

Maestría en

**CIENCIA DE DATOS Y MÁQUINAS DE APRENDIZAJE CON MENCIÓN
EN INTELIGENCIA ARTIFICIAL**

**Trabajo previo a la obtención del título de Magíster en Ciencia de Datos y
Máquinas de Aprendizaje con mención en Inteligencia Artificial.**

AUTORES:

Meza Chico Ariel Antonio

Navas Lema María Cristina

Yaguachi Barahona Paul Marcelo

TUTORES:

Navas Castellanos Andrés Roberto

Andrade Montalvo Pablo Xavier

TEMA

Estudio comparativo de redes neuronales profundas para la predicción de
tendencias de consumo energético en la provincia de Tungurahua

Certificación de autoría

Nosotros, **Meza Chico Ariel Antonio, Navas Lema María Cristina, Yaguachi Barahona Paul Marcelo**, declaramos bajo juramento que el trabajo aquí descrito es de nuestra autoría; que no ha sido presentado anteriormente para ningún grado o calificación profesional y que se ha consultado la bibliografía detallada.

Cedemos nuestros derechos de propiedad intelectual a la Universidad Internacional del Ecuador (UIDE), para que sea publicado y divulgado en internet, según lo establecido en la Ley de Propiedad Intelectual, su reglamento y demás disposiciones legales.



Firma

Meza Chico Ariel Antonio



Firma

Navas Lema María Cristina



Firma

Yaguachi Barahona Paul Marcelo

Autorización de Derechos de Propiedad Intelectual

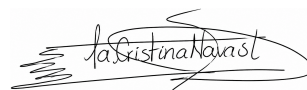
Nosotros, **Meza Chico Ariel Antonio, Navas Lema María Cristina, Yaguachi Barahona Paul Marcelo**, en calidad de autores del trabajo de investigación titulado *Estudio comparativo de redes neuronales profundas para la predicción de tendencias de consumo energético en la provincia de Tungurahua*, autorizamos a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que nos pertenecen o de parte de los que contiene esta obra, con fines estrictamente académicos o de investigación. Los derechos que como autores nos corresponden, lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

D. M. Quito, junio 2026



Firma

Meza Chico Ariel Antonio



Firma

Navas Lema María Cristina

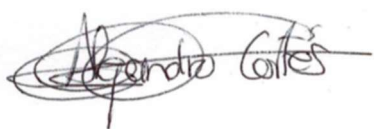


Firma

Yaguachi Barahona Paul Marcelo

Aprobación de dirección y coordinación del programa

Nosotros, **Director Alejandro Cortés López de EIG y Coordinadora Karla Estefanía Mora Cajas de UIDE**, declaramos que: **Meza Chico Ariel Antonio, Navas Lema María Cristina, Yaguachi Barahona Paul Marcelo** son los autores exclusivos de la presente investigación y que ésta es original, auténtica y personal de ellos.



Alejandro Cortés López
Director/a de la
Maestría en Ciencia de datos y máquinas
de aprendizaje con mención en Inteligencia
Artificial



Karla Estefanía Mora Cajas
Coordinador/a de la
Maestría en Ciencia de datos y máquinas
de aprendizaje con mención en Inteligencia
Artificial

DEDICATORIA

A mi familia.

Ariel Antonio Meza Chico

El presente proyecto se lo dedico a mi esposo, a mi familia pequeña, a mi familia grande y a todos los seres que me estiman y que de una u otra manera me han ayudado a crecer profesional y personalmente.

María Cristina Navas Lema

Con profundo amor y gratitud, dedico este trabajo a Dios, por brindarme la vida, la fortaleza, la sabiduría y la perseverancia necesarias para superar cada desafío y alcanzar esta importante meta académica.

A mi madre, quien desde el cielo sigue guiando mi camino. Gracias por tus enseñanzas, tu amor y por hacerme creer siempre que podía alcanzar cualquier meta que me propusiera.

A mi padre, por su apoyo incondicional, su amor y el ejemplo de esfuerzo y perseverancia que me ha brindado a lo largo de mi vida.

A mi pareja, por estar a mi lado con paciencia y amor durante todo este proceso. Gracias por ser mi fuente de motivación para alcanzar esta meta tan importante.

Y a mis hermanos, porque siempre han estado compartiendo conmigo los momentos de alegría y dificultad.

Paul Marcelo Yaguachi Barahona

AGRADECIMIENTOS

Agradezco a todas las personas que han sido parte de este proceso, especialmente a mi familia, que ha estado presente y me ha acompañado de distintas formas.

También agradezco a mis compañeros Paúl y Cristina, quienes asumieron esta responsabilidad con compromiso, eficiencia y seriedad.

Ariel Antonio Meza Chico

Agradezco a los docentes y directivos de la carrera por los conocimientos impartidos a lo largo de este proceso académico, los cuales constituirán una base fundamental para mi desarrollo y desempeño profesional.

Expreso también mi sincero agradecimiento a mis compañeros Paúl y Ariel, con quienes llevé a cabo este proyecto de manera eficiente y colaborativa. Su compromiso, intercambio de ideas y conocimientos, así como el trabajo en equipo, fueron esenciales para el desarrollo exitoso de este proyecto.

Finalmente, agradezco profundamente a mi esposo y a mi familia por su apoyo incondicional, comprensión y motivación constante durante todo este proceso educativo, siendo un pilar fundamental para alcanzar esta meta.

María Cristina Navas Lema

Expreso mi más sincero agradecimiento a los docentes de la maestría, quienes compartieron sus conocimientos, experiencia y orientación académica, contribuyendo significativamente a mi formación profesional.

Agradezco de manera especial a mis compañeros Cristina y Ariel; su compromiso y colaboración a lo largo de este proceso fueron fundamentales para alcanzar este importante logro académico.

Finalmente, agradezco a la Universidad Internacional del Ecuador por brindarme la oportunidad de continuar mi formación académica y profesional.

Paul Marcelo Yaguachi Barahona

RESUMEN

El presente proyecto compara el desempeño de las redes neuronales LSTM, GRU y Transformer para predecir la demanda energética en la provincia de Tungurahua. La investigación siguió la metodología CRISP-DM. Incluyó las fases de comprensión del negocio, comprensión de los datos, preparación de los datos, modelado y evaluación. Se construyó un conjunto de datos con información histórica del consumo eléctrico residencial de la Agencia de Regulación y Control de Electricidad (ARCONEL). Además, se incorporaron variables climáticas de NASA POWER y datos demográficos del Instituto Nacional de Estadística y Censos (INEC). El período analizado comprendió datos mensuales entre 2014 y 2025.

Durante la preparación de los datos se realizaron procesos de limpieza, tratamiento de valores inconsistentes, detección de anomalías e imputación de registros atípicos. Después, se diseñó un entorno experimental homogéneo para evaluar tres arquitecturas de aprendizaje profundo usadas en series temporales: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) y Transformer. Los modelos se entrenaron con distintas ventanas temporales. Su desempeño se evaluó mediante el Error Absoluto Medio (MAE), la Raíz del Error Cuadrático Medio (RMSE), el Error Porcentual Absoluto Medio (MAPE) y el coeficiente de determinación (R^2).

Los resultados demostraron que las redes neuronales profundas capturan adecuadamente los patrones temporales del consumo energético residencial. Entre los modelos evaluados, la arquitectura GRU obtuvo el mejor desempeño predictivo. Alcanzó un MAPE aproximado del 5.26% y los menores valores de error frente a las arquitecturas LSTM y Transformer.

Finalmente, se concluye que las arquitecturas de aprendizaje profundo constituyen una alternativa viable y efectiva para predecir la demanda energética en Tungurahua.

Palabras clave: consumo energético, series temporales, redes neuronales profundas, LSTM, GRU, transformer, predicción de demanda eléctrica, inteligencia artificial.

ABSTRACT

This project compares the performance of LSTM, GRU, and Transformer neural networks for forecasting energy demand in Tungurahua Province. The research followed the CRISP-DM methodology. It included the phases of business understanding, data understanding, data preparation, modeling, and evaluation. A dataset was built using historical residential electricity consumption data from the Electricity Regulation and Control Agency (ARCONEL). Climate variables from NASA POWER and demographic data from the National Institute of Statistics and Census (INEC) were also incorporated. The analysis covered monthly data from 2014 to 2025.

During the data preparation phase, data cleaning, inconsistent value treatment, anomaly detection, and outlier imputation were performed. Then, a homogeneous experimental environment was designed to evaluate three deep learning architectures commonly used for time series forecasting: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Transformer. The models were trained using different time window configurations. Their performance was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2).

The results showed that deep learning models effectively capture the temporal patterns of residential energy consumption. Among the evaluated models, the GRU architecture achieved the best predictive performance. It reached an approximate MAPE of 5.26 % and produced lower error values than the LSTM and Transformer architectures.

Finally, the study concludes that deep learning architectures are a viable and effective alternative for forecasting energy demand in Tungurahua.

Keywords: energy consumption, time series, deep neural networks, LSTM, GRU, transformer, electrical demand forecasting, artificial intelligence

TABLA DE CONTENIDOS (Índice)

1	CAPÍTULO 1: INTRODUCCIÓN	1
1.1	Definición del Problema	1
1.2	Justificación e importancia del trabajo de investigación	2
1.3	Alcance	3
1.4	Objetivos	4
1.4.1	Objetivo general	4
1.4.2	Objetivos específicos	4
2	CAPÍTULO 2: REVISIÓN DE LITERATURA	6
2.1	Estado del Arte	6
2.2	Consumo energético	9
2.2.1	Concepto de demanda energética	10
2.2.2	Tipos de demanda energética	11
2.2.3	Factores climáticos	11
2.2.4	Contexto energético del Ecuador	12
2.2.5	Contexto energético de la provincia de Tungurahua	16
2.3	Inteligencia Artificial y Machine Learning	17
2.4	Deep Learning	18
2.5	Redes Neuronales Profundas aplicadas a series temporales	19
2.6	Modelos a comparar	20
2.6.1	Long Short-Term Memory	20
2.6.2	Gated Recurrent Unit	21
2.6.3	Transformer	21

2.6.4	Criterio de comparación	22
2.7	Métricas de evaluación	22
2.8	Explicación conceptual de ACF y PACF	24
3	CAPÍTULO 3: DESARROLLO	26
3.1	Metodología de proyectos	26
3.1.1	CRISP-DM	27
3.1.1.1	Comprensión del negocio	27
3.1.1.2	Comprensión de datos	29
3.1.1.3	Preparación de datos	29
3.1.1.4	Modelado	30
3.1.1.5	Evaluación	31
3.1.1.6	Despliegue	32
3.1.2	Estructura y organización del proyecto planteado	33
3.1.2.1	Esquema General del Diseño	33
3.1.2.2	Diagrama de Gantt del proyecto	33
3.1.2.3	Flujo metodológico	35
3.2	Comprensión del negocio	36
3.2.1	Herramientas a utilizar	36
3.2.1.1	Python	36
3.2.1.2	Pandas	37
3.2.1.3	NumPy	37
3.2.1.4	Matplotlib/Seaborn	37
3.2.1.5	TensorFlow/Keras o PyTorch	37
3.2.1.6	Scikit-learn	38
3.3	Comprensión de los datos - Descripción del dataset	38
3.3.1	Fuente de datos energéticos	38
3.3.2	Datos climáticos	40
3.3.3	Datos poblacionales	42
3.4	Preparación y preprocesamiento de datos	42

3.4.1	Período y área de estudio	42
3.4.2	Preprocesamiento de datos	43
3.4.2.1	Limpieza de datos	43
3.4.2.2	Tratamiento de valores inconsistentes	46
3.4.2.3	Construcción de la serie temporal	49
3.4.2.4	Imputación eventos atípicos	50
3.4.2.5	Método de imputación con patrón estacional	51
3.5	Análisis Exploratorio de Datos (EDA)	53
3.5.1	Dataset procesado	53
3.5.2	Serie temporal mensual residencial	56
3.6	Configuración para el diseño experimental	58
3.6.1	Selección de Variables	59
3.6.2	Ventanas temporales	61
3.6.3	Análisis comparativo de ventanas temporales	62
3.6.4	División dataset	64
3.6.5	Escalado de Características	64
3.6.6	Resultado de la configuración experimental	65
3.7	Red Neuronal LSTM (Long Short-Term Memory)	66
3.7.1	Construcción del Modelo LSTM	66
3.7.2	Entrenamiento del Modelo LSTM	68
3.7.3	Evaluación	69
3.7.4	Visualización	72
3.7.5	Tabla Resumen de Métricas	77
3.8	Red Neuronal Transformer	78
3.8.1	Construcción del Modelo Transformer	78
3.8.2	Entrenamiento del Modelo Transformer	79
3.8.3	Evaluación	80
3.8.4	Visualización	82
3.8.5	Tabla Resumen de Métricas	86

	XI
3.9 Red Neuronal GRU	87
3.9.1 Construcción del Modelo GRU	88
3.9.2 Entrenamiento del Modelo GRU	89
3.9.3 Evaluación	90
3.10 Comparación de Eficiencia de Los Modelos Aplicados	91
4 CAPÍTULO 4: ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS	93
4.1 Resultados	93
4.1.1 Resultados del modelo GRU	93
4.1.1.1 Curva de pérdida durante el entrenamiento	93
4.1.1.2 Predicción frente a valores reales	93
4.1.1.3 Análisis en el conjunto de prueba	94
4.1.1.4 Dispersión real vs. predicción	95
4.1.1.5 Análisis de residuos	96
4.1.1.6 Tabla resumen de métricas	96
4.1.1.7 Comparación contra modelos de referencia (<i>baselines</i>)	97
4.1.1.8 Discusión	97
4.1.1.9 Observaciones y hallazgos	99
5 CAPÍTULO 5: CONCLUSIONES Y RECOMENDACIONES	101
5.1 Conclusiones	101
5.2 Recomendaciones	102
Referencias	103
Anexo A: Detalles de la experimentación con ventanas temporales alternativas	107
Anexo B: Repositorio y laboratorio práctico	111

LISTA DE TABLAS (Índice de tablas)

Tabla 1	Comparación de metodologías para proyectos de ciencia de datos	26
Tabla 2	<i>Características del conjunto de datos analizado</i>	43
Tabla 3	<i>Normalización de nombres de variables</i>	44
Tabla 4	<i>Homologación de nombre de cantones</i>	45
Tabla 5	<i>Renombrado y traducción de variables</i>	45
Tabla 6	<i>Tipos de datos</i>	46
Tabla 7	<i>Tratamiento de valores nulos</i>	47
Tabla 8	<i>Distribución temporal de los registros con inconsistencias</i>	48
Tabla 9	<i>Métodos de agregación por tipo de variable</i>	50
Tabla 10	<i>Estadísticas descriptivas claves del conjunto de datos.</i>	54
Tabla 11	<i>Especificación de la variable del estudio</i>	59
Tabla 12	<i>Variables independientes consideradas en la investigación</i>	60
Tabla 13	<i>Métricas de rendimiento según el tamaño de la ventana temporal en meses.</i>	63
Tabla 14	<i>Resumen del proceso de normalización de datos.</i>	65
Tabla 15	<i>Arquitectura del modelo LSTM</i>	67
Tabla 16	<i>Parámetros de entrenamiento del modelo LSTM</i>	68
Tabla 17	<i>Resumen de métricas del modelo LSTM</i>	70
Tabla 18	<i>Métricas de desempeño de los modelos baseline sobre el conjunto de prueba LSTM</i>	71
Tabla 19	<i>Comparación del modelo LSTM frente a los modelos baseline sobre el conjunto de prueba</i>	71
Tabla 20	<i>Métricas de evaluación del modelo LTSM</i>	77
Tabla 21	<i>Arquitectura del modelo Transformer</i>	79

Tabla 22	<i>Resumen de métricas del modelo Transformer</i>	81
Tabla 23	<i>Métricas de desempeño de los modelos baseline sobre el conjunto de prueba</i>	81
Tabla 24	<i>Comparación del modelo Transformer frente a los modelos baseline sobre el conjunto de prueba</i>	82
Tabla 25	<i>Métricas de evaluación del modelo Transformer</i>	86
Tabla 26	<i>Arquitectura del modelo GRU</i>	89
Tabla 27	<i>Resumen de métricas del modelo GRU</i>	90
Tabla 28	<i>Comparación del modelo GRU frente a los modelos baseline sobre el conjunto de prueba</i>	91
Tabla 29	<i>Métricas del modelo GRU por conjunto de datos</i>	97

LISTA DE FIGURAS (Índice de figuras)

Figura 1	<i>Comportamiento promedio anual de la temperatura en la ciudad de Guayaquil [°C]</i>	13
Figura 2	<i>Curva de carga diaria</i>	14
Figura 3	<i>Crecimiento mensual de la demanda de energía</i>	15
Figura 4	<i>Caudales afluentes del embalse Mazar</i>	15
Figura 5	<i>Evolución de las emisiones de CO₂ en el sector eléctrico (Miles TCO₂) .</i>	16
Figura 6	<i>Fases de la metodología CRISP-DM</i>	27
Figura 7	<i>Organigrama horizontal de la fase comprensión del negocio</i>	28
Figura 8	<i>Organigrama horizontal de la fase comprensión de los datos</i>	29
Figura 9	<i>Organigrama horizontal de la fase comprensión de los datos</i>	30
Figura 10	<i>Organigrama horizontal de la fase modelado</i>	31
Figura 11	<i>Organigrama horizontal de la fase de evaluación</i>	32
Figura 12	<i>Organigrama horizontal de la fase de despliegue</i>	33
Figura 13	<i>Cronograma de actividades del proyecto basado en la metodología CRISP-DM</i>	34
Figura 14	<i>Diagrama de flujo de las fases del proyecto</i>	35
Figura 15	<i>Pipeline de la metodología CRISP-DM</i>	36
Figura 16	<i>Página principal de la plataforma de reportes de ARCONEL</i>	38
Figura 17	<i>Diagrama de flujo para la descarga automatizada de reportes energéticos</i>	39
Figura 18	<i>Diagrama de flujo para la extracción y procesamiento de datos climáticos</i>	41
Figura 19	<i>Distribución estadística de la energía facturada (kWh) y el monto facturado (USD).</i>	49
Figura 20	<i>Distribución Principales Variables Categóricas</i>	54

Figura 21	<i>Distribución Variables Climáticas y Demográficas</i>	55
Figura 22	<i>Matriz de Correlación de Variables Numéricas</i>	55
Figura 23	<i>Análisis Temporal Variables Objetivo</i>	56
Figura 24	<i>Evolución Mensual Consumo Energético</i>	57
Figura 25	<i>Análisis de Dependencia Temporal mediante ACF y PACF</i>	58
Figura 26	<i>Resumen del entrenamiento del modelo LSTM</i>	69
Figura 27	<i>Curva de Pérdida LSTM</i>	73
Figura 28	<i>Predicción LSTM</i>	73
Figura 29	<i>Análisis en el Conjunto de Prueba LSTM</i>	74
Figura 30	<i>Dispersión Real vs Predicción LSTM</i>	75
Figura 31	<i>Análisis de Residuos LSTM</i>	77
Figura 32	<i>Resumen del entrenamiento del modelo Transformer</i>	80
Figura 33	<i>Curva de Pérdida Transformer</i>	83
Figura 34	<i>Predicción Transformer</i>	83
Figura 35	<i>Análisis en el Conjunto de Prueba Transformer</i>	84
Figura 36	<i>Dispersión Real Transformer</i>	85
Figura 37	<i>Análisis de Residuos Transformer</i>	86
Figura 38	<i>Diagrama de flujo del proceso de construcción y evaluación del modelo GRU</i>	88
Figura 39	<i>Resumen del entrenamiento del modelo GRU</i>	90
Figura 40	<i>Comparativa entre los Modelos Analizados</i>	92
Figura 41	<i>Curva de pérdida — Entrenamiento del modelo GRU</i>	93
Figura 42	<i>Predicción GRU vs valores reales — Consumo energético</i>	94
Figura 43	<i>Detalle de predicción GRU en el conjunto de prueba</i>	94
Figura 44	<i>Dispersión de valores reales vs. predicción — GRU</i>	95
Figura 45	<i>Análisis de residuos — GRU</i>	96
Figura 46	<i>Curvas de pérdida (MSE) de entrenamiento y validación para ventanas alternativas</i>	107

Figura 47	<i>Predicción vs Real en la línea de tiempo (look_back = 3 meses)</i>	108
Figura 48	<i>Predicción vs Real en la línea de tiempo (look_back = 6 meses)</i>	108
Figura 49	<i>Predicción vs Real en la línea de tiempo (look_back = 12 meses)</i>	109
Figura 50	<i>Gráficos de dispersión multimodelo para look_back = 3 meses</i>	109
Figura 51	<i>Gráficos de dispersión multimodelo para look_back = 6 meses</i>	110
Figura 52	<i>Gráficos de dispersión multimodelo para look_back = 12 meses</i>	110

CAPÍTULO 1: INTRODUCCIÓN

1.1 Definición del Problema

La gestión de los sistemas eléctricos en Ecuador atraviesa una etapa crítica. En los últimos años, la alta dependencia de la generación hidroeléctrica ha incrementado la vulnerabilidad frente a fenómenos climáticos, como la sequía. Esta situación ha aumentado el riesgo de racionamientos eléctricos. En este contexto, la provincia de Tungurahua constituye un caso de estudio relevante por su importancia en el desarrollo comercial, industrial y, especialmente, turístico de la Sierra central.

Sin embargo, en la actualidad la planificación energética en esta zona así como a nivel país tiene un problema importante; ya que resulta complicado prever con cierta exactitud cómo va a evolucionar el consumo eléctrico, puesto que no solo depende de la demanda residencial básica, sino que está marcado por el dinamismo de la economía local y por la gran cantidad de población que genera el turismo.

Ante esta realidad, el presente trabajo se orienta al desarrollo, evaluación e identificación de un modelo predictivo de tendencias de consumo energético residencial específicamente en la provincia de Tungurahua. El estudio se fundamenta en el análisis de datos históricos de consumo eléctrico, los cuales serán segmentados por cantones para identificar patrones y tendencias relevantes que permitan una predicción de la demanda energética.

El proyecto se formula como un estudio comparativo entre arquitecturas de redes neuronales que son ampliamente utilizadas en el análisis de series temporales, estas arquitecturas son: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) y Transformer.

La base de la comparación de los modelos se centrará en la implementación, entrenamiento y evaluación bajo un mismo conjunto de datos y condiciones experimentales, esto con el fin de analizar las diferencias en términos de capacidad, desempeño y escalabilidad predictiva de los modelos.

Para el desarrollo de la investigación se adoptará la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining), la cual proporciona un marco estructurado para

proyectos de análisis de datos (Haya, 2021; Saltz, 2021). Esta metodología permitirá organizar el proyecto en diferentes fases que son claramente definidas, lo que asegura la coherencia metodológica y el seguimiento a lo largo del proceso de investigación.

En conclusión, los resultados no solo permitirán determinar cuál de las arquitecturas posee una mayor precisión para capturar la realidad energética de la provincia, sino que aportarán evidencia técnica fundamental para la toma de decisiones informadas que promuevan el desarrollo sostenible y la seguridad energética en la región.

1.2 Justificación e importancia del trabajo de investigación

La predicción de tendencias de consumo energético mediante redes neuronales constituye un tema de gran importancia social, económica y estratégica para el Ecuador, y específicamente para la provincia de Tungurahua. La energía eléctrica es un recurso fundamental para el desarrollo, ya que sostiene el funcionamiento de los sectores productivos, los servicios básicos, la industria y la calidad de vida de la población. En este contexto, la determinación precisa de las tendencias de consumo resulta indispensable para garantizar la sostenibilidad, la seguridad energética y el progreso nacional.

Tungurahua se distingue por ser un polo de desarrollo comercial e industrial en la Sierra central, pero es su creciente dinamismo turístico lo que introduce una complejidad en la demanda energética. Cantones con alto volumen turístico, como Baños de Agua Santa, generan una población no residente pero recurrente que provoca picos de consumo no lineales, los cuales tienen origen en los calendarios festivos y estacionalidades climáticas. La provincia presenta una alta vulnerabilidad ante fenómenos naturales como la sequía debido a su dependencia de la generación hidroeléctrica. La disminución de los niveles de los embalses afecta directamente la capacidad de producción, generando riesgos de desabastecimiento, racionamientos y un incremento en los costos operativos que impactan negativamente en el bienestar social y la estabilidad empresarial de la zona (Agencia de Regulación y Control de Electricidad – ARCONEL, 2025).

A esta problemática se suma la deficiente planificación y administración del consumo energético, lo que dificulta la toma de decisiones oportunas y eficientes. La falta de herramien-

tas predictivas precisas limita la capacidad de anticipar escenarios de alta demanda, gestionar adecuadamente los recursos disponibles y diseñar políticas públicas que promuevan un uso racional de la energía. En este sentido, la aplicación de técnicas avanzadas de inteligencia artificial, como las redes neuronales, representa una alternativa innovadora y eficaz para abordar estos desafíos.

Las redes neuronales permiten analizar grandes volúmenes de datos históricos de consumo energético, variables climáticas y factores socioeconómicos, identificando patrones complejos y no lineales que los métodos tradicionales no logran captar. Gracias a su capacidad de aprendizaje y adaptación, estos modelos pueden generar predicciones más precisas sobre la demanda futura de energía eléctrica, contribuyendo a una mejor planificación operativa y estratégica del sistema eléctrico.

La implementación de modelos predictivos basados en redes neuronales puede apoyar a las entidades responsables de la gestión energética en la toma de decisiones informadas, optimizando la distribución de la energía, reduciendo pérdidas, previniendo crisis energéticas y fomentando el uso eficiente de los recursos. Este enfoque puede servir como base para el diseño de políticas públicas orientadas a la sostenibilidad, la diversificación de la matriz energética y la mitigación de los efectos del cambio climático en la provincia de Tungurahua.

En conclusión, el estudio de la predicción de tendencias de consumo energético mediante redes neuronales es pertinente y necesario para la provincia de Tungurahua, ya que aporta soluciones tecnológicas a una problemática actual y crítica. Este tema no solo tiene un alto valor académico y científico, sino que también contribuye al desarrollo sostenible, la estabilidad económica y el bienestar social.

1.3 Alcance

El alcance del presente proyecto se circunscribe al diseño, entrenamiento y evaluación comparativa de modelos predictivos basados en las arquitecturas de redes neuronales profundas, desarrollados en un entorno académico controlado. El estudio se delimita geográficamente a la provincia de Tungurahua, utilizando información histórica de consumo eléctrico recolectada desde el año 2014.

El proceso de modelado considerará variables temporales asociadas al consumo, y se tomará en cuenta variables climáticas y poblacionales como factores complementarios para analizar la volatilidad de la demanda energética en la provincia de estudio.

La predicción del consumo energético se plantea en horizontes de corto y mediano plazo, con el objetivo de analizar la capacidad de las redes neuronales para capturar tendencias generales, sin considerar eventos excepcionales o no recurrentes que no se encuentren reflejados en los datos históricos, tales como crisis económicas severas, cambios estructurales en la política energética o fenómenos climáticos extremos.

Finalmente, este trabajo se desarrolla con un enfoque académico y exploratorio, orientado a evaluar la viabilidad y el desempeño de las redes neuronales artificiales como herramienta para la predicción de tendencias de consumo energético. Los resultados obtenidos no pretenden sustituir los modelos oficiales utilizados por las entidades responsables de la planificación y gestión del sistema eléctrico nacional, sino aportar evidencia técnica y metodológica que contribuya al análisis y comprensión del comportamiento del consumo energético residencial en la provincia de Tungurahua.

1.4 Objetivos

1.4.1 Objetivo general

Contrastar el desempeño de modelos de redes neuronales profundas LSTM, GRU y Transformer para la predicción de la demanda energética en la provincia de Tungurahua, con el propósito de identificar cuál arquitectura ofrece mejores resultados frente a distintos patrones de consumo en el contexto de la Sierra ecuatoriana.

1.4.2 Objetivos específicos

- Recopilar los datos históricos de la demanda energética en la provincia de Tungurahua, considerando la variabilidad por cantones y tipos de consumo para identificar factores temporales y estacionalidades clave.
- Construir un pipeline de preprocesamiento basado en la fase de preparación de datos de

la metodología CRISP-DM, ejecutando la limpieza, transformación y normalización de la información recolectada.

- Implementar modelos de predicción basados en arquitecturas de redes neuronales Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Transformer, ajustando sus hiperparámetros para el análisis del caso de estudio.
- Validar las arquitecturas propuestas utilizando un conjunto de datos histórico en un entorno experimental homogéneo, asegurando la consistencia en la comparación de los resultados.
- Comparar el desempeño predictivo de los modelos LSTM, GRU y Transformer y evaluar los resultados mediante métricas cuantitativas apropiadas para series temporales.
- Determinar la arquitectura de red neuronal que presente mayores beneficios técnicos en términos de precisión y escalabilidad para apoyar la planificación energética en Tungurahua.

CAPÍTULO 2: REVISIÓN DE LITERATURA

2.1 Estado del Arte

Montero Laurencio et al. (2024), en su investigación *Predicción de consumo y demanda de electricidad mediante redes neuronales artificiales y algoritmo iterativo*, publicada en la Revista Politécnica en diciembre de 2024, señalaron que:

El objetivo principal de este estudio fue demostrar el potencial que tienen las redes neuronales artificiales (RNA) para pronosticar la demanda y el consumo de energía eléctrica. De manera específica, la investigación buscó predecir la demanda de potencia activa (kW) en un circuito de distribución primaria con alta presencia de edificios residenciales, así como estimar el consumo diario de energía eléctrica (kWh) en un edificio de oficinas.

Para alcanzar ese propósito, se optó por implementar las RNA del tipo perceptrón multicapa con algoritmo de retropropagación, combinadas con un mecanismo de aprendizaje automático basado en un algoritmo iterativo de funciones recursivas.

Los resultados obtenidos resultan significativos si se comparan con los métodos tradicionales evaluados en el mismo estudio. Para la predicción de demanda en el circuito de distribución primaria, el modelo basado en RNA alcanzó un ajuste del 94.5 %, superando notablemente el 38 % logrado con modelación paramétrica y el coeficiente de correlación de 0.64 obtenido mediante regresión múltiple. En cuanto al pronóstico del consumo diario del edificio de oficinas, el mejor modelo neuronal arrojó un coeficiente de correlación de $R = 0,9435$, frente a $R = 0,84$ conseguido con la regresión múltiple.

Jácome-Vélez et al. (2026), en su investigación *Predicción del consumo energético mediante un modelo matemático de regresión lineal a partir de la producción diaria en la planta envasadora Q'Agua*, publicada en la Revista Científica FINIBUS – Ingeniería, Industria y Arquitectura, señalaron que:

El propósito de este trabajo fue desarrollar un modelo matemático capaz de estimar la demanda de energía eléctrica de la fábrica tomando como base directa su volumen de produc-

ción cotidiano. La investigación se desarrolló en dos fases: primero, el registro sistemático del número de bidones envasados y del consumo eléctrico diario obtenido de medidores digitales; y segundo, el análisis exploratorio y modelado estadístico mediante regresión lineal simple con el método de mínimos cuadrados, utilizando R-Studio para validar los supuestos del modelo y calcular intervalos de confianza.

Los resultados obtenidos resultan particularmente sólidos desde el punto de vista estadístico. El modelo lineal ajustado, expresado en la Ecuación (2.1), presentó un coeficiente de determinación R^2 del 99,27 %:

$$Y = 4,9592 + 0,0403X \quad (2.1)$$

lo que indica que prácticamente toda la variabilidad del consumo energético queda explicada por el nivel de producción diario.

La importancia de este estudio radica en que demuestra que técnicas estadísticas relativamente accesibles —como la regresión lineal simple implementada en software de código abierto— pueden generar modelos predictivos de altísima precisión cuando existe una relación estructural fuerte entre la variable operativa y el consumo energético. Este aporte resulta especialmente relevante para pequeñas y medianas industrias que necesitan optimizar su gestión energética sin recurrir a métodos computacionalmente costosos.

Severiche-Maury et al. (2025), en su investigación *Predicting Energy Consumption and Time of Use of Home Appliances in an HEMS Using LSTM Networks and Smart Meters: A Case Study in Sincelejo, Colombia*, publicada en la revista *Sustainability*, señalaron que:

El objetivo central de este estudio radicó en desarrollar un modelo computacional basado en redes neuronales recurrentes de tipo LSTM (Long Short-Term Memory) capaz de pronosticar tanto el gasto energético, a nivel individual y total en una vivienda, como el tiempo exacto de uso de sus electrodomésticos.

Para concretar dichos objetivos, los investigadores diseñaron una metodología cimentada en la recolección continua de datos en un escenario real durante un periodo ininterrumpido de doce meses. El equipo instrumentó una vivienda típica de clase media ubicada en Sincelejo, Colombia, instalando una red de medidores inteligentes para registrar, con frecuencia horaria,

el comportamiento cotidiano de cinco dispositivos: un televisor, una computadora, un equipo de aire acondicionado, un ventilador y un sistema de iluminación. Este extenso volumen de información histórica fue preparado para alimentar una red neuronal profunda, es decir diseñaron un sistema que opera en paralelo para estimar de forma independiente el consumo de energía y el tiempo de funcionamiento.

Con el propósito de evaluar la viabilidad de la propuesta, la configuración del modelo fue sometida a validaciones con datos no vistos previamente y su rendimiento fue contrastado contra un modelo estadístico de referencia, empleando para ello métricas de error y coeficientes de correlación.

Los resultados derivados del entrenamiento y la validación evidenciaron un nivel de exactitud sobresaliente, confirmando la superioridad del enfoque de aprendizaje profundo. Las mediciones revelaron que la red neuronal logró asimilar las dependencias temporales complejas con índices de error muy bajos, destacando de manera particular en la predicción del tiempo de uso de los equipos. Al enfrentar el modelo propuesto con técnicas predictivas más simples e incluso con enfoques computacionales híbridos, la arquitectura LSTM demostró ser sustancialmente más precisa, logrando un buen balance al requerir un menor costo de procesamiento computacional. En definitiva, estos hallazgos brindan una base tecnológica altamente confiable para planificar el uso de los electrodomésticos, reducir el costo de las facturas eléctricas y promover hábitos de consumo orientados a la sostenibilidad ambiental a largo plazo.

Defaz Toapanta et al. (2025), en su investigación *Predicción de consumo de energía eléctrica en hogares utilizando modelos de aprendizaje automático*, desarrollada en la Universidad Internacional del Ecuador (UIDE), Quito, Ecuador, como parte de la Maestría en Ciencia de Datos y Máquinas de Aprendizaje con mención en Inteligencia Artificial, señalaron que:

El objetivo central del estudio fue identificar y determinar el mejor modelo de aprendizaje automático para pronosticar el gasto de energía eléctrica en el sector residencial de Ecuador, apoyándose en el análisis de datos históricos. La importancia de esta investigación radica en su capacidad para dotar al sector energético de un sistema predictivo que optimiza la planificación, previene fallos en la distribución y promueve un uso mucho más eficiente de los

recursos, impulsando indirectamente la sostenibilidad y la disminución del impacto ambiental de las ciudades.

Para llevar a cabo este análisis predictivo, los autores implementaron una metodología estructurada basándose en el marco de trabajo CRISP-DM, ideal para gestionar proyectos de ciencia de datos a través de fases ordenadas que van desde la comprensión del problema hasta la evaluación final. En su fase práctica, el equipo recopiló y depuró un conjunto de datos estadísticos del año 2023 proporcionado por la Agencia de Regulación y Control de Electricidad (ARCONEL), enfocándose particularmente en la información perteneciente al consumo de las viviendas. Posteriormente, segmentaron estos datos para someter a prueba a tres arquitecturas computacionales diferentes: Regresión Lineal, Random Forest y XGBoost. Esta metodología permitió contrastar el desempeño de un algoritmo tradicional frente a técnicas avanzadas de árboles de decisión, validando la fiabilidad de los resultados mediante métricas de error absoluto y coeficientes de determinación.

Los hallazgos derivados del entrenamiento de estos sistemas demostraron una abrumadora superioridad de las técnicas no lineales frente a los métodos convencionales. En concreto, el modelo de Random Forest se posicionó como el enfoque más eficiente, alcanzando una asombrosa precisión capaz de explicar más del 99% de la variabilidad del consumo eléctrico, superando ligeramente al algoritmo XGBoost y reduciendo drásticamente el margen de error de las predicciones. En conjunto, estos resultados pueden brindar a las autoridades y empresas ecuatorianas un instrumento informático para anticipar la demanda de la red, tomar decisiones operativas más certeras y diseñar políticas de consumo inteligente a nivel nacional.

2.2 Consumo energético

El consumo energético constituye un tema importante dentro de este trabajo. Antes de abordar sus características específicas, es importante comprender qué es la electricidad como fenómeno físico, debido a que sobre este concepto se construyen los aspectos relacionados con la medición y la gestión de la energía.

Según Universidad ORT Uruguay, Facultad de Ingeniería (s.f.), la electricidad es la forma de energía que surge a partir de los fenómenos físicos asociados a las cargas eléctricas

estacionarias o en movimiento. Más precisamente, al movimiento de los electrones en la capa exterior de los átomos de un material conductor.

Entonces, la electricidad es el movimiento y la presencia de cargas eléctricas. Cuando esas cargas se desplazan por un cable o se acumulan y luego se liberan, se genera una forma de energía que depende de cómo se comportan las cargas en los conductores. Por eso puede producirse y usarse para múltiples aplicaciones, como por ejemplo: luz, calor, uso de baterías o motores.

Con la definición de electricidad y la forma de energía que aparece a raíz de ella, y tomando como referencia a Soria Energía (s.f.), se puede establecer el consumo energético como la cantidad total de energía eléctrica que se utiliza para un proceso determinado en un periodo de tiempo.

Desde una perspectiva doméstica, el consumo energético está vinculado directamente con el número de miembros del hogar y sus edades, ubicación de la vivienda, estilo de vida con respecto al uso de aparatos electrónicos y eléctricos, entre otros.

2.2.1 Concepto de demanda energética

Se entiende por demanda energética la suma total de los consumos eléctricos individuales necesarios para satisfacer las necesidades de determinados clientes, procesos o sistemas, independientemente de su escala o naturaleza (Lozada Cabrera et al., 2022).

En la actualidad, con la automatización de procesos en las diferentes industrias, el crecimiento del uso de carros eléctricos y además el incremento de la infraestructura tecnológica para el desarrollo de inteligencia artificial han provocado un aumento significativo en la demanda energética. Por lo que, la planificación y gestión eficiente de los recursos electrónicos se han convertido en aspectos fundamentales para garantizar el correcto funcionamiento del sistema nacional de distribución eléctrica en el país.

Como investigadores, consideramos que la electricidad constituye un recurso fundamental para el desarrollo de la sociedad. Además de impulsar las actividades productivas, sostiene el funcionamiento de los sectores industrial, comercial y residencial.

2.2.2 *Tipos de demanda energética*

Según las diferentes actividades o sectores, existen algunos tipos relacionados con la demanda energética. A continuación se mencionan algunos de ellos.

- **Residencial:** La demanda energética residencial es toda la energía que consumen los hogares en el día a día: desde encender las luces hasta mantener la casa a una temperatura adecuada. Esta demanda depende de factores como qué tan grande es la vivienda, y cómo son los hábitos de quienes viven ahí.
- **Industrial:** La demanda energética industrial se refiere al consumo de energía dentro del sector industrial. Engloba el uso de electricidad para el funcionamiento de maquinaria, equipos de producción y diversos procesos propios de la industria. Este consumo puede cambiar dependiendo del tipo de industria o las características particulares de los procesos en la fabricación.
- **Comercial:** La demanda energética comercial abarca el uso de electricidad en espacios como oficinas, hoteles, restaurantes y locales minoristas. Este consumo se destina principalmente a la iluminación, la climatización y maquinaria de oficina. Cabe indicar que la energía requerida varía en función de las dimensiones del local y del tipo de negocio.
- **Transporte:** La demanda energética de transporte está vinculada con la energía requerida en vehículos y medios de transporte. La demanda energética en este sector varía según el tipo de vehículo o la distancia recorrida.
- **Agrícola:** La demanda energética agrícola se refiere a la energía en actividades agrícolas y agropecuarias. Incluye el uso de electricidad para sistemas de riego, maquinaria y sistemas de almacenamiento o procesamiento de alimentos. En este sector la demanda energética varía según el cultivo, la producción y las prácticas agrícolas empleadas.

2.2.3 *Factores climáticos*

Según Salinas Herrera y Pesántez Díaz (2025), para comprender cómo el clima influye en el consumo eléctrico, es importante considerar el concepto de sensación térmica. Este tér-

mino se refiere al nivel de frío o calor que percibe el cuerpo humano en un determinado entorno. Dicha percepción no depende únicamente de la temperatura registrada por el termómetro, sino también de la interacción con otros factores ambientales, como la temperatura, humedad y la velocidad del viento.

- **Temperatura:** Es el factor con mayor relevancia en el comportamiento del consumo eléctrico. Durante las temporadas de frío intenso, los equipos de calefacción funcionan a su máxima capacidad, mientras que en épocas de altas temperaturas, el uso del aire acondicionado se vuelve indispensable.
- **Humedad:** La humedad es un factor que altera la temperatura percibida por el cuerpo humano, debido a que la humedad presente en el aire influye en la forma en que el calor se almacena y se transfiere. Esto obliga a los sistemas de climatización a trabajar más para mantener condiciones ideales, lo que se traduce directamente en mayor consumo energético.
- **Velocidad del viento:** La velocidad del viento es un factor ambiental que, junto con la temperatura y la humedad, influye en la sensación térmica que percibe el cuerpo humano en su entorno. Sin embargo, este factor actúa de forma indirecta en el consumo energético, ya que el viento no cambia la temperatura del aire, pero sí intensifica la sensación de frío al acelerar la pérdida de calor corporal.

2.2.4 Contexto energético del Ecuador

El panorama energético en Ecuador se encuentra en un momento crítico y de transformación, fuertemente condicionado por factores climáticos y el incremento sostenido de la demanda eléctrica.

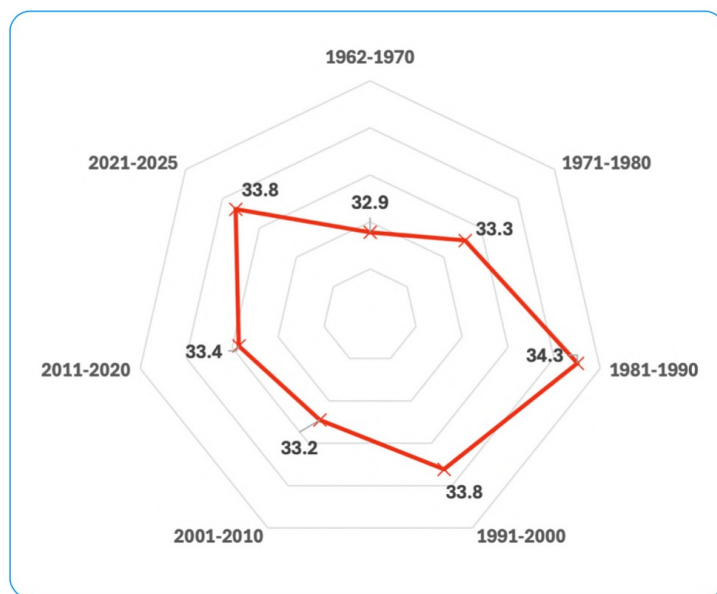
Durante los años 2024 y 2025, el sistema eléctrico ecuatoriano ha operado al límite de su capacidad debido a un problema estructural severo: el consumo de electricidad crece entre dos y cuatro veces más rápido que la capacidad de generación. En las horas pico (18:00 a 22:00), el país demanda entre 500.00 y 520.00 MW, cifra que el sistema a duras penas logra

cubrir de manera eficiente. A julio de 2025, Ecuador cuenta con una potencia efectiva nacional de 8.801 MW frente a una demanda regulada y no regulada en constante crecimiento.

Según la World Meteorological Organization (2025), los efectos del cambio climático alcanzaron niveles históricamente elevados durante 2024, registrándose anomalías térmicas globales sin precedentes respecto a la era preindustrial.

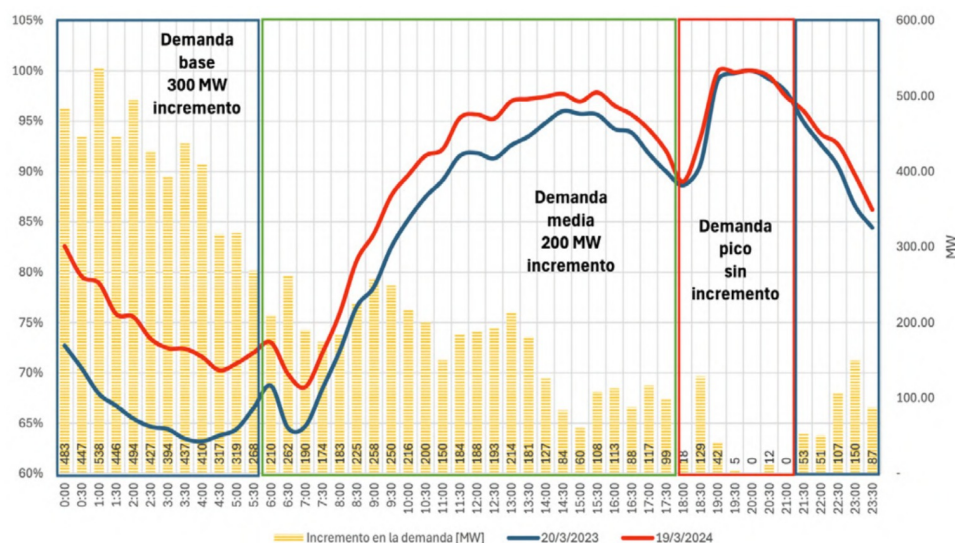
Figura 1

Comportamiento promedio anual de la temperatura en la ciudad de Guayaquil [°C]



Nota. Tomado de *Revista Panorama Eléctrico 2025*, por Agencia de Regulación y Control de Electricidad, 2025 (p. 6).

Figura 2
Curva de carga diaria



Nota. Tomado de *Revista Panorama Eléctrico 2025*, por Agencia de Regulación y Control de Electricidad, 2025 (p. 9).

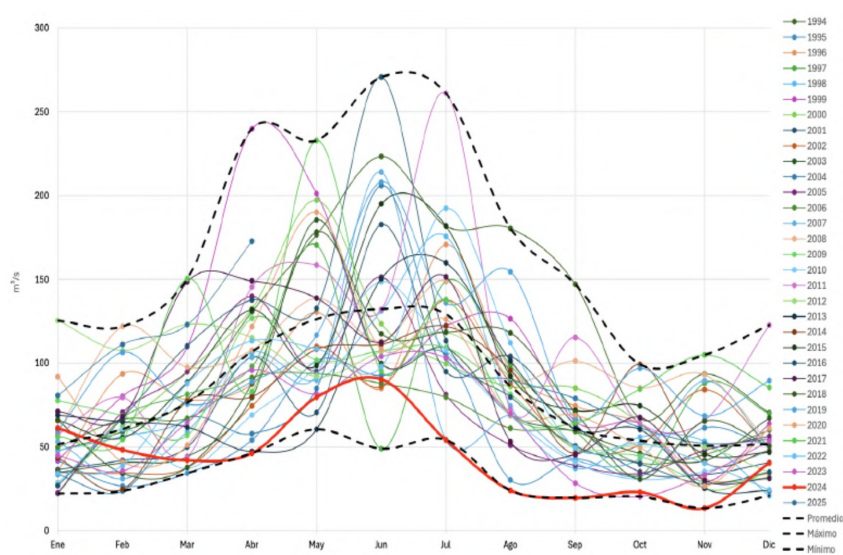
El impacto de los extremos climáticos en el año 2024 fue catalogado como históricamente anómalo para el sector eléctrico. En los meses de marzo y abril, temperaturas récord en la zona costera dispararon inusualmente la demanda en un 10% y 13% respectivamente, impulsada por el uso intensivo de climatización en el sector residencial. Al mismo tiempo, la cuenca del Amazonas, de donde proviene el grueso de la generación hidroeléctrica del país, registró caudales mínimos históricos entre julio y noviembre.

Figura 3
Crecimiento mensual de la demanda de energía



Nota. Tomado de *Revista Panorama Eléctrico 2025*, por Agencia de Regulación y Control de Electricidad, 2025 (p. 9).

Figura 4
Caudales afluentes del embalse Mazar



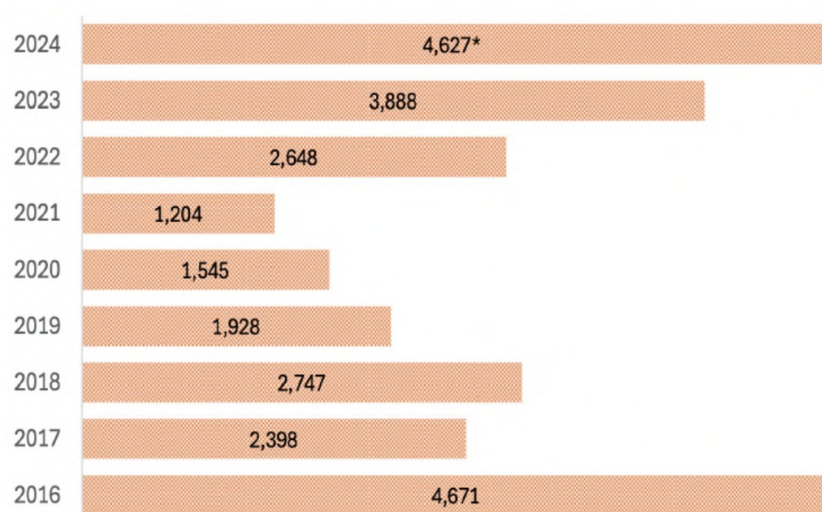
Nota. Tomado de *Revista Panorama Eléctrico 2025*, por Agencia de Regulación y Control de Electricidad, 2025 (p. 12).

La combinación del aumento del consumo y la drástica caída de la generación hidroeléctrica obligó a implementar racionamientos programados de energía en todo el país entre septiembre y diciembre de 2024. Para sostener la red eléctrica, se utilizó de manera intensiva el parque termoeléctrico. Esto provocó que los costos operativos se duplicaran y que las emisiones

de Gases de Efecto Invernadero (CO₂) se dispararan en 2024. Adicional a esto en el año 2025, el Gobierno Nacional se vio obligado a solicitar a las grandes industrias que se desconectarán de la red principal y utilizaran sistemas de autogeneración para brindar un alivio al sistema.

Figura 5

Evolución de las emisiones de CO₂ en el sector eléctrico (Miles TCO₂)



Nota. Tomado de *Revista Panorama Eléctrico 2025*, por Agencia de Regulación y Control de Electricidad, 2025 (p. 13).

2.2.5 Contexto energético de la provincia de Tungurahua

A pesar de que la provincia de Tungurahua aportó el 8,88 % de la generación eléctrica bruta nacional en el año de 2024, esta provincia no fue autosuficiente en horas de máximo consumo, registrando un déficit promedio de 48 MW en los momentos de mayor demanda eléctrica. Las principales centrales hidroeléctricas que abastecen la zona de Tungurahua, como Agoyán y La Península, estaban operando a tan solo el 62 % de su capacidad. Este bajo rendimiento en 2024 y 2025 se debe a una reducción del 14 % en los caudales de los ríos Pastaza y Patate por las fuertes sequías en esos años, esto sumado al deterioro tecnológico de las turbinas Pelton de Agoyán, que han perdido su eficiencia debido al desgaste por sedimentación y que requieren elevados costos de mantenimiento.

Los cortes de energía de finales de 2024 tuvieron consecuencias devastadoras para las actividades productivas de la provincia.

- **Sector Industrial:** El Parque Industrial de Ambato, que acoge a más de 100 empresas, sufrió apagones prolongados de hasta 64 horas entre octubre y noviembre de 2024. Esto generó pérdidas del 35 % en la producción y una caída del 40 % en las ventas de las industrias locales, además de daños en la maquinaria.
- **Sector Agrícola:** La falta de electricidad paralizó los sistemas de bombeo de riego en zonas como Pelileo y Tangaiche. Sumado a la sequía extrema, esto provocó una caída del 50 % en la producción de cultivos clave para la seguridad alimentaria.

En medio de este escenario complejo, en junio de 2025 el Gobierno Nacional anunció la Resolución Nro. ARCONEL-006/25 (Agencia de Regulación y Control de Electricidad, 2025), emitida por la Agencia de Regulación y Control de Electricidad, mediante la cual se estableció un incremento en las tarifas eléctricas para los sectores industriales. En Tungurahua, más de 90 empresas ubicadas en los parques industriales se vieron afectadas, enfrentando un alza tarifaria por kilovatio hora, lo que representó un costo no programado que incrementaron sus gastos operativos.

2.3 Inteligencia Artificial y Machine Learning

La Inteligencia Artificial (IA) corresponde al área de la computación orientada al desarrollo de sistemas capaces de realizar tareas que tradicionalmente requieren inteligencia humana, como el reconocimiento de patrones, la toma de decisiones, la clasificación de información y la predicción de comportamientos futuros. En el contexto energético, la IA permite analizar grandes volúmenes de datos históricos para identificar relaciones que no siempre son evidentes mediante métodos tradicionales.

Dentro de la Inteligencia Artificial, el *Machine Learning* o aprendizaje automático se enfoca en el desarrollo de algoritmos capaces de aprender a partir de datos. A diferencia de los enfoques basados en reglas fijas, los modelos de aprendizaje automático identifican patrones internos en la información disponible y los utilizan para realizar predicciones sobre nuevos registros. Bishop, 2006; Géron, 2022.

En este trabajo, el problema de predicción del consumo energético se enmarca dentro

del aprendizaje supervisado, debido a que se dispone de una variable objetivo conocida: el consumo o demanda energética observada en periodos anteriores. A partir de esta información histórica, los modelos aprenden la relación entre las variables de entrada y el comportamiento del consumo eléctrico, con el propósito de estimar valores futuros.

La aplicación de técnicas de *Machine Learning* resulta pertinente en este tipo de problemas porque el consumo eléctrico puede depender de múltiples factores, como el tiempo, la ubicación geográfica, el tipo de consumo, las condiciones climáticas y la dinámica poblacional. Por tanto, estos modelos permiten transformar datos históricos en herramientas de apoyo para la planificación energética y la toma de decisiones.

2.4 Deep Learning

El *Deep Learning* o aprendizaje profundo es una rama del *Machine Learning* basada en redes neuronales artificiales con múltiples capas de procesamiento. Estas arquitecturas permiten aprender representaciones complejas de los datos mediante transformaciones sucesivas, lo que las hace adecuadas para problemas donde existen relaciones no lineales y patrones difíciles de capturar mediante modelos estadísticos convencionales Goodfellow et al., 2016.

Las redes neuronales profundas están compuestas por una capa de entrada, una o varias capas ocultas y una capa de salida. Cada neurona procesa la información mediante pesos, sesgos y funciones de activación. Durante el entrenamiento, el modelo ajusta sus parámetros internos con el objetivo de reducir la diferencia entre los valores reales y los valores predichos.

En el caso de la predicción del consumo energético, el *Deep Learning* resulta especialmente relevante porque permite modelar comportamientos complejos asociados a la demanda eléctrica. El consumo puede presentar tendencias, estacionalidades, variaciones territoriales y relaciones no lineales con variables externas. Estas características justifican el uso de arquitecturas profundas capaces de aprender patrones temporales a partir de datos históricos.

Sin embargo, el uso de modelos de aprendizaje profundo también exige una preparación rigurosa de los datos. Es necesario ordenar correctamente la información en el tiempo, normalizar las variables numéricas, tratar valores faltantes, evitar fuga de información y definir adecuadamente los conjuntos de entrenamiento, validación y prueba. De lo contrario, el mode-

lo puede obtener resultados aparentemente favorables, pero poco útiles para predecir periodos futuros.

2.5 Redes Neuronales Profundas aplicadas a series temporales

Una serie temporal es una secuencia de observaciones ordenadas cronológicamente. En este estudio, el consumo energético se analiza como una serie temporal, debido a que cada observación está asociada a un periodo específico y su interpretación depende del orden en que ocurren los datos.

Las redes neuronales profundas aplicadas a series temporales buscan aprender patrones históricos para estimar valores futuros. A diferencia de un análisis aislado de registros individuales, este enfoque considera que el valor de un periodo puede estar relacionado con los periodos anteriores. En el caso del consumo eléctrico, esta relación puede manifestarse mediante tendencias de crecimiento, comportamientos estacionales, cambios en la demanda por cantón o variaciones asociadas al tipo de consumo.

Para aplicar modelos neuronales a series temporales, es común construir ventanas de observación. Esto significa que el modelo utiliza una secuencia de datos pasados para predecir un valor futuro. Por ejemplo, una ventana de varios meses puede aportar información sobre la evolución reciente del consumo y permitir que el modelo capture patrones recurrentes. En este sentido, la predicción de series temporales requiere respetar la estructura cronológica de los datos, ya que el objetivo es estimar valores futuros a partir de información histórica disponible Hyndman y Athanasopoulos, 2021.

En este tipo de problemas, la división de los datos debe respetar el orden temporal. No es adecuado separar los registros de forma aleatoria, ya que esto puede generar fuga de información desde el futuro hacia el pasado. Por ello, los datos deben organizarse cronológicamente, utilizando los primeros periodos para entrenamiento, los siguientes para validación y los últimos para prueba.

Las redes neuronales profundas son adecuadas para este propósito porque pueden modelar dependencias temporales y relaciones no lineales entre variables. No obstante, su desempeño depende de la calidad del conjunto de datos, de la correcta construcción de las secuencias

temporales y de la arquitectura seleccionada para el problema de predicción.

2.6 Modelos a comparar

El presente estudio compara tres arquitecturas de redes neuronales profundas aplicadas a la predicción del consumo energético: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) y Transformer. Estos modelos fueron seleccionados por ser algunas de las arquitecturas más utilizadas en la predicción de series temporales mediante aprendizaje profundo. Cada una presenta un enfoque diferente para modelar las dependencias temporales de los datos: LSTM incorpora mecanismos de memoria, GRU simplifica esta estructura para reducir la complejidad computacional y Transformer emplea mecanismos de atención para analizar la información de la secuencia. La comparación de estas arquitecturas permite identificar cuál ofrece el mejor desempeño en la predicción del consumo eléctrico residencial de la provincia de Tungurahua.

2.6.1 *Long Short-Term Memory*

Las redes Long Short-Term Memory (LSTM) son una variante de las redes neuronales recurrentes diseñadas para capturar dependencias temporales de largo plazo. Su arquitectura incorpora mecanismos internos de memoria que permiten conservar, actualizar o descartar información relevante a lo largo de una secuencia.

En el contexto de la predicción energética, una red LSTM puede utilizar información de periodos anteriores para estimar el consumo futuro. Esta característica resulta útil cuando existen patrones estacionales, tendencias acumuladas o comportamientos históricos que influyen sobre la demanda eléctrica. Por esta razón, las LSTM son ampliamente utilizadas en problemas de predicción de series temporales. La arquitectura LSTM fue propuesta precisamente para abordar limitaciones de las redes recurrentes tradicionales al trabajar con dependencias temporales extensas Hochreiter y Schmidhuber, 1997.

2.6.2 *Gated Recurrent Unit*

Las redes Gated Recurrent Unit (GRU) son una arquitectura recurrente similar a las LSTM, pero con una estructura más simple. Estas redes utilizan compuertas internas para controlar el flujo de información a través del tiempo, aunque con una menor cantidad de parámetros.

La principal ventaja de las GRU es su eficiencia computacional. Al ser una arquitectura más ligera, puede requerir menor tiempo de entrenamiento y menor costo de procesamiento, manteniendo una buena capacidad para capturar dependencias temporales. Esta arquitectura fue introducida dentro de modelos encoder-decoder recurrentes, como una alternativa capaz de representar secuencias mediante mecanismos de compuertas Cho et al., 2014. En este trabajo, su comparación con las LSTM permite evaluar si una arquitectura más simple puede ofrecer resultados predictivos similares en el análisis del consumo energético.

2.6.3 *Transformer*

Los modelos Transformer se basan en mecanismos de atención, los cuales permiten identificar qué partes de una secuencia son más relevantes para generar una predicción. A diferencia de las arquitecturas recurrentes, los Transformers no procesan la información estrictamente de forma secuencial, sino que pueden analizar relaciones entre distintos puntos temporales de manera más flexible.

En series temporales, esta capacidad permite que el modelo otorgue mayor importancia a determinados periodos históricos que aportan información relevante para la predicción. En el caso del consumo energético, esto puede ser útil para capturar patrones de largo plazo, relaciones no lineales y posibles efectos estacionales dentro de la demanda eléctrica. La arquitectura Transformer fue propuesta como un modelo basado únicamente en mecanismos de atención, prescindiendo de recurrencia y convoluciones en su diseño original Vaswani et al., 2017.

No obstante, los Transformers suelen requerir una configuración más cuidadosa y pueden demandar mayores recursos computacionales. Por ello, su inclusión en el estudio permite evaluar si su mayor complejidad se traduce en una mejora real del desempeño predictivo frente a modelos recurrentes como LSTM y GRU.

2.6.4 Criterio de comparación

Para garantizar una comparación homogénea, los tres modelos deberán evaluarse bajo condiciones experimentales equivalentes. Esto implica utilizar la misma variable objetivo, la misma estructura de datos, ventanas temporales comparables y una división cronológica consistente entre entrenamiento, validación y prueba.

La selección del modelo más adecuado no debe depender únicamente de la complejidad de la arquitectura, sino de su desempeño real en el problema planteado. Un modelo más complejo no necesariamente será superior si su mejora en precisión es marginal o si requiere un costo computacional considerablemente mayor. Por tanto, la comparación deberá considerar precisión, estabilidad, eficiencia y utilidad práctica para apoyar la planificación energética en la provincia de Tungurahua.

2.7 Métricas de evaluación

La evaluación del desempeño es una etapa importante en el desarrollo de modelos predictivos, especialmente en tareas o procesos aplicados a la predicción de consumo energético. Una métrica de evaluación permite medir qué tan cercanas son las predicciones de un modelo a los valores reales observados, cuantificando el error residual sobre un conjunto de datos de prueba.

■ Error Absoluto Medio (MAE)

Es una de las métricas más simples e intuitivas para la evaluación de modelos de regresión. Mide el promedio de las diferencias absolutas entre los valores predichos por el modelo y los valores reales observados.

Su fórmula es:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.2)$$

Donde n es el número total de observaciones, y_i representa el valor real de la i -ésima muestra y $\hat{y}_i$ es el valor predicho por el modelo para esa misma muestra.

Una característica clave del MAE es que representa un puntaje lineal, lo que significa que todas las diferencias individuales contribuyen equitativamente al cálculo del error medio. Gracias a esto, la métrica es altamente resistente a la presencia de valores atípicos (outliers) en los datos. Además, se expresa en las mismas unidades que la variable de respuesta, lo que la convierte en una métrica muy intuitiva y sencilla de interpretar para evaluar la magnitud del error.

■ Error Cuadrático Medio (MSE)

Es una métrica fundamental en el análisis de regresión que cuantifica el promedio de los cuadrados de las diferencias entre los valores observados y los valores predichos.

Su fórmula es:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.3)$$

Donde n es el número de observaciones, y_i es el valor real y $\hat{y}_i$ es el valor predicho.

Debido a que las diferencias se elevan al cuadrado, esta métrica penaliza de forma desproporcionada los errores grandes, otorgándoles un mayor peso en comparación con los errores pequeños, lo que la hace muy sensible a los valores atípicos.

Su principal limitación en cuanto a interpretabilidad es que sus resultados se expresan en las unidades originales al cuadrado, lo que no siempre resulta intuitivo.

■ Raíz del Error Cuadrático Medio (RMSE)

Se define como la raíz cuadrada del MSE y constituye una de las métricas más ampliamente utilizadas en la evaluación de modelos de regresión y pronóstico de series de tiempo.

Su fórmula es:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.4)$$

Al igual que el MSE, el RMSE penaliza con mayor fuerza los errores de gran magnitud, por lo que es sensible a los valores atípicos. Un valor más bajo de RMSE indica que el modelo genera predicciones que se acercan con gran exactitud a la realidad.

■ Coeficiente de Determinación (R^2)

El R^2 es una medida estadística que indica la proporción de la varianza en la variable dependiente (o variable objetivo) que puede ser explicada por las variables independientes presentes en el modelo.

Su fórmula es:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.5)$$

donde $\bar{y}$ es la media de los valores reales, el numerador representa la suma de los residuales al cuadrado y el denominador la suma total de cuadrados.

Por lo general, el rango de valores del R^2 oscila entre 0 y 1. Un valor de 1 indica que el modelo explica perfectamente la varianza de los datos, mientras que un valor de 0 indica que el modelo no aporta ninguna explicación sobre la variación.

■ Error Porcentual Absoluto Medio (MAPE)

Es una métrica que expresa el error de predicción como un porcentaje relativo respecto a los valores reales, lo que facilita su interpretación y comparación entre modelos o conjuntos de datos con diferentes unidades de medida.

Su fórmula es:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (2.6)$$

donde y_i es el valor real, $\hat{y}_i$ es el valor predicho y n el número total de observaciones.

Dado que se expresa como un porcentaje, es una métrica independiente de la escala, lo que resulta sumamente útil y fácil de interpretar, especialmente para comparar el rendimiento de un modelo a través de conjuntos de datos con magnitudes muy diferentes.

2.8 Explicación conceptual de ACF y PACF

El análisis de dependencia temporal constituye una etapa fundamental en el estudio de series temporales, ya que permite identificar la influencia que ejercen las observaciones pasadas sobre los valores futuros de la serie. Entre las herramientas más utilizadas para este

propósito se encuentran la función de autocorrelación (ACF) y la función de autocorrelación parcial (PACF), las cuales proporcionan información relevante sobre la estructura temporal de los datos y facilitan la selección de configuraciones adecuadas para los modelos predictivos.

- **Función de Autocorrelación (ACF):** Mide la relación existente entre los valores de una serie temporal y sus propios valores observados en períodos anteriores. Su principal utilidad es identificar la persistencia temporal de la serie y determinar cuántos rezagos mantienen una influencia significativa sobre las observaciones actuales.
- **Función de Autocorrelación Parcial (PACF):** Mide la correlación entre una observación y un rezago específico, eliminando el efecto de los rezagos intermedios. Esta herramienta permite identificar de manera más precisa cuáles son los rezagos que aportan información relevante para explicar el comportamiento de la serie, aislando el impacto directo de cada periodo.
- **Estructura Autorregresiva (AR):** Un modelo autorregresivo (AR, por sus siglas en inglés Autoregressive) describe una serie temporal cuyo valor actual depende de uno o varios valores observados en períodos anteriores. La notación $AR(p)$ indica que el valor presente está influenciado por los últimos p períodos. Por ejemplo, una estructura $AR(1)$ significa que el comportamiento actual de la serie depende principalmente del valor registrado en el período inmediatamente anterior, mientras que una estructura $AR(2)$ considera la influencia de los dos períodos previos.

El análisis conjunto de ACF y PACF es ampliamente utilizado en series temporales para comprender la estructura de dependencia temporal de los datos, identificar patrones autorregresivos y apoyar la selección de ventanas temporales adecuadas para los modelos predictivos.

CAPÍTULO 3: DESARROLLO

3.1 Metodología de proyectos

La metodología de proyectos corresponde al conjunto de procedimientos, fases y técnicas utilizadas para planificar, desarrollar, controlar y finalizar un proyecto de manera organizada optimizando recursos.

En el área de ciencia de datos e inteligencia artificial, las metodologías de proyectos facilitan la gestión del ciclo de vida de los modelos analíticos, desde la definición del problema hasta la validación de resultados.

La tabla siguiente presenta una comparación de las principales ventajas y desventajas de las metodologías más utilizadas en proyectos de ciencia de datos.

Tabla 1: *Comparación de metodologías para proyectos de ciencia de datos*

Metodología	Ventajas	Desventajas
CRISP-DM	Estructurada, iterativa, orientada al negocio y ampliamente aceptada en investigación e industria.	No define prácticas específicas para el desarrollo de software.
KDD	Pionera en minería de datos. Enfatiza la selección y transformación de datos.	Es menos flexible y da poca importancia a la comprensión del negocio.
SEMMA	Muy enfocada en el modelado y análisis estadístico.	Omite la comprensión del negocio y el despliegue. Además, está ligada a herramientas SAS.
TDSP	Diseñada para proyectos colaborativos y de producción. Incluye control de versiones y despliegue.	Es más compleja y está orientada a equipos de desarrollo.
OSEMN	Sencilla y fácil de aplicar en proyectos pequeños.	Carece de documentación formal y de una fase de comprensión del negocio.

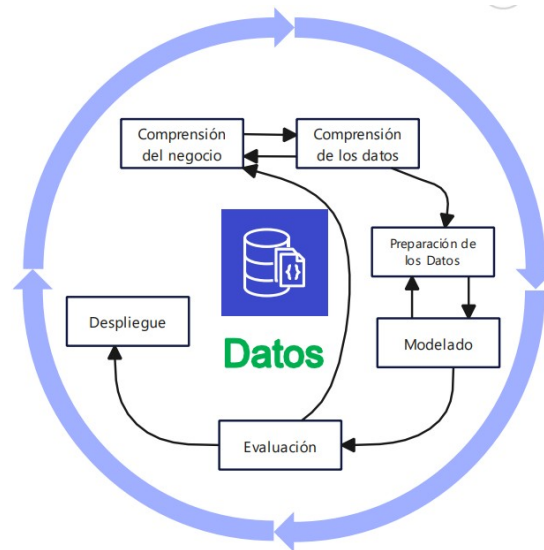
El presente proyecto utilizará la metodologías CRISP-DM debido a su enfoque flexible, iterativo y orientado al análisis de datos.

3.1.1 CRISP-DM

Son las siglas de Cross-Industry Standard Process for Data Mining, es un método orientado a trabajos de minería de datos, mismo que incluye un resumen del ciclo de vida del proyecto, describiendo las tareas necesarias en cada fase, optimizando los recursos.

Figura 6

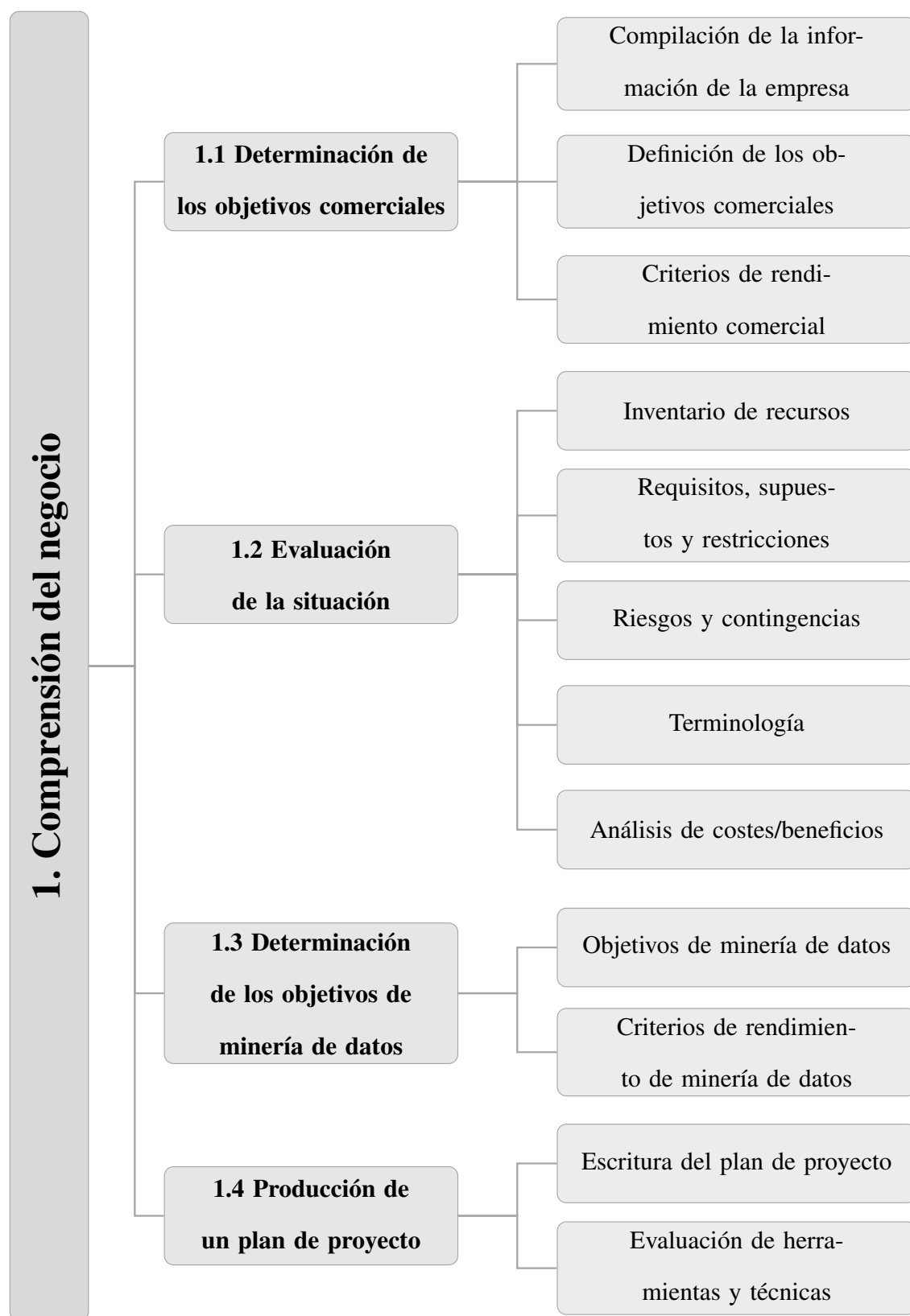
Fases de la metodología CRISP-DM



Nota. Elaboración propia basada en la metodología CRISP-DM.

3.1.1.1 Comprensión del negocio

Esta fase se caracteriza por la investigación y análisis que deben realizarse antes de aplicar la minería de datos como tal. Los analistas se dedican a indagar sobre el giro del negocio, personas involucradas y recursos ya que es necesario comprender los objetivos de la organización y asegurar que todos los involucrados estén de acuerdo, permitiendo planificar correctamente el proyecto y optimizar recursos. Esta fase comprende:

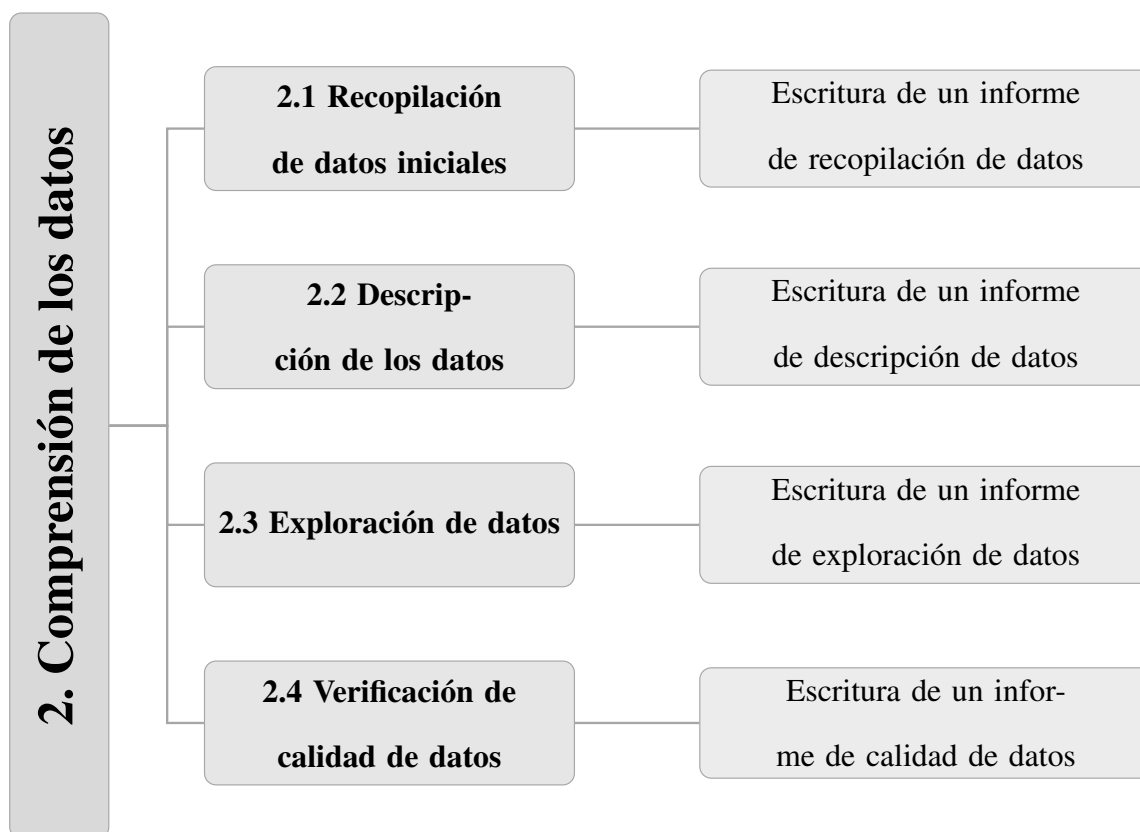
Figura 7*Organigrama horizontal de la fase comprensión del negocio**Nota.* Elaboración propia basada en la metodología CRISP-DM.

3.1.1.2 Comprensión de datos

En esta fase se analizan las fuentes de datos, la disponibilidad y calidad de los mismos. Este paso es esencial para evitar problemas posteriores, en donde los datos son la materia prima del análisis.

Figura 8

Organigrama horizontal de la fase comprensión de los datos



Nota. Elaboración propia basada en la metodología CRISP-DM.

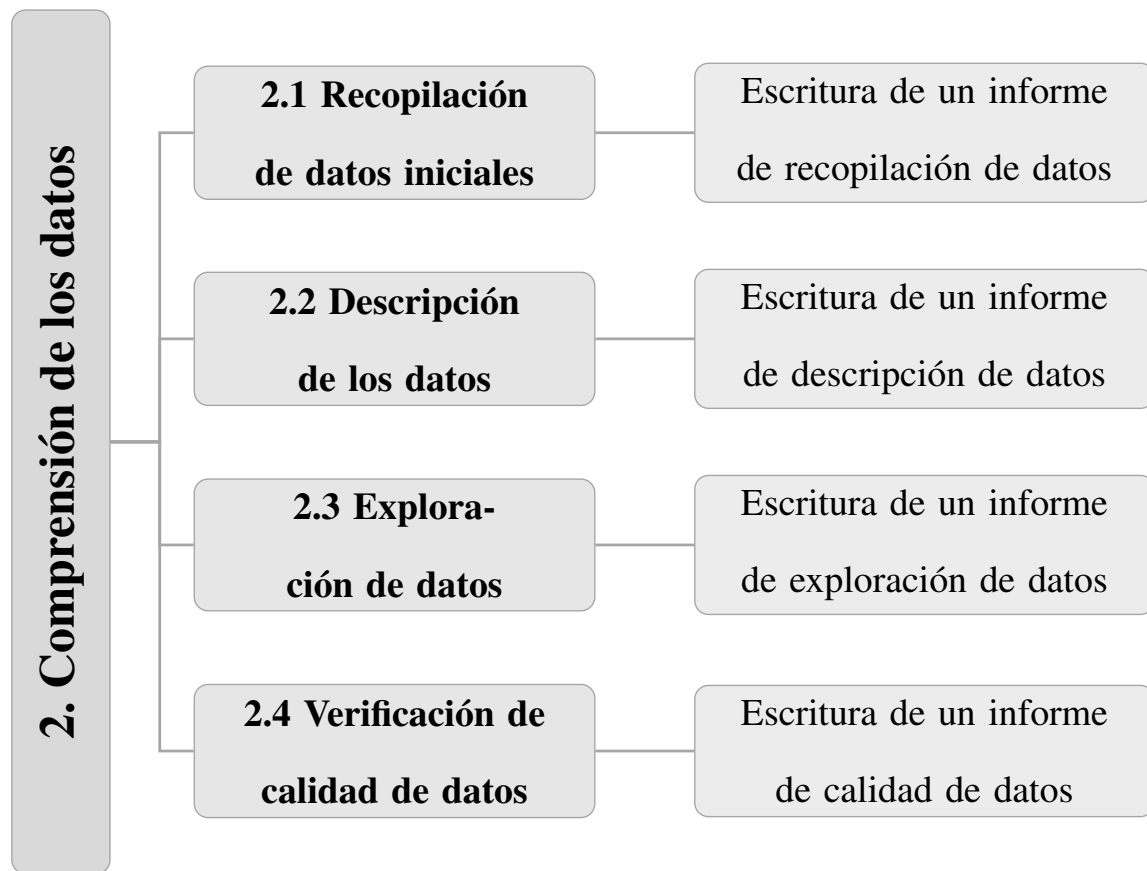
3.1.1.3 Preparación de datos

Esta fase representa la mayor cantidad de exigencia dentro del proyecto (entre un 50-70 %). Dependiendo el proyecto u organización esta fase incluye las siguientes tareas: Fusión de conjuntos o registros de datos, Selección de una muestra de un subconjunto de datos, Agregación de registros, Derivación de nuevos atributos, Clasificación de los datos para el modelado, Eliminación o sustitución de valores en blanco o perdidos, División en conjuntos de datos de

prueba y entrenamiento.

Figura 9

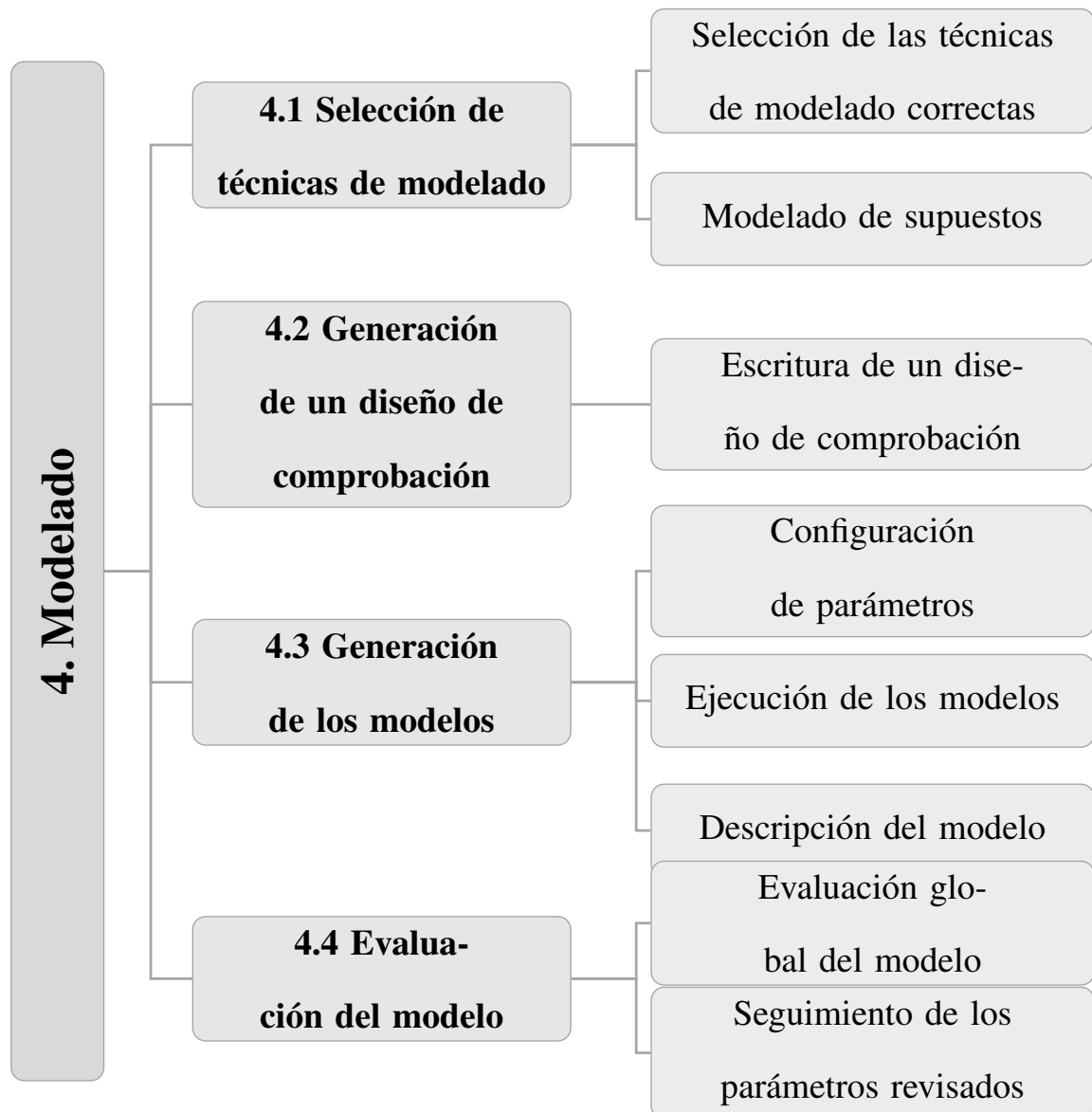
Organigrama horizontal de la fase comprensión de los datos



Nota. Elaboración propia basada en la metodología CRISP-DM.

3.1.1.4 Modelado

Esta fase representa propiamente al modelado de los datos. Los analistas ejecutan varios modelos definiendo datos, parámetros e hiperparámetros para obtener salidas corregidas a través de iteraciones.

Figura 10*Organigrama horizontal de la fase modelado*

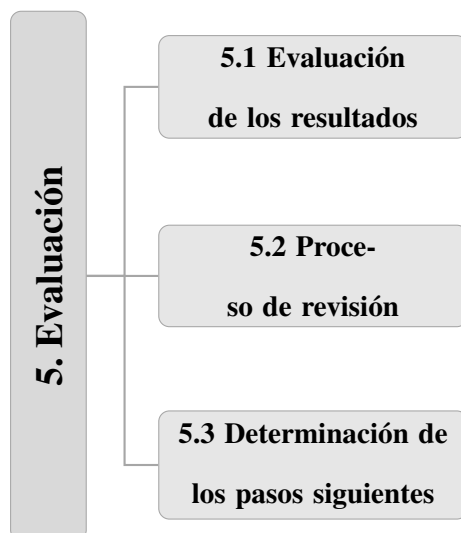
Nota. Elaboración propia basada en la metodología CRISP-DM.

3.1.1.5 Evaluación

En esta fase se evalúa los resultados utilizando los criterios ya establecidos en la fase inicial. En este punto se pueden generar modelos así como descubrimientos.

Figura 11

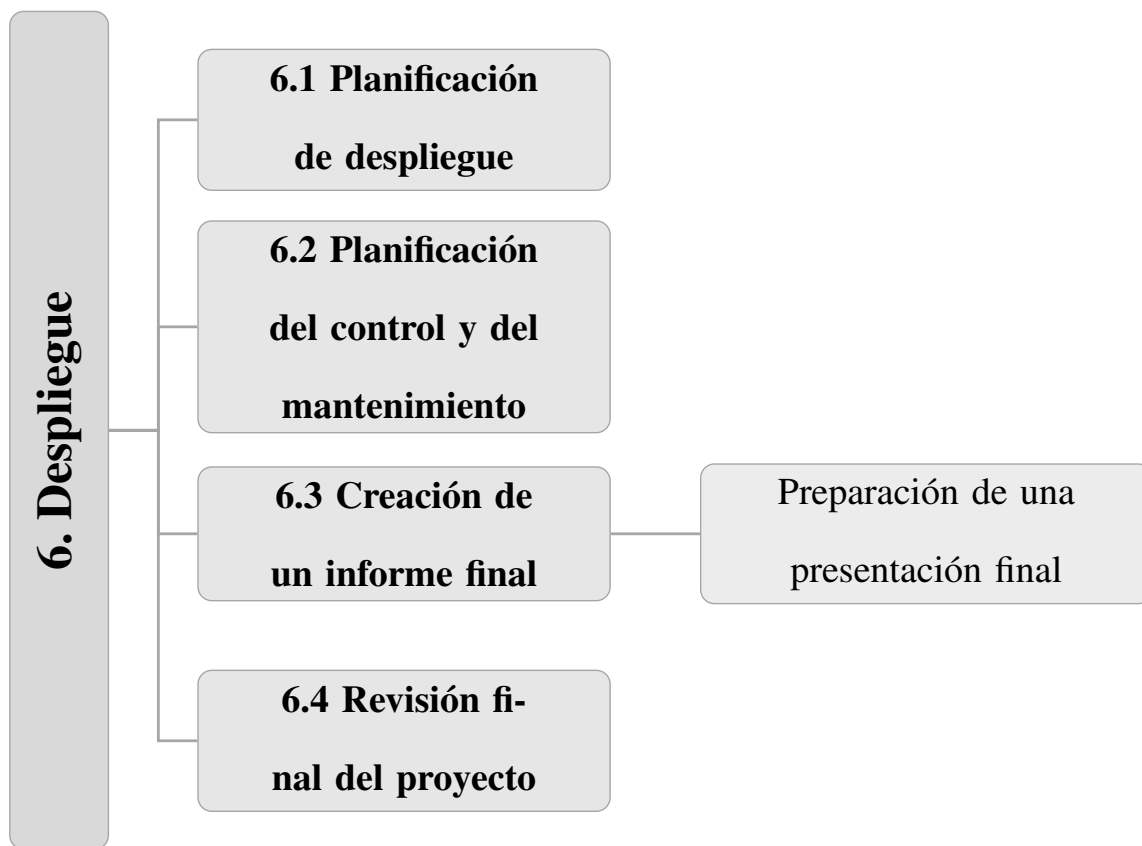
Organigrama horizontal de la fase de evaluación



Nota. Elaboración propia.

3.1.1.6 Despliegue

Esta etapa consiste en utilizar los conocimientos adquiridos e implementar mejoras en la organización.

Figura 12*Organigrama horizontal de la fase de despliegue**Nota.* Elaboración propia.

3.1.2 Estructura y organización del proyecto planteado

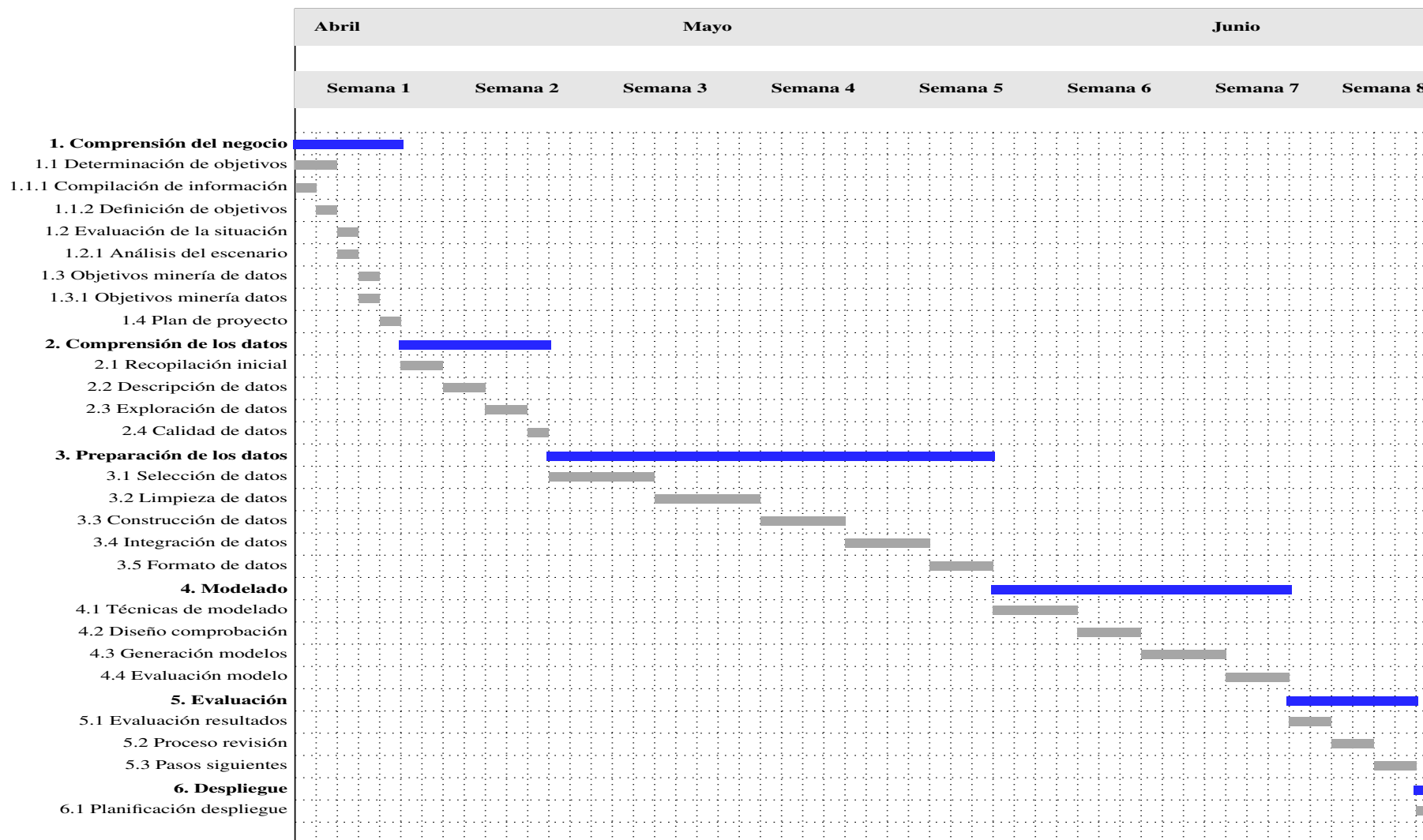
3.1.2.1 Esquema General del Diseño

Para el caso del presente proyecto el esquema General de fases y tareas se describe a continuación

3.1.2.2 Diagrama de Gantt del proyecto

Figura 13

Cronograma de actividades del proyecto basado en la metodología CRISP-DM

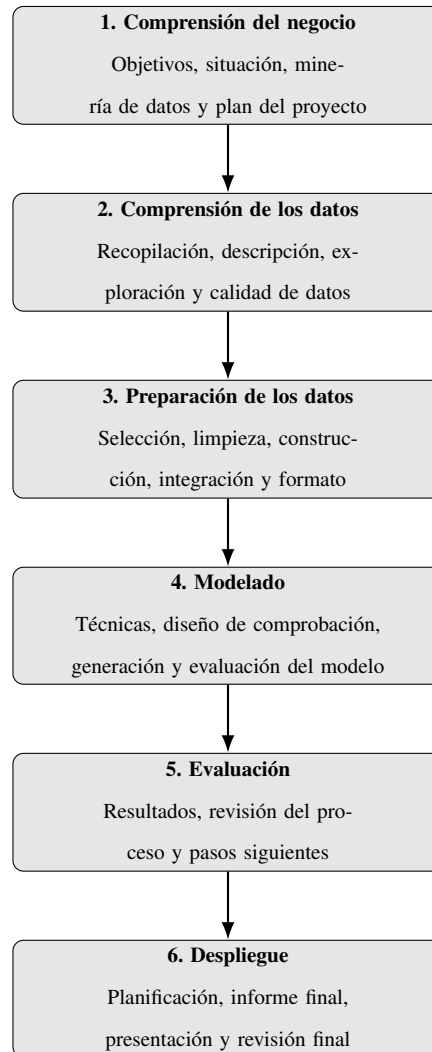


Nota. Elaboración propia. El cronograma muestra la planificación de actividades del proyecto siguiendo las fases de la metodología CRISP-DM.

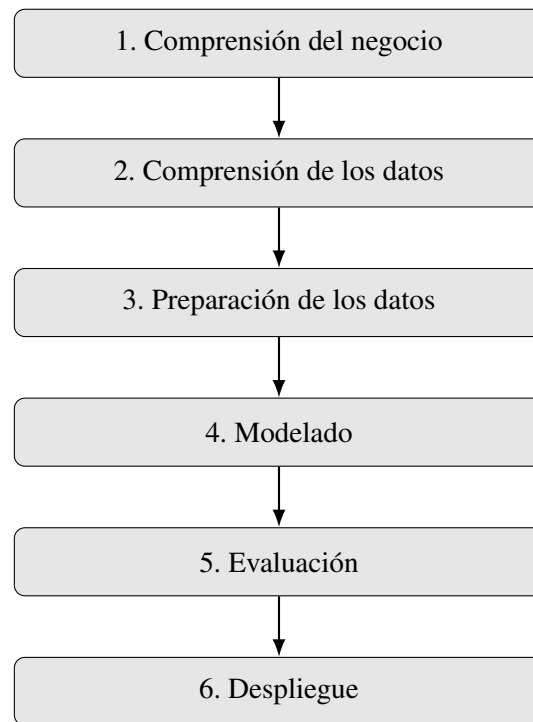
3.1.2.3 Flujo metodológico

Figura 14

Diagrama de flujo de las fases del proyecto



Nota. Elaboración propia.

Figura 15*Pipeline de la metodología CRISP-DM*

Nota. Elaboración propia.

3.2 Comprensión del negocio

La fase de comprensión del negocio nos permite entender los datos y los correspondientes objetivos de proyecto. En el capítulo 2, se describe a profundidad la descripción teórica de la temática planteada, en esta sección se analizará la parte práctica para el desarrollo del modelo.

3.2.1 Herramientas a utilizar

3.2.1.1 Python

"Python es un lenguaje de programación creado por Guido van Rossum a principios de los años 90 cuyo nombre está inspirado en el grupo de cómicos ingleses "Monty Python". Es un lenguaje similar a Perl, pero con una sintaxis muy limpia y que favorece un código legible. Se trata de un lenguaje interpretado o de script, con tipado dinámico fuerte, multiplataforma y

orientado a objetos"(González Duque, 2008).

3.2.1.2 *Pandas*

Pandas es una librería construida sobre NumPy que se utiliza para análisis y manipulación de datos. Su estructura principal es el DataFrame, similar a una tabla de Excel o una base de datos, permitiendo filtrar, ordenar, agrupar, limpiar y analizar información de manera sencilla y eficiente.

3.2.1.3 *NumPy*

NumPy es una librería de Python enfocada en cálculo numérico y manejo eficiente de arreglos multidimensionales (arrays). Permite realizar operaciones matemáticas rápidas, álgebra lineal, estadísticas, matrices y procesamiento numérico de alto rendimiento. Es la base de muchas librerías de ciencia de datos e inteligencia artificial.

“NumPy proporciona almacenamiento y manipulación eficiente de arreglos densos tipados en Python” (VanderPlas, 2016).

3.2.1.4 *Matplotlib/Seaborn*

Librerías para visualización de datos. Matplotlib permite crear gráficos básicos y personalizados, mientras que Seaborn facilita gráficos estadísticos más avanzados y visualmente atractivos. Matplotlib y Seaborn son librerías empleadas para la visualización de datos y representación gráfica de información estadística.

3.2.1.5 *TensorFlow/Keras o PyTorch*

Librerías utilizadas para desarrollar modelos de inteligencia artificial y redes neuronales profundas. Permiten entrenar modelos de machine learning y deep learning para clasificación, predicción, visión artificial y procesamiento de lenguaje natural. TensorFlow y Keras son librerías utilizadas para construir y entrenar modelos de aprendizaje profundo mediante redes neuronales artificiales.

3.2.1.6 Scikit-learn

Librería de Python enfocada en machine learning tradicional. Incluye algoritmos como regresión lineal, árboles de decisión, clustering, SVM y herramientas para entrenamiento y evaluación de modelos.

Scikit-learn es una librería de Python que proporciona herramientas simples y eficientes para análisis de datos y machine learning.

3.3 Comprensión de los datos - Descripción del dataset

3.3.1 Fuente de datos energéticos

Esta información se obtuvo de la plataforma del Sistema de Información de Datos Energéticos (SISDAT) de la Agencia de Regulación y Control de Electricidad (Agencia de Regulación y Control de Electricidad (ARCONEL), 2026).

Figura 16

Página principal de la plataforma de reportes de ARCONEL

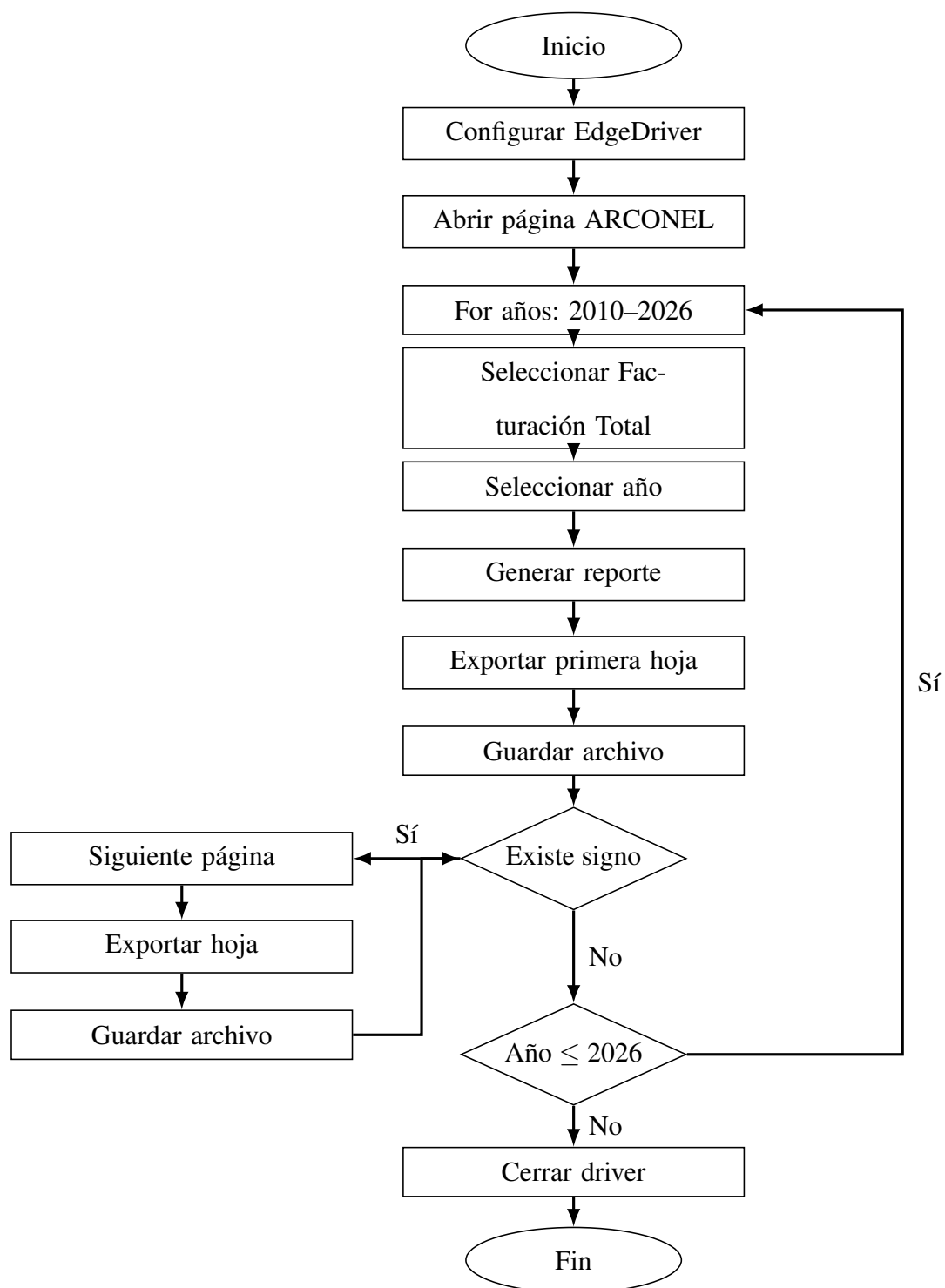
Empresa	Corriente	Año	Fecha	Mes	Grupo Consumo	Tipo Cliente	Clientes	Energía Facturada (kWh)	Facturación (Servicio)
87 CIEL-Quayas Los Rios		2010	7 Jul	Jul	Otros	Regulado	2072	15775.230906	
87 CIEL-Quayas Los Rios		2010	8 Ago	Ago	Otros	Regulado	2568	6172.296832	
87 CIEL-Quayas Los Rios		2010	8 Ago	Ago	Residencial	Regulado	235588	42713.239676	
87 CIEL-Quayas Los Rios		2010	10 Oct	Oct	Alumbrado Público	Regulado	69	2035.763985	

Nota. Captura de pantalla obtenida de la plataforma de reportes de ARCONEL (Agencia de Regulación y Control de Electricidad (ARCONEL), 2026).

Se desarrolló un programa con Selenium con el objetivo de extraer los datos de manera automática (ETL). A continuación se muestra el diagrama de flujo.

Figura 17

Diagrama de flujo para la descarga automatizada de reportes energéticos



Nota. Elaboración propia.

3.3.2 *Datos climáticos*

Los datos climáticos se obtuvieron de NASA POWER (Prediction Of Worldwide Energy Resources) es una plataforma gratuita de datos climáticos y meteorológicos creada por NASA. Proporciona información histórica y en tiempo casi real obtenida de satélites y modelos climáticos globales.

Permite descargar variables como: temperatura, humedad, precipitación, radiación solar, velocidad del viento, presión atmosférica, entre otras.

Es muy usada en:

Ciencia de Datos, Machine Learning, Agricultura, Energía solar, Investigaciones académicas, predicción climática.

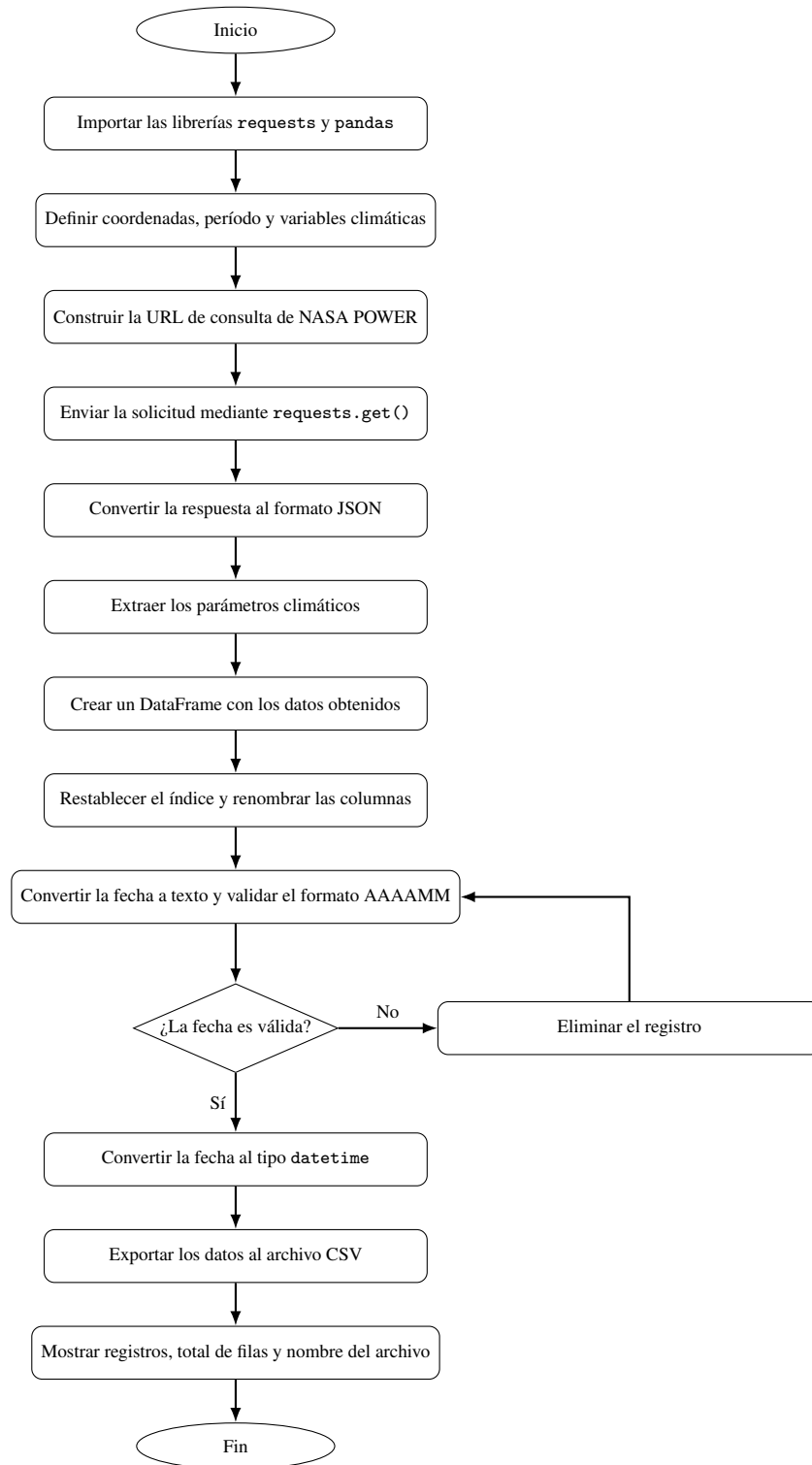
La ventaja principal es que:

no se necesitan estaciones meteorológicas físicas, se puede consultar cualquier coordenada del mundo, tiene datos desde aproximadamente 1981, y posee API directa para Python.

Se hace uso de esta librería para generar los datos correspondientes. A continuación se muestra el diagrama de flujo seguido para extraer la data:

Figura 18

Diagrama de flujo para la extracción y procesamiento de datos climáticos



Nota. Elaboración propia. Código utilizado para la transformación, limpieza y exportación de los datos climáticos obtenidos desde la API NASA POWER.

3.3.3 *Datos poblacionales*

Los datos de población se obtuvieron del INEC.

El Instituto Nacional de Estadística y Censos (INEC) es la institución oficial del Ecuador encargada de recopilar, procesar y publicar estadísticas nacionales sobre población, economía, empleo, salud, educación, vivienda, migración, entre otros temas. Realiza censos, encuestas y registros estadísticos para apoyar la toma de decisiones del gobierno, universidades y empresas.

El INEC maneja una gran base de datos estadística del país, pero no es una “base de datos pública con toda la información personal de todos los ecuatorianos”. Lo que posee son:

censos poblacionales, registros administrativos, encuestas nacionales, proyecciones poblacionales, estadísticas de nacimientos, defunciones, empleo, pobreza, etc.

Uno de sus sistemas más importantes es el: Registro Estadístico Base de Población del Ecuador (REBPE), que consolida información demográfica nacional proveniente de distintas fuentes estatales. Según el INEC, el REBPE 2023 contiene información estadística de más de 17 millones de personas registradas en Ecuador.

Sin embargo, los datos personales individuales están protegidos, el INEC publica principalmente datos anonimizados y agregados, y existe “secreto estadístico”, por lo que no cualquiera puede acceder a información privada de ciudadanos específicos.

3.4 Preparación y preprocesamiento de datos

3.4.1 *Período y área de estudio*

El presente estudio se desarrolló utilizando información histórica correspondiente al período comprendido entre enero de 2014 y diciembre de 2025. Este intervalo de tiempo abarca un total de 144 observaciones mensuales, lo que permite analizar la evolución del consumo energético e identificar variaciones relevantes para la construcción de los modelos predictivos.

La investigación se enfocó en el sector residencial de la provincia de Tungurahua. Para ello, se consolidó la información proveniente de los distintos cantones y parroquias de la provincia mediante un proceso de agregación mensual, obteniendo una representación del com-

portamiento energético provincial.

La frecuencia temporal utilizada es mensual, debido a que las variables de consumo eléctrico, facturación, condiciones climáticas y características demográficas presentan una disponibilidad consistente en este nivel de agregación.

Tabla 2: *Características del conjunto de datos analizado*

Característica	Descripción
Período analizado	Enero 2014 – Diciembre 2025
Número de observaciones	144 meses
Cobertura geográfica	Provincia de Tungurahua, Ecuador
Sector analizado	Residencial
Nivel de agregación	Provincial
Frecuencia temporal	Mensual
Unidad de análisis	Consumo energético residencial provincial

3.4.2 Preprocesamiento de datos

3.4.2.1 Limpieza de datos

- **Normalización de nombres de variables.**- Con el objetivo de garantizar la consistencia y uniformidad de los datos, se realizó un proceso de normalización de los nombres de las variables. Este procedimiento permitió estandarizar la nomenclatura de las columnas aplicando las siguientes transformaciones:

- Conversión de todos los nombres a minúsculas siguiendo el formato `snake_case`.
- Eliminación de tildes, caracteres especiales y símbolos no alfanuméricos.
- Sustitución de símbolos específicos, como % por pct.
- Reemplazo de espacios y separadores por guiones bajos (_).

Esta etapa facilita la manipulación de los datos durante las fases posteriores de análisis, transformación y modelado.

Tabla 3: *Normalización de nombres de variables*

Nombre original	Nombre normalizado
Energía Facturada (kWh)	energy_billed_kwh
Facturación Servicio Eléctrico (USD)	billing_usd
Número Clientes	customer_count
Humidity %	humidity_pct
Población Total años	total_population_years

- **Estandarización Geográfica.**- Durante la exploración de los datos se identificaron diferencias en la nomenclatura de algunos cantones, ocasionadas por variaciones en la digitación. Por lo tanto con el propósito de garantizar la consistencia de la información geográfica, se realizó un proceso de homologación de la variable canton.

Para corregir estas inconsistencias, se unificaron denominaciones equivalentes bajo su nombre oficial:

- Los registros PELILEO fueron estandarizados como SAN PEDRO DE PELILEO.
- Los registros PILLARO y SANTIAGO DE PILLARO fueron consolidados como SANTIAGO DE PÍLLARO.

Como resultado de este proceso, la variable quedó conformada exclusivamente por los nueve cantones oficiales de la provincia de Tungurahua, asegurando uniformidad en los análisis posteriores y evitando la generación de categorías duplicadas que pudieran afectar los resultados estadísticos y predictivos.

Tabla 4: Homologación de nombre de cantones

Valor original	Valor estandarizado
PELILEO	SAN PEDRO DE PELILEO
PILLARO	SANTIAGO DE PÍLLARO
SANTIAGO DE PILLARO	SANTIAGO DE PÍLLARO

Nota. Esta tabla muestra solo los cantones con variaciones en la digitación.

- **Renombrado y Traducción de Variables.**- Con el fin de adoptar una nomenclatura consistente y alineada con las buenas prácticas utilizadas en proyectos de Ciencia de Datos y Machine Learning, se realizó un proceso de traducción y renombrado de las variables.

Este procedimiento permitió convertir los nombres originales de las columnas desde español hacia inglés, utilizando una terminología descriptiva.

Tabla 5: Renombrado y traducción de variables

Nombre original	Nombre traducido
anio	year
mes	month
empresa	company
grupo_consumo	consumption_group
provincia	province
canton	canton
parroquia	parish
numero_clientes	customer_count
energia_facturada_kwh	energy_billed_kwh
facturacion_servicio_electrico_usd	electric_service_billing_usd

Nota. La tabla presenta una muestra representativa del proceso de traducción y renombrado de variables. El diccionario completo de variables se encuentra disponible en el Anexo A.

- **Conversión de Tipos de Datos.-** Se realizó la conversión explícita de los tipos de datos con el fin de garantizar una representación adecuada de cada variable y optimizar su utilización durante las etapas de análisis exploratorio, procesamiento y modelado.

Tabla 6: *Tipos de datos*

Tipo	Variables
int64	year, customer_count, energy_billed_kwh, pop_0_4, pop_5_9, pop_10_14, pop_15_19, pop_20_24, pop_25_29, pop_30_34, pop_35_39, pop_40_44, pop_45_49, pop_50_54, pop_55_59, pop_60_64, pop_65_69, pop_70_74, pop_75_79, pop_80_84, pop_85_89, pop_90_94, pop_95_99, pop_100_plus, total_population_years
float64	electric_service_billing_usd, collection_usd, firefighter_tax_usd, trash_collection_usd, senior_citizen_discount_usd, dignity_tariff_usd, subsidized_energy_pec_kwh, subsidized_value_pec_usd, temp_avg_c, temp_max_c, temp_min_c, precipitation_mm, humidity_pct, wind_speed_m_s
category	canton_standardized, parish, canton, province, consumption_group, company, month

Impacto en el modelado: La homologación y estandarización de los nombres y tipos de las variables constituye una etapa fundamental para el modelado multivariable. Al garantizar un formato uniforme y libre de inconsistencias, se evitan pérdidas de información temporales durante la integración de las series de consumo eléctrico con las variables demográficas y climáticas. Como resultado, la matriz de entrada utilizada por las redes neuronales contiene registros completos y correctamente alineados para cada período mensual.

3.4.2.2 *Tratamiento de valores inconsistentes*

- **Tratamiento de valores nulos.-** Al analizar el dataset, se detectó que los registros que contienen valores nulos pertenecen a ciertas observaciones correspondientes al año 2026. Dado que este subconjunto representa una proporción mínima, se optó por eliminar directamente los registros del año 2026.
- **Tratamiento de valores negativos.-** Durante el proceso de validación de los datos se identificaron registros con valores negativos en las variables `energy_billed_kwh` y `billing_usd`, por lo que se decide aplicar un ajuste a estas variables.

Tabla 7: *Tratamiento de valores nulos*

Métrica	Valor
Registros con nulos	261
Porcentaje de nulos	0.75 %
Período afectado	Año 2026
Acción aplicada	Eliminación de registros

Con el propósito de preservar la trazabilidad de la información original, las variables iniciales no fueron modificadas. En su lugar, se generaron las variables ajustadas `energy_billed_kwh_adjusted` y `billing_usd_adjusted`, sobre las cuales se aplicaron reglas de corrección definidas según la combinación de signos observada en cada registro:

- **Caso A:** Cuando tanto la energía facturada como la facturación presentaban valores negativos, para este caso se interpretó el registro como un ajuste administrativo de una transacción previamente registrada. Bajo este supuesto, se consideró que el consumo efectivamente ocurrió y que el signo negativo correspondía únicamente a un ajuste. En consecuencia, se aplicó el valor absoluto a ambas variables.
- **Caso B:** Cuando la energía facturada presentaba un valor positivo y la facturación un valor negativo, se asumió que el consumo energético registrado era válido, mientras que el valor monetario correspondía a un ajuste financiero. En estos casos se conservó el valor original de energía y se asignó un valor de cero a la facturación ajustada.
- **Caso C:** Cuando la energía facturada era negativa y la facturación positiva, se asumió una combinación inconsistente, ya que implica la existencia de un cobro asociado a un consumo energético negativo. Por lo tanto ambas variables fueron ajustadas en cero.

Adicional, se incorporaron las variables `adjustment_type`, que registra el tipo de ajuste aplicado, e `is_corrected`, que identifica si un registro fue modificado durante esta etapa.

- **Registros con energía >0 pero 0 clientes.**- Se identificaron registros de demanda de energía en lugares donde la cantidad de clientes era 0.

A primera vista, esta situación podría interpretarse como una anomalía; sin embargo, una revisión del contexto de los datos permitió identificar que dichos registros correspondían principalmente a servicio de alumbrado público.

Debido a esto se decidió conservar los registros dentro del conjunto de datos y en lugar de eliminarlos o modificarlos, se incorporó una variable denominada `is_public_lighting`, la cual asigna el valor de 1 cuando se identifica un registro con consumo de energía positivo y cero clientes asociados, y 0 en caso contrario.

- **Inconsistencias temporales.**- Se identificó una inconsistencia en los registros correspondientes a los años 2011, 2012 y 2013. A diferencia del resto del conjunto de datos, que presentaba una frecuencia mensual, la información de estos años se encontraba disponible únicamente para el mes de diciembre.

Con el fin de evitar sesgos en los modelos de redes neuronales, se decidió excluir permanentemente todos los registros correspondientes a los años 2011, 2012 y 2013.

Tabla 8: *Distribución temporal de los registros con inconsistencias*

Año	Meses registrados	Meses faltantes	Mes disponible
2011	1	11	Diciembre
2012	1	11	Diciembre
2013	1	11	Diciembre

- **Análisis de valores extremos.**- Durante el análisis exploratorio, se identificó la presencia de valores extremos en las variables de consumo y facturación. Estas observaciones evidenciaron distribuciones con asimetría positiva y colas largas, lo cual indica períodos con niveles de consumo significativamente superiores al promedio.

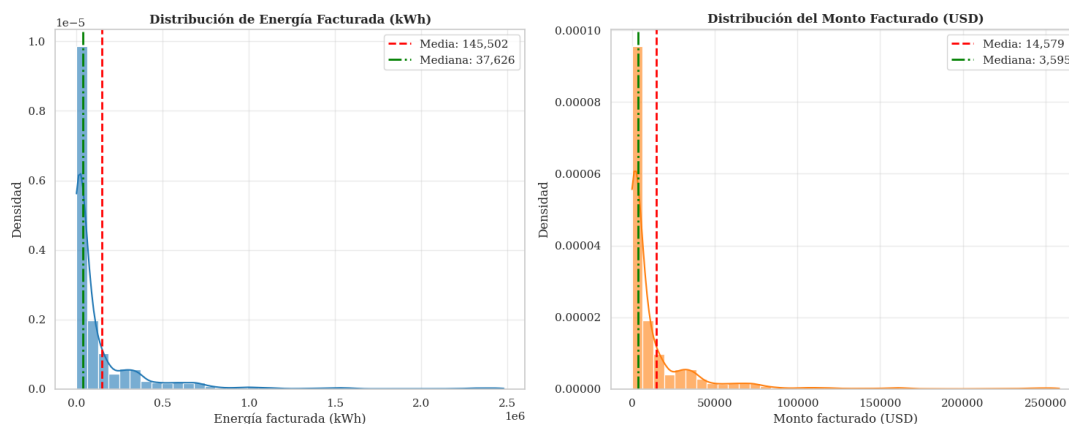
Tras un análisis contextual para determinar si dichos valores respondían a errores de registro o a comportamientos propios del sistema, se confirmó que los registros representan consumos válidos asociados al servicio eléctrico.

En consecuencia, se determinó no aplicar técnicas de eliminación de valores atípicos;

remover estas observaciones habría supuesto una pérdida de información importante que afectaría la capacidad de los modelos de redes neuronales.

Figura 19

Distribución estadística de la energía facturada (kWh) y el monto facturado (USD).



Nota. Las distribuciones evidencian una asimetría positiva, con predominio de abonados de bajo consumo y pocos de alta demanda, justificando la normalización de las variables

Impacto en el modelado: Las redes neuronales profundas, como LSTM y GRU, son sensibles a valores atípicos y registros inconsistentes. La permanencia de observaciones con facturación en cero, originadas por errores del sistema o clientes inactivos, podría introducir distorsiones en el proceso de entrenamiento y afectar la estimación de los parámetros del modelo. La eliminación de estos registros reduce el ruido en los datos, favorece una convergencia más estable durante el entrenamiento y mejora la capacidad del modelo para aprender los patrones reales de la serie temporal

3.4.2.3 Construcción de la serie temporal

Se construyó una serie temporal mensual representativa del consumo energético residencial de la provincia de Tungurahua. Para ello, se filtraron únicamente los registros pertenecientes al grupo de consumo Residencial y posteriormente se agregó la información a nivel provincial con frecuencia mensual.

Algunas variables fueron sumadas para obtener los totales mensuales. Y otras variables fueron promediadas con el fin de representar las condiciones de cada período.

Tabla 9: *Métodos de agregación por tipo de variable*

Tipo de variable	Método de agregación
Energía facturada kWh (Ajustada)	Suma
Facturación usd (Ajustada)	Suma
Número de clientes	Suma
Temperatura promedio	Promedio
Temperatura máxima	Promedio
Temperatura mínima	Promedio
Precipitación	Promedio
Humedad	Promedio
Velocidad del viento	Promedio
Población total	Promedio

Como resultado, se obtuvo una serie temporal mensual consolidada que sirve como base para las etapas de imputación, análisis exploratorio y modelado.

Impacto en el modelado: La agregación de los datos a nivel mensual reduce la variabilidad de corto plazo asociada a factores administrativos, como desfases en los ciclos de lectura de medidores o diferencias en los días de facturación. Asimismo, esta escala temporal facilita la integración con las variables demográficas y climáticas, que también se encuentran representadas mensualmente. Como resultado, los modelos pueden concentrarse en aprender los patrones de tendencia y estacionalidad propios de la serie, favoreciendo un entrenamiento más estable y una mejor capacidad predictiva.

3.4.2.4 Imputación eventos atípicos

Se imputaron los meses de octubre y noviembre de 2022 debido a inconsistencias en la información registrada, así como mayo de 2020, cuyo consumo estuvo influenciado por las restricciones asociadas a la pandemia de COVID-19. Estos casos fueron considerados eventos aislados que no representan el comportamiento normal de la demanda eléctrica en la provincia.

Por otra parte, los efectos de la crisis energética y los cortes de energía registrados durante 2024 no fueron imputados. A pesar de que provocaron cambios importantes en los niveles de consumo, estos valores corresponden a situaciones reales que ocurrieron dentro del sistema eléctrico y que podrían volver a presentarse en el futuro. Por esta razón, se decidió

conservarlos para que los modelos pudieran aprender cómo se comporta la demanda ante este tipo de escenarios.

3.4.2.5 Método de imputación con patrón estacional

La imputación de valores se realizó mediante un método de imputación con patrón estacional. Para ello, se estimó la tendencia de corto plazo de la serie, denominada nivel local, y se ajustó utilizando la mediana del comportamiento histórico de la demanda del mes correspondiente, preservando el patrón estacional de la serie.

Matemáticamente, el valor estimado $\hat{Y}_t$ para una fecha t se define de la siguiente manera:

$$\hat{Y}_t = L_t \times S_m \quad (3.1)$$

Donde:

- **L_t (Nivel Local):** Representa la tendencia reciente de la serie temporal en el momento de la imputación. Se calcula como el promedio de las seis observaciones mensuales anteriores al período con el dato atípico o faltante:

$$L_t = \frac{1}{6} \sum_{i=1}^6 Y_{t-i} \quad (3.2)$$

- **S_m (Factor Estacional):** Representa la desviación estacional típica del mes m ($m \in \{1, 2, \dots, 12\}$). Para su obtención, se calcula primero la relación o ratio de cada observación histórica (Y_i) respecto a la media anual de su respectivo año ($\bar{Y}_{\text{año}}$):

$$R_i = \frac{Y_i}{\bar{Y}_{\text{año}}} \quad (3.3)$$

Posteriormente, el factor estacional S_m se calcula como la mediana del comportamiento histórico del mismo mes a lo largo de los años disponibles:

$$S_m = \text{Mediana}(\{R_i \mid \text{Mes}(i) = m\}) \quad (3.4)$$

Este procedimiento permite que los valores imputados preserven el comportamiento estacional del consumo de energía eléctrica, reduciendo las distorsiones que podrían introducir métodos de imputación más simples, como la interpolación lineal. La imputación se aplicó a las variables `energy_billed_kwh_adjusted` y `billing_usd_adjusted`, ya que constituyen las principales variables de consumo energético y facturación consideradas para el análisis y el entrenamiento de los modelos predictivos.

- **COVID-19 (mayo 2020)** .- Se realizó una imputación sobre el período correspondiente a mayo de 2020, identificado como un evento excepcional dentro de la serie temporal debido al contexto asociado a la pandemia de COVID-19.

Con el objetivo de reducir la influencia de esta perturbación extraordinaria sobre el proceso de modelado, se aplicó el método de imputación estacional descrito anteriormente. Este procedimiento permitió estimar un valor representativo basado en el comportamiento histórico, preservando tanto la tendencia local como el patrón estacional del mes.

- **Incidencia de octubre 2022.**- Se realizó una imputación sobre el período de octubre del 2022 debido a la presencia de una inconsistencia en los datos registrados para dicho mes. Durante el proceso de validación se identificó que los valores registrados eran inferiores al comportamiento esperado de la serie, evidenciando un problema de calidad de los datos que podía afectar el análisis temporal y el entrenamiento de los modelos.

Para corregir esta situación se aplicó el método de imputación estacional descrito previamente. El procedimiento estimó el valor esperado del mes utilizando el comportamiento histórico, y el patrón estacional de octubre.

- **Incidencia de noviembre 2022.**- La imputación de noviembre de 2022 se realizó utilizando el mismo procedimiento estacional aplicado previamente a octubre de 2022. Con el objetivo de mantener la coherencia temporal del proceso y garantizar la consistencia de las estimaciones, se empleó el mismo historial de referencia que la imputación anterior. De esta manera, los valores imputados se basaron en información del comportamiento histórico de la serie, evitando que las observaciones afectadas por las inconsistencias detectadas a finales de 2022 influyeran en el proceso de modelado.

Validación del proceso de imputación.- Con el fin de verificar que el método implementado no introduce sesgos significativos que puedan alterar las propiedades distributivas globales de la serie de tiempo, se evaluaron dos criterios principales:

1. **Baja proporción de observaciones imputadas:** La serie temporal mensual consolidada (2014–2025) consta de 144 observaciones. El proceso de imputación se aplicó únicamente sobre tres de ellas (mayo de 2020, octubre de 2022 y noviembre de 2022), lo que representa apenas el 2.08 % del conjunto total de datos. Dado que la proporción de observaciones modificadas es reducida, es poco probable que la imputación altere significativamente las propiedades estadísticas de la serie temporal, sus patrones de correlación o el proceso de aprendizaje de los modelos de aprendizaje profundo.
2. **Reducción del sesgo durante el entrenamiento:** La corrección de estos periodos atípicos mediante imputación tuvo como finalidad evitar que los modelos de aprendizaje profundo aprendieran patrones no representativos del comportamiento habitual de la demanda eléctrica. De no haberse realizado la imputación, las redes neuronales habrían ajustado sus parámetros a eventos excepcionales, como el confinamiento ocasionado por la pandemia de COVID-19 y las inconsistencias en el almacenamiento de datos registradas por la empresa eléctrica en 2022, incorporando ruido no representativo en el conjunto de entrenamiento y afectando su capacidad de generalización.

3.5 Análisis Exploratorio de Datos (EDA)

3.5.1 *Dataset procesado*

Una vez completado el proceso de preparación de los datos, se realizó un análisis exploratorio con el fin de evaluar la calidad del conjunto de datos, comprender el comportamiento de las variables y obtener evidencia que respaldara las decisiones de modelado.

A continuación, se presentan las principales visualizaciones obtenidas durante el análisis exploratorio de datos (EDA).

- **Dimensiones del conjunto de datos:** 33,670 registros y 54 variables.

■ **Composición de variables:**

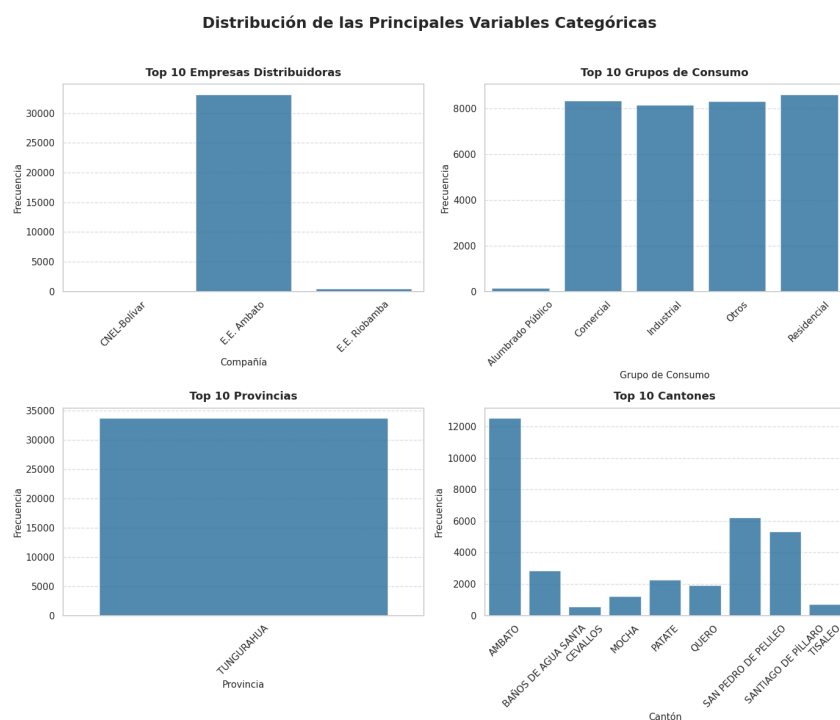
- Numéricas: 44
- Categóricas: 9
- Temporales: 1

Tabla 10: Estadísticas descriptivas claves del conjunto de datos.

Indicador	Valor
Energía media facturada	192,581 kWh
Energía mediana facturada	38,557 kWh
Asimetría (sesgo)	Presencia de sesgo positivo (derecha)
Facturación media (USD)	\$18,576.17
Cientes promedio por registro	922
Coefficiente de variación	314.2 % (alta heterogeneidad)

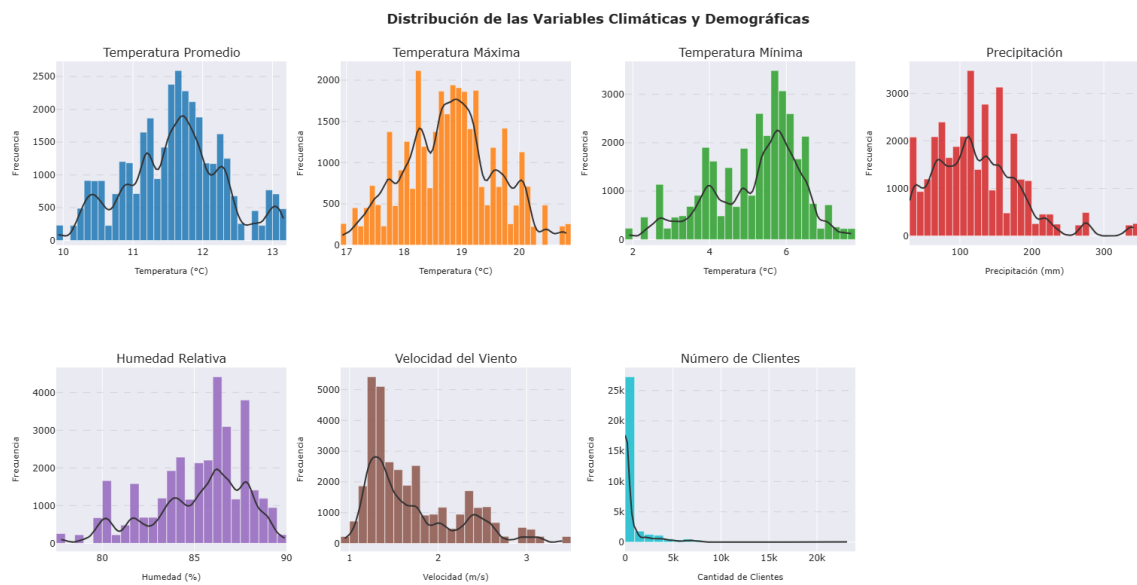
Figura 20

Distribución Principales Variables Categóricas



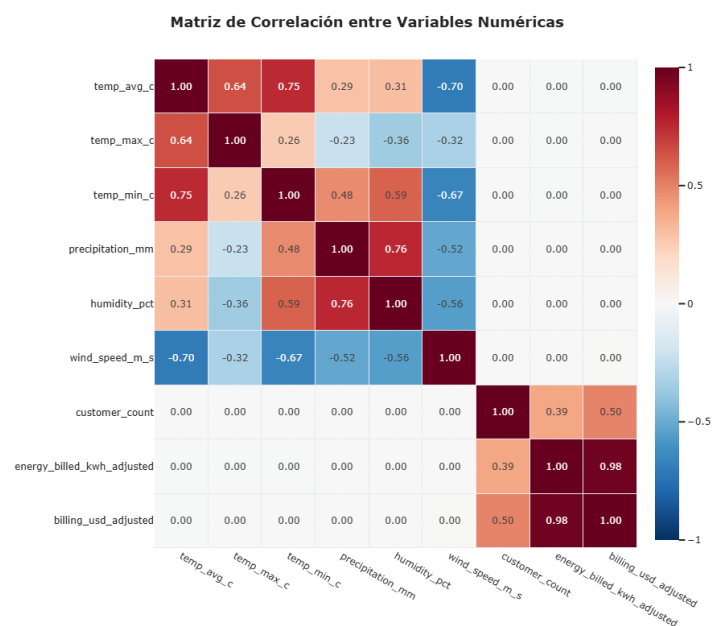
Nota. Se observa una mayor concentración de registros en el sector residencial, el cantón Ambato y la Empresa Eléctrica Ambato dentro de la provincia de Tungurahua.

Figura 21
Distribución Variables Climáticas y Demográficas

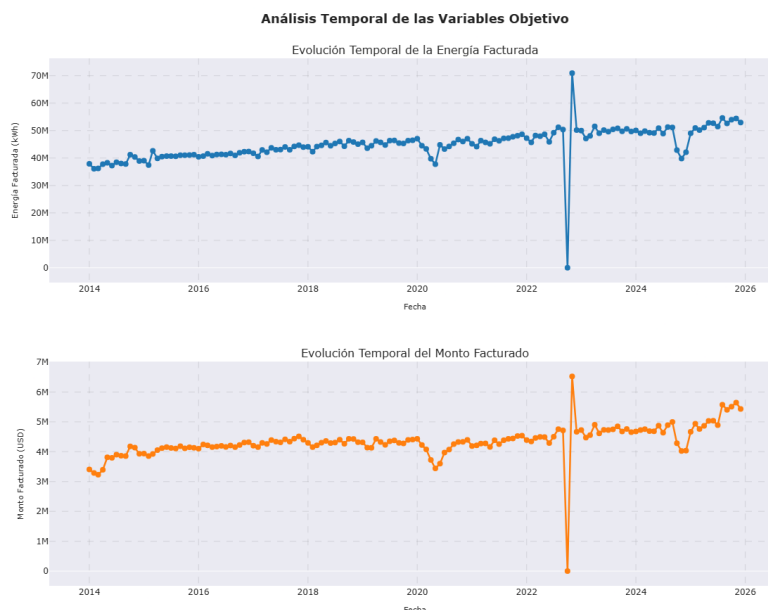


Nota. Las variables climáticas presentan distribuciones estables y concentradas en rangos específicos, mientras que el número de clientes muestra una distribución asimétrica con predominio de valores bajos y algunos registros de alta magnitud.

Figura 22
Matriz de Correlación de Variables Numéricas



Nota. Se observa una alta correlación entre energía facturada y facturación eléctrica, mientras que las variables climáticas presentan correlaciones moderadas entre sí y una relación débil con el consumo energético.

Figura 23*Análisis Temporal Variables Objetivo*

Nota. Se observa una tendencia creciente tanto en la energía facturada como en el monto facturado, junto con algunos valores atípicos identificados.

3.5.2 *Serie temporal mensual residencial*

Se realizó un análisis exploratorio específico sobre la serie temporal mensual correspondiente al consumo energético residencial de la provincia de Tungurahua. Esta etapa permitió examinar la evolución histórica del consumo, identificar patrones temporales y evaluar la serie. Asimismo, se analizaron las dependencias temporales mediante las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF), obteniendo evidencia para la configuración del diseño experimental.

Serie Temporal Imputada del Consumo Energético Residencial

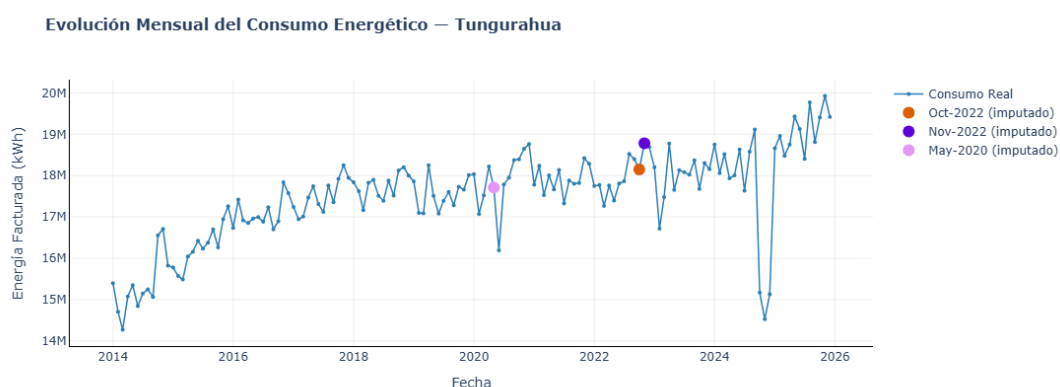
La figura presenta la evolución mensual del consumo energético residencial de la provincia de Tungurahua una vez aplicado el proceso de imputación de valores. Se observa una tendencia general creciente a lo largo del período analizado, pasando de aproximadamente 14 millones de kWh en 2014 a cerca de 20 millones de kWh hacia finales de 2025. Asimismo, la serie mantiene fluctuaciones mensuales moderadas propias del comportamiento del consumo

eléctrico.

Los puntos resaltados corresponden a los meses imputados (mayo de 2020, octubre de 2022 y noviembre de 2022), los cuales fueron corregidos para mantener la continuidad y consistencia de la serie temporal. Por otro lado, los meses afectados por la crisis de cortes eléctricos de 2024 no fueron imputados, ya que se consideró importante conservar dichos eventos dentro de la serie para que los modelos pudieran aprender, en la medida de lo posible, el comportamiento del consumo energético ante situaciones excepcionales y cambios abruptos en la demanda.

Figura 24

Evolución Mensual Consumo Energético



Nota. La serie temporal muestra una tendencia ascendente de largo plazo, correlacionada con el crecimiento poblacional regional, identificando las correcciones por imputación en 2020 y 2022, y la preservación de las fluctuaciones por la crisis energética de 2024.

Análisis de Dependencia Temporal mediante ACF y PACF

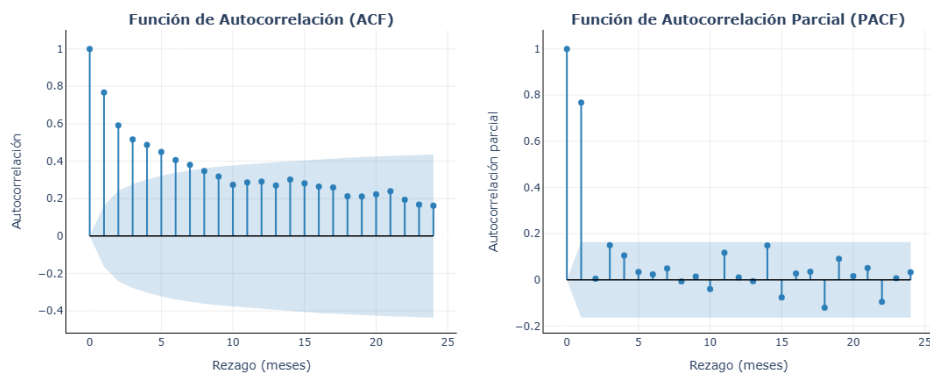
Se presentan las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) de la serie temporal de consumo energético residencial. En la gráfica ACF se observa un decaimiento gradual de las correlaciones a medida que aumenta el rezago, lo que indica una fuerte dependencia temporal entre observaciones consecutivas y sugiere la presencia de una tendencia en la serie.

La gráfica PACF muestra un pico significativo en el rezago 1, mientras que los rezagos posteriores presentan valores cercanos a cero y, en su mayoría, dentro de los intervalos de confianza. Este comportamiento es consistente con una estructura autorregresiva de primer orden (AR(1)), indicando que el consumo energético de un mes está influenciado principalmente por el valor observado en el mes inmediatamente anterior.

En conjunto, estos resultados respaldan el uso de ventanas temporales cortas para el modelado predictivo y justifican la selección de un *look_back* igual a 1 en los experimentos realizados.

Figura 25

Análisis de Dependencia Temporal mediante ACF y PACF



Nota. Las gráficas ACF y PACF evidencian una estructura autorregresiva de primer orden (AR(1)), donde el rezago 1 concentra la mayor dependencia temporal.

3.6 Configuración para el diseño experimental

Una vez construida la serie temporal mensual y aplicados los procesos de limpieza, transformación e imputación, se definió la configuración experimental para el entrenamiento y evaluación de los modelos predictivos.

El diseño experimental consideró la selección de la variable objetivo, el conjunto de variables predictoras, la definición de ventanas temporales, la división cronológica de los datos en entrenamiento, validación y prueba, así como la aplicación de técnicas de escalamiento necesarias para el adecuado funcionamiento de los modelos.

Esta configuración permitió evaluar de manera homogénea el desempeño de las arquitecturas LSTM, GRU y Transformer, asegurando que todas fueran entrenadas y comparadas bajo las mismas condiciones experimentales.

3.6.1 Selección de Variables

Para el desarrollo de los modelos de redes neuronales se utilizó una variable objetivo y un conjunto de variables predictoras relacionadas con factores energéticos, climáticos y demográficos. La selección de estas variables se realizó considerando su disponibilidad histórica y su influencia sobre el consumo eléctrico residencial en la provincia de Tungurahua.

- **Variable objetivo.**- La variable objetivo seleccionada para esta investigación fue la Energía Facturada kWh (Ajustada), la cual representa el consumo de energía eléctrica registrado en el sector residencial. En consecuencia fue utilizada como referencia principal para evaluar el desempeño predictivo de los modelos desarrollados.

Tabla 11: Especificación de la variable del estudio

Variable	Tipo	Unidad	Descripción
Energía Facturada kWh (Ajustada) energy_billed_kwh_adjusted	Objetivo	kWh	Consumo energético residencial mensual agregado a nivel provincial en Tungurahua.

- **Variables predictoras.**- Con el fin de capturar los diferentes factores que influyen en el comportamiento del consumo energético residencial, se consideró un conjunto de variables predictoras de naturaleza energética, climática y demográfica.

Estas variables fueron utilizadas como entradas de los modelos LSTM, GRU y Transformer, permitiendo evaluar la capacidad de cada arquitectura para identificar relaciones temporales y patrones relacionados al consumo energético.

Tabla 12: Variables independientes consideradas en la investigación

Categoría	Variable	Unidad	Descripción
Energética	Energía facturada	kWh	Valor de consumo energético del período anterior utilizado para capturar la dependencia temporal de la serie.
Energética	Facturación del servicio eléctrico	USD	Valor monetario facturado mensualmente al sector residencial.
Energética	Número de clientes	Clientes	Total de usuarios residenciales registrados en la provincia.
Climática	Temperatura promedio	°C	Temperatura media mensual registrada en la provincia.
Climática	Temperatura máxima	°C	Temperatura máxima mensual registrada.
Climática	Temperatura mínima	°C	Temperatura mínima mensual registrada.
Climática	Precipitación	mm	Precipitación acumulada mensual.
Climática	Humedad relativa	%	Humedad relativa promedio mensual.
Climática	Velocidad del viento	m/s	Velocidad promedio mensual del viento.
Demográfica	Población total	Habitantes	Población total de la provincia para cada período analizado.

Justificación de la selección de variables predictoras .- La selección de las variables independientes incluidas en este estudio responde a su relación con el consumo de energía eléctrica residencial en la provincia de Tungurahua. Para facilitar su interpretación, las variables se agrupan en las siguientes dimensiones:

■ Dimensión energética

- *Energía facturada del período anterior:* Representa la memoria temporal del sistema, permitiendo incorporar la persistencia del consumo observada entre períodos consecutivos.
- *Facturación del servicio eléctrico (USD):* Constituye un indicador económico relacionado con el nivel de consumo y el valor facturado del servicio eléctrico, aportando información complementaria sobre el comportamiento de la demanda.
- *Número de clientes:* Refleja la magnitud de la red de usuarios residenciales y el crecimiento del mercado eléctrico, siendo un factor directamente asociado al incremento de la demanda agregada.

■ Dimensión climática

- Las variables de temperatura media, máxima y mínima, humedad relativa, precipitación y velocidad del viento describen las condiciones ambientales que influyen en los hábitos de consumo de los hogares. Estas variables permiten representar el efecto del clima sobre la demanda eléctrica, asociado al uso de equipos eléctricos y a las variaciones estacionales del consumo.

■ Dimensión demográfica

- *Población total*: Representa el crecimiento estructural de la población de la provincia y constituye un factor determinante de la evolución de la demanda eléctrica a largo plazo, al reflejar el aumento potencial de usuarios y de las necesidades de consumo.

3.6.2 Ventanas temporales

Con el objetivo de preparar la serie temporal para modelos de aprendizaje profundo, los datos fueron transformados en secuencias mediante el uso de ventanas temporales deslizantes. Este procedimiento permite que los modelos LSTM, GRU y Transformer utilicen información histórica para predecir el consumo energético del siguiente período.

Se evaluaron diferentes tamaños de ventanas temporales, representados por los valores *1, 3, 6 y 12 meses*, con el propósito de analizar el efecto de distintos horizontes históricos sobre el desempeño de los modelos. Estos valores corresponden, respectivamente, al uso del mes anterior, el trimestre anterior, el semestre anterior y el año previo como contexto para la predicción.

Durante la construcción de las secuencias, cada muestra se formó utilizando los valores de consumo de los meses anteriores definidos por la ventana temporal. El valor objetivo correspondió al consumo del mes siguiente, permitiendo que el modelo aprendiera a predecir valores futuros a partir de información histórica, sin utilizar datos del futuro durante el entrenamiento.

Los resultados experimentales mostraron que la configuración con *look back igual a 1* obtuvo el mejor desempeño predictivo. Esta elección se encuentra respaldada por el análisis de autocorrelación parcial (PACF), el cual mostró una estructura predominante de tipo AR(1),

lo que confirma que el consumo de un mes depende básicamente del mes anterior. Asimismo, una ventana temporal corta simplifica la arquitectura del modelo, lo cual resulta adecuado considerando la extensión acotada de la serie de datos mensual disponible.

Las métricas detalladas de esta experimentación comparativa entre las diferentes ventanas temporales se presentan a continuación.

3.6.3 Análisis comparativo de ventanas temporales

Con la finalidad de validar empíricamente la ventana temporal óptima y descartar posibles sesgos por sobreajuste, se implementó un análisis de sensibilidad entrenando y evaluando las arquitecturas LSTM, GRU y Transformer bajo rezagos alternativos de 3, 6 y 12 meses. Esta experimentación permite justificar si una ventana más larga aporta información relevante o si introduce ruido en el modelo.

La Tabla 13 presenta la comparación de métricas de rendimiento (MAE, RMSE, MAPE y R^2) obtenidas en el conjunto de prueba para todos los modelos y ventanas evaluadas.

Tabla 13: Métricas de rendimiento según el tamaño de la ventana temporal en meses.

Meses	Modelo	MAE (kWh)	RMSE (kWh)	MAPE (%)	R^2
1	LSTM	901,295	1,327,930	5.37	0.1568
1	GRU	894,263	1,241,915	5.26	0.2625
1	Transformer	1,189,044	1,504,906	6.85	-0.0829
1	Persistencia (Baseline)	893,235	1,298,026	5.06	0.1940
3	LSTM	1,035,805	1,429,909	6.10	0.0223
3	GRU	3,110,596	3,355,695	16.52	-4.3846
3	Transformer	6,499,230	6,793,506	34.94	-21.0685
3	Persistencia (Baseline)	893,235	1,298,026	5.06	0.1943
6	LSTM	1,063,275	1,438,039	6.25	0.0550
6	GRU	3,310,178	3,582,679	17.55	-4.8656
6	Transformer	1,027,213	1,465,915	6.12	0.0180
6	Persistencia (Baseline)	913,918	1,324,793	5.18	0.1980
12	LSTM	1,185,635	1,520,579	6.91	-0.0091
12	GRU	1,154,853	1,498,646	6.75	0.0198
12	Transformer	2,097,449	2,191,552	11.42	-1.0961
12	Persistencia (Baseline)	930,320	1,351,172	5.28	0.2032

Los resultados evidencian que el incremento en el tamaño de la ventana temporal provoca una degradación generalizada de la precisión en todos los modelos predictivos. Para ventanas de 3 y 6 meses, los modelos GRU y Transformer experimentan una fuerte inestabilidad en el entrenamiento y sobreajuste, reflejado en valores de R^2 significativamente negativos y un incremento notable del error MAPE. Este comportamiento experimental ratifica la pertinencia del análisis de la Función de Autocorrelación Parcial (PACF), el cual sugería que la dependencia temporal del consumo se concentra principalmente en el mes inmediatamente anterior. Al ampliar las ventanas temporales se incorporan retardos que actúan como ruido, sobrecargando la optimización de pesos y provocando un bajo rendimiento en el conjunto de prueba. En

consecuencia, para los datos y las condiciones experimentales consideradas en este estudio, la configuración de $lookback = 1$ representa la única ventana capaz de superar al baseline de Persistencia y garantizar un ajuste positivo. El detalle visual de las curvas de entrenamiento para las ventanas alternativas se desglosa en el Anexo A.

3.6.4 División dataset

La división de los datos se realizó manteniendo el orden en que ocurrieron los registros en el tiempo. Esto se realizó así porque el objetivo del estudio es predecir valores futuros a partir de información pasada. Si los datos se mezclaran aleatoriamente, los modelo podría aprender de información futura.

Por esta razón, la serie temporal se dividió siguiendo su orden temporal en los siguientes conjuntos:

- **Entrenamiento (70 %):** Utilizado para el aprendizaje de los patrones históricos presentes en las variables energéticas, climáticas y demográficas.
- **Validación (15 %):** Empleado para monitorear el desempeño del modelo durante el entrenamiento y ajustar estrategias de regularización, tales como *Early Stopping* y *ReduceLROnPlateau*.
- **Prueba (15 %):** Reservado exclusivamente para la evaluación final del modelo, representando el período temporal más reciente de la serie.

3.6.5 Escalado de Características

Debido a que las variables utilizadas en el estudio presentan diferentes unidades y magnitudes, se aplicó un proceso de normalización previo al entrenamiento de los modelos. Esta etapa permitió llevar todas las variables a una escala común, evitando que aquellas con valores numéricamente más elevados ejercieran una influencia desproporcionada durante el aprendizaje.

Para ello se utilizó el método *Min-Max Scaling*, el cual transforma los datos al rango $[0, 1]$. Para realizar esta transformación, se tomó como referencia únicamente la información

del conjunto de entrenamiento. Luego, esa misma transformación se aplicó a los datos de validación y prueba. De esta manera, se evita que el modelo utilice información que no debería conocer de antemano, logrando una evaluación más realista de su desempeño.

Además, la variable objetivo también fue normalizado por separado. Esto permitió que, una vez obtenidas las predicciones, los resultados pudieran convertirse nuevamente a kilovatios-hora (kWh), facilitando su interpretación.

Al mantener el orden temporal de los datos y utilizar únicamente la información del conjunto de entrenamiento para realizar la normalización, se logró una evaluación más realista del rendimiento de los modelos en este estudio.

Tabla 14: *Resumen del proceso de normalización de datos.*

Elemento	Método aplicado
Variables predictoras	MinMaxScaler [0, 1]
Variable objetivo	MinMaxScaler [0, 1]
Ajuste del escalador	Solo conjunto de entrenamiento
Aplicación	Entrenamiento, validación y prueba
Objetivo	Evitar diferencias de escala y fuga de información

3.6.6 Resultado de la configuración experimental

Al concluir esta etapa, estructuramos la información en bloques secuenciales que cumplan con los requisitos de entrada de los modelos LSTM, GRU y Transformer. Esta configuración se consolidó de la siguiente manera:

- **Variables:** Contamos con una variable objetivo (energía facturada ajustada) y diez variables predictoras.
- **Estrategia temporal:** Generamos ventanas deslizantes para evaluar diversos periodos históricos, seleccionando finalmente una ventana de un mes para la predicción.
- **Gestión de datos:** Dividimos la serie cronológicamente en conjuntos de entrenamiento, validación y prueba, normalizando todas las características al rango [0, 1].

- **Procesamiento:** Aplicamos escalados independientes para la variable objetivo y mantuvimos la trazabilidad de las fechas, lo cual fue clave para graficar y comparar los resultados.

3.7 Red Neuronal LSTM (Long Short-Term Memory)

Se implementó una arquitectura basada en redes neuronales Long Short-Term Memory (LSTM) para la predicción del consumo energético mensual residencial en la provincia de Tungurahua. Este tipo de red neuronal está diseñada para trabajar con datos secuenciales y resulta adecuada para problemas de series temporales, donde los valores observados en períodos anteriores influyen en el comportamiento futuro de la variable de interés.

La elección de una arquitectura LSTM se fundamentó en su capacidad para capturar dependencias temporales dentro de la serie, permitiendo aprovechar la información contenida en observaciones pasadas para mejorar la precisión de las predicciones. Esta característica la convierte en una de las arquitecturas más utilizadas en tareas de pronóstico de demanda energética y análisis de series temporales.

3.7.1 Construcción del Modelo LSTM

Se implementó una arquitectura LSTM compuesta por dos capas recurrentes apiladas, diseñadas para capturar las dependencias temporales presentes en la serie de consumo energético.

La primera capa LSTM contó con 32 neuronas y tuvo la función de analizar la secuencia de datos mes a mes, identificando patrones y relaciones a lo largo del tiempo. Además, se configuró para enviar toda la información procesada a la siguiente capa LSTM, permitiendo que el modelo profundizara progresivamente en el aprendizaje de las tendencias presentes en la serie temporal.

A continuación, se añadió una segunda capa LSTM con 16 neuronas, cuya función fue resumir la información más relevante aprendida de la secuencia temporal. De esta manera, el modelo pudo concentrar los patrones identificados en una representación más compacta antes de realizar la predicción. Además, para evitar que el modelo se ajustara excesivamente a los

datos de entrenamiento, se incorporó una capa Dropout después de cada capa LSTM, desactivando aleatoriamente el 10 % de las neuronas durante el entrenamiento.

La información aprendida por las capas LSTM fue posteriormente procesada por una capa densa de 8 neuronas con función de activación ReLU. Esta capa permitió combinar y transformar los patrones identificados previamente, facilitando que el modelo construyera una representación más útil para realizar la predicción. Finalmente, se incorporó una capa de salida con una única neurona y activación lineal, encargada de generar el valor estimado de energía facturada para el mes siguiente. Esta configuración resulta adecuada para problemas de regresión, donde el objetivo es predecir una variable numérica continua.

Para el entrenamiento del modelo se utilizó el optimizador Adam con una tasa de aprendizaje de 0.0005, encargado de ajustar automáticamente los pesos de la red para mejorar sus predicciones. Como función de pérdida se empleó el error cuadrático medio (MSE), que penaliza con mayor intensidad los errores más grandes. Además, se utilizó el error absoluto medio (MAE) como métrica de seguimiento, ya que permite interpretar de forma más sencilla la diferencia promedio entre los valores reales y los predichos durante el proceso de entrenamiento.

Tabla 15: *Arquitectura del modelo LSTM*

Layer (type)	Output Shape	Param #	Connected to
input_layer_1 (InputLayer)	(None, 1, 10)	0	-
lstm (LSTM)	(None, 1, 32)	5504	input_layer_1[0][0]
dropout_1 (Dropout)	(None, 1, 32)	0	lstm[0][0]
lstm_1 (LSTM)	(None, 16)	3136	dropout_1[0][0]
dropout_2 (Dropout)	(None, 16)	0	lstm_1[0][0]
dense_1 (Dense)	(None, 8)	136	dropout_2[0][0]
dense_2 (Dense)	(None, 1)	9	dense_1[0][0]

Nota. Elaboración propia.

3.7.2 Entrenamiento del Modelo LSTM

El modelo se entrenó durante un máximo de 200 épocas utilizando un tamaño de lote de 8 muestras. Debido a que la serie temporal analizada cuenta con una cantidad limitada de observaciones mensuales, se optó por un tamaño de lote pequeño, lo que permitió realizar actualizaciones más frecuentes de los pesos de la red durante el entrenamiento. Esta estrategia favoreció el aprendizaje de los patrones presentes en los datos y contribuyó a una mejor adaptación del modelo al comportamiento de la serie.

El modelo fue entrenado utilizando únicamente los datos del conjunto de entrenamiento, mientras que el conjunto de validación se empleó para evaluar su desempeño durante el proceso de aprendizaje. Esto permitió verificar cómo evolucionaba la capacidad predictiva del modelo y detectar posibles problemas de sobreajuste. Además, se respetó en todo momento el orden cronológico de las observaciones, evitando mezclar los datos de forma aleatoria. De esta manera, se preservó la secuencia temporal de la serie y se garantizó una evaluación más realista del comportamiento del modelo frente a datos futuros.

Con el fin de lograr que el modelo pudiera realizar buenas predicciones sobre datos no vistos previamente, se aplicaron diferentes estrategias durante el entrenamiento para evitar el sobreajuste. Entre ellas, se incluyó la detención temprana del entrenamiento cuando el rendimiento dejaba de mejorar, la reducción automática de la tasa de aprendizaje para refinar el ajuste de los parámetros y el almacenamiento de la versión del modelo que obtuvo los mejores resultados en la fase de validación.

Tabla 16: *Parámetros de entrenamiento del modelo LSTM*

Parámetro	Valor
Épocas máximas	200
Batch Size	8
Optimizador	Adam
Learning Rate	0.0005
Función de pérdida	MSE
Métrica	MAE
Early Stopping	Sí
ReduceLROnPlateau	Sí
ModelCheckpoint	Sí
Shuffle	False

Figura 26*Resumen del entrenamiento del modelo LSTM*

```
Total params: 8,785 (34.32 KB)
Trainable params: 8,785 (34.32 KB)
Non-trainable params: 0 (0.00 B)
Épocas entrenadas : 60
Mejor val_loss : 0.013881
```

Nota. Elaboración propia.

3.7.3 Evaluación

Una vez finalizado el entrenamiento, se generaron predicciones sobre los conjuntos de entrenamiento, validación y prueba. Debido a que la variable objetivo fue normalizada durante la etapa de preprocesamiento, las predicciones obtenidas fueron transformadas nuevamente a su escala original mediante el proceso de desnormalización, permitiendo interpretar los resultados en unidades reales de kilovatios-hora (kWh).

Evaluación del Modelo LSTM

Para evaluar el desempeño del modelo se utilizaron cuatro métricas: el error absoluto medio (MAE), la raíz del error cuadrático medio (RMSE), el error porcentual absoluto medio (MAPE) y el coeficiente de determinación (R^2). Estas métricas permiten analizar la precisión de las predicciones desde diferentes perspectivas, considerando tanto el error absoluto como el error relativo y la capacidad explicativa del modelo.

Las métricas fueron calculadas sobre los conjuntos de entrenamiento, validación y prueba, siendo este último el más relevante para la evaluación final, ya que contiene observaciones no utilizadas durante el proceso de aprendizaje. Esta estrategia permitió analizar la capacidad de generalización del modelo frente a datos no observados.

Tabla 17: *Resumen de métricas del modelo LSTM*

Conjunto	MAE	RMSE	MAPE (%)	R^2
Entrenamiento	440 218,11	533 953,17	2.5993	0.6785
Validación	401 930.77	529 139.74	2.2402	-0.1978
Prueba	901 295.27	1 327 930.02	5.3659	0.1568

Al analizar los resultados de la Tabla 17, se observan comportamientos diferenciados entre los conjuntos que explican la dinámica del aprendizaje del modelo:

- **Precisión en entrenamiento y validación:** El modelo alcanza errores porcentuales sumamente bajos, con un MAPE de 2,60 % en entrenamiento y un 2,24 % en validación. Estos valores demuestran que la red LSTM logró asimilar de manera óptima las dinámicas cíclicas estacionales y la tendencia del consumo eléctrico mensual de Tungurahua.
- **R^2 en validación:** A pesar de que el modelo presenta un error porcentual bajo ($MAPE = 2,24\%$), el coeficiente de determinación ($R^2 = -0,1978$) resulta negativo. Esto ocurre porque, durante el período de validación, el consumo eléctrico varió muy poco. En estas condiciones, pequeñas diferencias entre los valores reales y los predichos afectan considerablemente el cálculo de R^2 , haciendo que tome valores negativos. Por ello, este resultado no significa que el modelo tenga un mal desempeño, sino que el coeficiente R^2 pierde capacidad para evaluar el ajuste cuando la variabilidad de los datos es muy baja.
- **Incremento del error en el conjunto de prueba:** En el conjunto de prueba, el MAPE se eleva al 5,37 % y el MAE a 901,295 kWh. Este incremento del error coincide con los meses finales de 2024, periodos en los cuales el sistema eléctrico del Ecuador sufrió racionamientos de energía programados. Al ser eventos no cíclicos y de fuerza mayor no registrados en los datos de entrenamiento, el modelo continuó proyectando un consumo normal, generando desviaciones inevitables debido a una situación excepcional que no estaba representada en los datos de entrenamiento

Comparación LSTM vs Baselines

Los resultados obtenidos por la arquitectura LSTM fueron comparados con dos modelos de referencia. El primero fue un modelo de persistencia, que asume que el valor del siguiente período será igual al observado en el período anterior. El segundo fue un modelo basado en el promedio histórico de los datos de entrenamiento. Estas comparaciones permitieron evaluar si la red neuronal era capaz de ofrecer predicciones más precisas y útiles que métodos mucho más simples, justificando así el uso de una arquitectura más compleja.

Tabla 18: Métricas de desempeño de los modelos baseline sobre el conjunto de prueba LSTM

Modelo	MAE	RMSE	MAPE (%)	R ²
Persistencia (naive)	893 235.41	1 298 026	5.0612	0.1943
Media histórica (naive)	1 705 382.91	1 810 176.14	9.4135	-0.5668

Tabla 19: Comparación del modelo LSTM frente a los modelos baseline sobre el conjunto de prueba

Modelo	MAE	MAPE (%)	R ²
LSTM	901,295	5.37	0.157
Persistencia (naive)	893,235	5.06	0.194
Media histórica (naive)	1 705 383	9.41	-0.567

La comparación de la arquitectura LSTM frente a los baselines revela aspectos críticos sobre la viabilidad del modelo y el comportamiento de la serie:

- **Modelo LSTM:** En el conjunto de prueba, el modelo LSTM obtuvo un MAE de 901,295 kWh y un MAPE de 5,37%, lo que indica una buena precisión en la estimación del consumo eléctrico. Sin embargo, el coeficiente de determinación fue bajo ($R^2 = 0,157$), lo que evidencia que, aunque el modelo reproduce adecuadamente la tendencia general de la demanda, presenta dificultades para representar las variaciones registradas durante los eventos excepcionales.

- **Modelo Persistencia:** El modelo de persistencia resulta sumamente competitivo, registrando un MAPE de 5,06 % frente al 5,37 % de la LSTM. Esto se debe a la elevada inercia mensual del consumo eléctrico residencial. No obstante, mientras la persistencia es un modelo *pasivo* incapaz de reaccionar ante cambios de entorno, la red LSTM es un modelo *activo* y multivariable. Su principal aporte consiste en la capacidad de simular distintos escenarios futuros, como el crecimiento demográfico o el incremento de la temperatura, proporcionando información útil para la planificación estratégica, una funcionalidad que el modelo de persistencia no ofrece.
- **Modelo de Media Histórica:** El modelo de media histórica presentó el menor desempeño, con un MAPE de 9,41 %, un MAE de 1,705,383 kWh y un R^2 de $-0,567$. Estos resultados indican que no representa adecuadamente el comportamiento del consumo eléctrico en Tungurahua, debido a la tendencia creciente de la serie temporal. En un contexto de planificación, su uso podría subestimar la demanda mensual en aproximadamente 1.7 millones de kWh, lo que podría generar estimaciones de demanda inferiores a las reales y afectar la planificación del abastecimiento de energía

3.7.4 Visualización

Además de las métricas numéricas, se analizaron diferentes gráficos para evaluar visualmente el comportamiento del modelo, su capacidad de generalización y la calidad de sus predicciones.

Curva de Pérdida

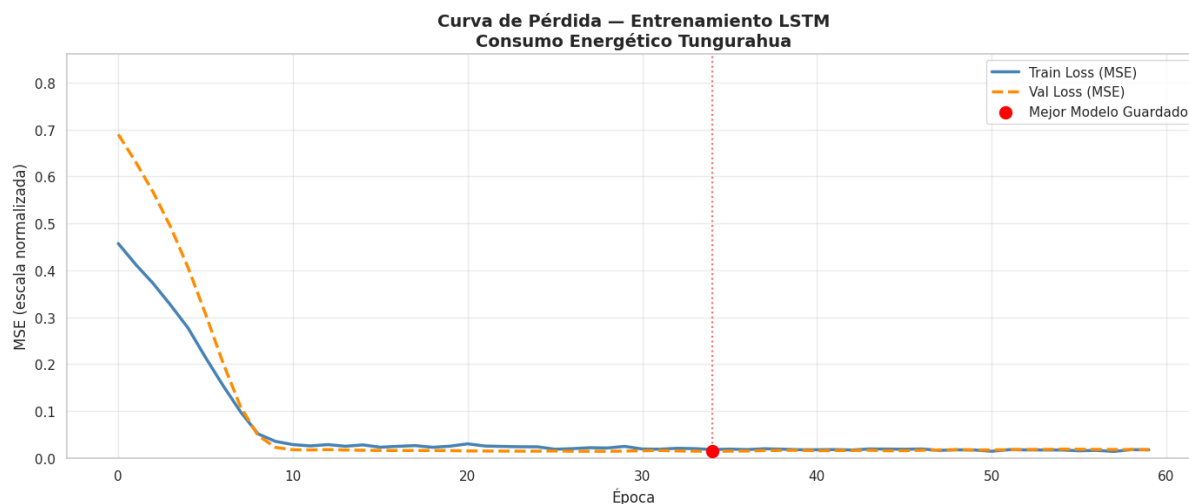
Como se observa en el gráfico, ambas curvas presentan una disminución pronunciada durante las primeras épocas y posteriormente se estabilizan en valores bajos y muy similares. Este comportamiento indica que el modelo logró aprender los patrones de la serie temporal de forma adecuada y mantuvo un desempeño consistente tanto en entrenamiento como en validación.

La escasa diferencia entre ambas curvas sugiere la ausencia de sobreajuste significati-

vo, evidenciando una adecuada capacidad de generalización. Asimismo, la mejor pérdida de validación se alcanzó aproximadamente en la época 34, momento en el cual se obtuvo la configuración con mejor desempeño predictivo.

Figura 27

Curva de Pérdida LSTM



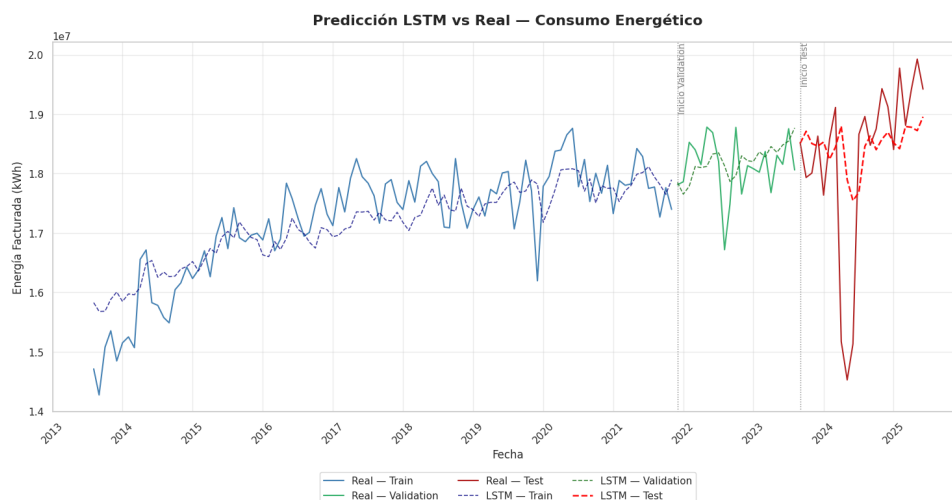
Nota. La convergencia de ambas curvas confirma un aprendizaje adecuado y que el modelo alcanzó su máxima capacidad de generalización.

Predicción vs Real

Durante las etapas de entrenamiento y validación, las predicciones siguen de manera cercana el comportamiento de los valores reales, evidenciando que la arquitectura fue capaz de aprender los patrones temporales presentes en la serie. En el conjunto de prueba, aunque se observan diferencias puntuales en algunos períodos de alta variabilidad, el modelo mantiene una buena aproximación a la tendencia general del consumo.

Figura 28

Predicción LSTM



Nota. El modelo LSTM logra reproducir la tendencia y los patrones cíclicos, con una ligera suavización de los picos de consumo.

Análisis en el Conjunto de Prueba

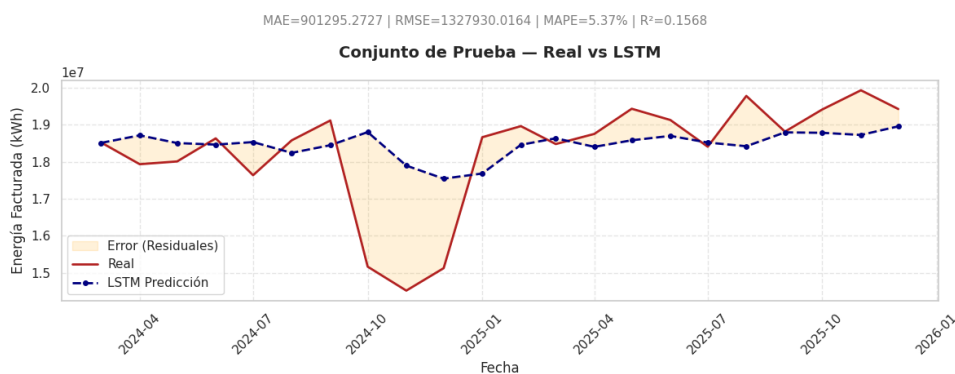
Se observa que el modelo logra reproducir adecuadamente la tendencia general del consumo energético, manteniendo predicciones cercanas a los valores reales durante la mayor parte del período evaluado.

Las mayores diferencias se presentan entre octubre y diciembre de 2024, donde la serie real registra una disminución abrupta del consumo que no es replicada completamente por el modelo. Como consecuencia, se genera un incremento temporal del error de predicción, evidenciado por el área sombreada correspondiente a los residuos. Fuera de este período, las predicciones muestran una buena alineación con la evolución observada de la serie.

Los resultados obtenidos en el conjunto de prueba evidencian que el modelo reproduce adecuadamente el comportamiento general del consumo eléctrico, manteniendo un error porcentual bajo y una precisión suficiente para aplicaciones de planificación energética. No obstante, el valor de R^2 indica que aún existe variabilidad de la demanda que el modelo no logra explicar, principalmente ante cambios abruptos o eventos extraordinarios en la serie temporal.

Figura 29

Análisis en el Conjunto de Prueba LSTM



Nota. El análisis evidencia que el modelo aprende la inercia temporal, pero presenta limitaciones para anticipar cambios bruscos.

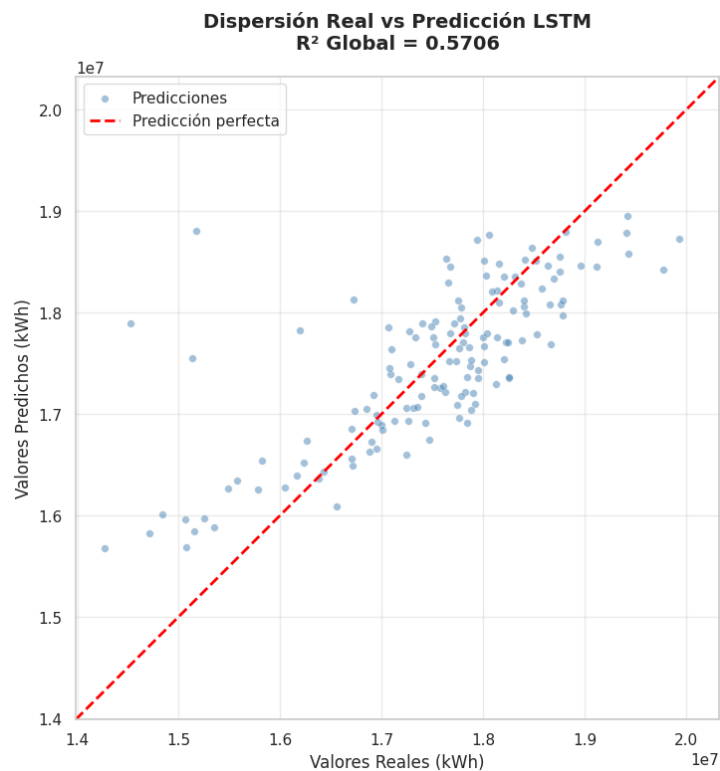
Dispersión Real vs Predicción

En el gráfico de dispersión entre lo real y las predicciones, se observa que la mayoría de los puntos se concentra alrededor de la línea diagonal, lo que indica una razonable correlación entre las predicciones y los valores reales de consumo energético. No obstante, también se aprecia cierta dispersión, especialmente en los extremos de la distribución, evidenciando que el modelo presenta mayores dificultades para reproducir algunos valores atípicos o cambios bruscos en la serie.

El coeficiente de determinación global obtenido fue $R^2 = 0,5706$, lo que indica que aproximadamente el 57% de la variabilidad observada en el consumo eléctrico es explicada por las predicciones del modelo.

Figura 30

Dispersión Real vs Predicción LSTM



Nota. La concentración de puntos en torno a la diagonal ideal representa una correlación positiva moderada. Se observa que el modelo muestra un ajuste estrecho en consumos intermedios, pero tiende a subestimar levemente los registros de demanda excepcionalmente altos.

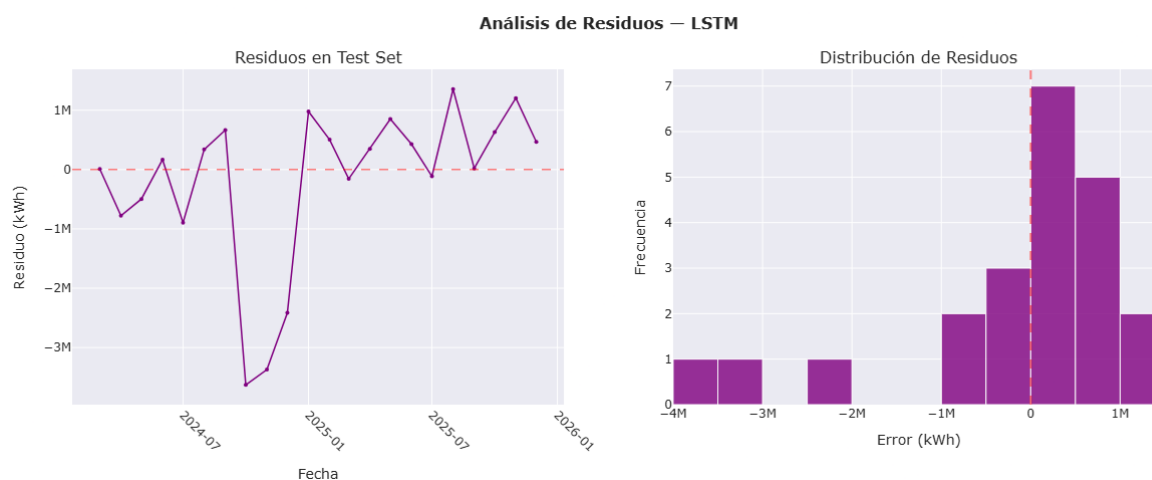
Análisis de Residuos LSTM

En la gráfica temporal se identifican algunos residuos negativos de gran magnitud a finales de 2024, evidenciando períodos en los que el modelo sobreestimó los valores reales. Estos errores coinciden con una caída abrupta observada en la serie temporal, la cual resultó difícil de reproducir para el modelo debido a su carácter excepcional.

Por su parte, el histograma muestra una distribución de residuos concentrada alrededor de cero, lo que sugiere que las predicciones fueron precisas para la mayoría de las observaciones. Aunque se observa una ligera asimetría hacia valores negativos causada por algunos errores extremos.

En conjunto, el análisis de residuos confirma que el modelo LSTM logró capturar adecuadamente el comportamiento general de la serie, concentrando los mayores errores en períodos específicos asociados a cambios abruptos del consumo energético.

Figura 31
Análisis de Residuos LSTM



Nota. Los residuos se distribuyen alrededor de cero y no evidencian sesgos sistemáticos; los mayores errores corresponden a anomalías externas.

3.7.5 Tabla Resumen de Métricas

Se presenta una tabla consolidada con las principales métricas de desempeño obtenidas por el modelo LSTM, sobre los conjuntos de entrenamiento, validación y prueba.

Estos resultados indican que el modelo fue capaz de predecir el consumo energético con un error porcentual relativamente bajo, capturando adecuadamente la tendencia general de la serie, aunque con dificultades para representar algunos cambios bruscos en los consumos.

Tabla 20: *Métricas de evaluación del modelo LSTM*

Conjunto	MAE (kWh)	RMSE (kWh)	MAPE (%)	R ²
Entrenamiento	440218.1117	533953.1655	2.5993	0.6785
Validación	401930.7695	529139.7423	2.2402	-0.1978
Prueba	901295.2727	1327930.0164	5.3659	0.1568

3.8 Red Neuronal Transformer

Se implementó una arquitectura basada en Transformer adaptada para la predicción de series temporales de consumo energético. El modelo emplea un mecanismo de autoatención multi-cabeza (Multi-Head Attention), el cual permite capturar dependencias temporales entre las observaciones de entrada sin necesidad de procesarlas de manera secuencial, facilitando el aprendizaje de patrones complejos y relaciones de largo alcance dentro de la serie temporal.

La arquitectura incorpora conexiones residuales y capas de normalización (Layer Normalization) para mejorar la estabilidad del entrenamiento y favorecer la propagación de gradientes. Posteriormente, las representaciones generadas por el mecanismo de atención son agregadas mediante una capa de Global Average Pooling, permitiendo obtener una representación compacta de la secuencia temporal. Finalmente, se utilizan capas densas completamente conectadas para realizar la predicción del consumo energético del mes siguiente.

Con el objetivo de garantizar una comparación equitativa con los modelos LSTM y GRU, se mantuvieron condiciones experimentales equivalentes, incluyendo la misma estrategia de partición de datos, el uso de ventanas temporales de entrada de diferentes tamaños (3, 6, 12 y 18 meses), la función de pérdida basada en el error cuadrático medio (MSE), el optimizador Adam con una tasa de aprendizaje de 0.001, la regularización mediante capas Dropout y la aplicación de Early Stopping para evitar el sobreentrenamiento.

3.8.1 Construcción del Modelo Transformer

Se implementa una arquitectura basada en Transformer adaptada para la predicción de series temporales. El modelo utiliza un bloque codificador compuesto por mecanismos de autoatención multi-cabeza y redes feed-forward, complementados con conexiones residuales y capas de normalización para mejorar la estabilidad del aprendizaje. Posteriormente, las representaciones generadas son agregadas mediante una capa de pooling global y procesadas por capas densas para producir la predicción final. Con el fin de realizar una comparación justa con las arquitecturas LSTM y GRU, se mantienen hiperparámetros equivalentes, incluyendo la tasa de aprendizaje, la regularización mediante Dropout y la configuración de la capa de salida para

regresión.

Tabla 21: *Arquitectura del modelo Transformer*

Layer (type)	Output Shape	Param #	Connected to
input_layer_2 (InputLayer)	(None, 1, 10)	0	-
layer_normalization (LayerNormalization)	(None, 1, 10)	20	input_layer_2[0][0]
multi_head_attention (MultiHeadAttention)	(None, 1, 10)	2762	layer_normalization[0][0], layer_normalization[0][0]
dropout_5 (Dropout)	(None, 1, 10)	0	multi_head_attention[0][0]
add (Add)	(None, 1, 10)	0	dropout_5[0][0], input_layer_2[0][0]
layer_normalization_1 (LayerNormalization)	(None, 1, 10)	20	add[0][0]
dense_4 (Dense)	(None, 1, 32)	352	layer_normalization_1[0][0]
dropout_6 (Dropout)	(None, 1, 32)	0	dense_4[0][0]
dense_5 (Dense)	(None, 1, 10)	330	dropout_6[0][0]
add_1 (Add)	(None, 1, 10)	0	dense_5[0][0], add[0][0]
global_average_pooling1d (GlobalAveragePooling1D)	(None, 10)	0	add_1[0][0]
dense_6 (Dense)	(None, 8)	88	global_average_pooling1d[0][0]
dropout_7 (Dropout)	(None, 8)	0	dense_6[0][0]
dense_7 (Dense)	(None, 1)	9	dropout_7[0][0]

Nota. Elaboración propia.

3.8.2 Entrenamiento del Modelo Transformer

El modelo Transformer se entrena utilizando los conjuntos de entrenamiento y validación bajo las mismas condiciones experimentales empleadas en los modelos anteriormente

descritos. Se implementan estrategias de Early Stopping para prevenir el sobreajuste, reducción adaptativa de la tasa de aprendizaje mediante ReduceLROnPlateau y almacenamiento automático del mejor modelo utilizando ModelCheckpoint. Estas técnicas permiten optimizar el proceso de entrenamiento y favorecer la capacidad de generalización del modelo.

Figura 32

Resumen del entrenamiento del modelo Transformer

```
Total params: 3,581 (13.99 KB)
Trainable params: 3,581 (13.99 KB)
Non-trainable params: 0 (0.00 B)
Épocas entrenadas : 26
Mejor val_loss : 0.013808
```

Nota. Elaboración propia.

3.8.3 Evaluación

Esta sección presenta la evaluación del modelo Transformer desarrollado para la predicción del consumo energético residencial. Para determinar su capacidad predictiva, se analizaron métricas de error ampliamente utilizadas en problemas de regresión, permitiendo cuantificar la diferencia entre los valores observados y las predicciones generadas por el modelo.

Evaluación del Modelo Transformer

Con el propósito de medir el desempeño predictivo del modelo Transformer, se calcularon las métricas MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) y MAPE (Mean Absolute Percentage Error). Estas métricas permiten evaluar tanto el error absoluto como el error relativo de las predicciones, proporcionando una visión integral de la precisión alcanzada por el modelo.

Tabla 22: *Resumen de métricas del modelo Transformer*

Conjunto	MAE	RMSE	MAPE (%)	R ²
Entrenamiento	348 185	436 617	2.0215	0.7850
Validación	441 110	527 758	2.4291	-0.1915
Prueba	1 185 052	1 505 002	6.8372	-0.0831

Nota. Elaboración propia.

Comparación Transformer vs Baselines

Con el objetivo de determinar si el modelo Transformer aporta una mejora real en la capacidad predictiva, se realizó una comparación contra dos modelos de referencia o *baselines*. El primer baseline corresponde al método de persistencia, el cual asume que el consumo del siguiente mes será igual al último valor observado. El segundo baseline utiliza como predicción la media histórica del conjunto de entrenamiento. Estos enfoques representan referencias simples ampliamente utilizadas en problemas de series temporales y permiten verificar si la complejidad adicional del modelo Transformer se traduce efectivamente en una mejora del desempeño predictivo. La comparación se realizó sobre el conjunto de prueba utilizando las métricas MAE, MAPE y coeficiente de determinación (R^2).

Tabla 23: *Métricas de desempeño de los modelos baseline sobre el conjunto de prueba*

Modelo	MAE	RMSE	MAPE (%)	R ²
Persistencia (naive)	893 235	1 298 026	5.0612	0.1943
Media histórica (naive)	1 705 383	1 810 176	9.4135	-0.5668

Nota. Elaboración propia.

Tabla 24: Comparación del modelo Transformer frente a los modelos baseline sobre el conjunto de prueba

Modelo	MAE	MAPE (%)	R ²
Transformer	1 185 052	6.84	-0.083
Persistencia (naive)	893 235	5.06	0.194
Media histórica (naive)	1 705 383	9.41	-0.567

Nota. Elaboración propia.

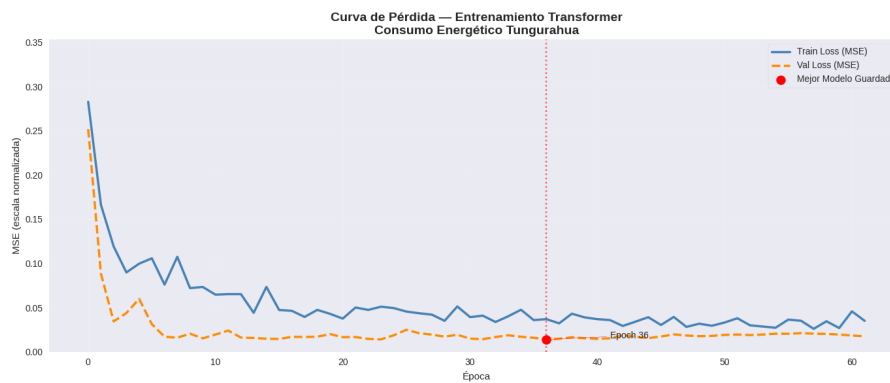
3.8.4 Visualización

Además de las métricas cuantitativas, se analizaron diversas representaciones gráficas con el fin de evaluar visualmente el comportamiento del modelo. Estas visualizaciones permiten identificar patrones de aprendizaje, capacidad de generalización y posibles errores sistemáticos presentes en las predicciones.

Curva de Pérdida

La curva de pérdida muestra la evolución de la función de error durante el proceso de entrenamiento y validación. Su análisis permite verificar la convergencia del modelo, identificar posibles problemas de sobreajuste o subajuste y evaluar la efectividad de la estrategia de entrenamiento implementada.

Figura 33
Curva de Pérdida Transformer

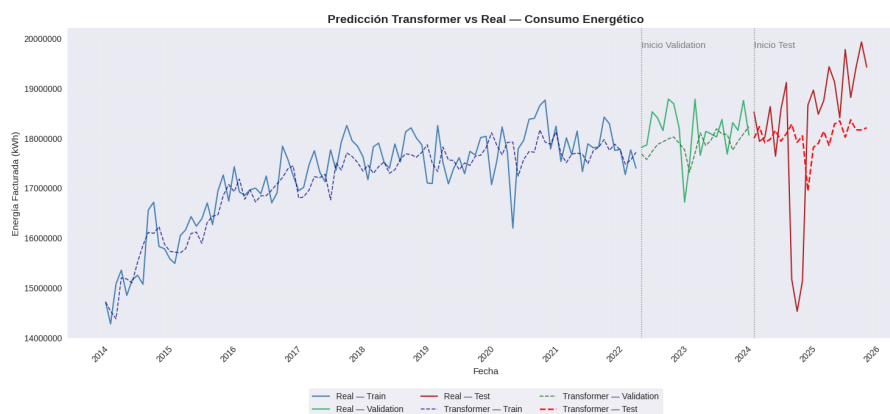


Nota. Elaboración propia.

Predicción vs Real

Esta visualización compara los valores reales del consumo energético con las predicciones generadas por el modelo Transformer. Su objetivo es evaluar la capacidad del modelo para seguir la tendencia general de la serie temporal y reproducir adecuadamente las variaciones observadas en los datos históricos.

Figura 34
Predicción Transformer



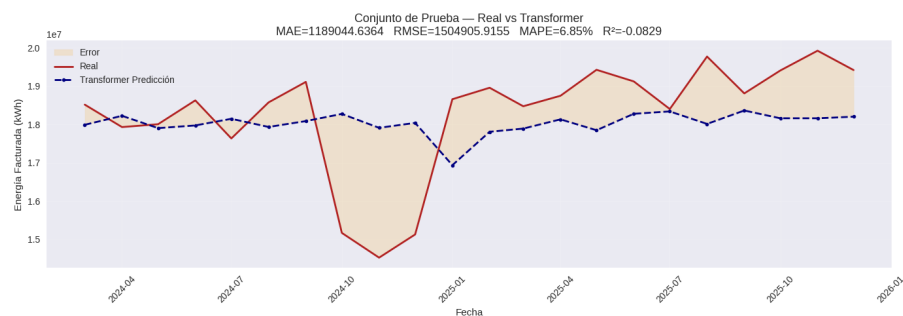
Nota. Elaboración propia.

Detalle en el Conjunto de Prueba

Con el fin de analizar el comportamiento del modelo en los datos no observados durante el entrenamiento, se presenta una visualización enfocada exclusivamente en el conjunto de prueba. Esto permite examinar con mayor detalle la precisión de las predicciones y detectar posibles desviaciones locales.

Figura 35

Análisis en el Conjunto de Prueba Transformer

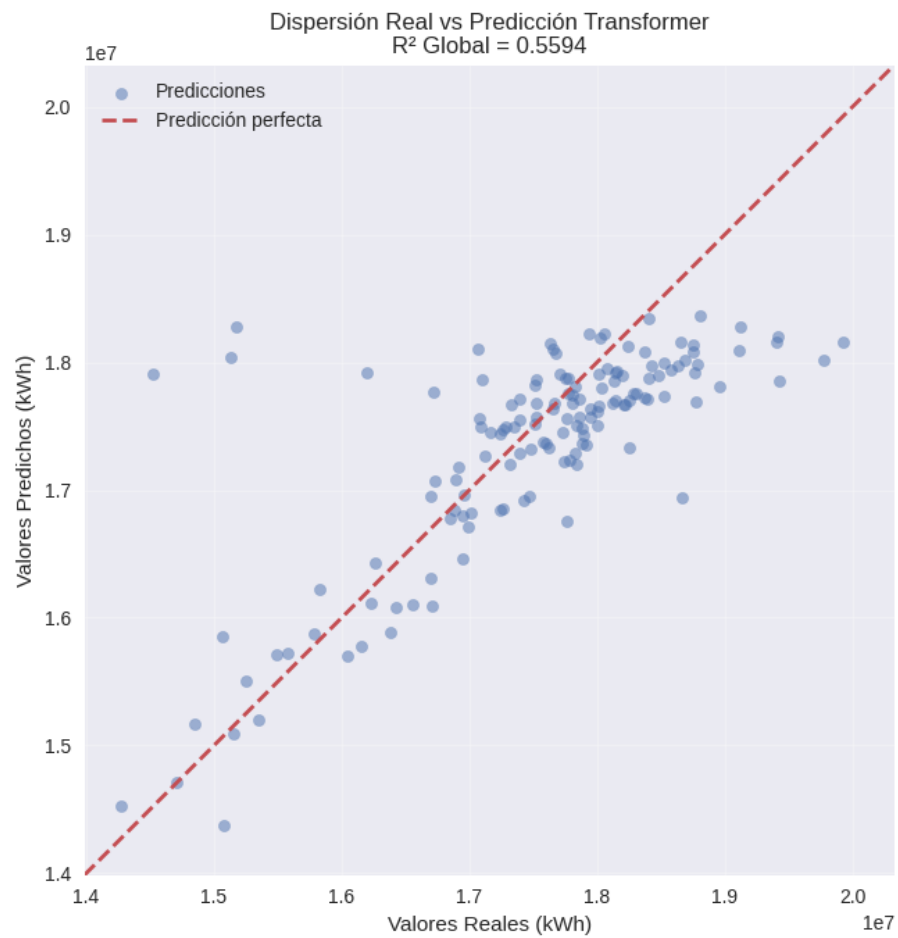


Nota. Elaboración propia.

Dispersión Real vs Predicción

El gráfico de dispersión entre valores reales y predichos permite evaluar el grado de correlación existente entre ambas variables. Cuanto más próximos se encuentren los puntos a la línea ideal de referencia, mayor será la capacidad del modelo para reproducir los valores observados.

Figura 36
Dispersión Real Transformer

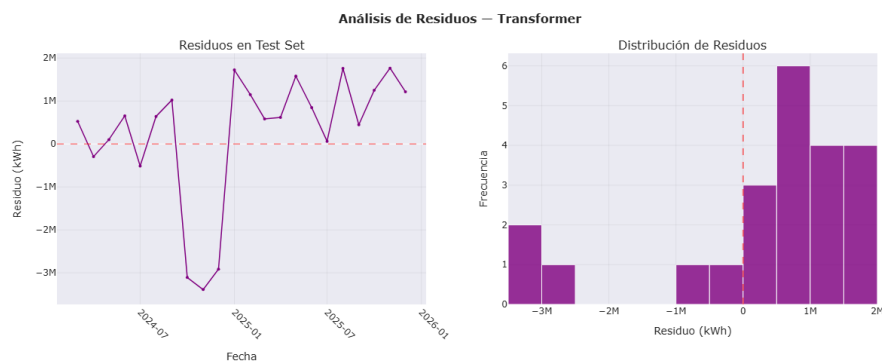


Nota. Elaboración propia.

Análisis de Residuos

El análisis de residuos examina las diferencias entre los valores reales y los predichos por el modelo. Este estudio permite identificar patrones sistemáticos de error, posibles sesgos en las predicciones y verificar si los errores se distribuyen de manera aleatoria, condición deseable para un modelo predictivo robusto.

Figura 37
Análisis de Residuos Transformer



Nota. Elaboración propia.

3.8.5 Tabla Resumen de Métricas

Finalmente, se presenta una tabla consolidada con las principales métricas de desempeño obtenidas por el modelo Transformer. Esta síntesis facilita la interpretación de los resultados y permite realizar comparaciones directas con las demás arquitecturas evaluadas en la investigación.

Tabla 25: *Métricas de evaluación del modelo Transformer*

Conjunto	MAE (kWh)	RMSE (kWh)	MAPE (%)	R ²
Entrenamiento	350 939.6717	439 701.7556	2.0369	0.7820
Validación	445 533.8171	531 297.6852	2.4528	-0.2076
Prueba	1 189 044.6364	1 504 905.9155	6.8546	-0.0829

Nota. Elaboración propia.

3.9 Red Neuronal GRU

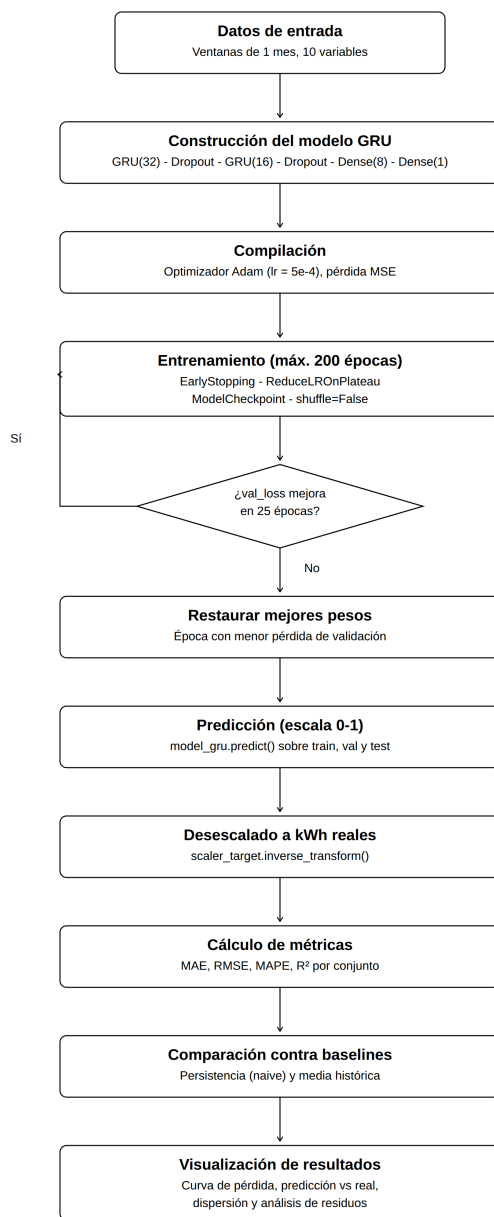
Se implementó una arquitectura basada en GRU (Gated Recurrent Unit) adaptada para la predicción de series temporales de consumo energético. El modelo emplea celdas recurrentes con mecanismos de puerta (*gates*) de actualización y de olvido, los cuales permiten retener información relevante a lo largo de la secuencia temporal de entrada con una estructura más simple y un menor número de parámetros que una celda LSTM equivalente.

Con el objetivo de garantizar una comparación equitativa con los modelos LSTM y Transformer, se mantuvieron condiciones experimentales idénticas, incluyendo la misma estrategia de partición de datos (70% entrenamiento, 15% validación, 15% prueba), la misma ventana temporal de entrada (`LOOK_BACK` = 1 mes), la función de pérdida basada en el error cuadrático medio (MSE), el optimizador Adam con una tasa de aprendizaje inicial de 5×10^{-4} , la regularización mediante capas Dropout y la aplicación de Early Stopping para evitar el sobreentrenamiento.

La Figura 38 resume el proceso completo seguido para la construcción, entrenamiento y evaluación del modelo, desde la recepción de los datos preparados hasta la generación de las visualizaciones finales, las cuales se presentan en el Capítulo 4.

Figura 38

Diagrama de flujo del proceso de construcción y evaluación del modelo GRU



Nota. Elaboración propia.

3.9.1 Construcción del Modelo GRU

Se implementa una arquitectura basada en GRU compuesta por dos capas recurrentes apiladas (32 y 16 unidades respectivamente), cada una seguida de una capa de regularización

Dropout (10 %), una capa densa intermedia de 8 neuronas con activación ReLU, y una capa de salida lineal de una sola neurona para la predicción de regresión.

Tabla 26: *Arquitectura del modelo GRU*

Layer (type)	Output Shape	Param #	Connected to
input_layer_1 (InputLayer)	(None, 1, 10)	0	-
gru (GRU)	(None, 1, 32)	4224	input_layer_1[0][0]
dropout_2 (Dropout)	(None, 1, 32)	0	gru[0][0]
gru_1 (GRU)	(None, 16)	2400	dropout_2[0][0]
dropout_3 (Dropout)	(None, 16)	0	gru_1[0][0]
dense_2 (Dense)	(None, 8)	136	dropout_3[0][0]
dense_3 (Dense)	(None, 1)	9	dense_2[0][0]

Nota. Elaboración propia.

3.9.2 Entrenamiento del Modelo GRU

El modelo GRU se entrena utilizando los conjuntos de entrenamiento y validación bajo las mismas condiciones experimentales empleadas en los modelos LSTM y Transformer. Se implementan estrategias de Early Stopping (paciencia de 25 épocas, restaurando los pesos de la mejor época) para prevenir el sobreajuste, reducción adaptativa de la tasa de aprendizaje mediante ReduceLROnPlateau (factor 0.5, paciencia de 10 épocas) y almacenamiento automático del mejor modelo utilizando ModelCheckpoint. Bajo estas condiciones, el entrenamiento se detuvo de forma anticipada en la época 52 (de un máximo de 200), restaurando los pesos correspondientes a la época 27, donde se registró la menor pérdida de validación.

Figura 39*Resumen del entrenamiento del modelo GRU*

```

Total params: 6,769 (26.44 KB)
Trainable params: 6,769 (26.44 KB)
Non-trainable params: 0 (0.00 B)
Épocas entrenadas : 52
Mejor val_loss : 0.013100

```

3.9.3 Evaluación

Una vez finalizado el entrenamiento, se generaron predicciones sobre los conjuntos de entrenamiento, validación y prueba. Debido a que la variable objetivo fue normalizada durante la etapa de preprocesamiento, las predicciones obtenidas fueron transformadas nuevamente a su escala original mediante el proceso de desnormalización, permitiendo interpretar los resultados en unidades reales de kilovatios-hora (kWh).

Para evaluar el desempeño del modelo se utilizaron las mismas cuatro métricas empleadas en los modelos LSTM y Transformer: el error absoluto medio (MAE), la raíz del error cuadrático medio (RMSE), el error porcentual absoluto medio (MAPE) y el coeficiente de determinación (R^2).

Tabla 27: *Resumen de métricas del modelo GRU*

Conjunto	MAE (kWh)	RMSE (kWh)	MAPE (%)	R^2
Entrenamiento	429 190.27	526 255.45	2.5197	0.6877
Validación	388 708.39	514 048.05	2.1617	-0.1304
Prueba	894 263.82	1 241 915.18	5.2609	0.2625

Nota. Elaboración propia.

Los resultados obtenidos por la arquitectura GRU fueron comparados con los mismos dos modelos de referencia utilizados para evaluar al LSTM y al Transformer: un modelo de

persistencia y un modelo basado en el promedio histórico del conjunto de entrenamiento.

Tabla 28: Comparación del modelo GRU frente a los modelos baseline sobre el conjunto de prueba

Modelo	MAE (kWh)	MAPE (%)	R ²
GRU	894 264	5.26	0.262
Persistencia (naive)	893 235	5.06	0.194
Media histórica (naive)	1 705 383	9.41	-0.567

Nota. Elaboración propia.

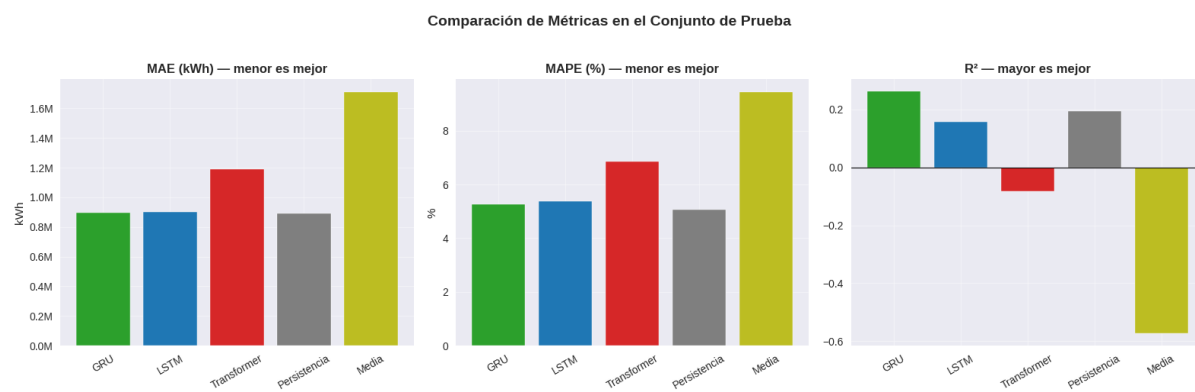
A nivel agregado, GRU obtuvo un desempeño ligeramente superior al LSTM en las cuatro métricas sobre el conjunto de prueba, con una arquitectura más liviana (6 769 parámetros frente a los 8 785 del modelo LSTM). El análisis visual del comportamiento del modelo, así como la discusión detallada de estos resultados frente a los baselines, se desarrolla en el Capítulo 4.

3.10 Comparación de Eficiencia de Los Modelos Aplicados

Una vez evaluadas las arquitecturas LSTM, GRU y Transformer, se realizó una comparación integral de su desempeño. Para ello se consideraron las métricas de precisión obtenidas, el comportamiento durante el entrenamiento y la capacidad de generalización sobre el conjunto de prueba. Este análisis permite identificar la arquitectura más adecuada para la predicción del consumo energético residencial en la provincia de Tungurahua y establecer recomendaciones para futuras investigaciones en el área.

Figura 40

Comparativa entre los Modelos Analizados



Nota. Elaboración propia.

CAPÍTULO 4: ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

4.1 Resultados

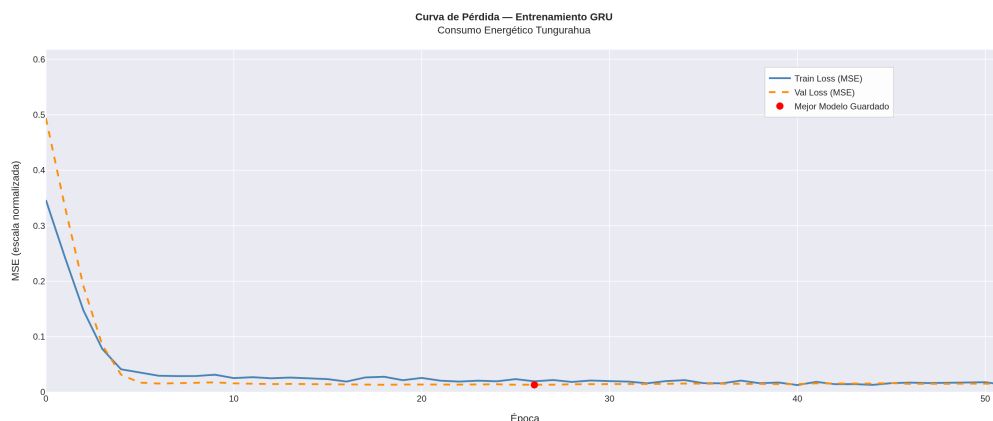
4.1.1 Resultados del modelo GRU

4.1.1.1 Curva de pérdida durante el entrenamiento

La Figura 41 muestra la evolución del error cuadrático medio (MSE, en escala normalizada) durante el entrenamiento. Ambas curvas —entrenamiento y validación— descienden de forma conjunta durante las primeras épocas y se estabilizan alrededor de la época 10, sin evidenciar una divergencia marcada entre ambas series que indique sobreajuste severo. El punto marcado corresponde a la época 27, donde se registró la menor pérdida de validación antes de la detención anticipada del entrenamiento.

Figura 41

Curva de pérdida — Entrenamiento del modelo GRU



Nota. Elaboración propia. MSE en escala normalizada (0-1) sobre los conjuntos de entrenamiento y validación.

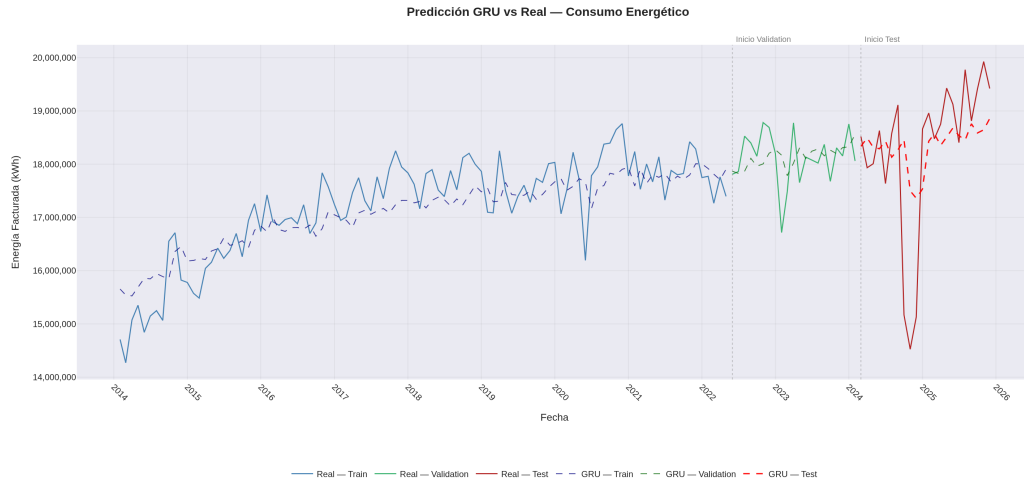
4.1.1.2 Predicción frente a valores reales

La Figura 42 contrasta la serie real de consumo energético con la predicción generada por el modelo GRU a lo largo de los tres conjuntos (entrenamiento, validación y prueba). El modelo logra seguir razonablemente bien la tendencia general creciente del consumo entre

2014 y 2022. Sin embargo, en el conjunto de prueba (2024-2025) se observa una caída abrupta del valor real que el modelo no anticipa, generando el mayor desajuste de toda la serie.

Figura 42

Predicción GRU vs valores reales — Consumo energético



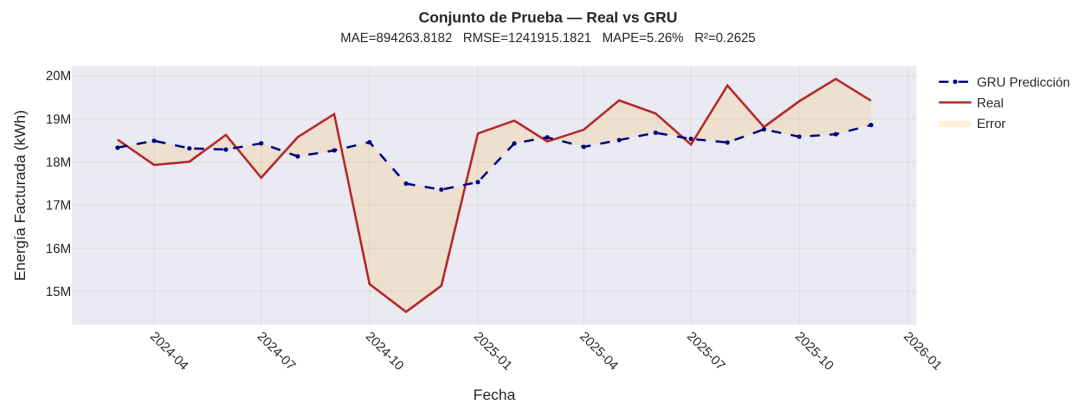
Nota. Elaboración propia.

4.1.1.3 Análisis en el conjunto de prueba

La Figura 43 presenta una vista ampliada del conjunto de prueba, mostrando la banda de error entre el valor real y la predicción del modelo GRU mes a mes.

Figura 43

Detalle de predicción GRU en el conjunto de prueba

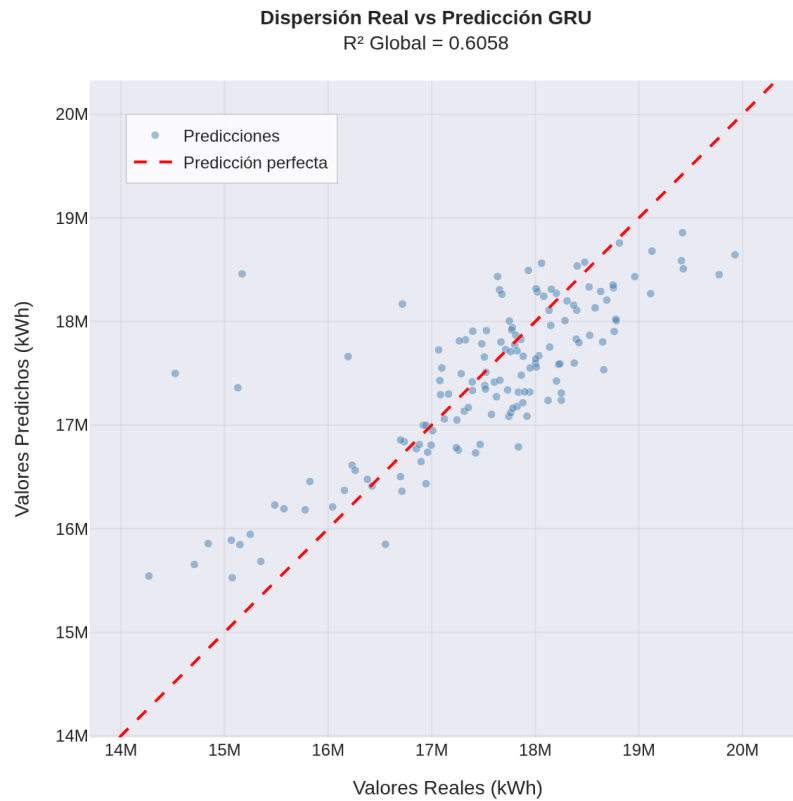


Nota. Elaboración propia.

4.1.1.4 Dispersión real vs. predicción

La Figura 44 presenta el diagrama de dispersión entre los valores reales y los predichos para el conjunto completo de datos (entrenamiento, validación y prueba), junto con la línea de predicción perfecta. El coeficiente de determinación global obtenido es $R^2 = 0,6058$. Es importante señalar que este valor corresponde al ajuste **global** del modelo sobre el total de las observaciones, y no debe confundirse con el R^2 calculado exclusivamente sobre el conjunto de prueba (ver Tabla 29), que es la métrica relevante para evaluar la capacidad real de generalización del modelo.

Figura 44
Dispersión de valores reales vs. predicción — GRU

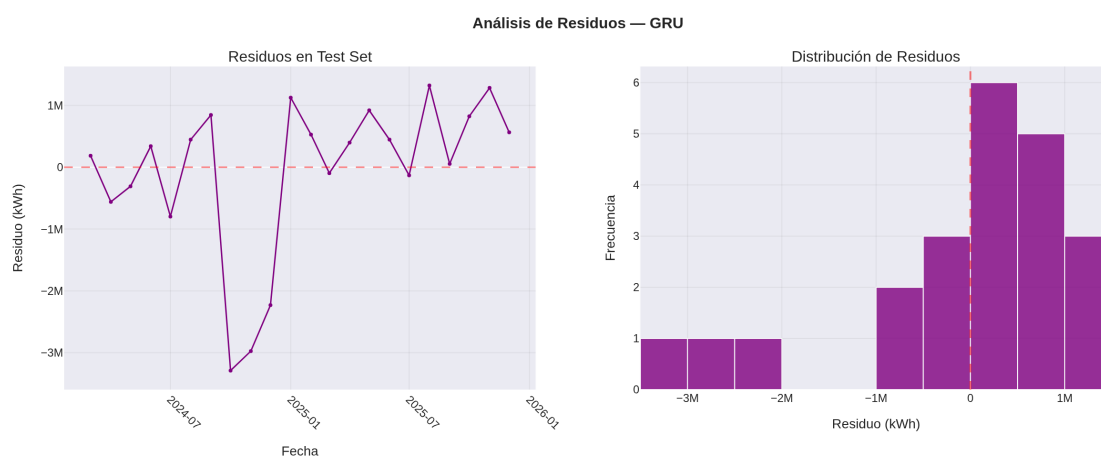


Nota. Elaboración propia.

4.1.1.5 Análisis de residuos

La Figura 45 muestra el comportamiento de los residuos (valor real menos valor predicho) sobre el conjunto de prueba. Se identifican dos residuos negativos considerablemente más grandes que el resto (cerca de -3.0 y -3.3 millones de kWh) durante el último trimestre de 2024, coincidentes con la caída abrupta observada en la Figura 42. El histograma de distribución de residuos muestra una concentración mayoritaria entre -1 y +1.5 millones de kWh, con los dos valores extremos mencionados como casos atípicos claramente diferenciados del resto de la distribución.

Figura 45
Análisis de residuos — GRU



Nota. Elaboración propia.

4.1.1.6 Tabla resumen de métricas

La Tabla 29 resume las métricas de error obtenidas por el modelo GRU en los tres conjuntos de datos.

Tabla 29: Métricas del modelo GRU por conjunto de datos

Conjunto	MAE (kWh)	RMSE (kWh)	MAPE (%)	R^2
Train	429,190	526,255	2.52	0.688
Validation	388,708	514,048	2.16	-0.130
Test	894,264	1,241,915	5.26	0.262

Nota. Elaboración propia. MAE y RMSE expresados en kWh; MAPE en porcentaje.

4.1.1.7 Comparación contra modelos de referencia (baselines)

La comparación numérica del modelo GRU frente a los dos baselines simples (persistencia y media histórica) ya fue presentada en la Tabla 28 del Capítulo 3; esta sección se centra en su interpretación. El modelo GRU **no logra superar de forma clara** al baseline de persistencia en términos de error absoluto: el MAE del GRU (894,264 kWh) es prácticamente idéntico —e incluso ligeramente superior— al del baseline naive (893,235 kWh), y su MAPE (5.26 %) es también levemente mayor (5.06 %). La ventaja del GRU sobre la persistencia naive se concentra en el coeficiente R^2 (0.262 frente a 0.194), lo que indica que el modelo explica una mayor proporción de la varianza total, pero no que sus errores puntuales sean menores. Ambos modelos superan ampliamente al baseline de media, que resulta el peor de los tres por un margen considerable.

4.1.1.8 Discusión

Los resultados obtenidos para el modelo GRU permiten matizar la expectativa inicial de que una arquitectura de aprendizaje profundo, por su mayor capacidad de representación, ofrecería necesariamente una mejora sustancial frente a métodos de pronóstico simples. Si bien el GRU presenta un mejor coeficiente de determinación que el baseline de persistencia (0.262 frente a 0.194), esta ventaja no se traduce en una reducción del error absoluto medio sobre el conjunto de prueba: ambos modelos cometen, en promedio, un error muy similar en kWh.

Dos factores explican en buena medida este comportamiento. En primer lugar, el tamaño del conjunto de entrenamiento (100 observaciones mensuales) es reducido para el potencial real de una red recurrente, lo que limita su capacidad de aprender patrones que vayan más allá

de lo que ya captura un modelo de persistencia. En segundo lugar, el conjunto de prueba contiene un evento de caída abrupta del consumo —concentrado en el último trimestre de 2024— que no tiene precedente en el periodo de entrenamiento ni en el de validación. Al no existir información histórica equivalente, ningún modelo entrenado únicamente sobre los datos disponibles, sea este una red neuronal o una regla de persistencia, puede anticipar dicho quiebre; el error que ambos cometen en esos meses concretos domina el MAE y el MAPE del conjunto de prueba en su totalidad.

Esta observación es relevante para la interpretación del desempeño relativo de los tres modelos comparados en este proyecto (LSTM, GRU y Transformer): dado que los tres se evalúan sobre el mismo conjunto de prueba y, por tanto, están expuestos al mismo evento atípico, las diferencias de desempeño entre ellos deben leerse como diferencias en su sensibilidad a dicho evento, y no exclusivamente como diferencias en su capacidad general de modelado de la serie.

Contrastando estos resultados con la literatura revisada en el Capítulo 2, la brecha de desempeño es considerable. Montero Laurencio et al. (2024) reportaron un coeficiente de correlación $R = 0,9435$ al pronosticar el consumo diario de un edificio de oficinas mediante redes neuronales tipo perceptrón multicapa; Jácome-Vélez et al. (2026) alcanzaron un $R^2 = 99,27\%$ estimando el consumo de una planta industrial a partir de su volumen de producción diario; y Defaz Toapanta et al. (2025), utilizando la misma fuente de datos de ARCONEL empleada en este estudio, reportaron un R^2 superior al 99% con un modelo Random Forest. Frente a estas cifras, el $R^2 = 0,262$ obtenido por el GRU en el conjunto de prueba de este trabajo resulta sustancialmente más modesto.

Esta diferencia no debe interpretarse como una limitación exclusiva del modelo GRU, sino que remite, al menos en parte, a diferencias estructurales entre los problemas abordados. El caso de Jácome-Vélez et al. (2026) es el más claro: la variable predictora (volumen de producción diario) mantiene una relación causal directa y de muy corto plazo con el consumo energético de la planta, lo que favorece un ajuste casi determinista y no representa un desafío de pronóstico equiparable al de una serie temporal mensual con eventos atípicos no vistos previamente. El estudio de Montero Laurencio et al. (2024) no detalla las variables predictoras

utilizadas para el pronóstico del consumo diario del edificio de oficinas, por lo que no es posible establecer con la misma certeza si su resultado responde a una relación igualmente directa; esto limita la comparabilidad directa de ese caso particular. En el de Defaz Toapanta et al. (2025), aunque la fuente de datos coincide con la de este estudio, cabe considerar si la validación se realizó mediante una partición estrictamente cronológica (entrenar con el pasado, evaluar sobre el futuro) o mediante una partición aleatoria de las observaciones; esta última práctica, frecuente en trabajos de aprendizaje automático que no distinguen explícitamente entre validación cruzada estándar y validación de series temporales, tiende a inflar el R^2 reportado al permitir que el modelo sea evaluado sobre observaciones estadísticamente similares a las vistas durante el entrenamiento. El presente estudio empleó una partición estrictamente cronológica, lo que expone al modelo a un escenario de evaluación más exigente y realista: predecir un periodo que no necesariamente comparte la misma distribución estadística que el histórico de entrenamiento, como efectivamente ocurrió con el evento atípico de finales de 2024. Esta distinción metodológica —y no únicamente la elección de arquitectura— puede explicar una parte importante de la brecha observada frente a los resultados reportados en la literatura.

4.1.1.9 Observaciones y hallazgos

- El modelo GRU alcanza su mejor desempeño relativo en el coeficiente de determinación ($R^2 = 0,262$ en prueba, frente a 0.194 de la persistencia naive), pero no logra una reducción del error absoluto medio respecto a dicho baseline.
- El entrenamiento se detuvo de forma anticipada en la época 52 de un máximo de 200, lo que sugiere que, dado el tamaño de la serie disponible, el modelo converge con relativa rapidez y no se beneficia de un entrenamiento prolongado.
- El análisis de residuos identificó dos valores extremos (cerca de -3.0 y -3.3 millones de kWh) concentrados en octubre y noviembre de 2024, los cuales explican una proporción desproporcionadamente alta del error total del modelo en el conjunto de prueba.
- Dichos meses corresponden a una caída de consumo de aproximadamente 25 % frente al nivel habitual de la serie, seguida de una recuperación abrupta en enero de 2025. Es-

te patrón —caída puntual y recuperación inmediata— es estructuralmente similar al de los valores que fueron identificados y corregidos como errores de captura en octubre y noviembre de 2022 durante la fase de preparación de datos, aunque a la fecha de este análisis no se ha aplicado un tratamiento equivalente sobre el periodo 2024, dado que no formaba parte del alcance original de la depuración de datos.

- La diferencia entre el R^2 global (0.6058, calculado sobre el conjunto completo) y el R^2 de prueba (0.262) evidencia que el ajuste del modelo es considerablemente más fuerte sobre los datos de entrenamiento que sobre los datos no vistos, lo cual debe tenerse presente al comunicar el desempeño del modelo de forma agregada.

CAPÍTULO 5: CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

A partir de los resultados obtenidos para el modelo GRU se concluye que, bajo las condiciones experimentales evaluadas, la arquitectura no ofrece una mejora práctica significativa frente a un modelo de referencia considerablemente más simple. La ventaja observada en el coeficiente de determinación no es suficiente, por sí sola, para justificar la mayor complejidad computacional y de mantenimiento que implica una red neuronal recurrente frente a una regla de persistencia, al menos para el horizonte de predicción y el tamaño de muestra disponibles en este estudio.

Se pudo establecer mediante este estudio que un modelo simple de persistencia ingenua obtuvo métricas de error altamente competitivas frente a las arquitecturas de Deep Learning evaluadas.

Este resultado demuestra que, para series temporales cortas y de frecuencia mensual como la analizada, los modelos tradicionales pueden constituir una alternativa más eficiente y práctica, ya que alcanzan niveles de precisión similares con una complejidad significativamente menor. En este contexto, la incorporación de arquitecturas más avanzadas no garantiza mejoras sustanciales en el desempeño predictivo bajo las condiciones de datos disponibles en este estudio.

El desempeño del modelo está condicionado de forma importante por la presencia de un evento no recurrente en el periodo de prueba, cuya magnitud no tiene precedente en el historial utilizado para el entrenamiento. Esto indica que las métricas de error reportadas no deben interpretarse como una medida estable del desempeño esperado del modelo en producción, sino como el resultado de una evaluación particular, sensible a la composición específica del conjunto de prueba seleccionado.

Por otra parte, el análisis de autocorrelación parcial (PACF) evidenció una estructura predominantemente AR(1), es decir, que el consumo de un mes depende principalmente del mes inmediatamente anterior y no de patrones cíclicos de mayor plazo. Los experimentos rea-

lizados confirmaron esta estructura: el uso de ventanas temporales más amplias (6 o 12 meses) no mejoró la predicción, sino que incorporó ruido que dificultó el aprendizaje de las redes neuronales. Este resultado es consistente con la ausencia, en la serie analizada, de una estacionalidad anual marcada que justifique ventanas más largas, aunque establecer la causa climática o socioeconómica exacta de este comportamiento excede el alcance de este estudio.

5.2 Recomendaciones

- Investigar el origen de la caída de consumo registrada entre octubre y diciembre de 2024 antes de utilizar este conjunto de prueba como referencia definitiva de desempeño, dado que su patrón es consistente con un posible error de captura o reporte de datos, similar a los ya identificados y corregidos para 2022.
- De confirmarse lo anterior, evaluar el efecto de aplicar el mismo procedimiento de imputación estacional utilizado en la fase de preparación de datos sobre dicho periodo, y reportar los resultados del modelo bajo ambos escenarios (con y sin tratamiento) para mayor transparencia.
- Ampliar, de ser posible, el horizonte temporal de los datos disponibles, dado que el tamaño actual de la muestra (144 observaciones mensuales) limita la capacidad de las arquitecturas recurrentes para diferenciarse de modelos de referencia más simples.
- Para trabajos futuros que busquen justificar el uso de arquitecturas de mayor complejidad, se recomienda reportar siempre el desempeño frente a baselines simples (persistencia y media histórica), tal como se hizo en este estudio, en lugar de evaluar el modelo de forma aislada.

Referencias

- Agencia de Regulación y Control de Electricidad. (2025, junio). *Resolución Nro. ARCONEL-006/25* (inf. téc.). Agencia de Regulación y Control de Electricidad. https://arconel.gob.ec/wp-content/uploads/downloads/2025/07/Resolucion-ARCONEL-006-25_PTSPEE-2025-codificado_certificado.pdf
- Agencia de Regulación y Control de Electricidad – ARCONEL. (2025). *Panorama eléctrico (Edición 30, septiembre 2025)* (inf. téc.) (PDF recuperado el 17 de enero de 2026). Agencia de Regulación y Control de Electricidad. https://arconel.gob.ec/wp-content/uploads/downloads/2025/10/PanoramaElectrico-30_v5-1.pdf
- Agencia de Regulación y Control de Electricidad (ARCONEL). (2026). *Plataforma del Sistema de Información de Datos Energéticos (SISDAT)* [Consultado el 15 de enero de 2026]. <https://reportes.arconel.gob.ec/>
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 1724-1734. <https://doi.org/10.3115/v1/D14-1179>
- Chollet, F. (2021). *Deep Learning with Python* (2.^a ed.). Manning Publications.
- Climent Diranzo, F. J., Valor i Micó, E., Torró i Enguix, H., & Caselles Miralles, V. (2003). Incidencia de la climatología en el consumo de gas y electricidad en España. *Información Comercial Española, ICE: Revista de Economía*, (808), 55-70. <https://www.divulgameteo.es/fotos/meteoroteca/Climatolog%C3%ADa-gas-electricidad.pdf>
- Deepchecks. (s.f.-a). Mean absolute error (MAE). <https://deepchecks.com/glossary/mean-absolute-error/>
- Deepchecks. (s.f.-b). Mean square error (MSE). <https://deepchecks.com/glossary/mean-square-error-mse/>
- Defaz Toapanta, V. E., Ortiz Velasco, L. V., Reyes Pérez, J. E., & Núñez Jácome, R. S. (2025). *Predicción de consumo de energía eléctrica en hogares utilizando modelos de apren-*

- dizaje automático* [Tesis de maestría, Universidad Internacional del Ecuador]. <https://repositorio.uide.edu.ec/handle/37000/8546>
- Gallo, V., Espejo, P., Alvarez, S., & Nuñez, J. (2025). Estudio del sistema de generación eléctrica mediante el análisis de centrales hidroeléctricas para denotar el déficit energético en la provincia Tungurahua. *CIYA. Ciencias de la Ingeniería y Aplicadas*, 9(2), 157-170. <https://doi.org/10.61236/ciya.v9i2.1109>
- Géron, A. (2022). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (3.^a ed.). O'Reilly Media.
- González Duque, R. (2008). *Python para todos* (1.^a ed.) [Libro en formato PDF]. Autoedición. <https://persoal.citius.usc.es/eva.cernadas/informaticaparacientificos/material/libros/Python%20para%20todos.pdf>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Haya, P. (2021, noviembre). *La metodología CRISP-DM en ciencia de datos* [Instituto de Ingeniería del Conocimiento (IIC)]. <https://www.iic.uam.es/innovacion/metodologia-crisp-dm-ciencia-de-datos/>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3.^a ed.). OTexts. <https://otexts.com/fpp3/>
- IBM. (2023). *Guía de CRISP-DM de IBM SPSS Modeler* [Versión 18.5.0]. IBM Corporation. https://www.ibm.com/docs/es/SS3RA7_18.5.0/nl/es/pdf/ModelerCRISPDM.pdf
- IBM. (s.f.). *R-squared statistics*. IBM Documentation. <https://www.ibm.com/docs/es/spss-statistics/32.0.0?topic=model-r-squared-statistics>
- IBM Corp. (s.f.). ¿Qué es la autocorrelación? <https://www.ibm.com/es-es/think/topics/autocorrelation>
- Jácome-Vélez, W., Vilela-Sabando, J., Cedeño-Hormaza, M., & Jácome-Duarte, C. (2026). Predicción del consumo energético mediante un modelo matemático de regresión lineal a partir de la producción diaria en la planta envasadora Q'Agua. *Revista Científica FI-*

- NIBUS – Ingeniería, Industria y Arquitectura*, 9(17), 129-138. <https://doi.org/10.56124/finibus.v9i17.011>
- Jansen, S. (2018). *Hands-on machine learning for algorithmic trading: Design and implement investment strategies with machine learning*. Packt Publishing.
- La Hora. (2026, abril). Consumo eléctrico crece hasta 4 veces más que la generación y pone al sistema al límite en Ecuador. <https://www.lahora.com.ec/economia/consumo-electrico-crece-hasta-4-veces-mas-que-la-generacion-y-pone-al-sistema-al-limite-en-ecuador-20260402-0006.html>
- Lozada Cabrera, C. H., Apolo, A., P., N., Paredes Moran, D. V., & Vique Salazar, G. I. (2022). Predicción de demanda eléctrica utilizando redes neuronales artificiales para un sistema de distribución de energía eléctrica. *Ciencia Latina Revista Científica Multidisciplinar*, 6(3), 3871-3894. https://doi.org/10.37811/cl_rcm.v6i3.2502
- Milvus. (s.f.). What is the mean absolute percentage error (MAPE), and how is it calculated? <https://milvus.io/ai-quick-reference/what-is-the-mean-absolute-percentage-error-mape-and-how-is-it-calculated>
- Montero Laurencio, R., Torres Breff, O., Marrero Ramírez, S., & Jiménez Jiménez, D. (2024). Predicción de consumo y demanda de electricidad mediante redes neuronales artificiales y algoritmo iterativo. *Revista Politécnica*, 54(3), 45-56. <https://doi.org/10.33333/rp.vol54n3.05>
- Müller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python*. O'Reilly Media.
- Red Eléctrica de España. (s.f.). Variación y componentes de la demanda de energía eléctrica. <https://www.sistemaelectrico-ree.es/es/informe-del-sistema-electrico/demanda/variacion-componentes-demanda-energia-electrica>
- Salinas Herrera, D. A., & Pesántez Díaz, F. G. (2025). Impacto del cambio climático en la planificación de la operación del Sistema Nacional Interconectado. En *Panorama Eléctrico* (pp. 8-20, Vol. 30). Agencia de Regulación y Control de Electricidad (ARCONEL).

- Saltz, J. S. (2021). CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps. *Proceedings of the 2021 IEEE International Conference on Big Data (BigData)*, 2337-2344. <https://doi.org/10.1109/BigData52589.2021.9671634>
- Schneider Electric. (2023, mayo). Factores clave para aumentar la eficiencia energética en edificios. <https://blogespanol.se.com/gestion-de-la-energie/2023/05/25/factores-clave-para-aumentar-la-eficiencia-energetica-en-edificios/>
- Scikit-learn. (s.f.). *mean_absolute_percentage_error*. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_absolute_percentage_error.html
- Severiche-Maury, C., et al. (2025). Predicting Energy Consumption and Time of Use of Home Appliances in an HEMS Using LSTM Networks and Smart Meters: A Case Study in Sincelejo, Colombia. *Sustainability*, 17(11), 4749.
- Soria Energía. (s.f.). Qué es el consumo energético. <https://soriaenergia.com/que-es-el-consumo-energetico/>
- Universidad ORT Uruguay, Facultad de Ingeniería. (s.f.). *Qué es la electricidad y quién la descubrió* [Consultado el 24 de mayo de 2026]. <https://fi.ort.edu.uy/blog/que-es-la-electricidad-y-quien-la-descubrido>
- VanderPlas, J. (2016). *Python Data Science Handbook: Essential Tools for Working with Data* [Consultado en línea]. O'Reilly Media. <https://jakevdp.github.io/PythonDataScienceHandbook/>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30, 5998-6008. <https://arxiv.org/abs/1706.03762>
- World Meteorological Organization. (2025). *State of the global climate 2024* (inf. téc.). WMO. <https://wmo.int/publication-series/state-of-global-climate/state-of-global-climate-2024>

Anexo A

Detalles de la experimentación con ventanas temporales alternativas

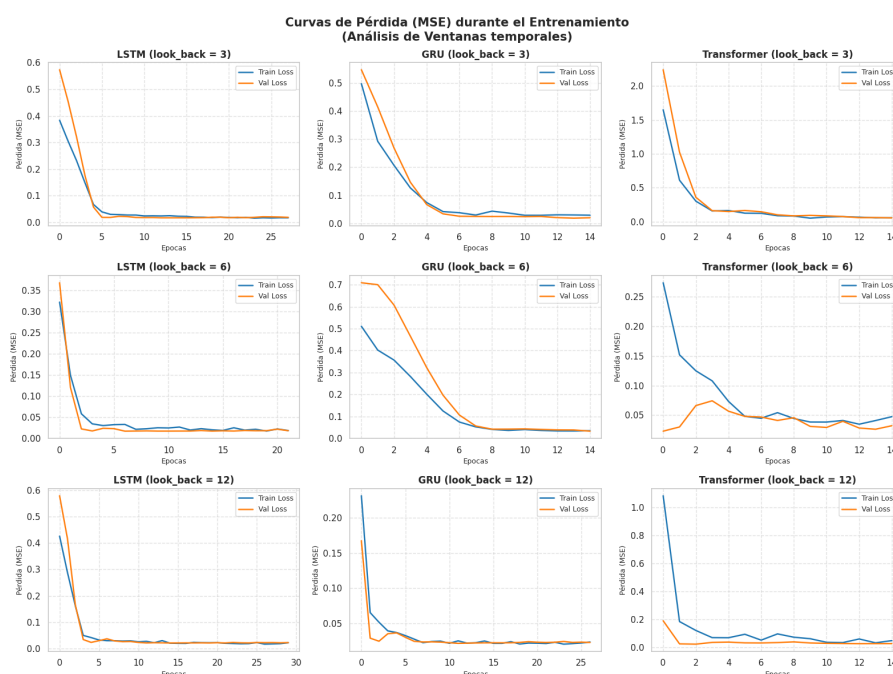
Este apéndice presenta el soporte visual y el análisis detallado del comportamiento de los modelos de aprendizaje profundo (LSTM, GRU y Transformer) entrenados utilizando ventanas temporales (look-back) alternativas de 3, 6 y 12 meses. Esta información complementa las justificaciones estadísticas y empíricas descritas en la Sección 3.6.3 de esta investigación.

Curvas de Pérdida en el Entrenamiento

La siguiente figura presenta el comportamiento del error cuadrático medio (MSE) durante el proceso de entrenamiento y validación para cada una de las épocas en todas las combinaciones de modelos y ventanas.

Figura 46

Curvas de pérdida (MSE) de entrenamiento y validación para ventanas alternativas

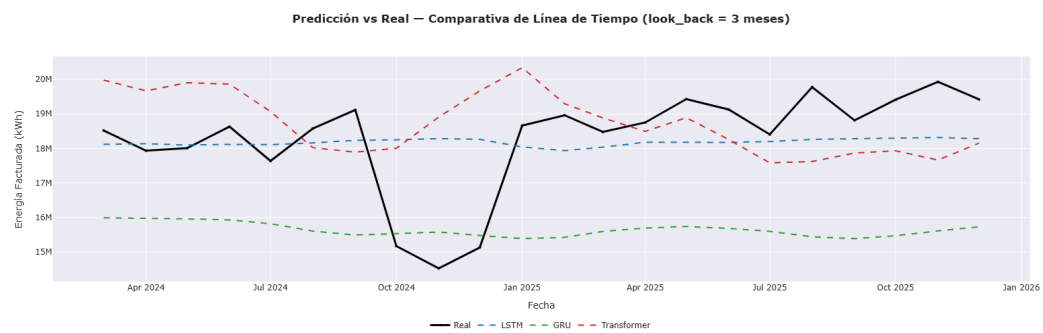


Nota. Se observa que en ventanas de 6 y 12 meses, la pérdida de validación (línea naranja) se desvía tempranamente de la curva de entrenamiento (línea azul) en los modelos GRU y Transformer, confirmando un sobreajuste severo debido a la baja cantidad de muestras por parámetro entrenable.

Comparativa de Predicción en el Conjunto de Prueba

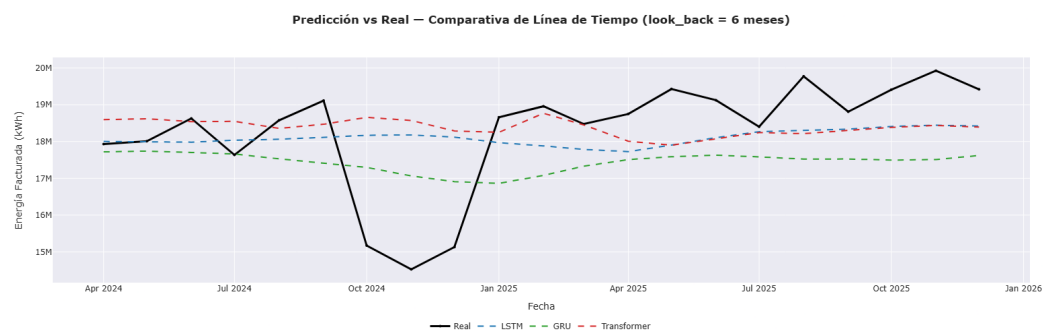
Las figuras a continuación muestran la trayectoria temporal del consumo real del conjunto de prueba frente a las estimaciones de los tres modelos de aprendizaje profundo evaluados para las ventanas alternas.

Figura 47
Predicción vs Real en la línea de tiempo (look_back = 3 meses)

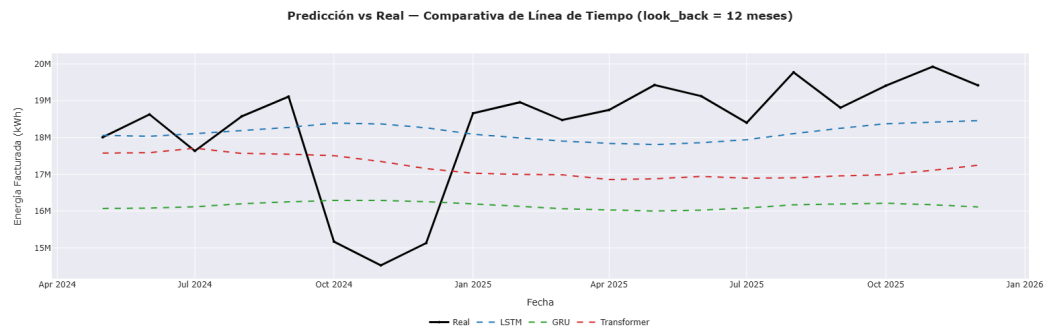


Nota. Con una ventana temporal de 3 meses, el modelo Transformer presenta un mayor desalineamiento respecto a la serie real, mientras que la GRU evidencia una subestimación sistemática del consumo durante el período evaluado..

Figura 48
Predicción vs Real en la línea de tiempo (look_back = 6 meses)



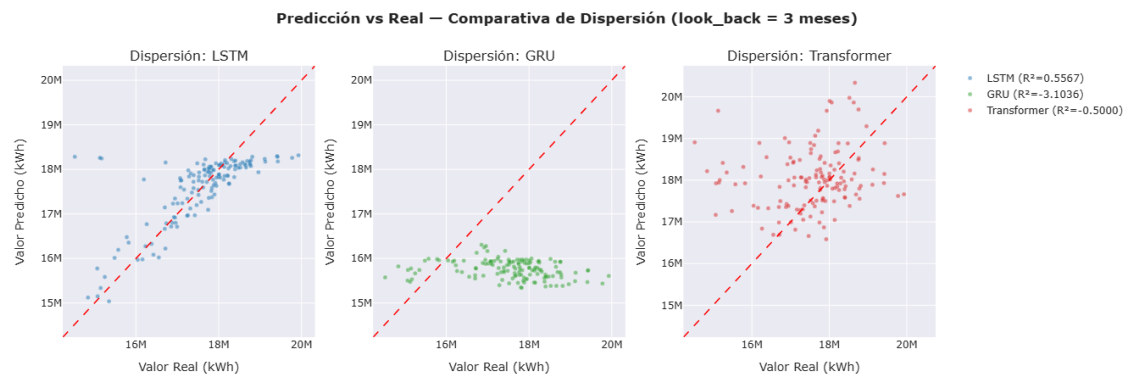
Nota. Con una ventana temporal de 6 meses, los modelos capturan la tendencia general de la serie, aunque presentan predicciones suavizadas que no reproducen con precisión las variaciones y cambios bruscos observados en la demanda real.

Figura 49*Predicción vs Real en la línea de tiempo (look_back = 12 meses)*

Nota. Con una ventana temporal de 12 meses, las predicciones de los tres modelos presentan un suavizado excesivo, lo que reduce su capacidad para representar la variabilidad y las fluctuaciones del consumo eléctrico mensual.

Comparativa de Dispersión Multimodelo

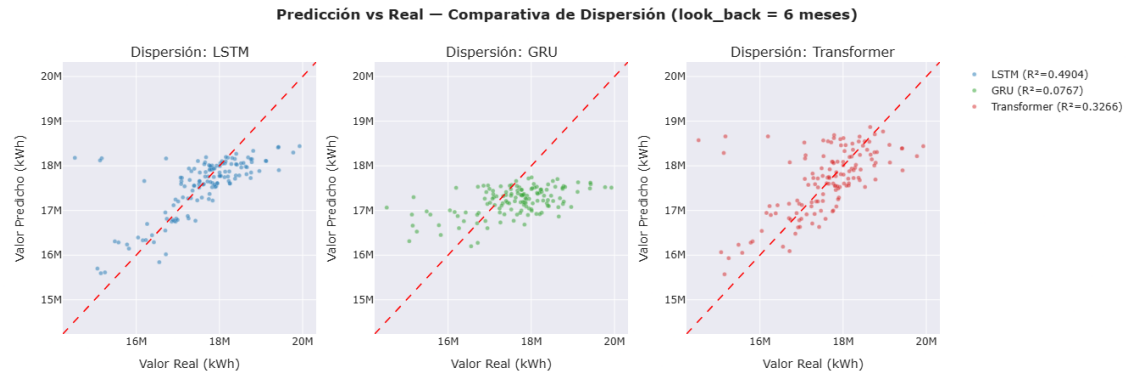
Los siguientes gráficos de dispersión evalúan la correspondencia de los modelos frente a la línea de predicción perfecta para cada una de las ventanas evaluadas.

Figura 50*Gráficos de dispersión multimodelo para look_back = 3 meses*

Nota. La amplia dispersión de los puntos respecto a la diagonal de referencia roja para GRU y Transformer evidencia la pérdida de ajuste y justifica los coeficientes R^2 negativos obtenidos.

Figura 51

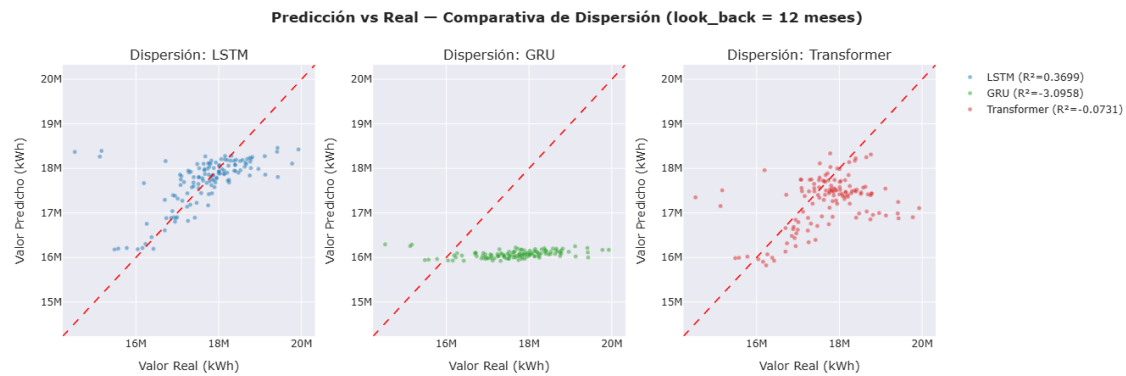
Gráficos de dispersión multimodelo para look_back = 6 meses



Nota. Se observa una mayor concentración de los puntos alrededor de la diagonal en el modelo LSTM. En contraste, GRU y Transformer mantienen una dispersión más elevada, lo que evidencia un menor ajuste entre los valores reales y predichos.

Figura 52

Gráficos de dispersión multimodelo para look_back = 12 meses



Nota. Con una ventana temporal de 12 meses, la dispersión de los puntos aumenta y la alineación con la diagonal de referencia disminuye, especialmente en los modelos GRU y Transformer, reflejando una pérdida del ajuste predictivo.

Anexo B

Repositorio y laboratorio práctico

El laboratorio práctico desarrollado durante la investigación se encuentran disponibles en el siguiente repositorio:

<https://github.com/PaulMarcelo/energy-consumption-project-tungurahua>

El repositorio contiene el notebooks de Google Colab con los procesos de preprocesamiento, el análisis exploratorio, el entrenamiento de modelos y los resultados experimentales obtenidos, además del conjunto de datos utilizado.