

Desarrollo de un modelo de inteligencia artificial de uso específico para estimación del estado de salud (SoH) de baterías de ion de litio

Autor (es)

Fabio Alexander Montenegro Tarazona

Universidad de Cundinamarca

Facultad de Ingeniería

Ingeniería Electrónica

Fusagasugá, Colombia

2025

Desarrollo de un modelo de inteligencia artificial de uso específico para estimación del estado de salud (SoH) de baterías de ion de litio

Trabajo de grado presentado como requisito parcial para optar por el título de ingeniero electrónico

Autor

Fabio Alexander Montenegro Tarazona

Director:

Dr. Ing. Andrés Felipe Guerrero Guerrero

Co-director:

Ing. Cristian Arenas Trujillo

Línea de investigación:

Aprendizaje, conocimiento, tecnologías, comunicación y digitalización

Universidad de Cundinamarca

Facultad de Ingeniería

Ingeniería Electrónica

Fusagasugá, Colombia

2025

Agradecimientos

En primer lugar, quiero manifestar mi profundo agradecimiento a Dios, por otorgarme la fortaleza, sabiduría y constancia necesarias para llevar a cabo este proyecto hasta su culminación. A mi padre y madre, quienes han sido el soporte fundamental a lo largo de este camino, brindándome no solo apoyo económico, sino también el respaldo emocional y motivacional durante toda mi etapa académica. Su apoyo incondicional ha sido esencial para lograr este importante objetivo.

De manera especial, agradezco a mis directores de proyecto, quienes mostraron una disposición excepcional para atender mis dudas, resolver inquietudes técnicas y ofrecer correcciones constructivas que enriquecieron la calidad y precisión del trabajo. Su experiencia y dedicación fueron clave para superar los desafíos y mejorar los resultados alcanzados.

Finalmente, extendiendo mi gratitud a todas las personas e instituciones que, de manera directa o indirecta, contribuyeron a la realización de este trabajo de grado, el cual representa un paso decisivo en mi formación profesional y en mi aporte al campo de la ingeniería y la inteligencia artificial aplicada a tecnologías de almacenamiento.

Resumen

Las baterías de ion de litio (LIBs, por sus siglas en inglés) hacen referencia a la tecnología de almacenamiento utilizada en transporte electrificado, electrónica portátil y sistemas de almacenamiento de energía estacionarios. De acuerdo a este concepto, son de gran utilidad por su capacidad de suministrar energía a sistemas portátiles durante un tiempo limitado, sin necesidad de una fuente eléctrica externa. Estos dispositivos son controlados por sistemas de gestión de baterías (BMS) que realizan funciones como protección, gestión térmica, balanceo de celdas, carga e indicación del estado, incluyendo el estado de carga (SoC, por sus siglas en inglés) y estado de salud (SoH, por sus siglas en inglés) de las mismas.

Para estudiar el SoH y predecir sus ciclos restantes de vida útil (RUL, por sus siglas en inglés), es fundamental contar con herramientas que permitan evaluar con precisión estas dos variables. En este contexto, mediante el desarrollo de un modelo de inteligencia artificial, es posible analizar el estado de salud de una batería de iones de litio, evaluando su comportamiento durante los ciclos de carga y descarga, lo cual facilita la toma de decisiones sobre el mantenimiento o reemplazo de dichos dispositivos.

En este trabajo se desarrolla un modelo de IA entrenado con diferentes algoritmos basados tanto en redes neuronales como de toma de decisiones, especializado para la predicción de los ciclos restantes de vida útil en las baterías de ion de litio, con un error cuadrático medio (RMSE, por sus siglas en inglés) inferior a cuatro, confirmando su precisión al momento de su ejecución.

El sistema emplea algoritmos de aprendizaje automático y profundo (Machine learning y Deep learning), que analizan datos operativos con los ciclos de carga y descarga, proporcionando estimaciones más rápidas y exactas en comparación a los métodos convencionales. Esta eficiencia es valiosa en entornos donde la disponibilidad de datos históricos es limitada, puesto que el sistema no depende de ciclos completos para generar resultados.

Finalmente, los resultados del modelo Stacking presentan un RMSE de 3,8284 al emplear una combinación de cuatro algoritmos de aprendizaje de máquina: Random Forest Regressor, Bagging Regressor, MLP Regressor y Decision Tree Regressor, cada uno de ellos evaluados con respecto a su comportamiento frente a los datos de validación no utilizados en el entrenamiento. Este mencionado apilamiento proporciona robustez, precisión y desempeño, conformando una estructura eficiente para la predicción del RUL (tiempo restante de vida útil) en las baterías de ion de litio.

Abstract

Lithium-ion batteries (LIBs) refer to the storage technology used in electrified transportation, portable electronics, and stationary energy storage systems. Based on this concept, they are highly valuable due to their ability to supply power to portable systems for a limited time without requiring an external power source. These devices are managed by Battery Management Systems (BMS), which perform functions such as protection, thermal management, cell balancing, charging, and status indication, including the State of Charge (SoC) and State of Health (SoH).

To study the SoH and predict their Remaining Useful Life (RUL), it is essential to have tools that allow for accurate evaluation of these two variables. In this context, by developing an artificial intelligence model, it is possible to analyze the health status of a lithium-ion battery by evaluating its behavior during charge and discharge cycles, thereby facilitating decision-making regarding the maintenance or replacement of these devices.

This work presents an AI model trained with different algorithms based on both neural networks and decision-making techniques, specifically designed to predict the remaining useful life cycles of lithium-ion batteries. The model achieves a Root Mean Square Error (RMSE) below four, confirming its precision during execution.

The system employs Machine Learning and Deep Learning algorithms that analyze operational data from charge and discharge cycles, providing faster and more accurate estimates compared to conventional methods. This efficiency is especially valuable in environments where historical data availability is limited, as the system does not rely on complete cycles to generate results.

Finally, the results of the Stacking model show an RMSE of 3.8284 when using a combination of four machine learning algorithms: Random Forest Regressor, Bagging Regressor, MLP Regressor, and Decision Tree Regressor, each evaluated based on their performance with validation data not used during training. This stacking approach provides robustness, accuracy, and performance, forming an efficient structure for RUL (Remaining Useful Life) prediction in lithium-ion batteries.

Contenido

Resumen	5
Abstract.....	6
INTRODUCCIÓN	11
PLANTEAMIENTO DEL PROBLEMA	12
Objetivo general	13
Objetivos específicos	13
JUSTIFICACIÓN	14
ALCANCES Y LIMITACIONES	15
Alcances.....	15
Limitaciones	15
1. MARCO REFERENCIAL	17
1.1. ESTADO DEL ARTE	17
1.2. FUNDAMENTOS TEÓRICOS	22
1.2.1. Batería de ion de litio (LIB, por sus siglas en inglés)	22
1.2.2. Inteligencia Artificial (Artificial Intelligence)	23
1.2.3. Aprendizaje automático (Machine learning)	23
1.2.4. Aprendizaje profundo (Deep Learning)	23
1.2.5. Modelos supervisados	23
1.2.6. Modelos No supervisados	29
1.2.7. Error cuadrático medio (MSE)	29
1.2.8. Raíz del error cuadrático medio (RMSE)	29
2. DISEÑO METODOLÓGICO	30
3. DESARROLLO DEL MODELO DE INTELIGENCIA ARTIFICIAL	32
3.1. ANÁLISIS	32
3.1.1. Identificación de parámetros de entrada según fuentes bibliográficas	32
3.1.2. Descripción de datos	34
3.1.3. Exploración de Datos (EDA):	35
3.2. DISEÑO	42

3.2.1. Estructura de entrenamiento:	42
3.2.2. Diagrama de Flujo	43
3.3. IMPLEMENTACIÓN	45
3.3.1. Carga y preparación de datos	45
3.3.2. Calidad del Dataset	45
3.3.3. Selección de algoritmos	46
3.4. PRUEBAS DE FUNCIONAMIENTO Y/O DESEMPEÑO	48
3.4.1. Modularización del código base	48
3.4.2. Evaluación final	49
4. ANÁLISIS DE RESULTADOS Y DISCUSIÓN	53
4.1. Comparación de resultados	53
5. CONCLUSIONES Y TRABAJOS FUTUROS	54
5.1. CONCLUSIONES	54
5.2. PERSPECTIVAS DE TRABAJO FUTURO	54
BIBLIOGRAFÍA	56

Índice de figuras

Figura 1. Celda de una batería de iones de litio [12].....	23
Figura 2. Ejemplo de funcionamiento del algoritmo KNN.....	25
Figura 3. Ejemplo de funcionamiento del SVM	26
Figura 4. Ejemplos de hiperplano tanto en R2 como en R3, respectivamente [7] .	26
Figura 5. Ejemplo del funcionamiento de regresión lineal.....	27
Figura 6. Esquema del funcionamiento del algoritmo Random Forest Regressor [7].	28
Figura 7. Fases de la Metodología.....	30
Figura 8. Diagrama de pastel de los parámetros de entrada comúnmente utilizados	34
Figura 9. Histograma RUL.	36
Figura 10. Tiempo de descarga (s) vs densidad	37
Figura 11. Decremento 3.6V – 3.4V vs densidad.....	37
Figura 12. Máximo voltaje de descarga vs densidad	38
Figura 13. Mínimo voltaje de descarga vs densidad	39
Figura 14. Tiempo en 4.15 (V) vs densidad	39
Figura 15. Tiempo de corriente constante (s) vs densidad	40
Figura 16. Tiempo de carga (s) vs densidad.....	41
Figura 17. Matriz de correlación.....	42
Figura 18. Diagrama de bloques respecto a la estructura de entrenamiento del modelo.	43
Figura 19. Diagrama de flujo del funcionamiento del modelo.	44
Figura 20. Carga y preparación del Dataset en el modelo.....	45
Figura 21. Representación de anomalía en el Dataset	45
Figura 22. Curva de pérdida durante el entrenamiento.....	47
Figura 23. Estructura modulada del código base del modelo.	49
Figura 24. Predicciones vs Datos reales RandomForest	50
Figura 25. Predicciones vs Datos reales MLP	50
Figura 26. Predicciones vs Datos reales Bagging.....	51
Figura 27. Predicciones vs Datos reales DecisionTree.....	51
Figura 28. Predicciones vs Datos reales Stacking final	52

Índice de tablas

Tabla 1. Análisis del coeficiente de Pearson en los dos tipos de batería [9].	18
Tabla 2. Funciones matemáticas utilizadas por el algoritmo de regresión líneal.	27
Tabla 3. Análisis de diferentes artículos científicos.	33
Tabla 4. Análisis de los parámetros de entrada más comunes en un modelo predictivo	34
Tabla 5. Comparación entre los distintos algoritmos utilizados por el modelo.	46
Tabla 6. Resultados con respecto a la prueba con el 20% de datos aislados.	52
Tabla 7. Comparación entre los valores de RMSE Inicial/Final	53

INTRODUCCIÓN

El mercado global de las baterías de iones de litio, teniendo en cuenta un estudio publicado por **Nature** en 2021, crecerá de 30.000 millones de dólares en 2017 a 100.000 millones en 2025 [1], por lo tanto, es importante resaltar todos los beneficios de estas tecnologías de almacenamiento en el mundo, como su alta densidad de energía y larga vida útil, características que las han posicionado como estándar en aplicaciones como vehículos eléctricos, dispositivos médicos portables y sistemas de sustento o respaldo energético en zonas con suministro eléctrico intermitente.

Sin embargo, este crecimiento acelerado revela una problemática central: la degradación irreversible de los componentes electroquímicos (litio, cobalto, polímero de litio, Manganeso, fosfato de hierro, entre otros), reduce su capacidad de almacenamiento y limita su vida útil operativa en implementaciones determinadas [1]. Este problema es especialmente crítico en entornos remotos o con poco acceso de una red eléctrica, donde al no realizar un análisis preventivo, cambio del dispositivo o mantenimiento del mismo, pueden comprometer sistemas esenciales, como la infraestructura de redes de interconexión.

Con respecto al proyecto marco “Desarrollo de Micro-redes y almacenamiento (ESS) prestadores de servicios complementarios para incrementar la cobertura, eficiencia y confiabilidad del servicio en el departamento de Cundinamarca” cuyo propósito es integrar fuentes de energías convencionales y no convencionales para contribuir en la mejora de la calidad de servicio especialmente en áreas afectadas por las fallas de la red de distribución [3]. Las baterías de iones de litio son componentes críticos de las tecnologías de almacenamiento (ESS, por sus siglas en inglés), al actuar como elementos de reserva de energía que garantizan la continuidad durante las funcionalidades del sistema, sin embargo, su capacidad de respuesta está limitada por la incertidumbre en el grado de deterioro interno, lo que dificulta predecir su estado actual en escenarios operativos.

De acuerdo a este contexto, el presente trabajo aborda el desarrollo de un modelo de inteligencia artificial, de uso específico, para predecir el estado de salud (SoH, por sus siglas en inglés) y el tiempo restante de vida útil (RUL, por sus siglas en inglés) de baterías de iones de litio, mediante el análisis de patrones de degradación en datos históricos de carga/descarga y la implementación de algoritmos avanzados de aprendizaje automático, con el objetivo de lograr estimaciones del RUL con un margen de error inferior al 5%, superando así los límites de los métodos convencionales utilizados hoy en día [2].

PLANTEAMIENTO DEL PROBLEMA

Las baterías de iones de litio son componentes críticos en las microrredes rurales, donde garantizan la continuidad energética durante las interrupciones de la red principal [3], sin embargo, la degradación que sufren, manifestada principalmente en la pérdida de capacidad debido al deterioro del material activo de la batería, genera incertidumbre en la determinación del RUL especialmente cuando carece de datos históricos completos sobre ciclos de carga/descarga o cuando las baterías operan en condiciones ambientales variables (temperaturas extremas, humedad). Esta problemática se agrava por la falta de métodos eficientes de predicción enfocados en el estudio del estado de salud de dicha tecnología de almacenamiento (SoH) en tiempos de procesamiento reducidos, especialmente en entornos remotos con limitado acceso a equipos de diagnóstico especializado.

Según estudios recientes, la degradación de las LIBs está directamente relacionada con parámetros como voltaje, corriente y temperatura, los cuales varían significativamente durante los ciclos de carga y descarga [4]. En condiciones reales, un error de estimación del SoH superior al 10% puede llevar a fallos críticos en sistemas dependientes de baterías, como los vehículos eléctricos o microrredes [4]. Además, métodos tradicionales de estimación basados en modelos físicos requieren tiempos prolongados de prueba (hasta 400 ciclos) y equipos especializados, lo que los hace inviables para aplicaciones remotas.

Con respecto a la incertidumbre en la estimación precisa del estado de salud (SoH) real de las baterías de iones de litio, esta representa un desafío importante para su gestión eficiente [4]. Esta problemática se origina principalmente en la naturaleza no lineal y compleja del proceso de degradación, que en sus etapas iniciales presenta cambios mínimos en la capacidad, dificultando la detección temprana de deterioro. Adicional, la falta de datos históricos completos y la variabilidad en las condiciones operativas, como fluctuaciones de temperatura y ciclos de carga/descargas irregulares [5], incrementan la dificultad para obtener predicciones fiables. Como consecuencia, los métodos tradicionales de estimación suelen ser poco precisos, generando incertidumbre en la predicción de la vida útil remanente (RUL) y afectando la toma de decisiones sobre mantenimiento y reemplazo.

Ante esta problemática, surge la pregunta problema ¿Cómo puede implementarse una solución viable por medio del aprendizaje autónomo, que no requiera equipos especializados, para predecir eficientemente la vida útil remanente de las baterías en las microrredes, utilizando únicamente tiempos de comportamiento de carga/descarga de las LIB?

OBJETIVOS DEL ESTUDIO

Objetivo general

Desarrollar un modelo de inteligencia artificial de uso específico, basado en redes neuronales, que permita la estimación del estado de salud (SoH) de baterías de ion de litio.

Objetivos específicos

1. Realizar una revisión bibliográfica que permita identificar los parámetros de entrada de un modelo de aprendizaje de máquina aplicado a la estimación del estado de salud de la batería de ion de litio.
2. Diseñar un modelo de aprendizaje automático adaptable a algoritmos de entrenamiento en función de la distribución de datos experimentales de baterías de ion de litio.
3. Evaluar los resultados de estimación obtenidos por el modelo de aprendizaje automático a partir de escenarios de prueba con datos experimentales.

JUSTIFICACIÓN

La implementación de un modelo eficiente para evaluar el estado de salud (SoH) y predecir la vida útil remanente (RUL) de baterías de iones de litio (una tecnología ampliamente usada en aplicaciones como vehículos eléctricos y sistemas de respaldo energético) ofrece una solución estratégica para dos problemas clave en microrredes remotas: los altos costos de mantenimiento y la dificultad para gestionar y planificar el almacenamiento energético de manera precisa. Este modelo proporciona una evaluación precisa del estado real de las baterías, evitando reemplazos preventivos innecesarios y costosos que se basan en estimaciones imprecisas.

En regiones remotas de Cundinamarca, donde el acceso a energía estable es limitado [6], anticipar fallas logra prevenir interrupciones prolongadas que afectan servicios esenciales, como la refrigeración de medicamentos en centros de salud, la continuidad educativa en escuelas y la productividad de pequeños negocios. La capacidad de prever y gestionar estos problemas contribuye directamente a mejorar la calidad de vida y la resiliencia de estas comunidades.

Este beneficio tiene aún más relevancia ante la creciente adopción de baterías de ion de litio en la tecnología moderna, donde se proyecta que el mercado global de estos elementos superará los 100 mil millones de dólares en 2025 [1], Este crecimiento destaca la importancia de estudiar y comprender el comportamiento de las baterías en entornos críticos, especialmente en zonas donde su correcto funcionamiento es vital para el suministro energético y el bienestar social.

Desde la perspectiva académica, el desarrollo y aplicación de modelos predictivos en el área de baterías de ion-litio representa un avance significativo en la Ingeniería electrónica y eléctrica. Esto proporciona a universidades y centros de investigación contar con herramientas de diagnóstico avanzadas, facilitando la formación de profesionales capacitados en tecnologías de almacenamiento energético y gestión inteligente de recursos. Además, la predicción eficaz del SoH y RUL aporta a la validación y mejora de metodologías de investigación, como el uso de espectroscopía de impedancia electroquímica y técnicas de inteligencia artificial para el análisis de datos de baterías [7]. Estudios recientes destacan que la predicción temprana y precisa de la vida útil, incluso a partir de los primeros ciclos de carga, abre nuevas oportunidades en la producción, uso y optimización de baterías, permitiendo clasificar rápidamente aquellas tecnologías según su vida útil esperada [8].

Por esta razón, el gran aporte de este proyecto en la ingeniería radica en proporcionar una herramienta clave para la predicción eficiente del estado de las baterías de ion de litio, lo que es fundamental en proyectos donde la portabilidad y la gestión eficiente de la energía son prioritarias.

ALCANCES Y LIMITACIONES

Alcances

El modelo desarrollado en este proyecto tiene como objetivo principal estimar el Estado de Salud (SoH) y la Vida Útil Restante (RUL) de baterías de iones de litio del tipo NMC-LCO 18650 con capacidad de 2.8 Ah, utilizando como datos de entrada únicamente parámetros temporales y de voltaje registrados durante los ciclos de carga y descarga. Es fundamental dejar claro que el modelo no realiza mediciones físicas directas de la batería, sino que opera exclusivamente con información previamente recolectada en condiciones de laboratorio, donde variables como el tiempo que tarda la batería en alcanzar ciertos voltajes críticos (por ejemplo, el tiempo que permanece en 4.15 V durante la carga o el tiempo que tarda en descender de 3.6 V a 3.4 V durante la descarga) pueden medirse con alta precisión mediante equipos especializados.

Las entradas del modelo consisten en registros estructurados que incluyen: los tiempos totales de carga y descarga, los intervalos de tiempo asociados a voltajes específicos (como los mencionados anteriormente), así como los valores máximos y mínimos de voltaje registrados en cada ciclo.

Adicionalmente, como salidas del modelo se identifican dos parámetros: en primer lugar, el Estado de Salud (SoH), expresado como un porcentaje que indica la capacidad remanente de la batería en comparación con su capacidad inicial; y en segundo lugar, la Vida Útil Restante (RUL), que representa el número estimado de ciclos de carga y descarga que la batería puede completar antes de que su SoH este por debajo del 80%, umbral comúnmente aceptado como fin de su vida útil en aplicaciones críticas.

Limitaciones

La efectividad del modelo está sujeta a la calidad y consistencia de los datos temporales recopilados, especialmente de las mediciones precisas de los tiempos de carga y descarga. En este sentido, el modelo ha demostrado un error cuadrático medio (RMSE) inferior al 4%, lo que indica un buen nivel de precisión, pero no garantiza una efectividad del 100%. Pequeñas variaciones, anomalías o imprecisiones en la instrumentación pueden afectar los resultados, puesto que el modelo es sensible a fluctuaciones en los datos de entrada.

Además, el modelo adaptativo se basa en la suposición de que los patrones temporales observados durante la fase de entrenamiento son representativos del comportamiento futuro de la batería. Sin embargo, esta premisa puede no cumplirse en situaciones donde la batería sufra una degradación acelerada o esté sometida a

condiciones de operación atípicas o extremas, como ciclos de carga irregulares, temperaturas muy variables o estrés mecánico. En estos casos, el comportamiento real de la batería podría desviarse considerablemente de los patrones aprendidos, lo que reduce la precisión del modelo.

Otra limitación importante es que el modelo no contempla posibles fluctuaciones en los tiempos de carga y descarga causadas por factores externos, como cambios bruscos de temperatura ambiental. Este factor puede introducir variabilidad en el rendimiento que no está reflejada en los datos de entrenamiento, afectando la capacidad del modelo para generalizar correctamente.

1. MARCO REFERENCIAL

1.1. ESTADO DEL ARTE

En esta sección se realizará una revisión y análisis de investigaciones previas enfocadas en el uso de algoritmos de aprendizaje automático para la predicción del estado de salud (SoH) y vida útil remanente (RUL) en baterías de iones de litio, con especial atención a métodos que integren datos operativos en tiempo real (voltaje, corriente, temperatura) y su procesamiento mediante arquitecturas eficientes adaptadas a entornos con recursos limitados, como microrredes rurales.

Tomando como referencia el proceso de estimación del SoH [9], se destacan varios aspectos clave. Durante la fase de relajación de la batería, se recopilan datos de voltaje que se convierten en indicadores de salud, los cuales caracterizan la degradación de la capacidad del componente. Posteriormente, se realiza el análisis de correlación utilizando el coeficiente de Pearson para identificar cuáles de estos indicadores están más correlacionados con la degradación de la capacidad de la batería.

Una vez hecho esto, los indicadores seleccionados se utilizan como características de entrada para el modelo SVR (Support Vector Regression), el cual consiste en encontrar un hiperplano en un espacio de alta dimensión que minimice el error entre los valores predichos y los reales [9], por medio de una función de pérdida, lo que significa que solo penaliza las predicciones que se desvían de los valores reales en más de un margen determinado. Esto permite que el modelo sea robusto a pequeñas variaciones en los datos.

En la **Tabla 1.** se observa el análisis de correlación de Pearson entre las características propuestas (indicadores de salud) y el Estado de Salud (SoH) de las baterías. En esta tabla, se presentan los coeficientes de correlación con respecto a los resultados obtenidos del proyecto referenciado [9], para las clases de baterías tanto de NCA como NCM.

Tabla 1. Análisis del coeficiente de Pearson en los dos tipos de batería [9].

Battery type	Cycling condition	HIs				
		Δv_{std}	Δv_{mean}	Δv_{max}	Δv_{min}	Δv_{diff}
NCA	CY25 – 0.5/1	-0.941	-0.865	-0.953	-0.800	-0.939
	CY35 – 0.5/1	-0.986	-0.992	-0.989	-0.937	-0.986
	CY45 – 0.5/1	-0.947	-0.993	-0.962	-0.875	-0.949
NCM	CY25 – 0.5/1	-0.914	-0.910	-0.915	-0.761	-0.912
	CY35 – 0.5/1	-0.949	-0.965	-0.952	-0.839	-0.950
	CY45 – 0.5/1	-0.956	-0.975	-0.960	-0.903	-0.957

En la **Tabla 1**, las variables V_{std} , V_{mean} , V_{max} , V_{min} y V_{diff} representan características extraídas del análisis de la fase de relajación del voltaje de la batería. Estas métricas se derivan específicamente de la secuencia de caída de voltaje durante dicha fase y se utilizan para caracterizar el estado de salud (SoH) de la batería.

- **V_{std} (desviación estándar del voltaje):** Mide la variabilidad o dispersión de los valores de voltaje durante la secuencia de caída en la fase de relajación. Indica cómo fluctúa el voltaje, lo que refleja cambios en el comportamiento de la batería.
- **V_{mean} (valor medio del voltaje):** Es el promedio de todos los valores de voltaje en la secuencia de caída durante la relajación. Representa el valor central o típico del voltaje en ese periodo.
- **V_{max} (valor máximo del voltaje):** Es el valor más alto registrado en la secuencia de caída del voltaje durante la fase de relajación.
- **V_{min} (valor mínimo del voltaje):** Es el valor más bajo registrado en la secuencia de caída del voltaje durante la fase de relajación.
- **V_{diff} (diferencia extrema del voltaje):** Calcula la diferencia entre el voltaje máximo y mínimo en la secuencia, mostrando la amplitud total de la caída de voltaje durante la relajación.

Los resultados indican que las características, excepto el valor mínimo, mantienen una fuerte correlación con el SoH, lo que sugiere que son útiles para caracterizar la degradación de la batería, también se menciona que la temperatura tiene un impacto significativo en la degradación del SoH, lo que resalta la importancia de considerar las condiciones operativas al analizar la salud de las baterías.

Finalmente se destacan algunas conclusiones de la práctica desarrollada, como la eficiencia del modelo propuesto, puesto que el modelo BO-SVR demuestra alta precisión en la estimación del SoH de las baterías de iones de litio. Los resultados muestran que las curvas de estimación del SoH generadas por el modelo se alinean

estrechamente con las curvas reales del estado de salud, lo que indica que el modelo es efectivo en diversas condiciones operativas.

Mencionando el desarrollo del algoritmo KNN (K-Nearest Neighbors) “Vecinos más cercanos”, el cual clasifica el número de ciclos utilizados de la batería [10], mediante la comparación de valores de un hardware de la nube con el conjunto de datos existentes a partir del número de ciclos utilizados.

Específicamente, extraen un indicador de salud (HI, por sus siglas en inglés) para medir el estado de salud de la batería (SoH) y luego construyen un modelo de aprendizaje automático que mapea la relación entre ellos para proporcionar una predicción exacta y así determinar el estado de salud de la batería de ion de litio [10].

El algoritmo KNN es ampliamente utilizado en aplicaciones de reconocimiento de patrones, sistemas de recomendación y análisis de datos [10], puesto que selecciona los valores más cercanos de un conjunto de entrenamiento y toma decisiones basadas en la mayoría de las clases de los mencionados vecinos cercanos (con respecto a la clasificación) o promedia los valores (con respecto a la regresión).

Para la creación del KNN se basaron en el lenguaje de alto nivel llamado Python. Dicho algoritmo entrena los datos y estima los puntos de prueba mediante un promedio ponderado de sus vecinos. Los gráficos resultantes del KNN revelan que la vida útil de una batería de ion de litio se mantiene al 100% hasta alcanzar 400 ciclos de carga y descarga.

Adicionalmente, se evidencia un proyecto basado en aprendizaje de máquina, el cual se centra en la evaluación del Estado de Salud (SoH) de las baterías de iones de litio [4]. Donde se establece un enfoque que incluye la recopilación de datos históricos de capacidad de las baterías, que se utiliza como respuesta para el modelo.

Los datos de entrada para el modelo de predicción, se fundamentan en características clave que reflejan su comportamiento y nivel de degradación a lo largo del ciclo de vida. En particular, el estudio se centró en variables con una relación con la capacidad y el rendimiento de la batería, tales como la tensión terminal, la resistencia interna, la temperatura de operación, el número de ciclos de carga y descarga, así como mediciones electroquímicas como la espectroscopía de impedancia electroquímica (EIS).

Estas características fueron seleccionadas mediante un análisis exhaustivo, identificándolas como los indicadores más representativos del deterioro progresivo de la batería. El uso de estos datos permite al modelo capturar la relación entre dichos parámetros y la pérdida de capacidad, facilitando predicciones eficientes del SoH en distintas etapas de uso [4]. Asimismo, la inclusión de datos históricos de capacidad y variables operativas contribuye a mejorar la precisión y robustez del modelo en la estimación del estado de salud de la batería.

Estos modelos consistieron en la evaluación de varios algoritmos de regresión, incluyendo Máquinas de Soporte Vectorial (SVM), Árboles de Regresión, Regresión por Aproximación de Núcleo y Modelos de Regresión Lineal [4]. Se utilizó validación cruzada de cinco pliegues para medir la calidad de los modelos, generando así métricas como el Error Cuadrático Medio y el Error Absoluto Medio.

En el desarrollo del proyecto de monitoreo del ciclo de carga y descarga de baterías de ion de litio, se partió del análisis de parámetros clave que permiten evaluar el comportamiento de las baterías durante su operación [5]. Entre los principales parámetros considerados se encuentran la tensión (voltaje) en distintas etapas del proceso, así como la corriente durante la carga y descarga. De manera específica, se registraron valores de voltaje en diferentes muestras para analizar su variación, capacidad y estado de salud (SoH) de las baterías.

La adquisición de estos datos se realizó mediante herramientas de monitoreo integradas en un Sistema de Gestión de Baterías (BMS), que facilitó la captura y análisis en tiempo real durante los ciclos de carga y descarga. Esta recopilación precisa en diversos puntos permitió identificar características fundamentales como la capacidad efectiva, las fluctuaciones en los valores máximos y mínimos de voltaje, y la estabilidad operativa del sistema, como se evidencia en la **Tabla 2**.

Tabla 2. Tabla de resultados de carga y descarga de las baterías. [5]

Parámetro	Valor promedio/ Valor Obtenido (Aproximado)	Descripción
Voltaje mínimo (antes de carga)	3.52V	Valor medio del voltaje en estado de descarga antes del proceso de carga
Voltaje después de carga	4.16V	Voltaje alcanzado tras finalizar el proceso de carga
Voltaje máximo durante la carga	3.75V	Valor medio del voltaje máximo durante la descarga
Voltaje después de descarga	2.99V	Voltaje final tras el proceso de descarga

Diferencia de voltaje de carga - descarga	0.64V (Carga) 0.76V (Descarga)	Variaciones en los voltajes que indican correcta transferencia y consumo de energía
Capacidad de la batería (Ah)	Entre 0.976 y 1.122 Ah	Capacidad efectiva de cada celda monitoreada
Tiempo de proceso de carga	Entre 3 y 4 horas	Tiempo estimado para cargar completamente las baterías
Tiempo de proceso de descarga	Variable, entre 2 y 4 horas	Tiempo de descarga registrado según los datos recopilados.

Esta **Tabla 2.** refleja cómo los parámetros monitorizados confirman que las baterías mantienen un buen rendimiento durante los ciclos, permitiendo determinar el SoH y la eficiencia del sistema de gestión de baterías. Además, evidencian que las variaciones de voltaje y capacidad están dentro de los rangos aceptables, asegurando un funcionamiento confiable.

Como resultado, el estudio proporcionó una comprensión detallada del comportamiento de las baterías bajo distintas condiciones de uso, estableciendo patrones de voltaje y capacidad que sirvieron de base para optimizar su control y gestión. Además, se confirmó que las variaciones en los parámetros monitoreados están estrechamente vinculadas con la salud y el rendimiento de las baterías, aportando información valiosa para el diseño y mejora de sistemas de gestión en aplicaciones prácticas, tales como vehículos eléctricos y sistemas de almacenamiento energético.

En otro estudio, de igual forma, se implementó un método basado en aprendizaje automático para evaluar la vida útil de baterías de iones de litio, utilizando específicamente el algoritmo K-Nearest Neighbors (KNN) para estimar el estado de salud (SoH) de las mismas [11]. Para ello, se consideraron variables esenciales como la corriente de carga y descarga, el voltaje y la temperatura, que son obtenidas a través de los sistemas de gestión de baterías (BMS) y enviadas a la nube para su análisis y procesamiento. Estos parámetros son fundamentales para describir el comportamiento y el envejecimiento de las baterías.

El objetivo central fue desarrollar un método eficiente y preciso que permita predecir en tiempo real la vida útil remanente de las baterías, con el fin de optimizar su gestión y mantenimiento en aplicaciones como vehículos eléctricos y sistemas de almacenamiento energético [11]. Los resultados obtenidos mostraron que, incluso utilizando datos parciales de los ciclos de carga y descarga, el algoritmo KNN

presentó una rápida capacidad de aprendizaje y buena capacidad de generalización, logrando estimaciones del SoH con un bajo margen de error. Así, este trabajo ofreció una solución práctica para el monitoreo y extensión de la vida útil de las baterías mediante la integración de técnicas avanzadas de análisis de datos y aprendizaje automático.

Finalmente, el KNN logró estimar la vida útil de la batería con un porcentaje de error menor al 10% [11], con un tiempo de ejecución de menos de 10 minutos.

1.2. FUNDAMENTOS TEÓRICOS

1.2.1. Batería de ion de litio (LIB, por sus siglas en inglés)

Estos componentes son reconocidos por tener un considerable valor de energía y eficiencia energética, puesto que, al no estar utilizándolas, retienen su carga por un periodo de tiempo mayor en comparación con otros tipos de baterías [12]. Otro aspecto de mayor relevancia es su velocidad de carga, lo cual es esencial para aplicaciones donde el tiempo sea un factor crítico. Para comprender el funcionamiento de una batería de ion de litio, es necesario examinar su estructura interna la cual se explicará a continuación.

1.2.1.1. Estructura de una batería de ion de litio

La estructura básica de una batería de iones de litio consta de un electrodo negativo, conocido como ánodo, que suele estar hecho de carbono. Incluye un electrodo positivo, llamado cátodo, que está compuesto por óxido de litio, de igual manera posee un separador que evita el corto circuito de los terminales, pero es penetrable por los iones de litio que flotan libremente entre los terminales en una solución llamada el electrolito [12]. Estos componentes internos de la batería de ion de litio se muestran en la **Figura 1**.

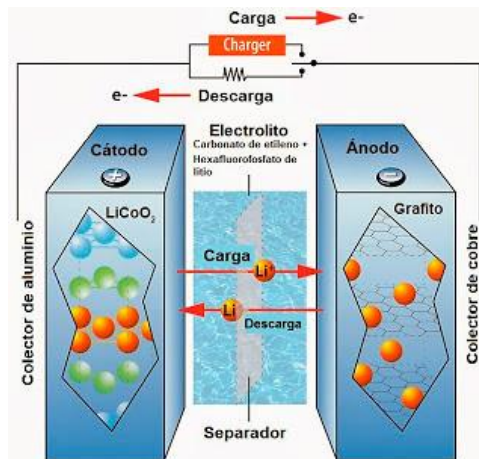


Figura 1. Celda de una batería de iones de litio [12]

1.2.2. Inteligencia Artificial (Artificial Intelligence)

Se presenta como un campo amplio que abarca técnicas que permiten a los ordenadores imitar o superar el comportamiento humano en tareas complejas. La IA incluye diferentes enfoques, como sistemas basados en reglas, representación del conocimiento, lógica, algoritmos genéticos, y también aprendizaje automático (ML), que es fundamental en la generación de sistemas inteligentes modernos [13].

1.2.3. Aprendizaje automático (Machine learning)

El machine learning se define como la capacidad de los sistemas para aprender de datos específicos de problemas, automatizando así la construcción de modelos analíticos y permitiendo que las computadoras descubran patrones y relaciones complejas sin programación explícita [13].

1.2.4. Aprendizaje profundo (Deep Learning)

El deep learning, en particular, se describe como un concepto dentro del machine learning que se basa en redes neuronales artificiales profundas, con múltiples capas, que generalmente superan a los modelos de machine learning más simples en muchas tareas, especialmente con datos de alta dimensión como imágenes, texto, vídeo, voz y audio [13].

1.2.5. Modelos supervisados

Este es el enfoque más comúnmente utilizado en el campo del aprendizaje automático. En este tipo de modelo, los datos empleados durante el entrenamiento consisten en pares de entradas y salidas previamente etiquetadas (X como entrada,

Y como salida). Dentro de esta categoría de aprendizaje supervisado se incluyen dos subtipos principales: clasificación y regresión.

En este método, la variable Y, también llamada salida o target, corresponde a una etiqueta conocida como clase. La clasificación, también denominada reconocimiento de patrones, consiste en que el modelo predice la clase a partir de un conjunto de datos de entrada x, asignando cada dato a una categoría específica. [7]. Existen diversos tipos de clasificadores, donde los más utilizados son: **k-NN (k-Nearest Neighbor)** y **SVM (Support Vector Machine)**.

1.2.5.1. K-Nearest Neighbors

El clasificador KNN parte de la premisa de que los nuevos datos comparten características y categorías similares a las de sus vecinos más próximos. Este método se basa en medir la distancia entre puntos dentro de un espacio gráfico, evaluando qué tan cerca está un nuevo punto respecto a los existentes. El algoritmo asigna al nuevo punto la etiqueta y atributos del vecino más cercano [7].

El funcionamiento del algoritmo es el siguiente, expresado en la **Figura 2**:

- Se determina el número K de vecinos con los cuales se comparará el nuevo punto.
- Se calcula la distancia euclidiana entre el nuevo punto y cada uno de los vecinos.
- Se seleccionan los K vecinos más próximos basándose en estas distancias.
- Se analizan las etiquetas de estos vecinos seleccionados.
- Finalmente, se asigna al nuevo punto la etiqueta que predomine entre sus vecinos más cercanos.

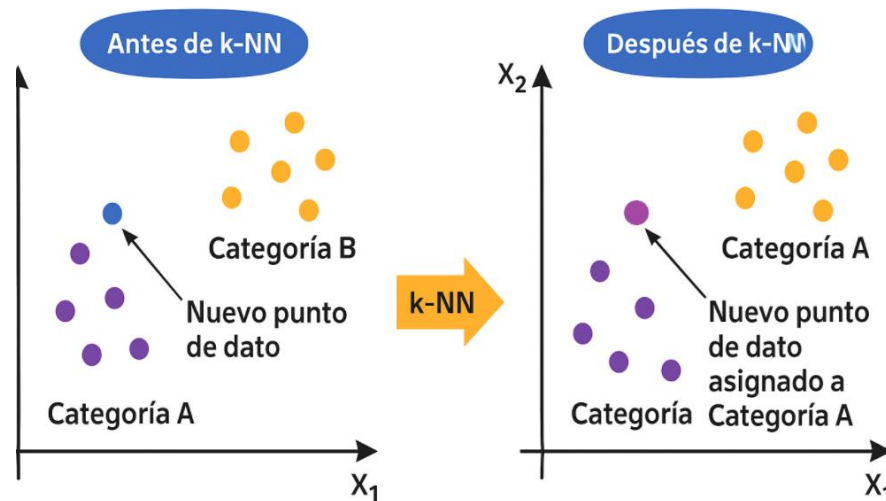


Figura 2. Ejemplo de funcionamiento del algoritmo KNN.

Al momento de elegir el valor de K, es importante considerar lo siguiente:

- No existe un método exacto para determinar el valor óptimo de K; generalmente se utiliza un proceso de prueba y error para identificar el valor que ofrece el mejor desempeño para el modelo en cuestión.
- Los valores de K pueden estar sujetos a fluctuaciones o ruido, lo que puede afectar la estabilidad del modelo.

1.2.5.2. Support Vector Machine

El propósito del algoritmo SVM (Support Vector Machine) es identificar un hiperplano en un espacio de múltiples dimensiones que permita separar y clasificar los datos en diferentes categorías [7].

Por otro lado, la regresión es una técnica estadística utilizada para analizar y modelar la relación entre una variable dependiente (y) y una o más variables independientes (x). Este enfoque de aprendizaje automático se emplea para predecir valores numéricos continuos o reales [7].

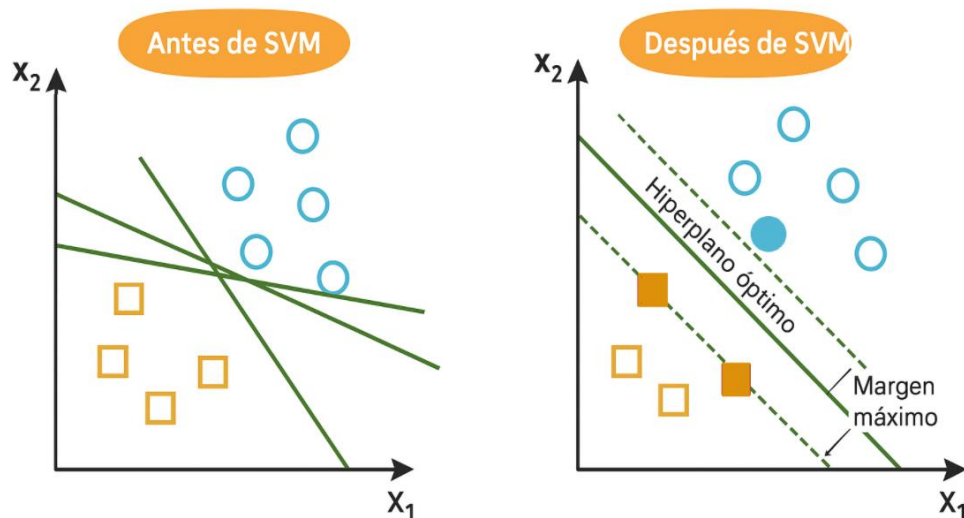


Figura 3. Ejemplo de funcionamiento del SVM

El propósito de este algoritmo es identificar un plano que maximice el margen, es decir, la mayor distancia posible entre los puntos pertenecientes a dos clases diferentes. Para lograr la separación de estas clases, se pueden utilizar varios hiperplanos. Estos hiperplanos actúan como fronteras de decisión que permiten distinguir entre los distintos grupos de datos. La dimensionalidad del hiperplano depende de la cantidad de variables de entrada o características; por ejemplo, en un espacio bidimensional el hiperplano es una línea, mientras que en un espacio tridimensional corresponde a un plano de dos dimensiones [7], como se evidencia en la **Figura 4**.

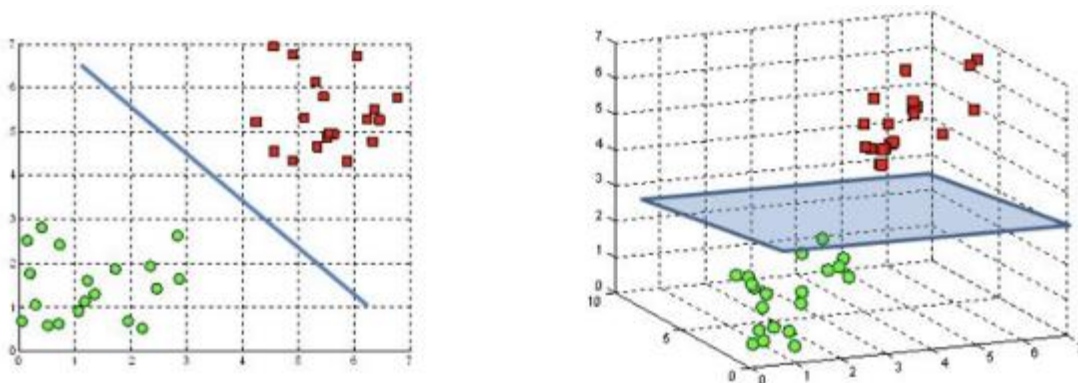


Figura 4. Ejemplos de hiperplano tanto en R2 como en R3, respectivamente [7]

Para ajustar la posición y orientación del hiperplano, el algoritmo utiliza los vectores soporte, que son los puntos de datos más cercanos al hiperplano. Estos vectores son fundamentales para maximizar el margen del clasificador. Cualquier modificación o eliminación de estos vectores soporte puede alterar la ubicación y la inclinación del hiperplano, ya que el SVM se basa en ellos para definir la frontera de separación entre clases.

1.2.5.3. Regresión lineal

Se trata de una regresión en la que existe una relación lineal entre las variables involucradas [7], donde su función principal es determinar los valores óptimos de los parámetros lineales que permitan ajustar la línea de regresión minimizando el error. Esto se puede evidenciar en la **Figura 5**.

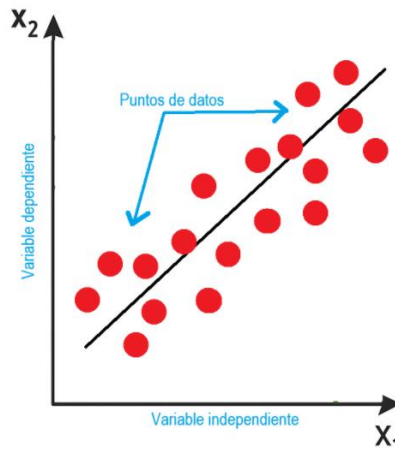


Figura 5. Ejemplo del funcionamiento de regresión lineal.

De acuerdo al funcionamiento de dicho algoritmo, es importante resaltar las funciones matemáticas utilizadas en él. Estas se observan en la Tabla

Tabla 2. Funciones matemáticas utilizadas por el algoritmo de regresión lineal.

Función Ecuación lineal	Función MSE (Mean Squared Error)
$y = a_0 + a_1 * x$	$MSE = \frac{1}{n} \sum_{i=1}^n (pred_i - y_i)^2$
Esta función se utiliza para encontrar relaciones lineales entre los datos de prueba.	Función que determina el error entre los datos que predice el algoritmo.

1.2.5.4. Random Forest Regressor

Este algoritmo de aprendizaje supervisado se fundamenta en el aprendizaje conjunto, que combina las predicciones de varios modelos de aprendizaje automático para obtener una predicción más precisa y robusta [7].

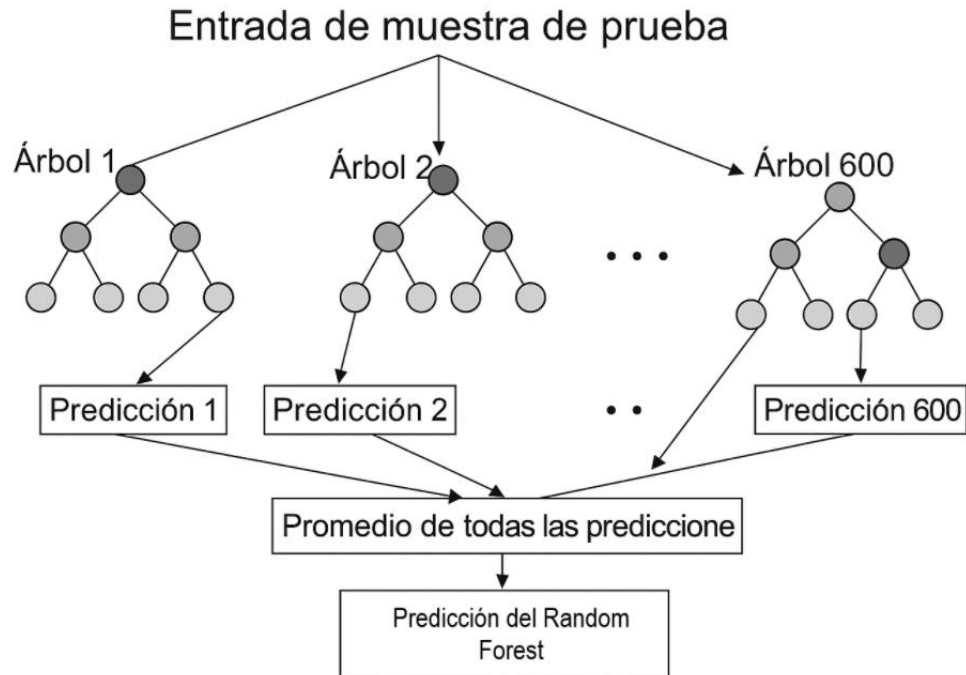


Figura 6. Esquema del funcionamiento del algoritmo Random Forest Regressor [7].

En la **Figura 6**, se ilustra el funcionamiento paso a paso del algoritmo mencionado anteriormente, cuyo proceso consta de varias etapas para garantizar su efectividad:

- Seleccionar aleatoriamente un conjunto de puntos para entrenar el modelo.
- Construir un árbol de decisiones con esos datos.
- Repetir este procedimiento para crear un número N de árboles, con el fin de obtener diversas predicciones.
- Cada árbol de decisiones genera una predicción individual, y una vez obtenidas todas las predicciones, se calcula su promedio. Este valor promedio se asigna como la predicción final correspondiente a la entrada o entradas X proporcionadas al algoritmo.

1.2.6. Modelos No supervisados

En estos modelos, los datos de entrada no cuentan con resultados o etiquetas predefinidas. En lugar de buscar una relación directa entre datos de entrada y salida, el objetivo es descubrir patrones, estructuras o agrupaciones dentro de la información. Este tipo de aprendizaje refleja más fielmente cómo las personas asimilan y organizan la información de manera natural [14].

1.2.7. Error cuadrático medio (MSE)

Denotado por sus siglas en inglés Mean Squared, es la media de los cuadrados de los errores o desviaciones, es decir, la diferencia entre el valor estimado por el algoritmo y el valor real [15]. Un valor más bajo de MSE indica un mejor rendimiento del algoritmo.

1.2.8. Raíz del error cuadrático medio (RMSE)

Denotado por sus siglas en inglés Root Mean Squared Error, es simplemente la raíz cuadrada de MSE. Esta métrica también mide el promedio de las desviaciones entre las predicciones del algoritmo y los datos reales, pero al tomar la raíz cuadrada, RMSE devuelve la desviación en las mismas unidades de los valores originales [15]. De igual forma al MSE, un valor más bajo de RMSE sugiere un mejor rendimiento del algoritmo.

2. DISEÑO METODOLÓGICO

La metodología se distribuye en tres etapas consecutivas las cuales permiten el desarrollo de cada objetivo específico planteado en el proyecto. Esto se evidencia en la **Figura 7**.



Figura 7. Fases de la Metodología.

Fase 1. En esta sección del proyecto se realizó una revisión bibliográfica respecto a los distintos parámetros de entrada de un modelo de aprendizaje de máquina, por medio de la consulta de bases de datos especializadas tanto en el estado de salud (SoH) de las baterías de ion de litio como en los modelos de estimación realizados en torno al Machine learning. El objetivo de esta sección fue la identificación de distintos datos de entrada utilizados en diferentes algoritmos de estimación. Como producto de esta fase se obtuvo un documento de estudio y análisis del estado del arte de cada valor de entrada utilizado en diferentes procesos de aprendizaje de máquina.

Fase 2. De acuerdo con la investigación y el análisis previo sobre los distintos parámetros de entrada utilizados en cada algoritmo creado para la estimación del SoH, se diseñó un modelo de aprendizaje automático validado en distintas etapas de comportamiento, mediante los indicadores de incertidumbre MSE (Mean Squared Error), RMSE (Root Mean Squared Error) y tiempo de preparación de cada uno de ellos. Al finalizar esta fase, se obtuvo un modelo stacking implementado con los algoritmos que presentaron un menor RMSE de estimación en función de la

distribución de datos experimentales de baterías de ion de litio, como en este caso: RandomForestRegressor, MLPRegressor, BaggingRegressor y DecisionTreeRegressor, ordenados respectivamente del mejor al menos preciso. Cabe aclarar que se realiza un proceso de modulación del código base, con respecto a la mejora en organización del modelo y facilitar su comprensión.

Fase 3. Finalmente, se verificó el correcto funcionamiento del modelo diseñado a través de pruebas de comparación y validación, utilizando baterías en tres diferentes estados de deterioro: uno totalmente nuevo, el segundo con un porcentaje de uso del 50% y un tercero completamente degradado. Cada uno se sometió a mediciones y análisis de su estado de salud por medio de equipos de diagnóstico del SoH en baterías. Este proceso tuvo el propósito de contrastar resultados predichos por el modelo con los datos experimentales previamente descritos. El producto final de la Fase 3 fue un modelo de inteligencia artificial de uso específico, operando con los algoritmos más eficientes en la estimación de baterías de ion de litio.

3.DESARROLLO DEL MODELO DE INTELIGENCIA ARTIFICIAL

En este capítulo se aborda el desarrollo completo del objetivo general del proyecto, estructurado en diversas etapas. Cada una de estas fases será explicada detalladamente, implementada y evaluada para concluir sobre su funcionamiento y contribución al logro de cada objetivo específico planteado.

3.1. ANÁLISIS

El análisis se llevará a cabo mediante la revisión y estudio de proyectos y artículos relacionados con la temática, con el fin de identificar y destacar las variables más relevantes en baterías de iones de litio. Una vez seleccionados los parámetros de entrada eficientes para el modelo, se procederá a un estudio detallado de cada uno de ellos, evaluando su comportamiento en el desempeño de las baterías.

3.1.1. Identificación de parámetros de entrada según fuentes bibliográficas

La revisión de los antecedentes muestra que la vida útil de estas baterías está vinculada a la gestión de sus ciclos, incluyendo factores como profundidad de descarga, la corriente de carga, el voltaje y la temperatura de cada fase.

Diversos estudios coinciden en que la duración y calidad de cada ciclo de carga/descarga impactan significativamente en la degradación y el estado de salud (SoH) de la batería. Por ejemplo, se ha evidenciado que estas tecnologías de almacenamiento pueden soportar entre 300 y más de 6,000 ciclos completos, dependiendo del tipo de batería y las condiciones de operación [16].

En cuanto a los parámetros monitoreados durante los ciclos, los valores de voltaje, corriente y tiempos de carga/descarga, son los más utilizados para caracterizar el comportamiento y la degradación de la batería. Estos datos, obtenidos a través de sistemas de gestión de baterías (BMS), logran medir datos que permiten reproducir o recrear la dinámica, detectar patrones relacionados con la pérdida de capacidad y eficiencia operativa [9]. Además, indicadores derivados del análisis de la fase de relajación de voltaje, como la desviación estándar, valores máximos y mínimos, y la diferencia extrema, se han mostrado altamente correlacionadas con el SoH, aportando valiosa información de entrenamiento para modelos predictivos [17].

La **Tabla 3.** resume los parámetros de entrada empleados por los autores analizados en sus modelos de inteligencia artificial y circuitos equivalentes. Esta comparación permite identificar estadísticas, cuyos resultados se visualizan en el diagrama de pastel de la **Figura 8.**

Tabla 3. Análisis de diferentes artículos científicos.

Artículo científico analizado	Año de publicación	Parámetro de entrada
Joint Estimation of SOC and SOE for Energy Storage Lithium-ion Batteries Based on KNN Algorithm	2024	Datos históricos, temperatura de operación.
Li-Ion Battery State of Health Assessment Using Machine Learning	2023	Datos históricos, corriente interna.
Li-ion Battery State of Health Estimation Based on an Improved Single Particle Model	2017	Resistencia interna, temperatura de operación.
Lithium-Ion Battery Life Estimation Using Machine Learning Algorithm	2022	Datos históricos, Voltajes en distintos puntos de operación.
Li-ion Battery SOH Estimation Based on BO-SVR Model	2023	Datos históricos.
Monitoring Process of Lithium-Ion during Charging and Discharging to Battery Management System Performance	2023	Datos históricos, corriente y voltaje en distintas fases de operación.
SOH analysis of Li-ion battery based on ECM parameters and broadband impedance measurements	2020	Resistencia interna, temperatura de operación.
Stock Prediction by Integrating Sentiment Scores of Financial News and MLP-Regressor: A Machine Learning Approach	2022	Datos históricos, Corriente, Voltaje de carga y descarga.

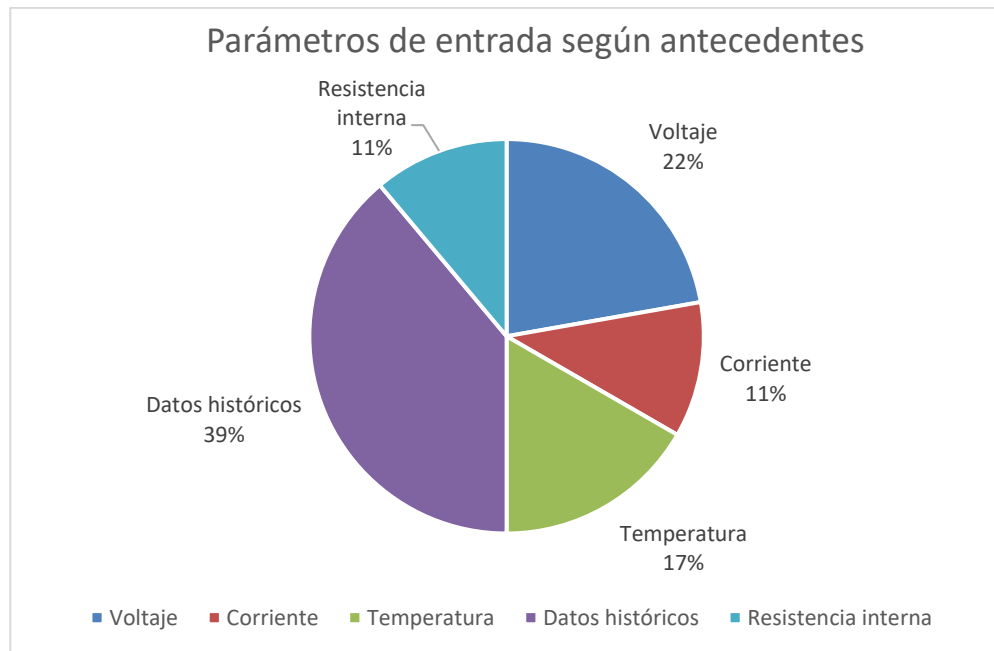


Figura 8. Diagrama de pastel de los parámetros de entrada comúnmente utilizados

En resumen, la **Tabla 4** se presenta el análisis de los parámetros evaluados, indicando para cada uno, su respectiva descripción.

Tabla 4. Análisis de los parámetros de entrada más comunes en un modelo predictivo

Parámetro de entrada	Porcentaje de uso	Descripción
Voltaje	22%	Voltaje en distintas fases del ciclo, incluyendo el de relajación y carga/descarga
Corriente	11%	Comportamiento de la corriente en el proceso de carga/descarga
Temperatura	17%	Temperatura de operación
Datos históricos	39%	Tiempos de carga/descarga en cada ciclo
Resistencia interna	11%	Resistencia interna de la batería

3.1.2. Descripción de datos

Al identificarse los parámetros de entrada usados por otros investigadores, el siguiente paso consiste en analizar datos experimentales con el fin de establecer aspectos comunes que permitan establecer correlaciones con el SoH y el RUL de la batería. Los datos para este proyecto se obtuvieron de Kaggle, un repositorio global reconocido por la comunidad científica que proporciona conjuntos de datos estandarizados y validados, muchos provenientes de investigaciones publicadas e

instituciones académicas [18]. Los datasets seleccionados cumplen con criterios esenciales de calidad: incluyen metadatos completos sobre su origen y metodología de recolección (trazabilidad), algunos sometidos a procesos de limpieza y normalización por la comunidad (preprocesamiento), y cuentan con referencias de uso en publicaciones indexadas y proyectos similares (validación práctica) [18], garantizando su confiabilidad para el desarrollo del modelo.

Es de gran importancia aclarar que el estudio empleó baterías NMC-LCO 18650 de 2.8Ah (14 unidades) sometidas a pruebas de ciclado durante aproximadamente 1000 ciclos en condiciones controladas (25°C), utilizando tasas de carga de 0.5C y descarga de 1.5C. Las variables registradas en el dataset son:

- ✓ **Índice del ciclo (Cycle_Index):** Número secuencial que identifica cada ciclo completo de carga-descarga realizado.
- ✓ **Tiempo de descarga (Discharge Time (s)):** Duración total (en segundos) del proceso de descarga en cada ciclo.
- ✓ **Decremento 3.6V – 3.4V (Decrement 3.6-3.4V (s)):** Tiempo requerido para que el voltaje de la batería disminuya de 3.6V a 3.4V durante la descarga - un indicador clave de degradación de los materiales del electrodo[18].
- ✓ **Voltaje máximo de descarga (Max. Voltage Discharge (V)):** Valor máximo de voltaje registrado durante la fase de descarga en cada ciclo.
- ✓ **Voltaje mínimo de descarga (Min. Voltage Charge (V)):** Valor mínimo de voltaje alcanzado durante el proceso de carga.
- ✓ **Tiempo en 4.15V (Time at 4.15V (s)):** Tiempo acumulado (en segundos) durante el cual el voltaje de la batería se mantuvo en 4.15V - punto crítico para evaluar la eficiencia de carga [18].
- ✓ **Tiempo de corriente constante (Time Constant Current (s)):** Duración de la fase de corriente constante durante el ciclo, donde la batería mantiene un flujo estable de electrones.
- ✓ **Tiempo de carga (Charging Time (s)):** Tiempo total invertido en completar el proceso de carga en cada ciclo.
- ✓ **Vida útil remanente (Remaining Useful Life):** Estimación de los ciclos de vida útil restante antes de que la batería alcance su fin de vida operativa (definido típicamente como 80% de su capacidad nominal) [18].

3.1.3. Exploración de Datos (EDA):

Con respecto a la exploración de datos, se realiza un estudio mediante gráficos del comportamiento del Dataset, como en este caso, la **Figura 9** revela el comportamiento del RUL (Ciclos restantes de carga) en distintas etapas de carga de la batería, donde se evidencia después de los 1100 ciclos aproximadamente, el considerable deterioro de la tecnología de almacenamiento.

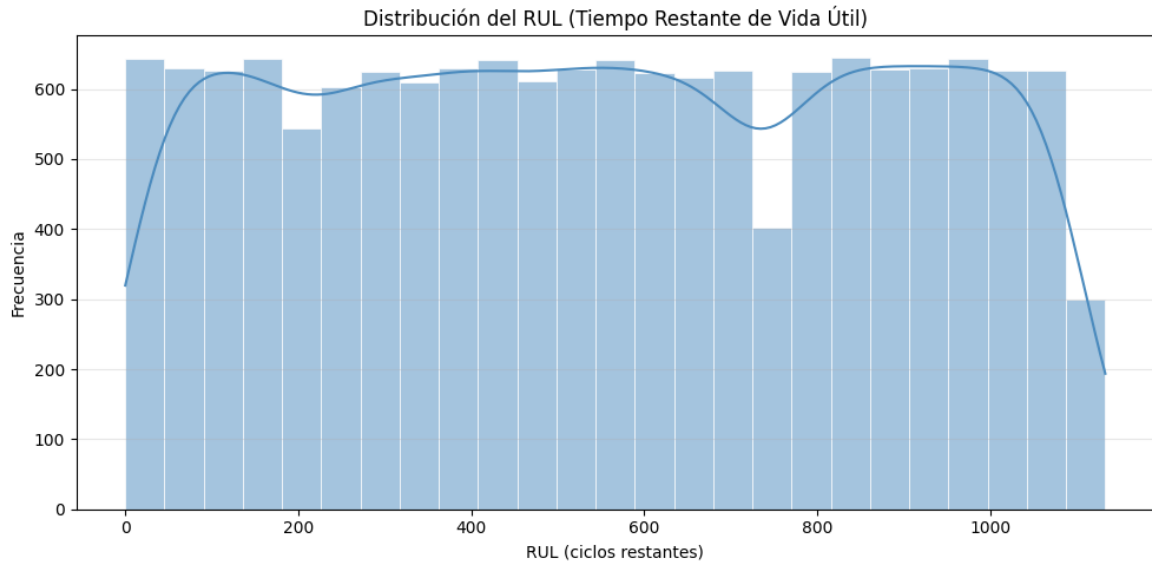


Figura 9. Histograma RUL.

Esta tendencia observada sugiere que las unidades analizadas mantienen un rendimiento adecuado durante la mayor parte de su ciclo de vida bajo las condiciones de prueba establecidas. Sin embargo, la ausencia de datos por debajo de los 200 ciclos representa una limitación para evaluar fases avanzadas de degradación, esto compromete la capacidad del modelo para predecir el fin de vida útil en escenarios reales con procesos de degradación acelerada. En conclusión, esta grafica se realiza con el objetivo de promediar el rango de vida útil de las baterías de ion de litio utilizadas en el Dataset.

Para analizar este comportamiento, se evalúa cada variable mediante su función de densidad de probabilidad. En el caso de la primera variable, tiempo de descarga (s), representado en la **Figura 10**, muestra una clara concentración de datos en valores bajos, indicando que la mayoría de las descargas ocurren en periodos cortos. La forma de la curva presenta una asimetría positiva, con una variación extendida hacia la derecha, esto sugiere la presencia de valores atípicos con tiempos de descarga considerablemente más altos. Esta distribución evidencia que, la mayoría de baterías se descarga entre el intervalo de $0,0 * 10^6$ a $0,1 * 10^6$ ó en otras palabras de 0 Horas a 27,777 Horas.

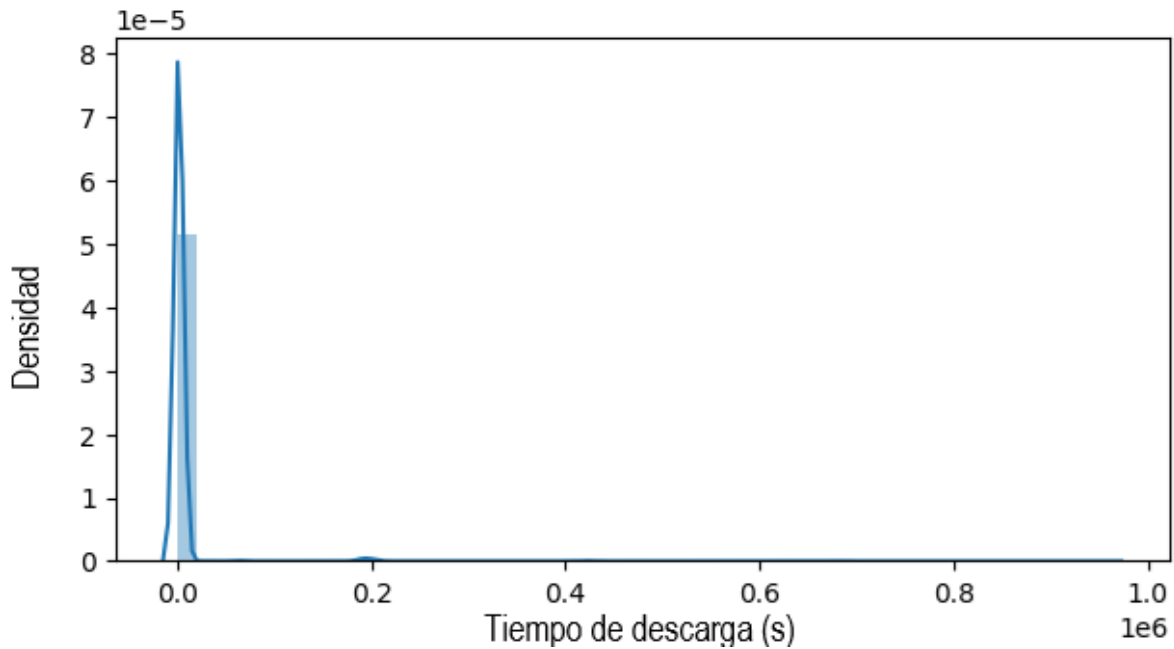


Figura 10. Tiempo de descarga (s) vs densidad

Seguidamente, se evidencia en la **Figura 11**, sobre el tiempo de decremento de 3.6 V a 3.4 V, donde se observa que la mayoría de los valores se concentran en un rango cercano a cero, formando una distribución estrecha con asimetría positiva. Esta forma indica que la mayoría de las descargas en ese rango de voltaje ocurren en un tiempo corto, mientras que existen algunos casos con tiempos altos o incluso negativos, esto se debe a algunas anomalías en el Dataset presentados por errores de medición o registros mal sincronizados. Esta concentración de datos cercanos a cero sugiere un comportamiento uniforme en ese rango de voltaje (Aproximadamente de 0 a 10.000 s) en la mayoría de ciclos.

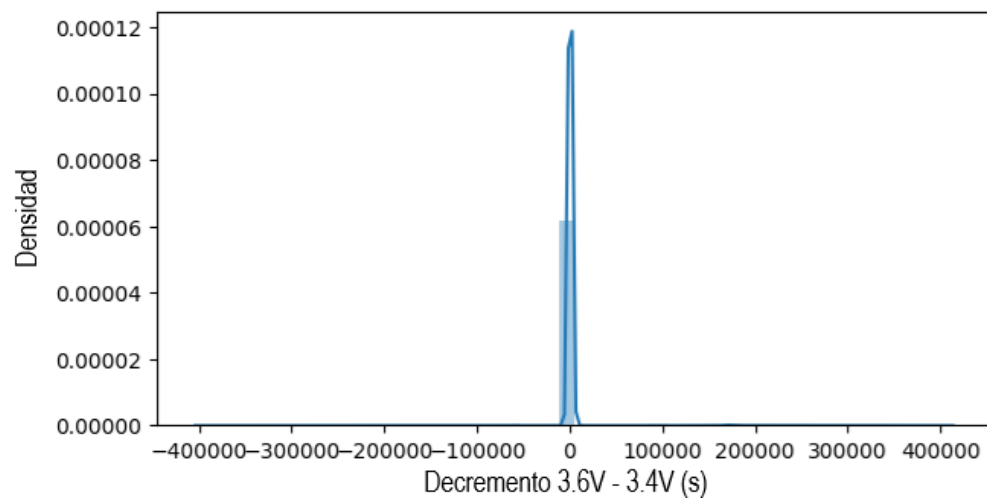


Figura 11. Decremento 3.6V – 3.4V vs densidad

Después, se elabora la gráfica de máximo voltaje de descarga, con respecto a la **Figura 12**, donde la densidad del voltaje máximo de descarga revela una distribución centrada alrededor de 3.8 V, con la mayoría de los valores concentrados entre 3.4 V y 4.2 V. Esta campana pronunciada muestra que las baterías operan predominantemente dentro del rango esperado para celdas de litio durante la fase de descarga. La densidad disminuye abruptamente por debajo de 3.4 V, marcando el límite inferior característico donde la batería se acerca al estado de descarga profunda, mientras que la caída progresiva después de 4.0 V refleja la transición hacia el final del ciclo. La presencia marginal de valores cercanos a 4.4 V podría indicar mediciones puntuales durante transiciones entre fases o dispositivos de medición.

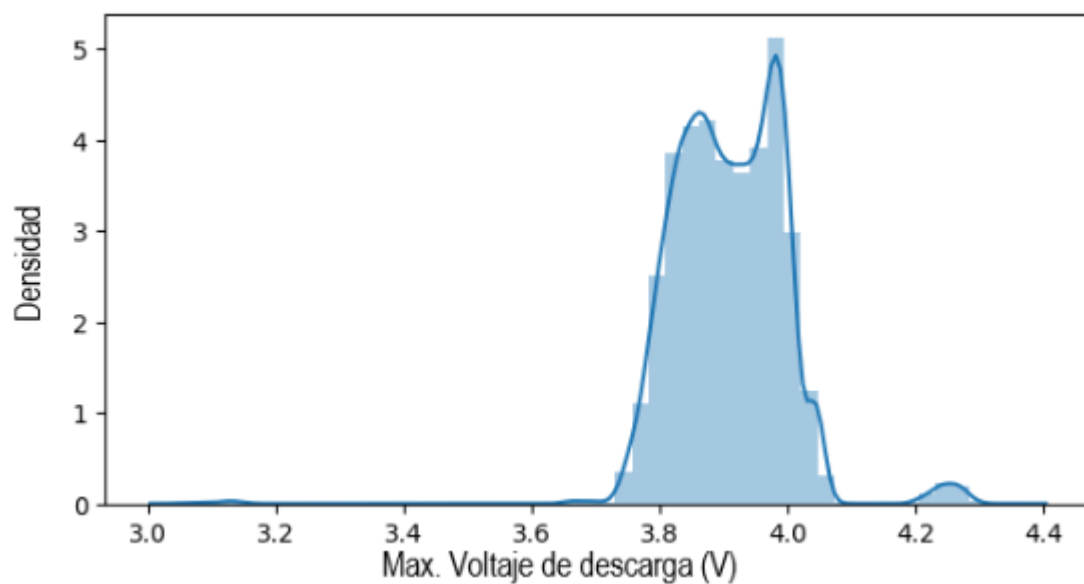


Figura 12. Máximo voltaje de descarga vs densidad

En la **Figura 13**, se muestra la distribución de densidad del voltaje mínimo registrado durante el proceso de carga de las baterías de ion de litio. Los valores se concentran principalmente en el rango de 3.4 V a 4.0 V, presentando un pico pronunciado alrededor de 3.6 V, que indica el voltaje mínimo más frecuente durante la fase de carga. La densidad disminuye gradualmente hacia voltajes más altos, con una caída abrupta por debajo de 3.2 V, esto sugiere que estos valores bajos son poco comunes en condiciones normales de operación.

Esta distribución refleja el comportamiento típico de las baterías de litio durante la carga, donde el voltaje mínimo tiende a estabilizarse en una meseta inferior antes de iniciar el ascenso hacia la carga completa. La presencia de valores cercanos a 4.2 V, aunque con baja densidad, corresponde a mediciones tomadas durante la fase final de carga constante a voltaje máximo.

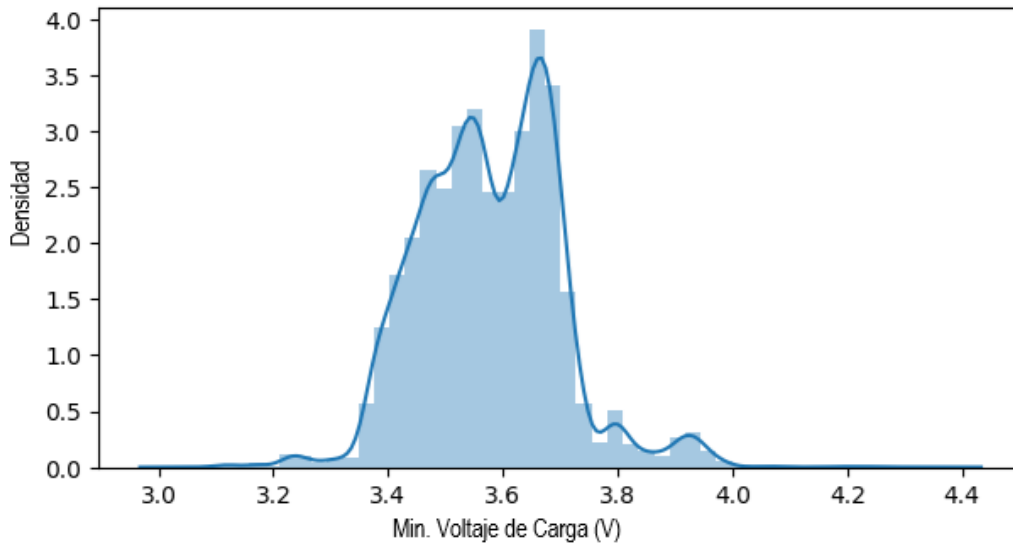


Figura 13. Mínimo voltaje de descarga vs densidad

Siguiendo el estudio de las gráficas del comportamiento de cada una de las variables de entrada, se observa en la **Figura 14**, la distribución de densidad del tiempo que las baterías de ion de litio permanecen a un voltaje de 4.15 V., se evidencia una fuerte concentración de datos en valores cercanos a cero, con una distribución notablemente asimétrica hacia la derecha. Esto indica que, en la mayoría de los casos, el tiempo en que la batería se mantiene en ese voltaje está en el rango de 0 a 10.000 Segundos, o en otras palabras, de 0 a 2.7 Horas. De igual forma, existen variaciones de casos aislados con tiempos más altos, que corresponde a errores en la medición del comportamiento de estas tecnologías de almacenamiento.

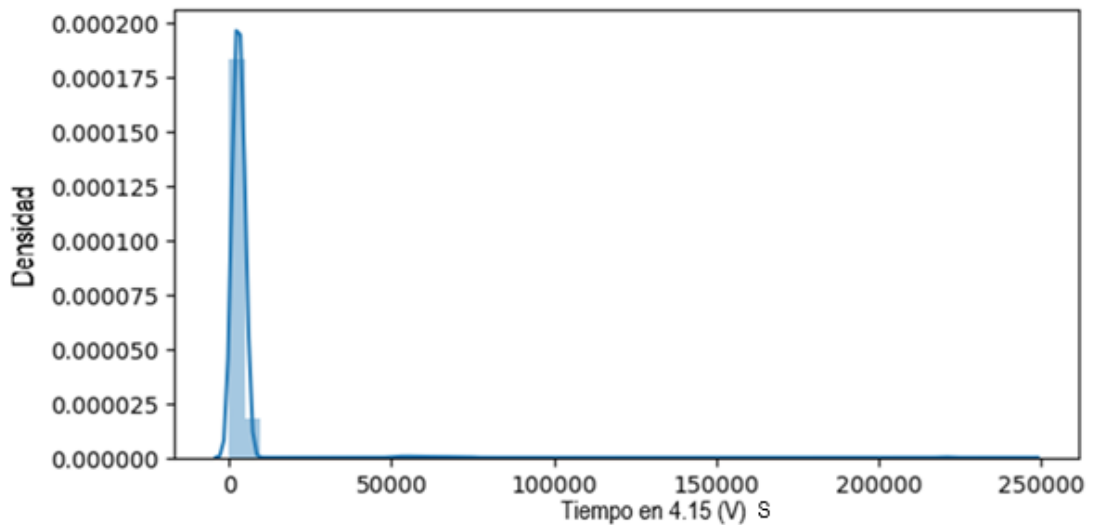


Figura 14. Tiempo en 4.15 (V) vs densidad

Adicionalmente, la **Figura 15**. muestra la distribución del tiempo en la fase de carga a corriente constante, se observa una clara concentración de datos en valores bajos donde la mayoría de los ciclos mantienen esta etapa por periodos relativamente cortos. El monitoreo de estos tiempos es importante, puesto que los tiempos anormalmente largos pueden señalar degradación o pérdida de eficiencia, y se recomienda correlacionar estos datos con variables como temperatura o resistencia interna para análisis futuros [5].

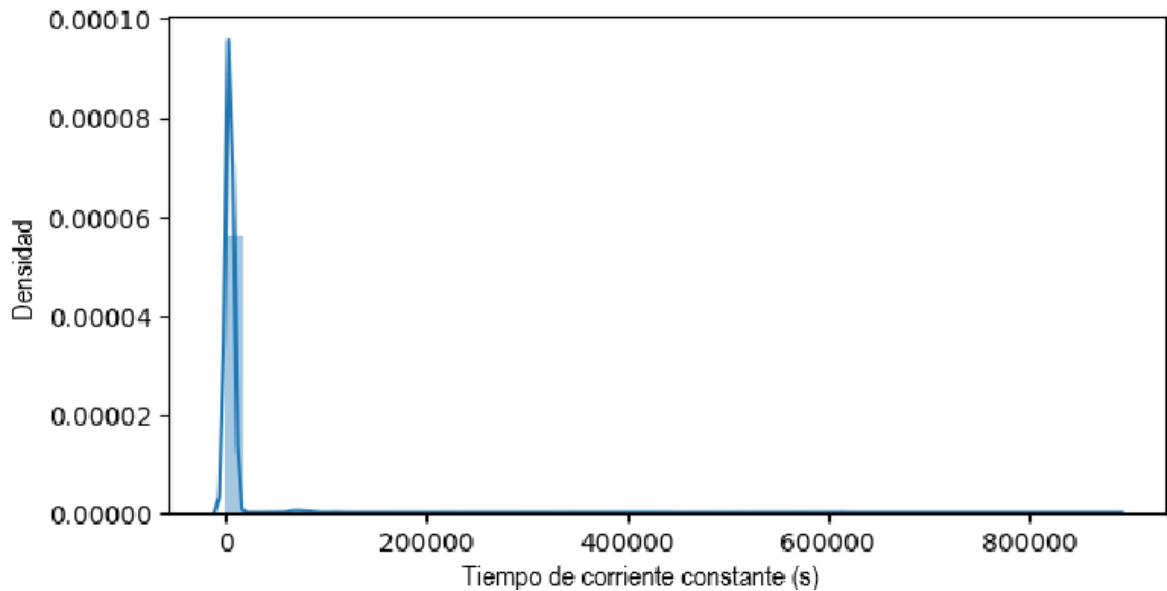


Figura 15. Tiempo de corriente constante (s) vs densidad

Finalmente, en la **Figura 16** se identifican patrones relevantes para el análisis del comportamiento de las baterías. Esta grafica muestra la distribución de densidad del tiempo de carga de las baterías de ion de litio, medida en segundos. Se observa una alta concentración de datos en el rango 0 a 10.000 Segundos, indicando que los valores no son mayores al mencionado intervalo. La forma representada es común en sistemas reales donde, aunque el comportamiento general es consistente, pueden existir condiciones particulares como degradación de la celda, bajas temperaturas o corrientes de carga reducidas [7].

La escala de densidad en el eje Y ($1e-5$) indica que las frecuencias absolutas son bajas, aunque la distribución relativa está claramente definida. La curva asimétrica, típica en datos de duración, confirma que el tiempo de carga sigue un patrón habitual, donde la mayoría de los casos se concentran cerca del valor modal, con una dispersión controlada hacia valores mayores.

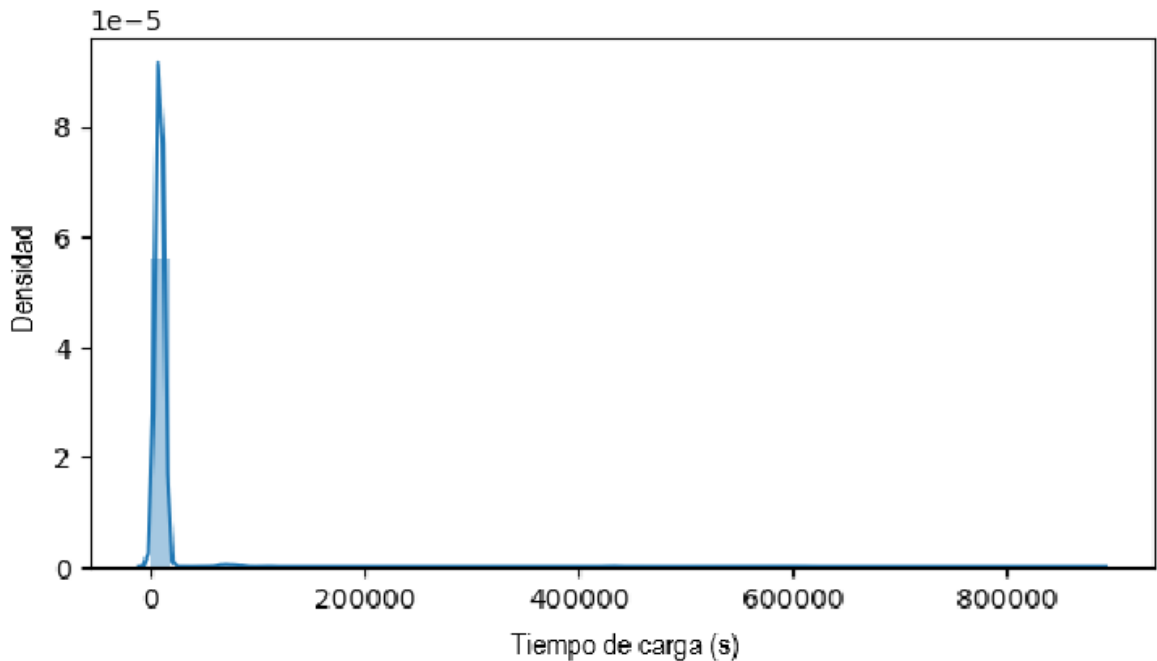


Figura 16. Tiempo de carga (s) vs densidad

Una vez analizadas y descritas las gráficas, se procede a construir la matriz de correlación entre las variables de entrada, la cual permite identificar las relaciones clave y similitudes entre los valores estudiados. En la **Figura 17**, se revela una correlación perfecta (-1.0) entre el índice de ciclos y el RUL, puesto que estos dos parámetros tienen una relación inversa proporcional, entre mayor sea el índice del ciclo, menor será el RUL total de la batería. Las variables clave para predecir la salud real de la batería son los voltajes: el voltaje máximo de descarga (correlación positiva fuerte de 0.78 respecto al RUL) y el voltaje mínimo en carga (correlación negativa fuerte de -0.76).

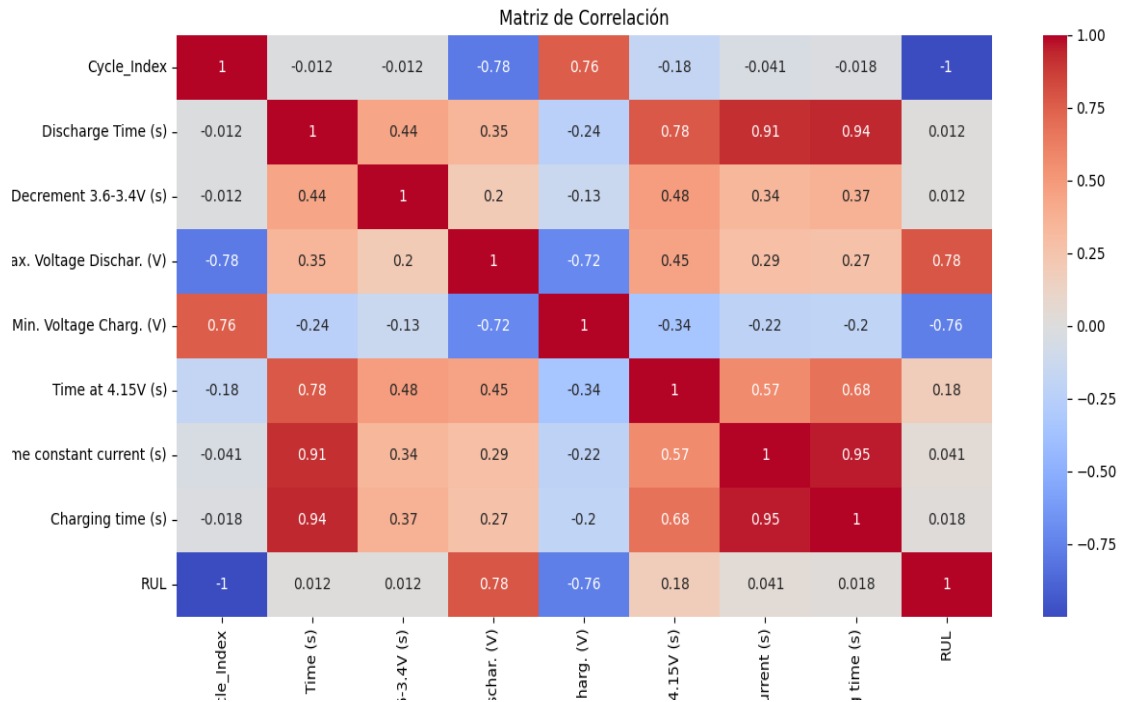


Figura 17. Matriz de correlación.

3.2. DISEÑO

Con respecto al diseño, este modelo se divide en dos grandes etapas, primero se tiene la carga de experiencias (Lectura del dataset) la cual se divide en dos conjuntos de meta-datos para entrenamiento y evaluación, después se inicia con la selección de algoritmos utilizados en el machine learning y Deep learning, se escogen los mejores con respecto a dos variables clave: El error cuadrático medio (MSE, por sus siglas en inglés) y la raíz del error cuadrático medio (RMSE, por sus siglas en inglés), se definen los más aceptables candidatos para después, realizar una técnica de apilamiento (Stacking) conformando un modelo robusto y preciso al momento de evaluación.

3.2.1. Estructura de entrenamiento:

El proceso llevado a cabo en el proyecto se resume en la **Figura 18** donde muestra el flujograma representativo de la estructura de entrenamiento, dividiéndose por etapas cada una de estas representadas por un bloque de proceso, donde se evidencia: la carga de datos junto al preprocesamiento de los mismos, el cual consiste en dividir estos valores del dataset en subconjuntos de train/test (Entrenamiento/Evaluación) con el objetivo de evitar un sobreajuste en la preparación de los datos y tomando como referencia el proceso de los diferentes autores mencionados en el documento [2] [9] [10]. Después se realiza la selección

de algoritmos dependiendo su rendimiento a la hora de ejecución, tanto en valores de RMSE y MSE, como en tiempo de entrenamiento.

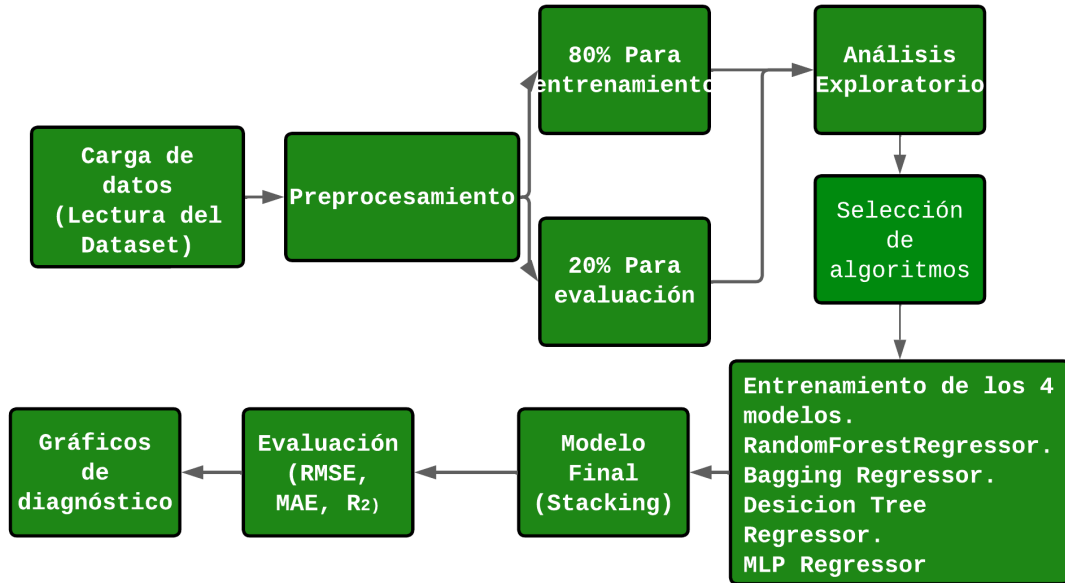


Figura 18. Diagrama de bloques respecto a la estructura de entrenamiento del modelo.

Al seleccionar los mejores algoritmos de machine learning, se realiza un proceso de apilamiento en cada uno de ellos, con el propósito de lograr un modelo final preciso para la estimación del SoH en las baterías de ion de litio y así, evaluarlo con los datos experimentales no utilizados en el entrenamiento (el subconjunto del 20% del dataset) mostrando su buen funcionamiento.

3.2.2. Diagrama de Flujo

A continuación, en la **Figura 19**, se evidencia el diagrama de flujo que describe el funcionamiento del modelo de inteligencia artificial. El objetivo es mostrar de forma clara el paso a paso de cada uno de los procesos realizados, desde el ingreso de los datos correspondientes a los parámetros de entrada del conjunto de entrenamiento (Dataset), hasta la generación de la salida representada por la variable RUL (Ciclos restantes de vida útil).

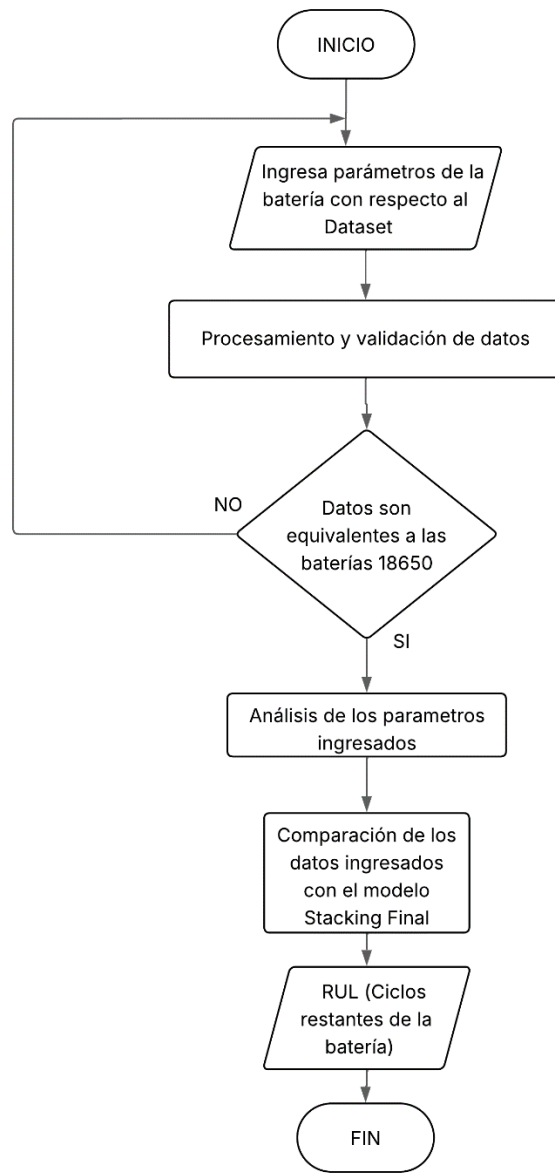


Figura 19. Diagrama de flujo del funcionamiento del modelo.

3.3. IMPLEMENTACIÓN

En esta sección se explica más a detalle cada una de las etapas resumidas anteriormente, definiendo su función y grado de impacto en el modelo de inteligencia artificial.

3.3.1. Carga y preparación de datos

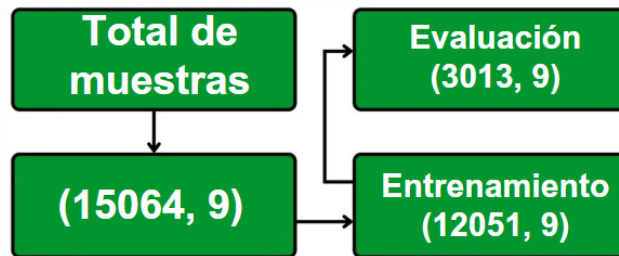


Figura 20. Carga y preparación del Dataset en el modelo.

El propósito de esta etapa es ingresar los datos brutos al pipeline para su posterior análisis y procesamiento. Este pipeline se destaca por utilizar la combinación de imputación de valores faltantes (SimpleImputer con estrategia de media) y escalado estándar (StandardScaler), garantizando que las transformaciones se aprendan solo el conjunto de entrenamiento para prevenir el sobreajuste. Además, la división 80/20 (train_test_Split) es eficiente para el tamaño de la base de datos, evitando el sacrificio de valores de gran importancia, esto se evidencia en la **Figura 20**.

3.3.2. Calidad del Dataset

Evaluando la calidad de los datos, se emplean diversas técnicas de visualización y exploración que facilitan las identificaciones de patrones, tendencias y anomalías presentes en la información, como se puede observar en la **Figura 21**.

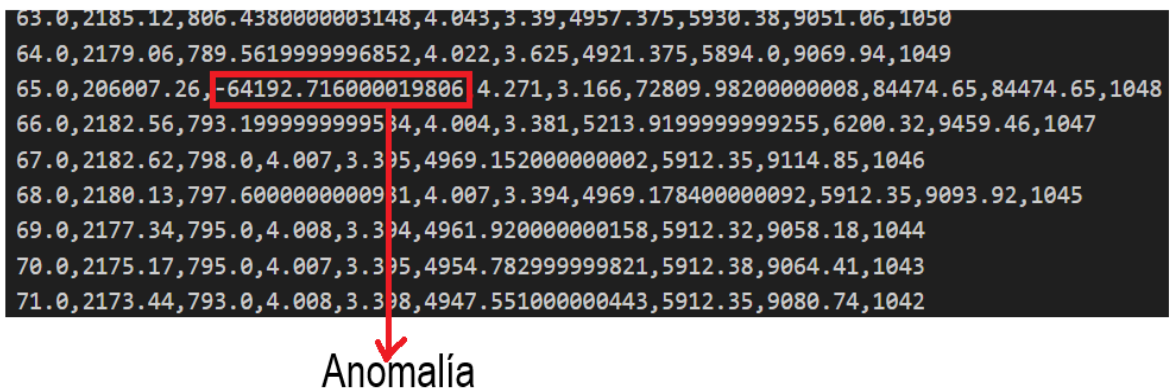


Figura 21. Representación de anomalía en el Dataset

Estas anomalías son eliminadas manualmente del Dataset garantizando la integridad de los datos de entrenamiento. Este preproceso garantiza el uso de mediciones confiables en el entrenamiento y así, obtener predicciones más acertadas en el modelo final.

3.3.3. Selección de algoritmos

3.3.3.1. Modelado Completo

La estrategia implementada consiste en realizar una evaluación comparativa entre nueve algoritmos de regresión lineal, correspondientes al machine learning y Deep learning (Creación de redes neuronales), esto para determinar cuál ofrece el mejor desempeño en la predicción de la vida útil remanente. Entre los modelos evaluados se encuentran RandomForestRegressor, MLPRegressor, BaggingRegressor, DecisionTreeRegressor, AdaBootsRegressor, SGDRegressor, LinearRegression, GradientBoostingRegressor, KNeighborsRegressor y SVR.

Los resultados obtenidos en la **Tabla 5**, revelan que el modelo Random Forest Regressor es el que mejor rendimiento presenta, con una raíz del error cuadrático medio (RMSE) de **23.20**, lo que indica una alta precisión en su funcionamiento. Por otro lado, el modelo con peor desempeño fue SGDRegressor, con un RMSE de 141.66, reflejando una menor capacidad para ajustarse a los datos.

Tabla 5. Comparación entre los distintos algoritmos utilizados por el modelo.

Algoritmo	MSE	RMSE	Tiempo
RandomForestRegressor	538.368288	23.202765	6.218842
MLPRegressor	566.6304	23.804	8.4563
BaggingRegressor	611.349974	24.725492	0.624315
DecisionTreeRegressor	955.865002	30.917067	0.103306
AdaBoostRegressor	960.425241	30.990728	0.033131
KNeighborsRegressor	1287.032924	35.875241	0.110578
GradientBoostingRegressor	2036.129743	45.123494	2.246836
SVR	10621.652715	103.061403	11.036764
LinearRegression	19248.315344	138.738298	0.023962
SGDRegressor	20069.864748	141.668150	0.051851

Con respecto al modelo MLPRegressor, este actúa como una red neuronal profunda que aprende patrones complejos en los datos de baterías mediante una arquitectura de dos capas ocultas (100 y 50 neuronas) [19]. Durante el entrenamiento, analiza los parámetros de entrada (voltajes, tiempos y ciclos) usando diferentes funciones

de activación ReLU, ajustando sus predicciones mediante retropropagación para minimizar el error cuadrático medio entre sus predicciones y el RUL real. Este algoritmo incorpora una parada temprana inteligente (`early_stopping=True`) que monitorea la pérdida en un 10% de datos de validación, esto se puede observar en la **Figura 22**.

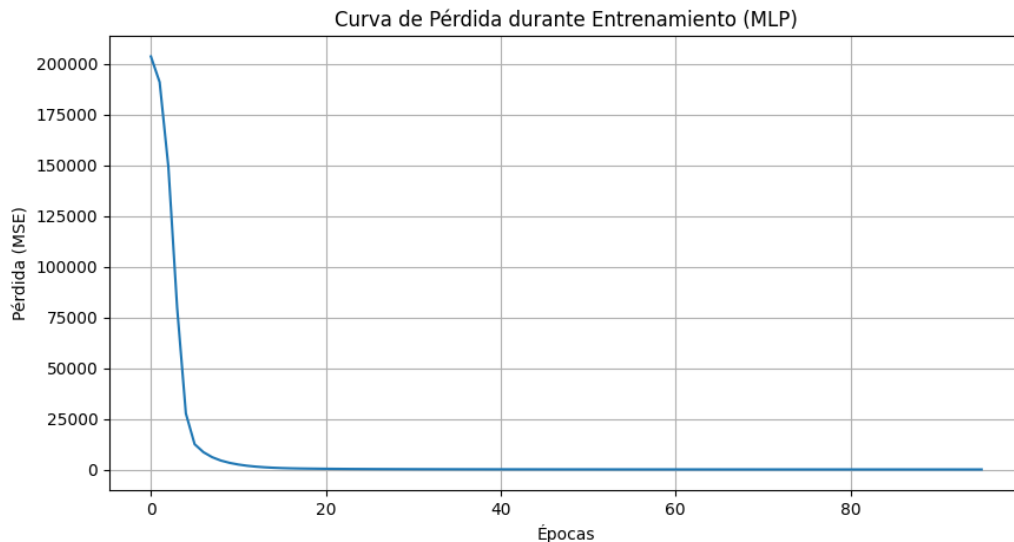


Figura 22. Curva de pérdida durante el entrenamiento.

Donde se evidencia la curva de pérdida durante el entrenamiento del Modelo MLP Regresor (Perceptrón Multicapa), utilizando el error cuadrático medio (MSE) como métrica de desempeño. Se observa una rápida disminución de la pérdida en las primeras épocas, pasando de valores superiores a 200.000 a valores cercanos a cero antes de las 20 épocas. A partir de este punto la curva tiende a estabilizarse y a mantenerse constante, es decir, el modelo ha alcanzado un estado de convergencia y ha aprendido adecuadamente los patrones del conjunto de entrenamiento. Este comportamiento indica que el modelo está bien ajustado y no presenta sobreentrenamiento.

3.3.3.2. Técnica Stacking (Apilamiento)

El modelo de **stacking** implementado opera mediante la arquitectura de dos niveles que combina múltiples algoritmos machine learning. En primer nivel, estos cuatro modelos base (RandomForest, Bagging, DecisionTree, MLP) generan predicciones independientes sobre los datos de las LIBs. Estas predicciones individuales se convierten en entradas para un metamodelo (regresión lineal) en el segundo nivel, que aprende a ponderarlas mediante coeficientes específicos [20]. Durante el entrenamiento, se emplea validación cruzada con 10 particiones para generar las predicciones base evitando así el sobreajuste y asegurando que el metamodelo

aprenda de patrones generalizables. La fusión final de predicciones no es un simple promedio de cada algoritmo, sino una combinación lineal calibrada [20] donde cada proceso contribuye según su desempeño medido, permitiendo que los modelos más precisos tengan mayor peso en la estimación final del RUL. Esta estrategia aprovecha las fortalezas de los distintos algoritmos, mitigando debilidades individuales y produciendo estimaciones ideales y precisas que cualquier modelo trabajando de forma aislada. Este proceso se concluye en la **Figura 23**.

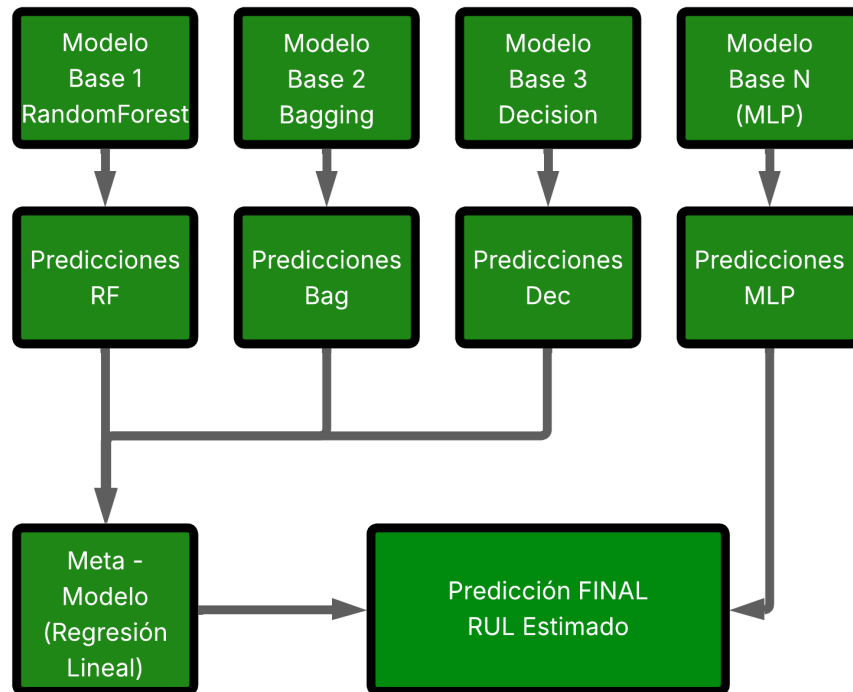


Figura 23. Diagrama de bloques de la estructura del modelado Stacking.

3.4. PRUEBAS DE FUNCIONAMIENTO Y/O DESEMPEÑO

En la siguiente sección se explica a detalle las validaciones correspondientes del modelo con respecto a los datos que no se tomaron en cuenta en el proceso de entrenamiento, siendo así una evaluación confiable.

3.4.1. Modularización del código base

Después de identificar los algoritmos con mejor desempeño y realizar la limpieza del Dataset, el sistema se implementa siguiendo una arquitectura modular que organiza el código en secciones ejecutables, esta estructura se evidenciada en la **Figura 24**, está compuesta por módulos independientes que se encargan de tareas específicas, tales como la carga de datos (load_data.py), el preprocesamiento

(preprocess.py), el entrenamiento (train.py), la evaluación (evaluate.py) y la visualización (visualize.py), entre otros.

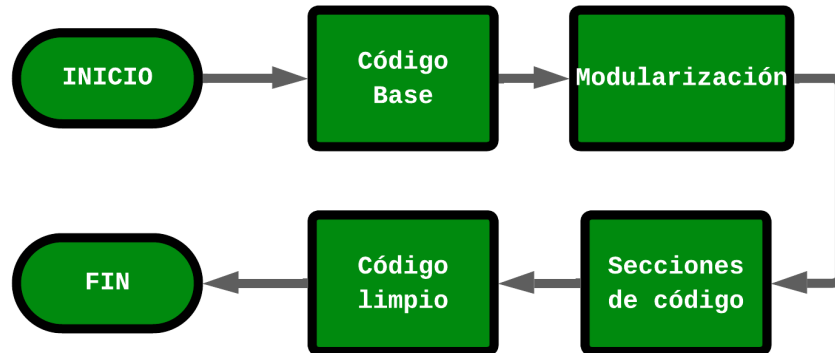


Figura 24. Estructura modular del código base del modelo.

Cada módulo está cuidadosamente documentado, especificando sus entradas, salidas y acompañado de pruebas unitarias que garantizan la calidad, limpieza y reproducibilidad del código.

3.4.2. Evaluación final

Para esta sección, se realizan las gráficas correspondientes a la validación del comportamiento de cada algoritmo utilizado en el modelo stacking, las gráficas contrastan los valores predichos de RUL (eje Y) con los valores reales medidos (eje X), la proximidad de los puntos a la línea diagonal indica la alta precisión del modelo. La dispersión homogénea de la línea sugiere un error constante en todo el rango de medición, sin sesgos sistemáticos. El coeficiente de determinación (R^2) confirma que el modelo captura un porcentaje de la variabilidad en la vida útil restante de las baterías de ion de litio.

Además, los gráficos de distribución de errores de predicción muestran que el modelo no presenta un sesgo direccional. Una distribución estrecha indica que los errores son pequeños y consistentes, lo cual refleja una alta precisión del modelo. En cambio, una distribución más ancha revela una mayor variabilidad en las predicciones y así, determina una menor confiabilidad del algoritmo.

A continuación, se presentan los resultados obtenidos de cada uno de los algoritmos del modelo **Stacking**:

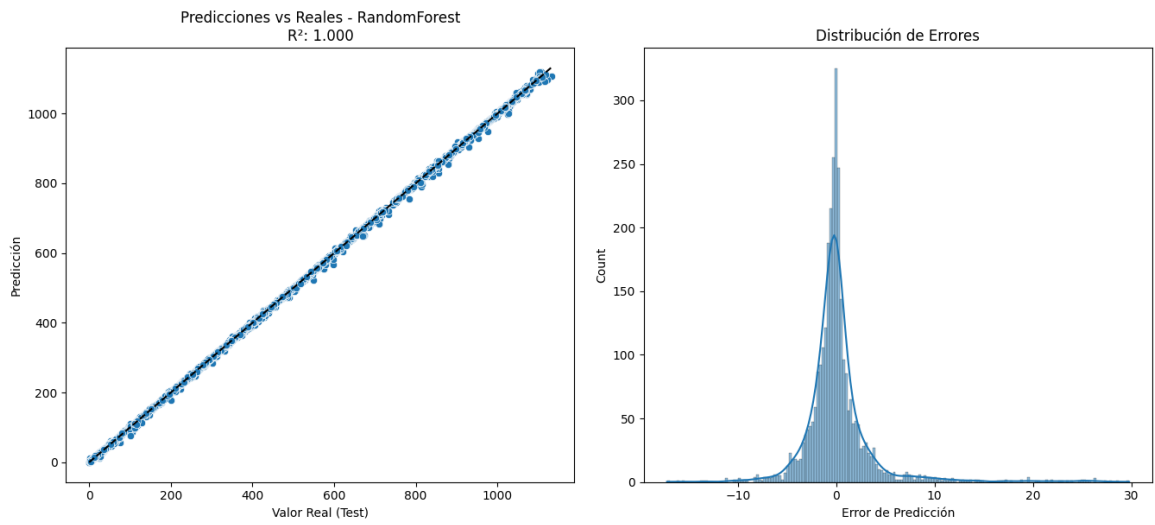


Figura 25. Predicciones vs Datos reales RandomForest

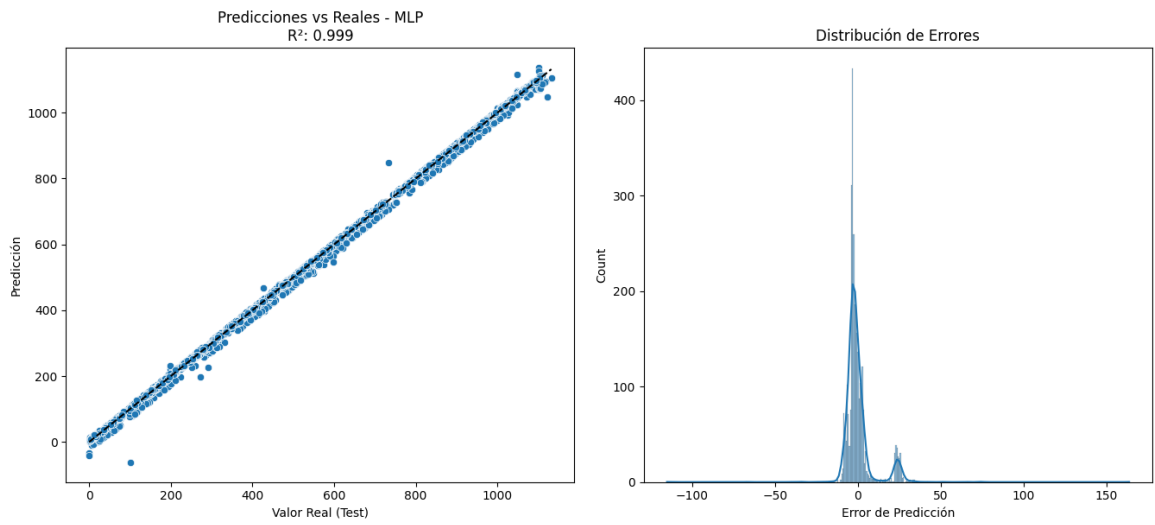


Figura 26. Predicciones vs Datos reales MLP

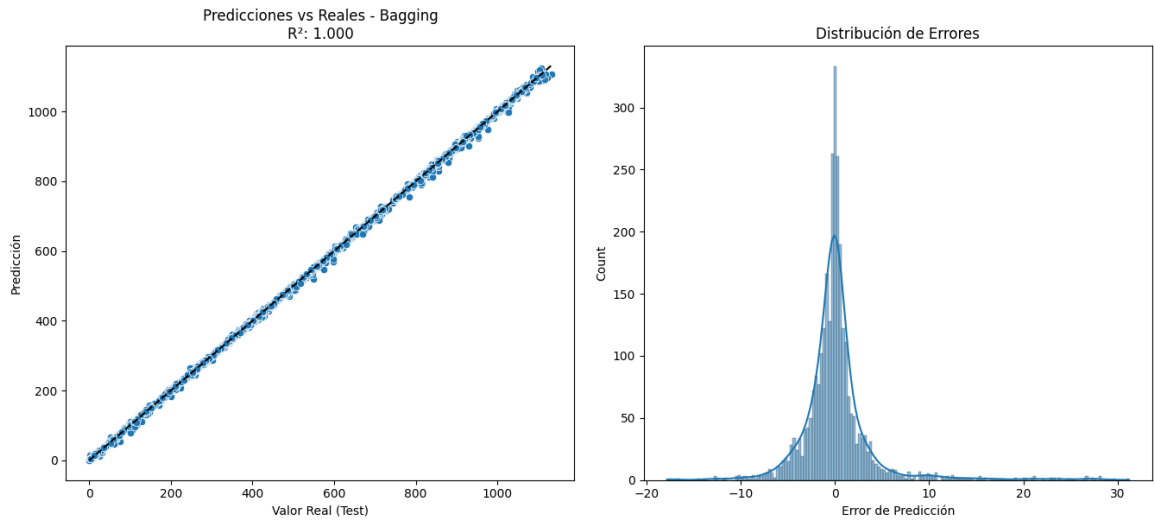


Figura 27. Predicciones vs Datos reales Bagging

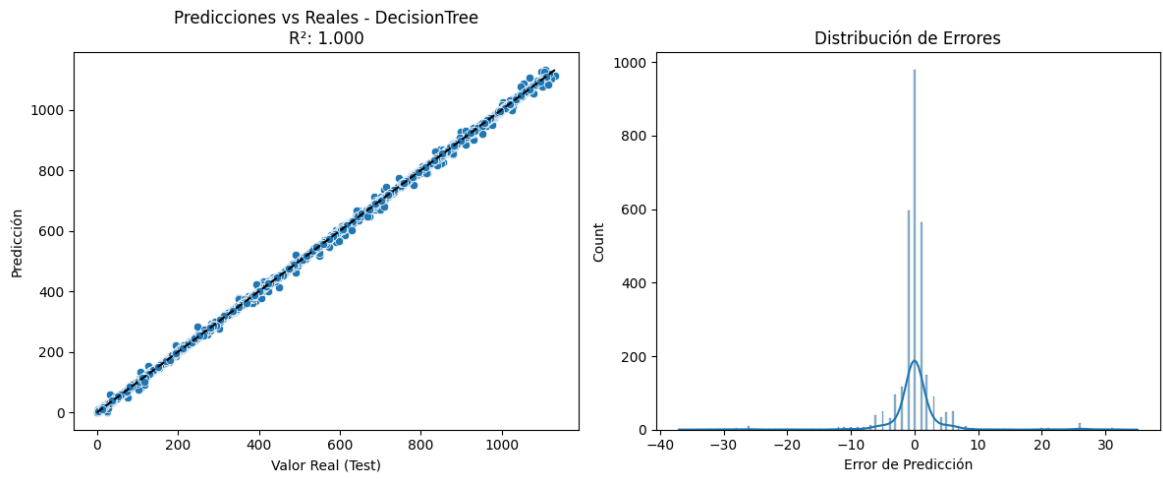


Figura 28. Predicciones vs Datos reales DecisionTree

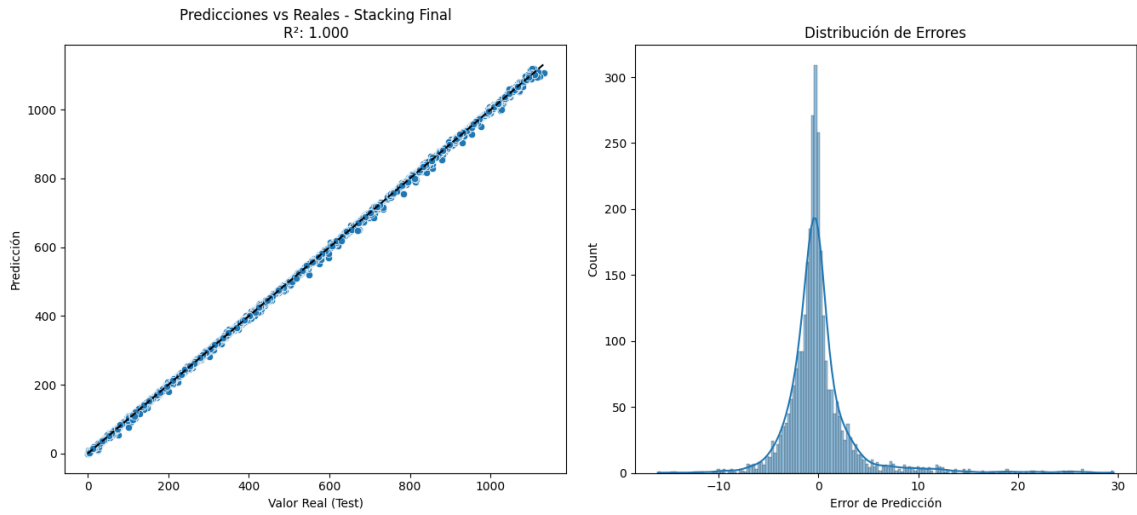


Figura 29. Predicciones vs Datos reales Stacking final

Finalmente, en las gráficas anteriores se evidencia la mejora que presenta el modelo stacking frente a cada algoritmo combinado, mostrando una salida final de RMSE: 3,828. Sin embargo, cada resultado se puede evidenciar en la **Tabla 6**, donde cada variable mencionada anteriormente se evalúa en cada segmento del modelo Stacking.

Tabla 6. Resultados con respecto a la prueba con el 20% de datos aislados.

Modelo Evaluado	RMSE	MAE	R²
RANDOMFORESTREGRESSOR	3.9246	1.9622	0.9999
BAGGINGREGRESSOR	4.0371	2.1636	0.9998
DECISIONTREEREGRESSOR	5.0463	2.1693	0.9998
MLPREGRESSOR	9.1937	5.1296	0.9992
STACKING FINAL	3.8280	2.0839	0.9999

4. ANÁLISIS DE RESULTADOS Y DISCUSIÓN

Finalizadas las pruebas de desempeño con respecto al 20% de datos aislados, se continúa la comparación de resultados con respecto a la función de mejora relativa.

4.1. Comparación de resultados

A continuación, se realiza la función de mejora con respecto al índice del error cuadrático medio en dos etapas, la primera en el primer proceso de entrenamiento y validación de los algoritmos individuales con el RMSE final del modelado Stacking (3,83).

$$Mejora \% = \left(1 - \frac{RMSE_{Final}}{RMSE_{Inicial}}\right) \times 100$$

Se evalúa dicha función con el objetivo de observar un cambio considerable en la precisión del algoritmo. En la **Tabla 7** se presenta dicha función con respecto al RMSE obtenido por cada algoritmo de aprendizaje de máquina.

Tabla 7. Comparación entre los valores de RMSE Inicial/Final

Algoritmo	RMSE Inicial	Mejora Relativa (%)
<i>RandomForestRegressor</i>	23.20	83.5%
<i>BaggingRegressor</i>	24.73	84.5%
<i>DecisionTreeRegressor</i>	30.92	87.6%
<i>AdaBoostRegressor</i>	30.99	87.6%
<i>KNeighborsRegressor</i>	35.88	89.3%
<i>GradientBoostingRegressor</i>	45.12	91.5%
<i>SVR</i>	103.06	96.3%
<i>LinearRegression</i>	138.74	97.2%
<i>SGDRegressor</i>	141.67	97.3%

El patrón observado indica que los algoritmos con un desempeño inicial más bajo, como el SGDRegressor que presentó un error cuadrático medio (RMSE) de 141.67, experimentan las mejoras más significativas, alcanzando una reducción del 97.3% en dicho error. Por otro lado, la técnica de Stacking actúa como un mecanismo de homogeneización del rendimiento, logrando que todos los modelos converjan a un RMSE aproximado de 3.83, lo que refleja una mejora sustancial y consistente en la precisión de las predicciones.

5. CONCLUSIONES Y TRABAJOS FUTUROS

5.1. CONCLUSIONES

En cuanto al primer objetivo, la revisión bibliográfica permitió identificar los parámetros de entrada más relevantes para los modelos de aprendizaje automático aplicados a la estimación del estado de salud (SoH) de baterías de ion de litio. Se destacaron variables operativas en tiempo real como el voltaje en diferentes fases del ciclo, la corriente de carga y descarga, y tiempos de respuesta.

Se diseñó un modelo de aprendizaje automático adaptable, que integra algoritmos de Machine learning y Deep learning, entrenados mediante técnicas como preprocesamiento, modelado y apilamiento. Este modelo es capaz de ajustarse a diferentes distribuciones de datos experimentales y procesar información parcial de ciclos de carga y descarga. La arquitectura propuesta combina datos operativos con indicadores derivados para mejorar la robustez y precisión del modelo.

Se implementa un pipeline que integra el preprocesamiento de los datos con el entrenamiento de los modelos, asegurando que cada algoritmo reciba datos correctamente preparados y que el flujo de trabajo sea eficiente y reproducible. El resultado de este proceso es una comparación entre los valores de MSE, RMSE y tiempo de ejecución, dando como conclusión que el algoritmo Random Forest alcanza el mejor rendimiento, con un RMSE de 23.20, en contraste, el SGDRegressor obtiene el peor desempeño, con un RMSE de 141.66, indicando una menor capacidad de ajuste a los datos. Estos hallazgos definen por categorías los mejores algoritmos en función al objetivo del proyecto.

Al utilizar la técnica de Stacking (Apilamiento), se combinan los algoritmos tanto de machine learning como de Deep learning para una salida final, esto se utiliza principalmente para reducir la varianza y evitar problemas de sobreajuste y así, obtener una predicción confiable, donde la base de este apilamiento consiste en utilizar las predicciones de múltiples algoritmos estándar para entrenar un modelo superior o metamodelo que combine las predicciones de los distintos algoritmos utilizados para obtener un resultado final más preciso.

Finalmente, el modelo Stacking obtuvo un RMSE de 3,8284 al combinar cuatro algoritmos de aprendizaje automático: Random Forest Regressor, Bagging Regressor, MLP Regressor y Decision Tree Regressor. Cada uno fue evaluado en función de su desempeño frente a los datos de validación no utilizados en el proceso de entrenamiento. Esta estrategia de apilamiento aporta robustez, precisión y alto rendimiento, conformando una estructura eficiente para la predicción del RUL (Remaining Useful Life) en baterías de ion de litio.

5.2. PERSPECTIVAS DE TRABAJO FUTURO

A partir de la formulación de este modelo, surgen diversas oportunidades de desarrollo para futuros trabajos relacionados con las baterías de ion de litio. Una de las principales opciones de avance, consiste en el diseño e implementación de un dispositivo, que permita medir en tiempo real las variables de entrada, asegurando que estas sean comparables con las utilizadas en el conjunto o Dataset de entrenamiento y así, logrando la automatización del proceso predictivo junto con la portabilidad del sistema de medición.

De acuerdo a lo anterior, este desarrollo garantiza la aplicación en diversos proyectos y sistemas que requieran un monitoreo continuo y gestión inteligente de las tecnologías de almacenamiento, permitiendo así anticipar fallas, evitar pausas inesperadas en cualquier sistema eléctrico y el estudio de comportamientos de los mismos.

Teniendo en cuenta que en este proyecto se desarrolla un modelo de inteligencia artificial, adaptativo y preciso con respecto a la estimación del SOH en las baterías de ion de litio, en continuidad a esto, no solo se podría adaptar a estas clases de baterías, sino a otros tipos de diferente material interno, como las elaboradas de plomo-ácido, las cuales son consideradas por su alta capacidad de tracción en aplicaciones de vehículos eléctricos y sistemas de almacenamiento de energía. [21] donde también es posible desarrollar modelos dinámicos con respecto a otras variables de entrada, como su resistencia y polarización interna.

BIBLIOGRAFÍA

- [1] Nature, "Lithium-ion batteries need to be greener and ethical," vol. 595, p. 1, Jul. 2021, Accessed: Dec. 13, 2024. [Online]. Available: <https://www.nature.com/articles/d41586-021-01735-z>
- [2] Luis Orlando Pérez Monsalve and Guillermo Andrés Jiménez Estévez, "Predicción de la vida útil de las baterías de un Nissan Leaf usando datos reales de conducción mediante el uso de Machine Learning," p. 36, 2024.
- [3] N. Lotfi, J. Li, R. G. Landers, and J. Park, *Li-ion Battery State of Health Estimation Based on an Improved Single Particle Model*, American Conference. Seattle, USA: IEEE, 2017.
- [4] K. K. Mishra and A. K. Singh, "Li-Ion Battery State of Health Assessment Using Machine Learning," in *India International Conference on Power Electronics, IICPE*, IEEE Computer Society, 2023. doi: 10.1109/IICPE60303.2023.10474671.
- [5] R. Badarudin *et al.*, "Monitoring Process of Lithium-Ion during Charging and Discharging to Battery Management System Performance," in *Proceedings - ICCTEIE 2023: 2023 International Conference on Converging Technology in Electrical and Information Engineering*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 154–159. doi: 10.1109/ICCTEIE60099.2023.10366752.
- [6] A. M. Ortiz *et al.*, "Análisis de la situación energética de Bogotá y Cundinamarca," Bogotá D.C., Jul. 2013. Accessed: Jan. 21, 2025. [Online]. Available: <https://www.repository.fedesarrollo.org.co/handle/11445/154>
- [7] A. M. Lacosta and T. Cortes Arcos, "Battery Characterization on the Basis of Open Data Using Artificial Intelligence Techniques," Zaragoza, Jun. 2022. Accessed: Jan. 31, 2025. [Online]. Available: <https://zaguan.unizar.es/record/117777/files/TAZ-TFG-2022-2252.pdf>
- [8] G. Flores Muñoz, J. Manuel, and E. González, "Modelos de predicción de la vida útil de baterías de iones de Litio," Sevilla, 2020. Accessed: Feb. 01, 2025. [Online]. Available: <https://biblus.us.es/bibing/proyectos/abreproy/92969/fichero/TFG-2969+FLORES+MU%C3%91OZ%2C+GUILLERMO.pdf>

- [9] Z. Li, D. Chen, J. Lu, W. Zhang, and C. Deng, "Li-ion Battery SOH Estimation Based on BO-SVR Model," in *2023 IEEE 7th Conference on Energy Internet and Energy System Integration, EI2 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 3737–3742. doi: 10.1109/EI259745.2023.10513320.
- [10] G. Liu *et al.*, "Joint Estimation of SOC and SOE for Energy Storage Lithium-ion Batteries Based on KNN Algorithm," in *2024 5th International Conference on Power Engineering, ICPE 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 501–505. doi: 10.1109/ICPE64565.2024.10929186.
- [11] P. Leninpugalhanthi *et al.*, "Lithium-Ion Battery Life Estimation Using Machine Learning Algorithm," in *8th International Conference on Advanced Computing and Communication Systems, ICACCS 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 573–576. doi: 10.1109/ICACCS54159.2022.9785066.
- [12] A. Castro Díaz and R. Wanton Suárez, "Lithium ion batteries: structure, operation, failures and their causes," Cuba, Jan. 2023. doi: CCBYNCSA4.0.
- [13] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," vol. C. Janiesch *et al.*, no. C6 . C8 . M15 . O3, pp. 686–695, Apr. 2021, doi: 10.1007/s12525-021-00475-2/Published.
- [14] Kevin P. Murphy, "Probabilistic Machine Learning: An introduction," The MIT Press.
- [15] Aurélien Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd Edicion., vol. O'Reilly Media. 1098125975, 2022.
- [16] Cambio Energético, "¿Cuál es la vida útil de una batería de litio para paneles solares?," <https://www.cambioenergetico.com/blog/vida-util-bateria-litio/>.
- [17] J. Sihvo, D.-I. Stroe, and T. Roinila, *SOH analysis of Li-ion battery based on ECM parameters and broadband impedance measurements*. Tampere, Finland - Aalborg, Denmark: IEEE, 2020.
- [18] Y. Seokcho, "Battery Remaining Life Prediction." Accessed: Jun. 11, 2024. [Online]. Available: <https://www.kaggle.com/code/yeonseokcho/battery-remaining-life-prediction>
- [19] J. Maqbool, P. Aggarwal, R. Kaur, A. Mittal, and I. A. Ganaie, "Stock Prediction by Integrating Sentiment Scores of Financial News and MLP-Regressor: A

Machine Learning Approach,” in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 1067–1078. doi: 10.1016/j.procs.2023.01.086.

- [20] D. Khyani, S. Jakkula, S. N. Gowda C, A. K. J, and S. K. R, “An Interpretation of Stacking and Blending Approach in Machine Learning,” *International Research Journal of Engineering and Technology*, vol. 08, no. 7, Jul. 2021, [Online]. Available: www.irjet.net
- [21] S. Kai and S. Qifang, “Overview of the Types of Battery Models,” Yantai, China in School of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255049, P. R. China, Jul. 2011.