Variables meteorológicas: las variables meteorológicas, como la temperatura, la humedad, la velocidad del viento, la radiación solar, son importantes para modelar el comportamiento térmico del edificio y para ajustar el modelo a las condiciones climáticas locales. Las siguientes variables atributo se han de usar para construir el modelo (22):
 - Temperatura ambiente (°C) ○ Radiación Solar (W/m²) ○ Velocidad media del viento (m/s) ○ Humedad relativa media (%) ○ Consumo de agua (m³) ○ Temperatura media (°C) ○ Consumo eléctrico (kWh/d)
 - Tipo de día (Semana, entre semana y/o fiestas nacionales)
- Técnicas de preprocesamiento de datos: se emplean técnicas de preprocesamiento de datos, tales como la normalización, la reducción de la dimensionalidad y la selección de variables, se realiza la

caracterización de las señales, el estudio de la estacionalidad y la detección de eventos anómalos, entre otros, con el objetivo de mejorar la calidad de los datos de entrada y disminuir la complejidad del modelo.

- Modelos de aprendizaje automático: se utilizan modelos de aprendizaje automático, como las redes neuronales, las máquinas de vectores de soporte, los árboles de decisión, para modelar las relaciones entre las variables de entrada y salida y para realizar la predicción del consumo energético.
- Validación y evaluación del modelo: es importante validar y evaluar el modelo de predicción para determinar su precisión y su capacidad para generalizar a nuevos conjuntos de datos. Se utilizan técnicas de validación cruzada, métricas de evaluación, como el error cuadrático medio y el coeficiente de determinación en los datos usado es test y train, para evaluar la calidad del modelo (22)

1.5 HERRAMIENTAS UTILIZADAS

En el proyecto, se emplearon diversas herramientas y librerías en Python para el análisis y modelado de datos. Entre ellas se encuentran pandas, matplotlib, seaborn, numpy, keras, scikit-learn, scipy y tensorflow.

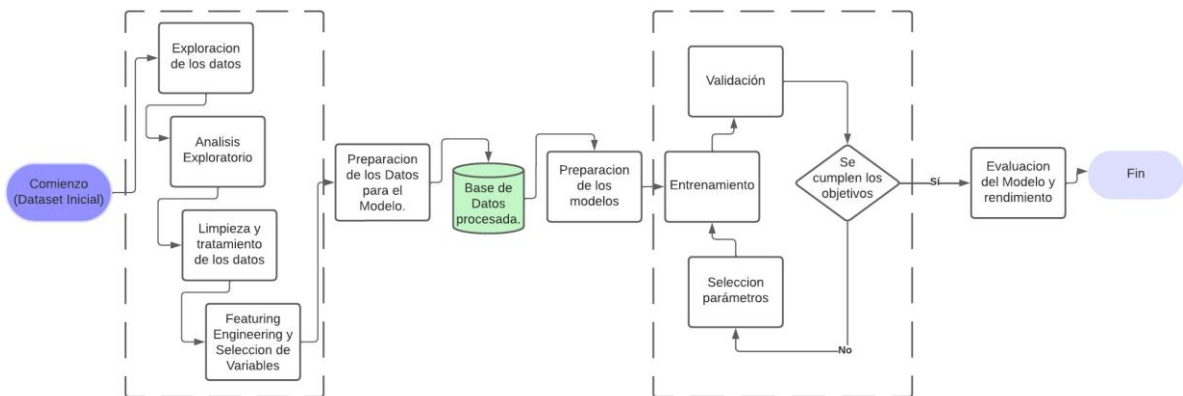
Pandas fue utilizado para el procesamiento y manipulación de datos tabulares, ofreciendo funciones para filtrado, transformación y limpieza de datos. Matplotlib y seaborn se utilizaron para la visualización de datos, creando gráficos informativos y visualmente atractivos.

En cuanto al modelado de datos, keras y tensorflow se utilizaron para el desarrollo de redes neuronales y aprendizaje automático, permitiendo la construcción y entrenamiento de modelos eficientes. Además, scikit-learn proporcionó herramientas para el preprocesamiento de datos, imputación de valores faltantes y evaluación de modelos.

Por otro lado, scipy ofreció funcionalidades estadísticas, como pruebas y cálculos de correlación. El código relacionado con el documento se subió a un repositorio de Github.¹

1.6 DIAGRAMA DE FLUJO

Se presenta en la siguiente figura el proceso que se rige este proyecto, desde el comienzo hasta el final con el dataset.



2. ANALISIS Y ESTRUCTURA DE LA BASE DE DATOS.

La estructura de los Datos presentados en este proyecto proviene de la empresa ASHRAE, una organización internacional dedicada a la tecnología y la investigación en calefacción, refrigeración, ventilación y calidad del aire en edificios, establece estándares y guías para la eficiencia energética y el diseño sostenible en la industria de la construcción. El conjunto de datos está disponible para el público y consta de más 3,053

medidores de energía (energy meters) recolectado de 1636 edificios. El rango temporal de los datos en la serie temporal es de dos años completos (2016 y 2017) tomando las muestras en una frecuencia horaria por hora para cada sensor que se nombra a continuación:

- Electricidad (Electricity)
- Calor (Heating)
- Agua de Enfriamiento (Cooling Water)
- Vapor (Steam)
- Solar (Solar)
- Medidores de riego (Irrigation Meters)
- Gas (Gas)

Cada sensor posee sus propias bases de datos identificada en el tiempo, asociada al edificio que se midió. A continuación, también se encuentra la base de datos llamada Metada, la descripción de la estructura de la base de datos es la siguiente:

Tabla No. 1

Descripción de Variables Metadata

Nombre de la Variable	Descripción de la Variable
building_id	Código o nombre del edificio con la estructura [SiteID_SimplifiedUsage_UniqueName.]
site_id	Código o nombre del sitio según la nomenclatura animal.
building_id_kaggle	ID del edificio utilizado en la competición de Kaggle (valor numérico).
site_id_kaggle	ID del sitio utilizado en la competición de Kaggle (valor numérico).
primaryspaceusage	Uso principal del espacio del edificio asignado según el esquema de descripción de tipos de edificios Energystar.
sub_primaryspaceusage	Subcategoría de los tipos de descripción de edificios Energystar.
sqm	Área del suelo del edificio en metros cuadrados (m2).
lat	Latitud de la ubicación del edificio a nivel de ciudad.
lng	Longitud de la ubicación del edificio a nivel de ciudad.
timezone	Zona horaria del sitio.
electricity	Presencia de medidor de electricidad en el edificio. "Sí" si es afirmativo, "NaN" si es negativo.

hotwater	Presencia de medidor de agua caliente en el edificio. "Sí" si es afirmativo, "NaN" si es negativo.
chilledwater	Presencia de medidor de agua fría en el edificio. "Sí" si es afirmativo, "NaN" si es negativo.
steam	Presencia de medidor de vapor en el edificio. "Sí" si es afirmativo, "NaN" si es negativo.
water	Presencia de medidor de agua en el edificio. "Sí" si es afirmativo, "NaN" si es negativo.
irrigation	Presencia de medidor de riego en el edificio. "Sí" si es afirmativo, "NaN" si es negativo.
solar	Presencia de medidor solar en el edificio. "Sí" si es afirmativo, "NaN" si es negativo.
gas	Presencia de medidor de gas en el edificio. "Sí" si es afirmativo, "NaN" si es negativo.
industry	Tipo de industria correspondiente al edificio.
subindustry	Desglose más detallado del tipo de industria correspondiente al edificio.
heatingtype	Tipo de calefacción en el edificio.
yearbuilt	Año en que se construyó el edificio por primera vez (formato YYYY).
date_opened	Fecha en que se abrió el edificio para su uso (formato D/M/YYYY).
numerooffloors	Número de pisos del edificio.
occupants	Número habitual de ocupantes en el edificio.
energystarscore	Calificación del edificio según el esquema de Energystar (Energy Star Score).
eui	Intensidad de uso de energía del edificio en kWh/año/m2.
site_eui	Intensidad de uso de energía (consumida/comprada) del sitio en kWh/año/m2.
source_eui	Consumo total de energía primaria normalizado por área, teniendo en cuenta la eficiencia de conversión de energía primaria a energía secundaria.
leed_level	Clasificación LEED del edificio (Leadership in Energy and Environmental Design).

Por último, está la base de datos llamada Weather, que refleja las condiciones climáticas para cada edificio durante los años 2016 al 2017, la descripción de la estructura de datos es la siguiente:

TABLA No. 2

Descripción de Variables Weather.

Nombre de la Variable	Descripción de la Variable
timestamp	fecha y hora en formato AAAA-MM-DD hh:mm: ss. Zona horaria local.

site_id	código/nombre animal para el sitio.
airTemperature	La temperatura del aire en grados Celsius (°C).
cloudCoverage	porción del cielo cubierta por nubes, en octas.
dewTemperature	el punto de rocío (la temperatura a la cual un determinado volumen de aire debe enfriarse a presión constante y contenido de vapor de agua constante para que ocurra la saturación) en grados Celsius (°C).
precipDepth1HR	la cantidad de precipitación líquida medida durante un período de acumulación de una hora (mm).
precipDepth6HR	la cantidad de precipitación líquida medida durante un período de acumulación de seis horas (mm).
seaLvlPressure	la presión atmosférica relativa al nivel medio del mar (mbar o hPa).
windDirection	el ángulo, medido en dirección de las agujas del reloj, entre el norte verdadero y la dirección desde la cual sopla el viento (grados).
windSpeed	la velocidad de desplazamiento horizontal del aire en relación a un punto fijo (m/s).

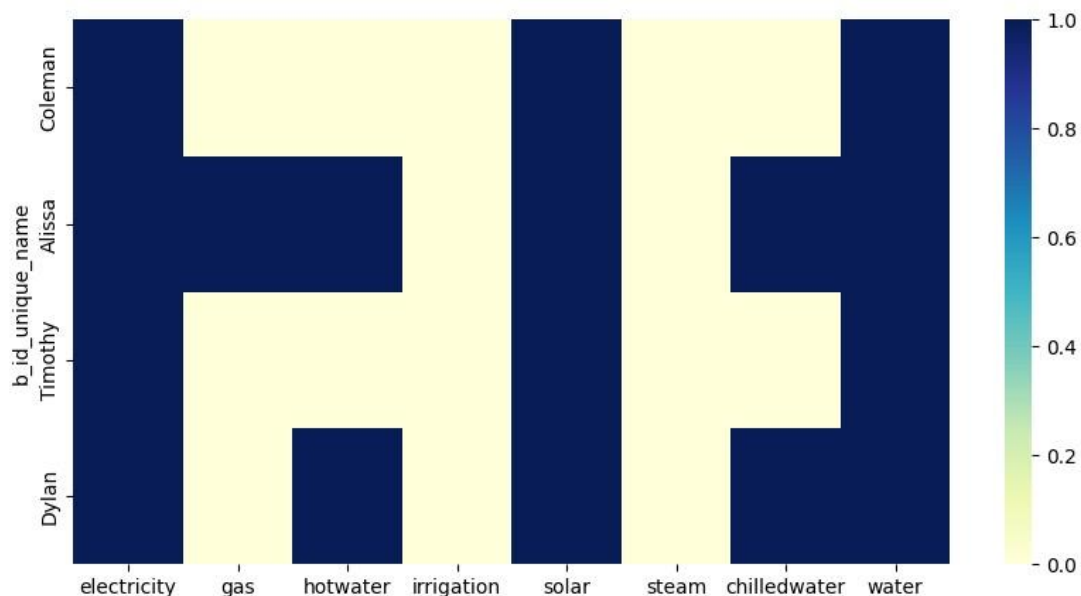
2.1. PROCESAMIENTO DE LA BASE DE DATOS:

Se llevó a cabo un análisis utilizando los datos de la metada con el objetivo de determinar el edificio que presentaba la mayor cantidad de medidores de energía. Tras examinar los registros, se identificaron cuatro edificios distintos: ['Alissa', 'Timothy', 'Dylan', 'Coleman'].

En la imagen No.1 representa cada sensor, donde estar en blanco significa no tener lecturas sobre el sensor y azul, tener lecturas para el sensor, esto por los edificios seleccionados.

Imagen No. 1

Mapa de Calor de Sensores.



Posteriormente, se procedió a rastrear cada uno de estos edificios en las diversas bases de datos que albergaban información relacionada con los diferentes tipos de medidores de energía utilizados en el estudio. Además, se accedió a la base de datos del clima (weather) con el fin de obtener datos sobre las condiciones climáticas asociadas a cada uno de los edificios durante el periodo de tiempo analizado. Como resultado de este proceso, se generó un nuevo conjunto de datos que se centraba en un edificio de Timothy seleccionado previamente de entre los cinco identificados. Este enfoque nos permitió obtener una perspectiva más específica y detallada de dicho edificio en particular, lo cual resulta relevante para el análisis y estudio ulterior de los datos recopilados. El Dataset recopila las siguientes variables en la unificación de los tres dataset: 'timestamp', 'electricity', 'solar', 'water',

'airTemperature', 'cloudCoverage', 'dewTemperature', 'precipDepth1HR', 'seaLvIPressure', 'windDirection', 'windSpeed'

2.2. ANALISIS ESTADISTICO DE LOS DATOS

Después de la creación del dataset se procede a una exploracion de los datos para el Edificio Timothy. El DataFrame tiene las siguientes características:

- Número de filas: 17525

- Número de columnas: 11

Resumen de las columnas más relevantes:

- timestamp: marca de tiempo (no nula en todas las filas)
- electricity: consumo de electricidad (nulos en 25 filas)
- solar: generación de energía solar (nulos en 20 filas)
- water: consumo de agua (nulos en 104 filas)
- airTemperature: temperatura del aire (nulos en 50 filas)
- cloudCoverage: cobertura de nubes (nulos en 4949 filas)
- dewTemperature: temperatura de rocío (nulos en 53 filas)
- precipDepth1HR: profundidad de precipitación en 1 hora (nulos en 61 filas)
- seaLvlPressure: presión al nivel del mar (nulos en 270 filas)
- windDirection: dirección del viento (nulos en 1069 filas)
- windSpeed: velocidad del viento (nulos en 165 filas)

2.3. EXPLICACION DEL PROCESO DE LIMPIEZA

La preparación de la data para un modelo con buenas métricas es importante, en esta etapa se identificaron los valores atípicos, valores duplicados, valores faltantes y cualquier otro problema con respecto a tipo de dato que se presentara.

- Valores Faltantes:

Al analizar los valores faltantes en el DataFrame proporcionado, se observa que existen columnas con diferentes cantidades de valores no nulos. Estos valores faltantes pueden afectar el análisis estadístico de las variables y deben ser tratados adecuadamente. Se pueden utilizar técnicas de imputación, como la sustitución por la media o la mediana, para llenar los valores faltantes y garantizar la integridad de los datos antes de realizar cualquier análisis estadístico. Además, es importante considerar la naturaleza y el contexto de los datos para determinar la estrategia de imputación más adecuada en cada caso. Observando la variable "Water", se identifica que tiene un 15.11% de valores nulos en el conjunto de datos y la VARIABLES "precipDepth1HR" con un 10.48%

Tabla No. 3

Porcentaje de Valores Nulos

Porcentaje de valores nulos	
Variable	Porcentaje
Electricity	0.13%
Solar	7.26%
Water	15.11%
AirTemperature	0.00%
CloudCoverage	9.62%
DewTemperature	1.05%
PrecipDepth1HR	10.48%
SeaLvlPressure	2.34%
Wind X	2.08%
Wind Y	3.53%

- Valores Atípicos

Los valores atípicos pueden influir en el rendimiento y la precisión de los modelos de análisis y predicción. La representatividad de los valores se ve reflejada en el Anexo (Tabla de Valores Atípicos).

Basado en los resultados del análisis, se puede concluir que la representatividad de los valores atípicos en varias columnas del conjunto de datos es relativamente baja. Las columnas "electricity", "airTemperature" y "windDirection" tienen una representatividad de valores atípicos del 0.0% relativamente baja, por lo que se decide mantener los valores atípicos.

2.4. DEPENDENCIA ENTRE LAS VARIABLES

El análisis de la matriz de correlación revela las relaciones lineales entre las variables estudiadas. Se utilizó tanto un diagrama de dispersión como el diagrama de correlación para evaluar estas relaciones. El diagrama de dispersión permite visualizar la distribución conjunta de dos variables y detectar patrones generales. Sin embargo, el diagrama de correlación ofrece una medida cuantitativa de la fuerza y dirección de la relación lineal entre las variables.

En base a la matriz de correlación, se identificaron las siguientes relaciones significativas:

- La variable "electricity" tiene una correlación negativa fuerte (-0.5) con la variable "solar" y una correlación negativa débil (-0.47) con la variable "airTemperature". Esto indica que a medida que aumenta la generación de electricidad solar, disminuye el consumo de electricidad. Además, a medida que aumenta la temperatura del aire, disminuye el consumo de electricidad.
- La variable "solar" muestra una correlación positiva débil (0.36) con la variable "water" y una correlación positiva moderada (0.5) con la variable "airTemperature". Esto sugiere que a medida que aumenta la generación de energía solar, también aumenta el consumo de agua y la temperatura del aire.
- La variable "airTemperature" tiene una correlación positiva fuerte (0.68) con la variable "dewTemperature". Esto indica una relación directa entre la temperatura del aire y la temperatura de rocío, lo cual es esperado ya que la temperatura de rocío depende de la humedad presente en el aire.
- La variable "sealLvlPreassure" tiene una correlación negativa moderada fuerte (-

0.55) con la variable "airTemperature". Esto indica una relación directa entre el nivel de sellado y la temperatura del aire, sugiriendo que es probable que la presión de nivel de sellado disminuya. Con "dewTemperature" tiene una correlación moderada a fuerte negativa (-0.54) sugiere que a medida que aumenta esta variable la presión de nivel de sellado disminuya.

- Las demás correlaciones son relativamente bajas y no presentan patrones notables

2.5. FEATURING ENGINEERING

En el proceso de construcción de modelos de aprendizaje, es común buscar formas de mejorar el rendimiento del modelo mediante la creación de nuevas variables o la transformación de variables existentes en un conjunto de datos. En este caso particular, se busca lograr una representación adecuada de las variables "windSpeed" y "windDirection" al combinarlas en un vector bidimensional. Esta representación en forma de vector proporcionará una visión más completa y concisa de la información relacionada con el viento.

En la Figura no. 5, la representación de "windSpeed" vs "windDirection" se observa de manera inconsistente y no logramos obtener mucha información clara sobre la relación entre estas variables. La dispersión de los puntos puede dificultar la interpretación y comprensión de los patrones o tendencias.

Por otro lado, en la Figura no. 6, se representa el vector de viento utilizando las variables "Wx" y "Wy". Estas variables se interpretan como las componentes de velocidad del viento en dirección x (horizontal) y dirección y (vertical), respectivamente. Esta representación en forma de vector simplifica la visualización y comprensión de la relación entre la velocidad y la dirección del viento. Proporciona una representación más clara y concisa de la información del viento, lo que puede ser beneficioso para el modelo y facilitar su interpretación.

Figura No. 5

Wind Direction vs Wind Velocity

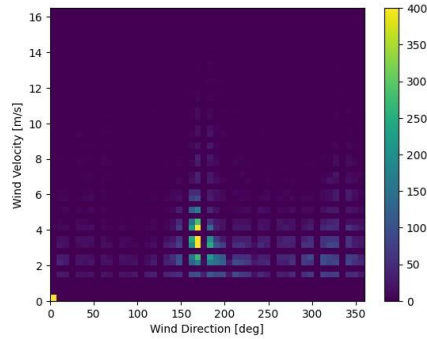
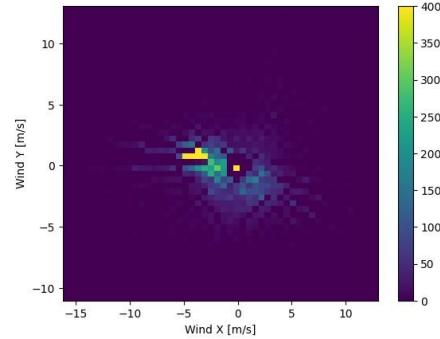


Figura No. 6

Componentes de "Wind" en X y Y



2.6. COMPONENTE DE SERIE-TEMPORAL:

El análisis de las variables reveló que no cumplen con la propiedad de varianza constante. Esto implica que la varianza de estas variables no se mantiene constante a lo largo del tiempo y presentan fluctuaciones. La falta de varianza constante puede deberse a diversos factores, como cambios estacionales, patrones cíclicos o tendencias en los datos, como se muestra en los Anexos. Al descomponer la serie temporal se comprende lo siguiente de cada variable:

- 'electricity':
 - Tendencia: La tendencia de la serie es no lineal, se observan cambios abruptos en la tendencia en momentos específicos, además de existir puntos de inflexión o momentos significativos que podrían estar relacionados con el cambio de estación y/o vacaciones en ese periodo.
 - Estacionalidad: sigue un patrón regular y se repite cada hora, la amplitud de la estacionalidad se mantiene constante entre -5 y 5, lo cual indica una variación periódica en la serie de tiempo.
 - Residuos: No se observa patrón o comportamiento sistemático en los residuos, lo que sugiere que son impredecibles y no están relacionados con la tendencia o la estacionalidad. Los residuos son particularmente bajos en el periodo donde hay cambios bruscos de la serie temporal.

- 'airTemperature'
 - Tendencia: No es lineal y presenta cambios ascendentes y descendente a lo largo del tiempo, especialmente en los meses de enero a junio (ascienden) en los años 2016 y 2017.
 - Estacionalidad: Sigue un patrón regular, con fluctuaciones dentro del rango de -5 a 5, la amplitud es constante y no varía en el tiempo.
 - Residuos: parece seguir un patrón aleatorio dentro del rango de -10 a 10, aun que hay valores atípicos o anomalías en los residuos durante los períodos de agosto del 2016 y enero del 2017.

- 'solar'
 - Tendencia: No es lineal y presenta cambios conforme al transcurso del año, especialmente en los primeros meses del año 2017, se observa los cambios en especial cuando es de noche.
 - Estacionalidad: Sigue un patrón regular, con fluctuaciones dentro del rango de mayor a 5 a un menos a 0, la amplitud es constante y no varía en el tiempo.

-
- Residuos: parece seguir un patrón aleatorio dentro del rango de -5 a 5, aunque hay valores atípicos o anomalías en los residuos durante los meses de junio y Julio en ambos años, tiene a tener una forma aleatoria.
- 'cloudCoverage' ○ Tendencia: No es lineal y presenta cambios bruscos con punto de inflexión en aquellos meses que se presenta mayores nubes como enero y febrero.
 - Estacionalidad: Sigue un patrón regular, con fluctuaciones dentro del rango de mayor a 0.1 a un menos a -0.1, la amplitud es constante y no varía en el tiempo.
 - Residuos: Sin tomar en cuenta los meses con mayor presencia de nubes, parecieran se aleatorias, cuando existen esos cambios bruscos los puntos tiende a subir por encima de 5 o a bajar muy cercano a -5, mantiene aleatoriedad, pero se ve afectado por la serie temporal.
- 'dewTemperature' ○ Tendencia: no es lineal a largo plazo, tiene una tendencia al alza, aumentando gradualmente en el tiempo, especialmente aumenta en los meses de junio y julio de ambos años, y desciende en los meses de diciembre y enero de ambos años.
 - Estacionalidad: patrones regulares y predecibles que se repiten hora con hora. Hay una clara variación estacional en la temperatura del rocío, durante los meses de invierno, la temperatura será baja y a medida que cambia de estación esta aumentará,
 - Residuos: se comporta de manera impredecible, por lo que se observa no afecta de manera significativa, aumenta en los meses de junio y julio de ambos años, pero mantiene la aleatoriedad.

-
- 'precipDepth1HR' ○ Tendencia: en este caso, la componente de tendencia es igual a la serie temporal, se comporta de manera no lineal y presenta cambios conforme al transcurso del año.
 - Estacionalidad: Existe un patrón estacional bien definido, este componente capturará la variación en la precipitación con los cambios de hora.
 - Residuos: representa fluctuaciones con variabilidad no sistemática, estos cambios pueden reflejar factores externos, resaltando la importancia de comprender esos eventos inusuales.
- 'seaLvlPressure' ○ Tendencia: en este caso, la componente de tendencia es igual a la serie temporal, se comporta de manera no lineal y presenta cambios conforme al transcurso del año.
 - Estacionalidad: Existe un patrón estacional bien definido, este componente capturará la variación en la precipitación con los cambios de hora.
 - Residuos: representa fluctuaciones con variabilidad no sistemática, estos cambios pueden reflejar factores externos, resaltando la importancia de comprender que pasa en ese tiempo en la serie temporal.
- 'water'

Tendencia: No es lineal y presenta cambios conforme al transcurso del tiempo, cambios abruptos, tiene a tener subida y bajas en periodos cortos en el tiempo.

 - Estacionalidad: Existe un patrón estacional bien definido, este componente capturará la variación en la precipitación con los cambios de hora.
 - Residuos: Los errores no siguen una distribución, se observa un comportamiento cíclico, indicando la presencia de patrones o estructura sistemáticas en los datos que no son aleatorias.

○

2.7. PREPARACION DE LOS DATOS PARA LOS MODELOS

Para preparar los datos de la serie temporal de consumo energético, se realiza un proceso de preprocesamiento y transformación. Primero, se aplica el escalado de los datos utilizando MinMaxScaler, que normaliza los valores de consumo en un rango específico, preservando así la estructura relativa de la serie. Esto es importante para garantizar que todas las características de la serie tengan un impacto similar durante el entrenamiento del modelo. A continuación, se crea un tensor 3D utilizando ventanas deslizantes de tamaño 6, 12 y 24. Estas ventanas representan el número de pasos de tiempo utilizados para predecir el siguiente valor en la serie. Finalmente, se divide el conjunto de datos en dos conjuntos: entrenamiento el 75% de los datos y prueba el 25% de los datos. El conjunto de entrenamiento se utiliza para ajustar los parámetros del modelo, mientras que el conjunto de prueba se utiliza para evaluar el rendimiento y la generalización del modelo en datos no vistos.

2.8. ARQUITECTURA DE LOS MODELOS

Se presenta las diferentes arquitecturas para la red neuronal recurrentes utilizadas para pronosticar el consumo energético en un edificio no residencial. El modelo genérico, utilizado como base es el siguiente, donde la mayoría se utilizará

como optimizador Adam ya que combina los beneficios de descenso de gradiente estocástico con la idea de adaptación de tasas de aprendizaje, teniendo control de la misma siendo adaptativa para cada red neuronal.

Tabla No. 3

Parámetros del Modelo Base

Modelo Base		
Layer	Units	Activación
LSTM (1)	75	tanh
LSTM (2)	50	tanh
Dense	1	lineal*
Loss	MeanSquaredError	
Optimizador	Adam	(lr: 0.001, 0.0001)
*por defecto en la función		

Desarrollando Este como modelo base, definiendo los criterios de selección como se observa en la Tabla No. 3, como se observa solamente contiene dos capas de LSTM, cada una con distinta unidad, los parámetros locales compartidos para cada capa son

- Función de Activación para LSTM: Tangente hiperbólica (tanh)
- Profundidad de Salida: Dense con función de activación “Lineal”, esto tendrá solo una salida.
- Optimizador: Adam con una tasa de aprendizaje de 0.001 y 0.0001

Este modelo esta diseñado para adaptarse a diferentes horizontes, en este caso se estarán incluyendo ventanas de 6 horas, 12 horas y 24 horas.

Tabla No. 4

Parámetros del Modelo LSTM con diferentes unidades

Modelo LSTM arq 2		
Layer	Units	Activation
LSTM (1)	64	tanh
LSTM (2)	64	tanh
Dense	1	lineal*

Loss	MeanSquaredError
Optimizador	Adam (lr: 0.001 y 0.0001)
*por defecto en la función	

El siguiente modelo a utilizar es una arquitectura con LSTM cuadrada como se observa en la Tabla no. 4, compuesta por dos capas LSTM. Cada capa LSTM cuenta con 64 unidades, lo que significa que hay un total de 128 unidades LSTM en el modelo. Estas unidades LSTM tienen la capacidad de capturar y recordar información relevante de las secuencias de entrada a largo plazo, los parámetros locales compartidos en cada capa son:

- Función de Activación para LSTM: Tangente hiperbólica (tanh)
- Profundidad de Salida: Dense con función de activación “Lineal”, esto tendrá solo una salida.
- Optimizador: Adam con una tasa de aprendizaje de 0.001 y 0.0001 Este modelo solo se adaptará a un horizonte.

Tabla No. 5

Modelo LSTM con capa Bidireccional

Modelo Bidireccional		
Layer	Units	Activation
LSTM Bi (1)	64	tanh
Dense	1	lineal*

Loss	MeanSquaredError
Optimizador	Adam (lr: 0.001, 0.0001)
*por defecto en la función	

La arquitectura de este modelo se basará en un modelo bidireccional, como se muestra en la Tabla No. 5. Estará compuesto por una sola capa con 64 unidades y con una capa de salida, se especifican los siguientes parámetros locales dentro de la arquitectura:

- Función de Activación para LSTM: Tangente hiperbólica (tanh)
- Profundidad de Salida: Dense con función de activación “Lineal”, esto tendrá solo una salida.
- Optimizador: Adam con una tasa de aprendizaje de 0.001 y 0.0001 Este modelo solo se adaptará a un horizonte.

Con esta configuración, el modelo bidireccional se encargará de capturar patrones complejos y dependencias tanto en el pasado como en el futuro de los datos de entrada. La función de activación tangente hiperbólica (tanh) se encargará de introducir no linealidad en el aprendizaje de las LSTM.

Tabla No. 6

Modelo LSTM con capa de Attention

Modelo Attention (+)LSTM		
Layer	Units	Activation
LSTM (1)	128	tanh
LSTM (2)	64	tanh
Attention	[LSTM (1), LSTM (2)]	
Dense	1	lineal*

Loss	MeanSquaredError
Optimizador	Adam (lr: 0.001 y 0.0001)
*por defecto en la función	

La siguiente arquitectura seleccionara el modelo de la serie temporal esta conformada por dos capas de LSTM donde la primera LSTM tiene 128 unidades y la segunda posee 64 unidades, seguidas de una capada de atención (Attention) puesto el enfoque en ambas capas y, por último, una capa de salida densa, se especifican los siguientes parámetros locales para la arquitectura:

- Función de Activación para las capas LSTM: Tangente hiperbólica (tanh)
- Profundidad de Salida: Dense con función de activación “Lineal”, esto tendrá solo una salida.
- Optimizador: Adam con una tasa de aprendizaje de 0.001 Este modelo solo se adapta para un horizonte.

Al aplicar la atención en la secuencia de valores, se puede enfocar la red en los momentos críticos o relevantes dentro de la serie, mejorando así la capacidad de predicción.

Tabla No. 7

Modelo LSTM con más capas.

Modelo (+)LSTM		
Layer	Units	Activation
LSTM (1)	256	tanh
LSTM (2)	128	tanh
LSTM (3)	64	tanh
LSTM (2)	32	tanh
Dense	1	lineal*

3. RESULTADOS

Esta sección se centra en el análisis de evaluación del rendimiento de los modelos propuestos en el contexto de un Trabajo Fin de Máster. Con el fin de realizar una evaluación exhaustiva, se llevarán a cabo varias comparaciones métricas para cada modelo. Además, se mostrarán las funciones de pérdida correspondientes a los conjuntos de entrenamiento y prueba de cada modelo, considerando los horizontes establecidos en el estudio.

El estudio de los horizontes propuestos es fundamental en esta evaluación, ya que permite considerar diferentes plazos de predicción y evaluar cómo los modelos se desempeñan en cada uno de ellos. Para este caso, se ha considerado el modelo base para los horizontes de pronóstico de 6 horas, 12 horas y 24 horas. Se ha evaluado el rendimiento tanto en entrenamiento como los datos en test de cada modelo utilizando la función de pérdida MAE.

Para cada uno de los modelos considerados, se llevó a cabo el entrenamiento durante 50 épocas y se realizó un seguimiento de la función de pérdida en cada época. En el caso del primer horizonte de 6 horas y el segundo horizonte de 12 horas, se observó que la red neuronal estaba aprendiendo de los datos de entrenamiento, lo que se manifestaba en una disminución gradual de la función de pérdida a lo largo de las épocas.

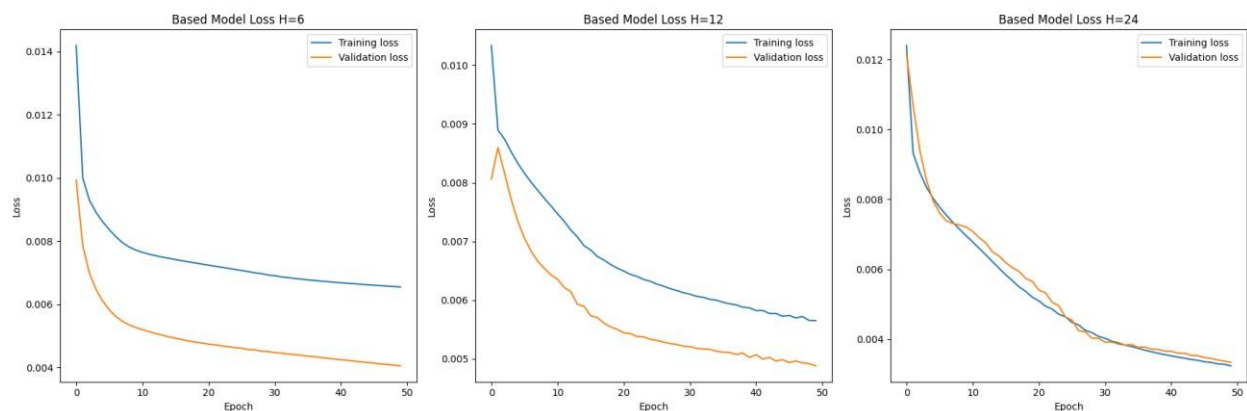
Sin embargo, se encontró una discrepancia interesante al evaluar los datos de validación. El valor del MAE en los datos de validación resultó ser inferior al obtenido en los datos de entrenamiento. Esta diferencia podría indicar que la red neuronal está sufriendo del fenómeno conocido como sobreajuste (Overfitting). En consecuencia, el modelo puede tener dificultades para generalizar y no está logrando un buen desempeño en datos nuevos o no vistos mostrando este comportamiento en la Figura No. 7, como consecuencia no generaliza bien los datos nuevos o no vistos.

Para el horizonte de pronóstico de 24 horas, se observó un comportamiento distinto en comparación con los horizontes anteriores. Durante el entrenamiento de la red neuronal, se evidenció un aprendizaje efectivo, reflejado en una disminución gradual de la función de pérdida MAE a lo largo de las épocas.

Además, la validación de los datos mostró un MAE cercano a los datos de entrenamiento, lo que sugiere que el modelo está generalizando bien a datos nuevos o no vistos. Sin embargo, en ocasiones, se pudo observar que la validación presentaba algunas disminuciones ligeras en comparación con los datos de entrenamiento, aunque en una escala de milésimas, observada en la tercera grafica de la figura no. X. Los resultados obtenidos en términos de la función de pérdida son los siguientes:

Figura No. 7

Grafica de diferentes Horizontes para el Modelo Base.



En la siguiente figura No. 8 se muestra la comparación entre los valores predichos y los valores reales en un rango de 100 horas. Esta visualización permite evaluar para cada horizonte la capacidad de capturar patrones y tendencias presente en los datos reales. Al analizar la figura, se pueden observar diferentes comportamientos en los valores predichos en comparación con los valores reales para cada horizonte de pronóstico.

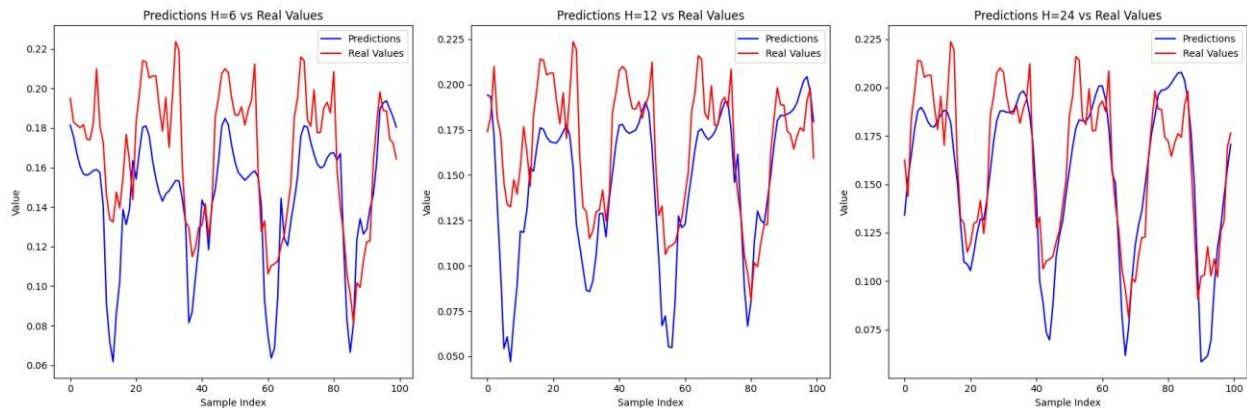
En el horizonte de 6 horas, se nota una tendencia irregular en los valores predichos en relación con los valores reales. Esto indica una imprecisión en las predicciones del modelo para este horizonte, ya que no logra capturar de manera adecuada los patrones y tendencias presentes en los datos reales.

En el horizonte de 12 horas, los valores predichos también muestran una tendencia irregular, pero se acercan más a los valores reales en comparación con el horizonte de 6 horas. Sin embargo, se observan variaciones y desviaciones, especialmente durante las horas de mayor actividad o picos, lo que indica que el modelo aún tiene dificultades para capturar completamente los patrones en esos momentos específicos.

Por otro lado, en el horizonte de 24 horas, se observa que el modelo ha logrado capturar de manera más precisa los patrones y tendencias presentes en los datos. Aunque existen desviaciones y variaciones en ciertos momentos, en general, los valores predichos están más alineados con los valores reales en comparación con los horizontes anteriores. Esto sugiere que el modelo ha mejorado su capacidad de predicción y puede ofrecer resultados más confiables en este horizonte. Observar la Figura No. 8 se observa el desempeño de los modelos para los horizontes de 6 horas, 12 horas y 24 horas de izquierda a derecha.

Figura No. 8

Predicción de valores vs los valores reales según horizontes.



Durante el entrenamiento de las diferentes arquitecturas propuestas, se dividió el dataset en aproximadamente un 75% para entrenamiento y un 25% para datos de evaluación. Se observó que el horizonte de pronóstico de 24 horas mostraba un mejor rendimiento en comparación con otros horizontes.

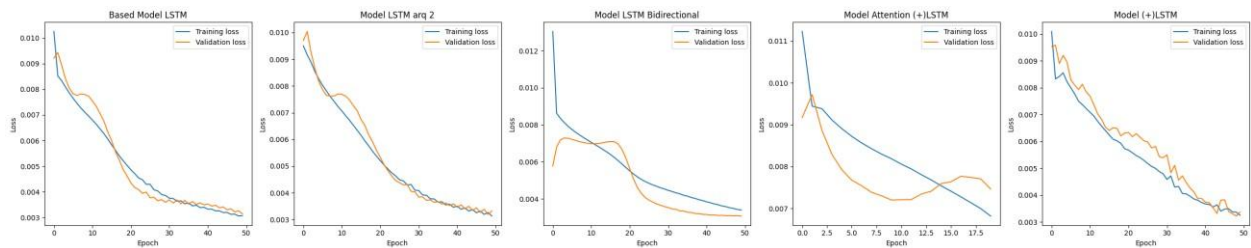
Inicialmente, se configuró el optimizador Adam con una tasa de aprendizaje de 0.001, lo cual permitió que los ajustes de los pesos durante el entrenamiento fueran más rápidos. Sin embargo, se observó que esta configuración resultó en un ajuste brusco y errores de convergencia en el entrenamiento de las redes neuronales.

Posteriormente, se realizó una segunda corrida utilizando una tasa de aprendizaje de 0.0001. Esta configuración ajusta los pesos de manera más lenta durante el entrenamiento, lo cual condujo a un mejor aprendizaje de las redes neuronales. Se observó una convergencia más estable y resultados de pérdida más satisfactorios con esta tasa de aprendizaje.

Por lo tanto, se concluyó que una tasa de aprendizaje de 0.0001 era más adecuada para lograr un entrenamiento eficiente y obtener mejores resultados en términos de rendimiento y convergencia de las redes neuronales en este caso particular.

Figura No. 9

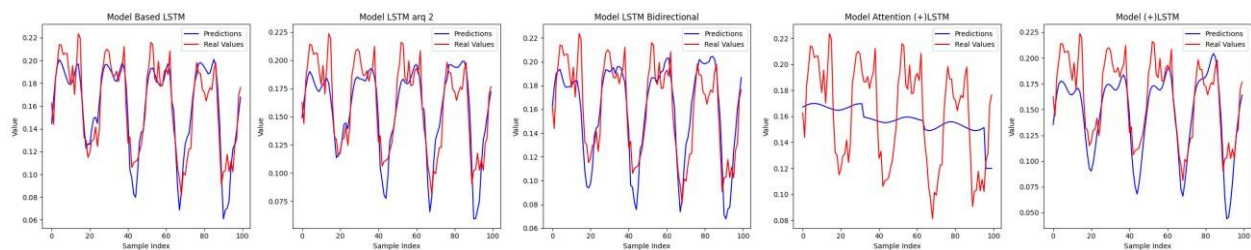
Desempeño de las diferentes Arquitectura LSTM



Al analizar la Figura No. 10, que muestra los resultados de las 100 horas de predicción, se puede observar la comparación entre los datos de prueba (representados por la línea roja) y los datos predichos por el modelo (representados por la línea azul). Al examinar la figura se destacan los modelos “Base LSTM”, “LSTM”, “LSTM Bidireccional” y “(+) LSTM” se ajustan mejor a los patrones presentes en los datos, mostrando una mejor capacidad para capturar y predecir los patrones y tendencias de la serie temporal. A diferencia de los otros modelos mencionados “Attention (+) LSTM” no logra capturar patrones, ni tendencia de manera efectiva, lo que indica que es un mal modelo para el problema en cuestión.

Figura No. 10

Valores predichos para cada arquitectura de redes neuronales vs el valor real.



Para concluir con la verificación de los modelos, se muestra la tabla no. X recopilando todas métricas del modelo para el conjunto de datos de prueba. Los valores que se muestran corresponden a los objetivos normalizados de la serie temporal.

Tabla No. 8

Representación de Métricas para cada Modelo.

Modelo	MAE	RMSE	NRMSE	$\rho_{y,\hat{y}}$
Base LSTM	0.0350	0.0560	0.7495	0.696
LSTM	0.0367	0.0576	0.7707	0.685
LSTM Bidireccional	0.0364	0.0553	0.7409	0.703
Attention (+) LSTM	0.0665	0.0863	1.1556	0.210
(+) LSTM	0.0665	0.0863	1.1556	0.210

De acuerdo con las métricas presentadas, los modelos “Base LSTM”, “LSTM Bidireccional” y “LSTM” obtienen resultados similares en términos de MAE, RMSE y NRMSE y coeficiente de correlación de Pearson. Por lo tanto, se puede decir que estos tres modelos tienen un rendimiento comparable.

Por otro lado, los modelos "Attention (+) LSTM" y "(+) LSTM" muestran un desempeño inferior en todas las métricas en comparación con los otros modelos. Estos modelos tienen un MAE y un RMSE más altos, un NRMSE más alto y una correlación más baja. Por lo tanto, estos modelos no son tan efectivos como los otros en la tarea evaluada.

4. DISCUSIÓN

Basándonos en los resultados obtenidos, podemos concluir que el modelo LSTM Bidireccional ha mostrado el mejor rendimiento en comparación con los otros modelos evaluados. Esto se debe a su bajo valor de RMSE, lo que indica que las predicciones tienen errores cuadráticos promedio menores en comparación con los valores reales. Además, el NRMSE también es bajo, lo que implica que el modelo tiene un buen ajuste en relación con la variabilidad de los datos. El coeficiente de correlación de Pearson es alto, lo que indica una relación positiva y fuerte entre las predicciones y los valores reales, sugiriendo que las predicciones siguen una tendencia similar a los valores reales.

Sin embargo, es importante mencionar que, al analizar el comportamiento del modelo en los conjuntos de entrenamiento y prueba, se observa un posible problema de sobreajuste. Aunque las métricas son favorables, es necesario tener precaución debido a que el modelo puede estar ajustándose demasiado a los datos de entrenamiento y no generalizar bien para nuevos datos. Esto podría afectar su capacidad para hacer predicciones precisas en situaciones reales.

5. CONCLUSIONES Y TRABAJO FUTURO

- Se identificó en la investigación y estado del arte que variables son relevantes para los datos en cuestión, se encontró la limitante que no existían esos datos específicos en nuestro dataset. A pesar de eso se realizó un análisis de correlación con las variables disponible y el consumo energético del edificio. Se recomienda obtener datos adicionales como la ocupación del edificio, Humedad Relativa del ambiente, Temperatura media dentro del edificio, como el promedio de objetos que depende del consumo energético en el edificio.
- Se recopilaron datos históricos precisos del consumo energético del edificio, permitiendo realizar un análisis detallado de cada variable y su descomposición en componentes dentro de las series temporales. Esta descomposición de las series temporales permitió una mejor comprensión de los factores que afectan el

consumo energético en diferentes momentos del año, identificando patrones, tendencias y estacionalidades en los datos. Para el futuro, se propone recopilar datos quince minútales en lugar de datos de mayor intervalo de tiempo. Esto proporcionará una granularidad más fina en la recopilación de datos, lo que permitirá capturar variaciones más pequeñas y detalles más precisos en el consumo energético del edificio. La recopilación de datos quince minútales brindará una visión más detallada de los patrones de consumo y permitirá una predicción más precisa a corto plazo.

- Se implementó un modelo de predicción basado en redes neuronales LSTM para predecir el consumo energético. Durante este proceso, se realizaron ajustes en los parámetros del modelo teniendo en cuenta el escalado de los datos. Esto garantizó que el modelo estuviera adecuadamente adaptado a los datos y mejoró su precisión. Como trabajo futuro, se propone implementar la predicción del consumo energético utilizando una ventana de 24 horas y un offset de 24. Esta estrategia permitirá aprovechar la dependencia temporal de los datos al utilizar los valores del día anterior antes de la hora objetivo para predecir la hora objetivo del siguiente día. Se espera que esta consideración adicional de la información temporal mejore la precisión del modelo y su capacidad para capturar patrones a lo largo del tiempo.
- En resumen, se exploraron arquitecturas adicionales, como LSTM Bidireccional y atención, para mejorar el rendimiento del modelo LSTM. Se llevó a cabo un estudio de horizontes para determinar la mejor configuración de predicción, y se utilizaron técnicas de validación cruzada y métricas para evaluar la precisión del modelo. Para trabajo futuro explorar técnicas de ensemble, como el promedio de múltiples modelos o el uso de métodos de combinación de modelos, para mejorar aún más la precisión de la predicción. Estas técnicas aprovechan la diversidad de modelos y pueden ayudar a mitigar el sesgo y la variabilidad inherentes a los modelos individuales

Bibliografía

1. Directiva (UE) 2018/844 del Parlamento Europeo y del Consejo. EUR-Lex . [En línea] 30 de Mayo de 2018. <https://eur-lex.europa.eu/legalcontent/ES/TXT/?uri=CELEX:32018L0844>.
2. Nguyen, A. T., & Haghighat, F. A review of energy consumption calculation in bulding: part 1 - Load calculation methods. *Renewable and Sustainable Energy Reviews*. [En línea] 2014. <https://doi.org/10.1016/j.rser.2014.05.056>.
3. *A review of building energy regulation an policy for energy conservation in developing countries*. Zhang, Y., Jiang, Y., Yu, B. 2012-223, s.l. : Energy and Buildings, 2015, Vol. 107. DOI: <https://doi.org/10.1016/j.enbuild.2015.09.002>.
4. IBM. *Neural Network Modeling*. [En línea] 2017 de Agosto de 2021. [Citado el: 23 de Abril de 2023.] <https://www.ibm.com/docs/es/spss-modeler/saas?topic=networksneural-model>.
5. Xeridia. Redes neuronales artificiales: qué son y cómo se entrenan (Parte I). [En línea] 2022 de Junio de 06. [Citado el: 2023 de Abril de 27 .] <https://www.xeridia.com/blog/redes-neuronales-artificiales-que-son-y-como-seentrenan-parte-i>.
6. Ortiz, Juan Antonio Perez. *Modelos Predictivos Basados en Redes Neuronales Recurrentes de Tiempo Discreto*. 2002.
7. Hastie, T., Tibshirani, R., & Friedman, J. *An Introduction to Statistical Learning: With Applications in R*. s.l. : Springer, 2009.
8. Arana, Carlos. *Redes neuronales recurrentes: Análisis de los modelos especializados en datos secuenciales*. Buenos Aires : Cema, 2021. 797.
9. Ian Goodfellow, Yoshua Bengio, Aaron Courville. *Deep Learning* . s.l. : The MIT Press, 2016.
10. *Speech recognition with deep recurrent neural networks*. Graves, A., Mohamed, A. R., & Hinton, G. s.l. : IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013.
11. colah's blog. Understanding LSTM Networks. [En línea] 2015. [Citado el:

2023 de Junio de 20.] <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>.

12. ProgrammerSought. Machine Learning/Deep Learning/NLP-5-RNN, LSTM, GRU. [En línea] [Citado el: 2023 de Junio de 20.] <https://programmersought.com/article/78575166382/>.

13. Graves, A., & Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *ScienceDirect*. [En línea] [Citado el: 2023 de Junio de 21.] <https://doi.org/10.1016/j.neunet.2005.06.042>.

14. *Attention-based LSTM for target-dependent sentiment classification*. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Zhou, L., Feng, J., & Zhao, D. s.l. : EMNLP, 2016.

15. Springer. *The Concise Encyclopedia of Statistics*. New York, NY : Springer, 2088. 978-0-387-32833-1.

16. Commission., European. Root Mean Square Error (RMSE). [En línea] 2019. [Citado el: 2023 de Junio de 21.] https://cros-legacy.ec.europa.eu/content/root-meansquare-error-rmse_en.

17. Otto, Saskia A. Marine Data Science. [En línea] 2019 de Enero de 7. [Citado el: 2023 de junio de 21.] <https://www.marinedatascience.co/blog/2019/01/07/normalizing-the-rmse/>.

18. Turney, Shaun. Pearson Correlation Coefficient (r) | Guide & Examples. [En línea] Scribbr, 13 de Mayo de 2022. [Citado el: 2023 de Junio de 21 .] www.scribbr.com/statistics/pearson-correlation-coefficient/.

19. *A hybrid model for short-term electricity demand forecasting*. Zhang, G. & Qi, M. 1, s.l. : IEEE Transactions on Power Systems , 2007, Vol. 22. URL: <https://ieeexplore.ieee.org/document/4424201>.

20. *Optimal Allocation of electric Vehicle Charging Infrastrucutre Bases on Multiobjective Genetic Algorithm*. Zhang, J., Zhu, W., & Zhu, J. 10, s.l. : Energies, 2019, Vol.

12. DOI: <https://doi.org/10.3390/en12101934>.

21. Hu, J., Li, J., & Li, H. Deep Learning for Predicting Building Energy Consumption: A comprehensive Review. *Energy and AI*. [En línea] 2019. <https://doi.org/10.1016/j.egyai.2019.100005>.
22. Duarte, A., Kazacos, M., Winter, I., & Pa[is, A. Analisis y prediccion de Consumo eléctrico en edificios mediante técnicas de Inteligencia Artificial. . [En línea] 2016. https://oa.upm.es/48458/1/PFC_Duarte_Alexis_Kazacos_Winter_Pais_Brito.pdf.
23. Santamouris, M. Energy an Climate in the Urban Built Enviroment. 2001.
24. Freedman, D. A. *Statistical Models: Theory and Practice*. s.l. : Cambridge University Press, 2009.
25. Bonet Cruz, Isis, y otros. Redes neuronales recurrentes para el análisis de secuencias. Ciudad de la Habana, Cuba : Revista Cubana de Ciencias Informáticas, 2007. Vol. 1, 4. 1994-1536.