



Escuela de Administración de Tecnologías de Información

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor
de recuperación de los vehículos eléctricos en Costa Rica

Trabajo Final de Graduación para optar al grado de Licenciatura en Administración
de Tecnología de Información
Modalidad Proyecto de Graduación

Elaborado por: Mauricio Loría Rodríguez

Prof. Tutor: Dr. Federico Torres Carballo

Cartago, Costa Rica

Semestre I

Mayo, 2025



Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica © 2025 por Mauricio Loría Rodríguez este trabajo tiene licencia de CC BY-NC-SA 4.0. Para ver una copia de esta licencia, visite <https://creativecommons.org/licenses/by-nc-sa/4.0/>

INSTITUTO TECNOLÓGICO DE COSTA RICA
ESCUELA DE ADMINISTRACIÓN DE TECNOLOGÍAS DE INFORMACIÓN
GRADO ACADÉMICO: LICENCIATURA

Los miembros del Tribunal Examinador de la Escuela de Administración de Tecnologías de Información, recomendamos que el siguiente informe del Trabajo Final de Graduación del estudiante Mauricio Loría Rodríguez sea aceptado como requisito parcial para obtener el grado académico de Licenciatura de Tecnología de Información.

Profesor Tutor

Lector externo

Lector académico

Coordinadora de Trabajo Final de Graduación

Dedicatoria

A mis papás Eladio y Lidiette, quienes siempre han creído en mí, apoyándome e impulsándome en cada paso, sin importar el tiempo o dificultad, demostrándome su amor, confianza, cariño y perseverancia para lograr mis objetivos.

A mis hermanos Karla y David, quienes son mis guías, motivándome, enseñándome una mejor forma de hacer las cosas, a partir de su experiencia y aconsejándome para crecer y ser mejor persona cada día, les agradezco siempre.

A mi novia, Silvana, quien me apoyó de forma incondicional, me inspiró y tuvo mucha paciencia y cariño conmigo durante el desarrollo de mi trabajo final de graduación.

Agradecimientos

A Hernán, quién ha sido un amigo y un mentor durante todo mi desarrollo profesional y agradezco increíblemente la primera oportunidad que me brindó, así como elaborar este trabajo final de graduación.

A mi profesor tutor Federico, por toda la guía y ayuda brindada durante la realización del proyecto.

A mis amigos Meg, Estef, Mari J, Mari, Tommy, Carlos, Campos, Sergio, Vini y Steven agradezco cada uno de los momentos que compartimos en la universidad y todos los momentos que reímos, tuvimos que correr y estudiar juntos, no hubiera sido lo mismo sin ustedes.

A mis amigos desde la escuela Juan Da, Meo y Emilio, aprecio mucho nuestra amistad, y el apoyo en todo momento y espero siga siendo así por toda la vida.

Resumen

Loría Rodríguez, Mauricio (2025). *Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica*. [Trabajo Final de Graduación, Tecnológico de Costa Rica]

Este Trabajo Final de Graduación desarrolla y valida un modelo predictivo que estima el valor residual de los vehículos eléctricos financiados por la empresa Áltica en Costa Rica, con el fin de reducir las pérdidas económicas derivadas de la alta variabilidad de precios al final del contrato.

La investigación es de tipo aplicada y cuantitativa; se siguió la metodología CRISP-DM: recopilación y depuración de 1 318 registros procedentes de cuatro fuentes (ventas históricas, avalúos y dos *web-scrapings*), análisis exploratorio y estandarización de variables clave (edad, kilometraje, autonomía, química de batería), comparación de algoritmos lineales, de ensamble y de series temporales, y validación mediante validación cruzada estratificada y un conjunto de prueba.

El modelo seleccionado, un *Stacking Regressor* que combina *Random Forest* y *Gradient Boosting*, alcanza $R^2 = 0,83$ y $MAE \approx \$6\,900$, cumpliendo los umbrales de negocio. La herramienta se integrará en el flujo de tasación como apoyo al juicio experto.

Palabras clave: valor residual, vehículos eléctricos, modelo predictivo, aprendizaje automático, riesgo financiero, ciencia de datos, analíticas de negocios

Abstract

Loría Rodríguez, Mauricio (2025). *Predictive model to reduce financial losses caused by variability in the residual value of electric vehicles in Costa Rica*. [Final Graduation Project, Costa Rica Institute of Technology]

This Final Graduation Project develops and validates a predictive model that estimates the residual value of electric vehicles financed by the company Ática in Costa Rica, aiming to lessen the economic losses stemming from high price variability at contract termination.

The research is applied and quantitative; it follows the CRISP-DM methodology: collection and cleansing of 1,318 records from four sources (historical sales, appraisals, and two web-scrapings); exploratory analysis and standardization of key variables (age, mileage, range, battery chemistry); comparison of linear, ensemble, and time-series algorithms; and validation through stratified cross-validation and a hold-out test set.

The selected model is the Stacking Regressor combining Random Forest and Gradient Boosting achieving $R^2 = 0.83$ and $MAE \approx \$6\,900$, satisfying the business thresholds. The tool will be integrated into the valuation workflow to support expert judgment.

Keywords: residual value, electric vehicles, predictive model, machine learning, financial risk, data science, business analytics

Tabla de Contenidos

1. Introducción	1
1.1. Descripción general	1
1.2. Antecedentes	3
1.2.1. Descripción de la organización	3
1.2.2. Misión	4
1.2.3. Visión	4
1.2.4. Valores	4
1.2.5. Trabajos similares realizados dentro y fuera de la organización	5
1.3. Planteamiento del problema	8
1.3.1. Situación problemática	8
1.3.2. Justificación del proyecto	10
1.3.3. Beneficios esperados o aportes del Trabajo Final de Graduación	12
1.4. Objetivos del Trabajo Final de Graduación	15
1.4.1. Objetivo general	15
1.4.2. Objetivos específicos	15
1.5. Alcance	16
1.6. Supuestos	17
1.7. Entregables	17
1.7.1. Entregables del producto	17
1.7.2. Entregables de gestión	19
1.8. Limitaciones	19
2. Marco Conceptual	21
2.1. Analítica de negocios	21
2.1.1. Tipos de analítica	22
2.2. Inteligencia de negocios	22
2.2.1. Datos, información y conocimiento	23
2.2.2. Relación y diferencias con analítica de negocios	23
2.2.3. Extracción, transformación y carga (ETL)	25
2.3. Inteligencia de negocios en las finanzas	27
2.3.1. Conceptos fundamentales	27
2.3.2. Gestión de riesgos financieros	29
2.3.3. Valoración predictiva de activos	30
2.3.4. Factores determinantes del valor de recuperación de activos	30
2.4. Ciencia de datos	32
2.4.4. Técnicas de modelado predictivo	34
2.4.5. Aprendizaje automático y técnicas de ensamblaje	38
2.4.6. Evaluación y validación de modelos predictivos	38
2.5. Herramientas tecnológicas	39
2.6. Vehículos eléctricos	40
2.6.1. Tipos de baterías	40
2.6.2. Diferencia entre las baterías basadas en litio y de estado sólido	41
2.6.3. Contexto global	42
2.6.4. Mercado costarricense	43

3. Marco metodológico	46
3.1. Tipo de investigación.....	46
3.1.1. Investigación aplicada	46
3.1.2. Investigación cuantitativa	47
3.2. Enfoque y diseño de la investigación	47
3.2.1. Enfoque de la investigación.....	47
3.2.2. Diseño de la investigación.....	50
3.3. Alcance de la investigación	53
3.4. Fuentes de datos e información.....	55
3.4.1. Fuentes primarias	55
3.4.2. Fuentes secundarias	56
3.5. Sujetos de investigación	57
3.6. Variables de la investigación	58
3.7. Técnicas e instrumentos de recolección de datos.....	61
3.7.1. Entrevistas	61
3.7.2. Revisión documental.....	61
3.7.3. Observación	63
3.8. Procedimiento metodológico de la investigación	63
3.8.1. Fase 1: Adquisición de la información	64
3.8.2. Fase 2: Análisis exploratorio	65
3.8.3. Fase 3: Diseño de los modelos predictivos	65
3.8.4. Fase 4: Validación del modelo	66
3.9. Operacionalización de las variables o categorías	66
3.10. Tabla resumen del procedimiento metodológico de la Investigación	70
4. Análisis de resultados	75
4.1. Resultados de la Fase 1: Adquisición de la información.....	75
4.1.1. Fuentes consolidadas	75
4.1.2. Rendimiento de las baterías	76
4.1.3. Criterios de depuración y enriquecimiento	76
4.1.4. Validación experta	76
4.1.5. Resultado de la fase y alienación con los objetivos	77
4.2. Resultados de la Fase 2: Análisis exploratorio (EDA)	77
4.2.1. Introducción y delimitación	77
4.2.2. Calidad y consistencia de los datos.....	77
4.2.3. Panorama del mercado.....	78
4.2.4. Relación entre la capacidad útil de la batería y la autonomía	79
4.2.5. Antigüedad del vehículo y su impacto.....	80
4.2.6. Factores técnicos y estructurales	81
4.2.7. Síntesis e integración.....	82
4.3. Resultados de la Fase 3: Diseño de los modelos predictivos	83
4.3.1. Delimitación del problema y criterios de modelado	83
4.3.2. Arquitectura de la canalización de datos.....	83
4.3.3. Selección de algoritmos y justificación.....	84
4.3.4. Modelos ARIMA y SARIMA.....	85
4.3.5. Diseño del modelo ensamblado (<i>Stacking</i>).....	86

4.4.	Resultados de la Fase 4: Validación del modelo	86
4.4.1.	Exactitud y confiabilidad	87
5.	Propuesta de solución	90
5.1.	Descripción de la solución.....	90
5.1.1.	Arquitectura analítica del modelo	90
5.1.2.	Flujo de datos y proceso de tasación	91
5.1.3.	Infraestructura tecnológica	92
5.1.4.	Gobierno y ciclo de vida del modelo.....	93
5.2.	Costos de elaboración del modelo	93
5.2.1.	Costos salarios de los involucrados	93
5.2.2.	Costos herramienta utilizada	95
5.2.3.	Costos totales	95
6.	Conclusiones	97
6.1.	Conclusión: Objetivo específico uno	97
6.2.	Conclusión: Objetivo específico dos	97
6.3.	Conclusión: Objetivo específico tres	98
6.4.	Conclusión: Objetivo específico cuatro	98
6.5.	Conclusión: Objetivo general	99
7.	Recomendaciones	100
8.	Referencias	101
9.	Apéndices.....	107
10.	Anexos	124

Índice de Figuras

Figura 1. <i>Organigrama de la organización</i>	3
Figura 2. <i>Acumulado de vehículos eléctricos en Costa Rica año 2024</i>	8
Figura 3. <i>Diagrama de Ishikawa - Planteamiento del problema</i>	10
Figura 4. <i>Metodología de Ralph Kimball</i>	24
Figura 5. <i>Diagrama del proceso de extracción, transformación y carga (ETL)</i>	26
Figura 6. <i>Fases de CRISP-DM modelo de referencia</i>	33
Figura 7. <i>Ejemplo del proceso iterativo en el análisis exploratorio de datos</i>	33
Figura 8. <i>Inventario global de vehículos eléctricos 2013-2023</i>	42
Figura 9. <i>Evolución y proyección acumulada de los automóviles eléctricos en Costa Rica</i>	44
Figura 10. <i>Fases metodológicas de la investigación</i>	64
Figura 11. <i>Subactividades de la fase uno</i>	65
Figura 12. <i>Subactividades de la fase dos</i>	65
Figura 13. <i>Subactividades de la fase tres</i>	66
Figura 14. <i>Subactividades de la fase cuatro</i>	66
Figura 15. <i>Descripción del conjunto de datos</i>	78
Figura 16. <i>Distribución de precios de vehículos eléctricos en Costa Rica</i>	79
Figura 17. <i>Relación entre capacidad de la batería y la autonomía</i>	79
Figura 18. <i>Precio vs año de fabricación</i>	80
Figura 19. <i>Distribución de vehículos por antigüedad</i>	81
Figura 20. <i>Modelo ARIMA predicción</i>	85
Figura 21. <i>Modelo SARIMA predicción</i>	85
Figura 22. <i>Desempeño de los modelos evaluados</i>	87
Figura 23. <i>Valor real vs valor predicho</i>	88
Figura 24. <i>Flujo de predicción del valor de recuperación</i>	91
Figura 25. <i>Flujo de trabajo con el modelo de predicción</i>	92

Índice de Tablas

Tabla 1. <i>Proyectos internos</i>	5
Tabla 2. <i>Proyectos externos</i>	6
Tabla 4. <i>Tipos de analítica</i>	22
Tabla 5. <i>Cuadro comparativo de técnicas de modelado predictivo</i>	35
Tabla 6. <i>Cuadro comparativo de tipos de baterías</i>	40
Tabla 7. <i>Enfoques de investigación</i>	47
Tabla 8. <i>Diseño de la investigación</i>	51
Tabla 9. <i>Alcances de la investigación</i>	53
Tabla 10. <i>Fuentes de información primarias</i>	55
Tabla 11. <i>Fuentes de información secundarias</i>	56
Tabla 12. <i>Sujetos de investigación</i>	57
Tabla 13. <i>Variables de investigación</i>	59
Tabla 14. <i>Operacionalización de las variables</i>	67
Tabla 15. <i>Resumen del procedimiento metodológico</i>	70
Tabla 16. <i>Tabla de fuentes consolidadas</i>	75
Tabla 17. <i>Familias de algoritmos</i>	84
Tabla 18. <i>Infraestructura tecnológica del servidor</i>	92
Tabla 19. <i>Costo por salario de Estudiante investigador</i>	94
Tabla 20. <i>Costo por salario de Gerente de TI</i>	94
Tabla 21. <i>Costo por salario de Analista financiero del sector automotriz</i>	95
Tabla 22. <i>Costo por herramienta</i>	95
Tabla 23. <i>Costos totales</i>	95

Índice de Apéndices

Apéndice A. Cronograma de actividades.....	107
Apéndice B. Plantilla de minuta de reuniones	108
Apéndice C. Plantilla hoja de registro observacional	108
Apéndice D. Plantilla de entrevista abierta	108
Apéndice E. Plantilla de revisión documental	109
Apéndice F. Minuta de Reunión #1	109
Apéndice G. Minuta de Reunión #2	110
Apéndice I. Minuta de Reunión #3	111
Apéndice J. Minuta de Reunión #4	112
Apéndice K. Minuta de Reunión #5	113
Apéndice L. Minuta de Reunión #6	114
Apéndice M. Minuta de Reunión #7	116
Apéndice N. Minuta de Reunión #8	116
Apéndice O. Entrevista #1	117
Apéndice P. Revisión documental #1	119
Apéndice Q. Revisión documental #2	120
Apéndice R. Revisión documental #3	120
Apéndice S. Revisión documental #4	121
Apéndice T. Entrevista #2.....	122

1. Introducción

1.1. Descripción general

En el contexto actual de transformación digital y transición hacia energías limpias, los vehículos eléctricos se han consolidado como una alternativa sostenible y eficiente en el sector automotriz. No obstante, el mercado de vehículos eléctricos en Costa Rica enfrenta desafíos significativos, especialmente en términos de la incertidumbre relacionada con su valor de reventa. Esta falta de predictibilidad dificulta el desarrollo de herramientas analíticas robustas que permitan estimar su depreciación a lo largo del tiempo. Como consecuencia, las entidades financieras, como Áltica, enfrentan riesgos adicionales al financiar estos vehículos, lo que complica la gestión efectiva de sus activos y podría traducirse en pérdidas financieras.

El presente proyecto tiene como objetivo abordar esta problemática mediante el desarrollo de un modelo predictivo analítico capaz de estimar el valor futuro de los vehículos eléctricos financiados por la entidad. A través del uso de técnicas avanzadas de análisis estadístico y aprendizaje automático, se busca garantizar la precisión y confiabilidad del modelo, asegurando su viabilidad técnica durante el primer semestre de 2025.

La relevancia de este proyecto radica en la creciente necesidad de herramientas predictivas en el sector financiero como lo son herramientas de aprendizaje automático o *machine learning*, donde la dispersión de datos históricos y la falta de modelos establecidos para la depreciación de vehículos eléctricos dificultan la evaluación precisa de su valor de recuperación. Este esfuerzo no solo contribuirá a la estabilidad financiera de la entidad al proporcionar una herramienta analítica confiable, sino que también fortalecerá la confianza en el mercado de vehículos eléctricos, promoviendo su adopción y sostenibilidad a largo plazo.

Para alcanzar este objetivo, el proyecto sigue un enfoque metodológico estructurado que se divide en cuatro fases clave. En primer lugar, se realiza la recopilación y análisis de datos históricos y actuales del mercado costarricense, identificando factores determinantes como ventas, depreciación y rendimiento de baterías. Posteriormente, se procede al análisis exploratorio de los datos para garantizar su calidad y representatividad. En la tercera fase, se diseña el modelo predictivo utilizando algoritmos de aprendizaje automático, seleccionando el más adecuado en función de su desempeño. Finalmente, se valida el modelo mediante pruebas con datos reales y

técnicas de validación cruzada, utilizando métricas específicas para garantizar su precisión y confiabilidad.

El presente proyecto representa una oportunidad significativa para aplicar conocimientos avanzados en ciencia de datos y aprendizaje automático, con el propósito de aportar valor tangible tanto a la entidad financiera como al sector automotriz en general. Al desarrollar una herramienta predictiva precisa y funcional, se contribuye a mitigar la incertidumbre financiera, mejorar la toma de decisiones estratégicas y fomentar prácticas más sostenibles y eficientes en el mercado de vehículos eléctricos.

1.2. Antecedentes

1.2.1. Descripción de la organización

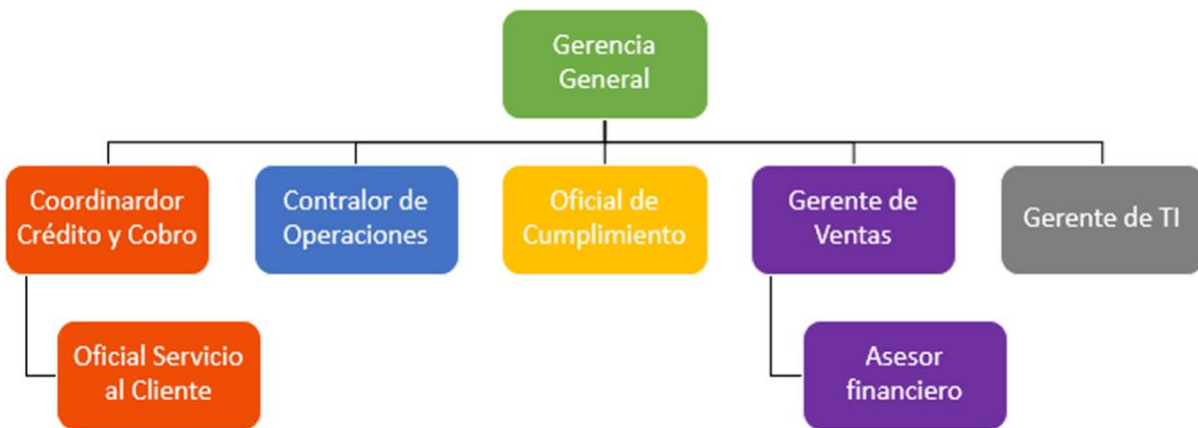
En el año 2005, los miembros de la familia Volio se convierten en los representantes oficiales de BMW Motorrad en Costa Rica; de esta forma, se empezó la incursión en el área automotriz y venta de estos vehículos a nivel nacional. En 2007, adquirieron Fischel, la cual es la cadena de farmacias líder en Costa Rica, con más de 70 puntos de venta en ese momento. Para el año 2008, se convirtieron en los representantes oficiales de BMW para Nicaragua, extendiendo su presencia a nivel centroamericano. Luego, en 2009, se involucraron en el área de bienes raíces, comprendiendo que puede ser un nuevo mercado el cual pueden explotar.

Durante el año 2010, incluyeron la marca MINI a su catálogo, convirtiéndose así en los representantes oficiales de esta en Costa Rica. En 2015, tomaron la decisión de vender su participación mayoritaria en CEFA Fischel a Cuestamoras, y en poco tiempo añadieron una nueva marca a su catálogo, la cual fue KTM. Por último, se centralizaron las diferentes empresas en su casa matriz, denominada Grupo 823, y las marcas automotrices se consolidaron bajo el portafolio de Red Motors en el año 2016.

Los clientes presentaron interés en las marcas ofrecidas en el portafolio de Red Motors y se identificó una oportunidad de ofrecer un nuevo servicio, por lo que la familia Volio decidió formar Áltica, el socio financiero para adquirir activos de forma rápida a sus clientes, y cuyos servicios contemplan alquiler, *leasing*, créditos prendarios, entre otros. De esta forma, su alianza estratégica con Red Motors les permite proveer financiamiento para la adquisición de vehículos y motos de tipo premium. No obstante, no se detuvieron ahí y hoy en día ofrecen estos servicios de financiamiento a otras marcas establecidas en el territorio nacional.

Seguidamente, en la **Figura 1**, se puede apreciar el organigrama actual de la empresa, para comprender mejor su composición.

Figura 1. *Organigrama de la organización*



Fuente: Elaboración propia a partir de la información brindada por el gerente de TI (2024).

1.2.2. Misión

La entidad financiera Áltica define la siguiente misión para cumplir con sus operaciones:

Proveer soluciones financieras que apoyen el cumplimiento de sueños en comunicados donde operemos y que generen resultados de excelencia (Áltica, s.f.).

1.2.3. Visión

Por su parte, Áltica define la siguiente visión para cumplir con sus operaciones: Ser parte de experiencias excepcionales para nuestros *stakeholders* (Áltica, s.f.).

1.2.4. Valores

Los valores presentes dentro de la organización son detallados a continuación:

- Pasión: Emprendemos cada proyecto con pasión.
- Integridad: La integridad es nuestro pilar de trabajo.
- Competitividad: Transformamos estrategia en competitividad.
- Compromiso: Crecemos a base de esfuerzo y compromiso.

1.2.5. Trabajos similares realizados dentro y fuera de la organización

Después de la revisión documental, mediante el uso de bases de datos internas y externas, esta sección incluye trabajos similares y proyectos relacionados con el tema del valor de recuperación de vehículos eléctricos y los modelos de predicción.

1.2.5.1. Trabajos dentro de la organización

En la **Tabla 1**, se detallan los proyectos internos que han sido realizados dentro de la organización y que actuarán como insumo para el desarrollo del presente proyecto.

Tabla 1. *Proyectos internos*

Autor/Encargado	Título	Objetivo	Referencia
Luis David Acuña Campos	Propuesta de solución de inteligencia de negocios para automatizar la generación de reportes	En este documento se detallan dos propuestas de solución frente a las necesidades y problemática presentes en la empresa Ática Socios Financieros, S.A para la generación de reportes. La organización tiene necesidades específicas de reportes, integrando información de tres sistemas; además, solo un colaborador es responsable y tiene los conocimientos para realizarlos. Este proceso es manual, lo que implica una serie de desventajas como el tiempo para generar los reportes y la entrega oportuna de estos.	https://repositoriotec.tec.ac.cr/handle/2238/11045

Fuente: Elaboración propia (2025).

1.2.5.2. Trabajos fuera de la organización

En la **Tabla 2**, se detallan los proyectos externos que han sido realizados dentro de la organización y que actuarán como insumo para el desarrollo del presente proyecto.

Tabla 2. *Proyectos externos*

Autor/Encargado	Título	Objetivo	Referencia
Laura Roberson	Accelerating Electric Vehicle Adoption in the United States: The Impact of Experience, Incentives, and Resale Value	En los últimos años, con el creciente enfoque en el cambio climático y las tecnologías beneficiosas para el medio ambiente, la industria del transporte ha centrado su atención en el desarrollo de vehículos eléctricos enchufables (PEV). Los vehículos eléctricos modernos se reintrodujeron en el mercado estadounidense en 2011, pero, a pesar de sus beneficios tanto para el medio ambiente como para los consumidores, la adopción de esta tecnología de vehículos sigue siendo baja.	https://www.proquest.com/docview/3014023436/abstract/4D8FF66370BA4C5DPQ/1?accountid=27651&sourcetype=Dissertations%20%20Theses
Patil Gautam, Gayatri Pode, Ramchandra B. Pode, Godwin Kafui Ayetor and Boucar Diouf	Depreciation in the Electric Vehicle Transition: Sustainability of the Second-Hand Electric Vehicle Market	Su objetivo consiste en analizar los factores que impulsan la rápida depreciación de los vehículos eléctricos, en especial la evolución tecnológica, el papel de las baterías y la influencia de los subsidios, y cómo estos aspectos afectan el mercado de vehículos eléctricos de segunda mano.	https://www.researchgate.net/publication/386474117_Depreciation_in_the_Electric_Vehicle_Transition_Sustainability_of_the_Second-Hand_Electric_Veh

			icle_Market
Stella Domínguez, Samuel Pedreros, David Delgadillo and John Anzola	A Depreciation Method Based on Perceived Information Asymmetry in the Market for Electric Vehicles in Colombia	El objetivo del artículo es proponer un nuevo método de depreciación para vehículos eléctricos que mejore la precisión de la valoración de los vehículos eléctricos usados. Este método, basado en un modelo de selección adversa que incorpora percepciones de calidad extraídas de plataformas de comercio electrónico, busca mitigar la asimetría de información y la selección adversa, permitiendo así establecer precios de mercado más justos y confiables.	https://www.scilit.com/publications/3d39f0589990adede172c1bff28464b9

Fuente: Elaboración propia (2025).

1.3. Planteamiento del problema

En esta sección, se describe la situación problemática hallada dentro del entorno de la organización, la cual motiva el desarrollo del proyecto, así como la mención de los beneficios esperados de su ejecución. Adicionalmente, se mencionan los beneficios esperados con la posible solución del problema identificado.

1.3.1. Situación problemática

En los últimos años, Costa Rica ha experimentado un crecimiento significativo en la adopción de vehículos eléctricos, el cual refleja los esfuerzos del país por avanzar hacia la movilidad sostenible. Este aumento no solo incluye modelos de marcas reconocidas internacionalmente como BMW, Toyota y Nissan, sino también una gran cantidad de vehículos provenientes de marcas chinas emergentes. Esta sobreoferta ha generado un mercado saturado donde los consumidores tienen una amplia gama de opciones para elegir al momento de adquirir un vehículo eléctrico. Según los datos del Ministerio de Ambiente y Energía (MINAE) (2024), se ha registrado un crecimiento acumulado en las diferentes categorías de vehículos eléctricos, tal como se muestra en la **Figura 2**.

Figura 2. Acumulado de vehículos eléctricos en Costa Rica año 2024

ACUMULADO DE VEHÍCULOS ELÉCTRICOS					
TIPO	Automó,	Motos	Especial	Trabajo	TOTAL
<=2010	72	34	113	14	233
2011	85	39	165	15	304
2012	96	56	245	15	412
2013	103	89	290	40	522
2014	115	144	376	43	678
2015	119	170	463	43	795
2016	130	195	648	43	1.016
2017	156	278	763	43	1.240
2018	349	356	910	46	1.661
2019	806	476	1063	48	2.393
2020	1417	659	1162	62	3.300
2021	2443	846	1263	81	4.633
2022	4007	1069	1326	124	6.526
2023	8889	1242	1712	375	12.218
2024	18536	1358	1980	858	22.732

Fuente: Ministerio de Ambiente y Energía (2024).

Este crecimiento demuestra el impacto de las políticas públicas y los incentivos fiscales, así como un aumento de la conciencia ambiental entre los costarricenses.

Sin embargo, esta diversidad y abundancia de opciones ha traído consigo desafíos importantes. Tras los primeros dos a tres años de uso, se ha observado que los vehículos eléctricos presentan condiciones muy variables en cuanto a su estado general y, específicamente, en la salud de sus baterías. Algunos vehículos mantienen sus baterías en excelente estado, conservando una alta cantidad de ciclos de vida útil, lo que garantiza un rendimiento óptimo y una estabilidad en su valor de recuperación. Por el contrario, otros experimentan una degradación acelerada de sus baterías, reduciendo significativamente su autonomía y funcionalidad, lo que impacta negativamente en su valor de mercado.

Esta disparidad en el estado de los vehículos usados plantea un reto considerable para las entidades financieras involucradas en su financiamiento. Estas instituciones enfrentan dificultades para estimar con precisión el valor de recuperación de los vehículos eléctricos en un momento determinado. El proceso actual no considera variables críticas como el estado de la batería, la marca, el modelo, el historial de mantenimiento y las tendencias del mercado, lo que puede llevar a sobreestimar el valor del activo financiado.

La consecuencia directa de esta imprecisión es el incremento en el riesgo financiero. Si una entidad financiera calcula incorrectamente el valor de rescate de un vehículo eléctrico, podría enfrentar pérdidas significativas debido a la estimación imprecisa de su valor de recuperación. Esto no solo afecta la rentabilidad de la financiera, sino que también puede influir en las políticas de crédito, haciéndolas más restrictivas y limitando el acceso de los consumidores a opciones de financiamiento asequibles.

Además, el mercado de vehículos eléctricos está influenciado por factores adicionales como la rápida evolución tecnológica —en poco tiempo los modelos pueden encontrarse obsoletos— y la falta de infraestructura adecuada para el mantenimiento y carga de estos vehículos en ciertas regiones. Asimismo, existe una variabilidad en la calidad y durabilidad de las baterías entre diferentes marcas y modelos, especialmente entre las marcas menos reconocidas o de reciente ingreso en el mercado costarricense.

Otro aspecto crítico es la escasez de datos históricos y estadísticos sobre el desempeño y depreciación de vehículos eléctricos en el país. La escasa disponibilidad de información consolidada sobre importaciones, ventas, rendimiento de baterías y comportamiento del mercado

recuperación de dichos vehículos y como consecuencia del uso de métodos tradicionales no adecuados para estimar con precisión su valor de recuperación, la entidad financiera enfrenta un riesgo considerable de pérdidas financieras. Para optar por el grado de Licenciatura en Administración de Tecnología de Información, el presente Trabajo Final de Graduación (TFG) aborda esta problemática mediante el desarrollo de un modelo predictivo basado en inteligencia de negocios, aplicando modelos de aprendizaje automático para la estimación del valor de recuperación de los activos.

En la actualidad, el entorno financiero requiere herramientas que permitan optimizar la rentabilidad y mejorar la eficiencia operativa. El uso de un modelo predictivo, apoyado por conceptos de inteligencia de negocios, facilita la anticipación de la depreciación de los activos y la toma de decisiones informadas. El enfoque principal del proyecto está alineado con la aplicación de tecnologías de ciencia de datos, análisis predictivo y *machine learning*, lo cual se ajusta a las tendencias actuales de transformación digital en la industria financiera.

La investigación se fundamenta en la alta volatilidad y la incertidumbre en la valoración de vehículos eléctricos, lo que expone a las entidades financieras a riesgos significativos. En Costa Rica, la variabilidad en el estado de los vehículos debido a factores como el deterioro de baterías, el uso que reciben y las fluctuaciones en las condiciones del mercado complica la estimación precisa del valor de recuperación. Estudios recientes indican que los métodos tradicionales de valoración no capturan la complejidad dinámica de estos activos, generando posibles sobrevaloraciones y, en consecuencia, pérdidas financieras importantes (Mora, 2021). Por lo tanto, resulta crucial desarrollar soluciones predictivas que permitan una estimación más ajustada y confiable, optimizando la asignación de recursos y reduciendo el riesgo de impagos.

Adicionalmente, el acelerado proceso de transformación digital en el sector financiero, impulsado por la creciente disponibilidad de grandes volúmenes de datos y el desarrollo de nuevas tecnologías, abre la puerta a enfoques innovadores para la toma de decisiones. La integración de técnicas de *machine learning* y análisis estadístico avanzado en un modelo predictivo no solo facilitará la anticipación de la depreciación de los vehículos, sino que también fortalecerá la competitividad de la entidad. Al combinar datos de diversas fuentes y aplicar modelos que se adapten a las condiciones cambiantes del mercado, la solución propuesta permitirá a la institución financiera responder de manera proactiva a las tendencias emergentes y optimizar sus políticas de

crédito, contribuyendo así a la estabilidad financiera y a una mejora continua en la gestión de activos (Damodaran, 2012; Hyndman & Athanasopoulos, 2018).

La participación de un administrador de Tecnologías de Información (ATI) es crucial para el éxito de este proyecto, debido a la combinación de habilidades técnicas y conocimientos del negocio requeridos. El ATI posee un conocimiento profundo del negocio y una comprensión sólida del área financiera, lo cual es esencial para traducir las necesidades de la entidad en requerimientos tecnológicos. Además, el manejo efectivo de los involucrados o *stakeholders* y la capacidad de adquirir información relevante y necesaria permiten definir con precisión los objetivos y alcances del proyecto.

En términos técnicos, el ATI tiene la capacidad de liderar la exploración de datos, así como el modelado y la aplicación de técnicas de aprendizaje automático. Asimismo, cuenta con la experiencia necesaria para determinar y validar el modelo predictivo correcto que brinde resultados precisos y beneficiosos para la entidad financiera. Esta combinación de competencias asegura que el modelo predictivo desarrollado no solo sea técnicamente sólido, sino también alineado con los objetivos estratégicos del negocio, garantizando un impacto positivo en la toma de decisiones y en la eficiencia de la gestión de activos por parte de la entidad financiera Áltica.

1.3.3. Beneficios esperados o aportes del Trabajo Final de Graduación

El desarrollo e implementación de un modelo predictivo para estimar el valor futuro de vehículos eléctricos financiados en Costa Rica trae consigo una serie de beneficios que se clasifican en directos e indirectos. Estos beneficios impactan tanto a la entidad financiera como al mercado y a la sociedad en general.

1.3.3.1. Beneficios directos

A continuación, se presentan los beneficios directos identificados:

1. Mejora en la precisión de la valoración de activos: Desarrollo e implementación de un modelo funcional del modelo predictivo en un periodo de seis meses, el cual permitirá a la financiera comenzar a estimar con mayor precisión el valor futuro de los vehículos eléctricos en segmentos específicos. Esto posibilitará la ejecución de ajustes iniciales en

las condiciones de financiamiento y una reducción inmediata del riesgo de pérdidas por fluctuaciones en el valor de recuperación.

2. Evaluación de escenarios financieros en solicitudes de crédito: El modelo predictivo facilitará la evaluación de diferentes escenarios financieros relacionados con el financiamiento de vehículos eléctricos, ayudando a minimizar incertidumbres en áreas clave.
3. Integración de alianzas estratégicas con concesionarias: Permitirá establecer alianzas estratégicas con concesionarias, posibilitando el acceso a datos más precisos y actualizados sobre rendimiento y depreciación. Esto mejorará la precisión del modelo predictivo y fortalecerá la posición de la financiera en el mercado.
4. Mejora en la gestión del portafolio de activos: La entidad financiera podrá monitorear y evaluar continuamente el valor de los activos financiados, lo cual le permitirá una gestión más dinámica y proactiva de su cartera.

1.3.3.2. Beneficios indirectos

A continuación, se presentan los beneficios indirectos identificados:

1. Fomento de la confianza en el mercado de vehículos eléctricos: Al proporcionar estimaciones más confiables sobre el valor de los vehículos eléctricos, se incrementará la confianza de los consumidores y las empresas de este mercado emergente, promoviendo su adopción.
2. Contribución al desarrollo sostenible: Al facilitar el financiamiento y aumentar la accesibilidad de los vehículos eléctricos, se apoyará la reducción de emisiones contaminantes y se contribuye a los objetivos ambientales y de sostenibilidad del país.
3. Mejora de las decisiones financieras: La información proporcionada por el modelo predictivo permitirá mejorar la calidad de las decisiones financieras, facilitando un análisis más detallado para la aprobación y gestión de créditos, contribuyendo a una asignación de recursos más efectiva a mediano plazo.
4. Impulso a la innovación tecnológica: El uso de técnicas avanzadas de análisis y ciencia de datos estimulará la innovación dentro de la entidad financiera y podrá inspirar a otras organizaciones a adoptar tecnologías similares.

5. Generación de conocimiento y experiencia: El proyecto aportará conocimientos valiosos al sector financiero y académico sobre la depreciación y valoración de activos tecnológicos, lo que puede ser utilizado en futuras investigaciones y desarrollos.
6. Fortalecimiento de la competitividad de la financiera: Al adoptar herramientas analíticas avanzadas, la entidad se posicionará como líder en innovación y gestión de riesgos, mejorando su reputación y atrayendo a nuevos clientes.
7. Mejora en la satisfacción del cliente: Al ofrecer condiciones de financiamiento más justas y ajustadas a la realidad del mercado, los clientes percibirán un mayor valor y confianza en los servicios de la financiera.
8. Desarrollo del mercado secundario: Una valoración más precisa de los vehículos eléctricos usados podrá dinamizar el mercado de reventa, ofreciendo más opciones a los consumidores y prolongando el ciclo de vida útil de los vehículos.

1.4. Objetivos del Trabajo Final de Graduación

A continuación, se detallan el objetivo general y los objetivos específicos del proyecto, los cuales tienen el propósito de resolver la problemática planteada anteriormente.

1.4.1. Objetivo general

Desarrollar un modelo funcional de un modelo predictivo analítico del valor de recuperación de los vehículos eléctricos financiados por la entidad, asegurando su exactitud y confiabilidad mediante pruebas y validación técnica, durante el primer semestre de 2025.

1.4.2. Objetivos específicos

1. Recolectar datos históricos y relevantes sobre el mercado de vehículos eléctricos en Costa Rica, incluyendo factores como ventas, valor actual del mercado y rendimiento de las baterías, a partir de una base sólida de información confiable.
2. Analizar los datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica, considerando los factores mencionados, los cuales influyen en el valor de recuperación, para su integración en el modelo.
3. Diseñar un modelo estadístico y de aprendizaje automático que integre los factores identificados, para la predicción del valor de recuperación de los vehículos eléctricos financiados por la entidad.
4. Validar el modelo predictivo mediante pruebas con datos reales y técnicas de validación cruzada, para el aseguramiento de su exactitud y confiabilidad en la estimación del valor de recuperación de los vehículos eléctricos.

1.5. Alcance

El alcance del proyecto se define considerando los objetivos generales y específicos planteados. Este abarca desde la adquisición de información necesaria para la elaboración del modelo predictivo hasta su implementación y validación dentro de la entidad financiera. Para su ejecución, se desarrolla un modelo predictivo que permite estimar el valor futuro de los vehículos eléctricos financiados, con el objetivo de reducir los riesgos financieros y optimizar la rentabilidad.

Las actividades por realizar en el proyecto incluyen:

- Recolección de la información:

En la primera fase del proyecto, se procede con la adquisición de información, de manera que se recopilan datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica. Esto incluye información sobre ventas, importaciones, valor actual del mercado y rendimiento de baterías. La recopilación de datos es crucial para identificar los factores clave que influyen en el valor de reventa de los vehículos eléctricos, estableciendo la base para el análisis posterior.

- Análisis exploratorio de datos:

La segunda fase comprende el análisis exploratorio de datos y su preparación. En esta fase, se aplican técnicas de análisis exploratorio que permiten comprender mejor la naturaleza de los datos, detectar patrones relevantes y limpiar la información para asegurar la calidad y consistencia de los datos utilizados. Esta fase es esencial para garantizar que el modelo predictivo cuente con datos sólidos y representativos.

- Diseño del modelo predictivo:

En la tercera fase, se elabora el diseño del modelo predictivo. Para ello, se diseñan y desarrollan diversos modelos estadísticos y de aprendizaje automático, evaluando el rendimiento de cada uno para determinar cuál se ajusta mejor a la predicción del valor futuro de los vehículos eléctricos. Esto incluye la aplicación de diferentes técnicas de regresión y la selección del modelo más adecuado a partir del análisis de métricas de desempeño como la precisión y la capacidad de generalización.

- Validación del modelo:

La validación del modelo es la cuarta fase del proyecto. En este punto, se llevan a cabo pruebas de confiabilidad y desempeño del modelo para asegurar su exactitud y precisión. Esta fase permite ajustar y mejorar el modelo predictivo, garantizando que sus predicciones sean precisas y útiles para la entidad financiera.

Cada una de estas fases está diseñada para cumplir con los objetivos específicos del proyecto, garantizando que el modelo predictivo sea preciso, confiable y alineado con las necesidades de la entidad financiera.

1.6. Supuestos

A continuación, se especifican aspectos que no forman parte del alcance de la investigación:

- Disponibilidad y calidad de datos: Se asume que los datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica están disponibles y son suficientes y de alta calidad para entrenar y validar el modelo predictivo. Esto es fundamental para asegurar la precisión del modelo.
- Participación activa de los *stakeholders*: Se considera que los *stakeholders* de la entidad financiera participarán de manera activa durante todo el ciclo de vida del proyecto, proporcionando información y retroalimentación relevante que permita alinear el desarrollo del modelo con las necesidades del negocio.

1.7. Entregables

A continuación, se presentan los entregables que serán producto del desarrollo de la metodología del proyecto.

1.7.1. Entregables del producto

En esta sección, se detallan los entregables del presente proyecto para Áltica, los cuales están organizados según los objetivos específicos planteados.

- **Objetivo específico n°.1:** Recolectar datos históricos y relevantes sobre el mercado de vehículos eléctricos en Costa Rica, incluyendo factores como ventas, valor actual del mercado y rendimiento de las baterías, a partir de fuentes confiables.
 - Informe de adquisición de datos: Este informe documenta las fuentes de datos, los métodos de adquisición y la calidad de los datos recopilados, asegurando que se cuenta con la información necesaria y relevante para el desarrollo del modelo predictivo.
 - Reporte de análisis exploratorio de datos: Se generó un informe detallado del análisis exploratorio de datos realizado. Este entregable describe los patrones identificados, la calidad de los datos y las modificaciones realizadas para la limpieza y preparación de los datos.

- **Objetivo específico n°.2:** Analizar los datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica, considerando los factores mencionados, los cuales influyen en el valor de recuperación, para su integración en el modelo.
 - Informe de adquisición de datos: Este informe documenta las fuentes de datos, los métodos de adquisición y la calidad de los datos recopilados, asegurando que se cuenta con la información necesaria y relevante para el desarrollo del modelo predictivo.
 - Reporte de análisis exploratorio de datos: Se generó un informe detallado del análisis exploratorio de datos realizado. Este entregable describe los patrones identificados, la calidad de los datos y las modificaciones realizadas para la limpieza y preparación de los datos.

- **Objetivo específico n°.3:** Diseñar un modelo estadístico y de aprendizaje automático que integre los factores identificados, para la predicción del valor de recuperación de los vehículos eléctricos financiados por la entidad.
 - Modelo predictivo: Se desarrolló un modelo predictivo que se ajusta al objetivo de diseñar un modelo estadístico y de aprendizaje automático. Este modelo incluye el código y los algoritmos empleados, con explicaciones sobre el enfoque utilizado

para integrar los factores relevantes y predecir el valor futuro de los vehículos eléctricos.

- **Objetivo específico n°.4:** Validar el modelo predictivo mediante pruebas con datos reales y técnicas de validación cruzada, para el aseguramiento de su exactitud y confiabilidad en la estimación del valor de recuperación de los vehículos eléctricos.
 - Informe de validación del modelo: En la fase de validación, se generó un informe que documenta las pruebas y validaciones cruzadas realizadas, así como los resultados obtenidos y los ajustes realizados para mejorar la exactitud y confiabilidad del modelo.

1.7.2. Entregables de gestión

En esta sección, se describen los artefactos asociados a la gestión del proyecto.

1.7.2.1. Minutas

La plantilla para las minutas que se realizaron a lo largo de la ejecución del proyecto se encuentra en el **Apéndice B**.

1.7.2.2. Gestión del cambio

La plantilla para el control de los cambios efectuados durante la ejecución del proyecto se encuentra en el **Anexo I**.

1.7.2.3. Cronograma

El cronograma posee la finalidad de informar y gestionar, entre la organización, la academia y el autor, el desarrollo del presente proyecto, incluyendo las fases y fechas relevantes. El cronograma planeado se encuentra en el **Apéndice A**.

1.8. Limitaciones

Seguidamente, se enlistan algunos elementos o factores que podrían llegar a restringir el desarrollo de la investigación:

- **Calidad y disponibilidad de datos:** La precisión del modelo depende en gran medida de la calidad y cantidad de datos disponibles sobre las transacciones de vehículos eléctricos en el país. Actualmente, no existe una base de datos centralizada y accesible que contenga

información detallada sobre precios de reventa, historial de mantenimiento, uso del vehículo y estado de la batería. Esto puede generar sesgos en el modelo al basarse en conjuntos de datos parciales o inconsistentes.

- Fuentes de datos limitadas: La recopilación de datos se apoya en plataformas de comercio electrónico como CRAutos y bases de datos internas de la entidad financiera, lo que puede no representar el mercado en su totalidad. Esto podría introducir sesgos en la estimación del valor residual, ya que ciertos segmentos del mercado podrían estar subrepresentados, como transacciones privadas o ventas directas en concesionarios.
- Uso de variables indirectas: Dado que no siempre se tiene acceso a datos específicos sobre el estado real de la batería y otros componentes críticos, el modelo podría utilizar variables indirectas, como el año del modelo y el kilometraje acumulado, para inferir la degradación del vehículo. Sin embargo, estas variables pueden no ser totalmente representativas del estado real del vehículo eléctrico, lo que podría afectar la precisión de las predicciones.
- Impacto de factores externos no modelados: Elementos como fluctuaciones en la economía, políticas gubernamentales, crisis de abastecimiento de baterías o incluso la percepción de los consumidores sobre la confiabilidad de ciertas marcas y modelos pueden influir en la depreciación de los vehículos eléctricos de maneras que el modelo no puede prever completamente.

2. Marco Conceptual

El marco conceptual constituye una sección fundamental en toda investigación académica, pues ofrece la base teórica y conceptual que sustenta el estudio. Este apartado permite al investigador precisar sus objetivos, refinar las hipótesis y establecer las relaciones entre las variables de interés, garantizando la solidez teórico-metodológica del proyecto (Hernández Sampieri et al., 2014, p. 149).

Además, el marco conceptual orienta la recolección y el análisis de datos al identificar las variables y relaciones que se examinarán en el estudio, garantizando que la investigación esté enfocada y que los datos recopilados sean pertinentes y fiables. Del mismo modo, facilita la comunicación y la colaboración entre los miembros del equipo de investigación al proporcionar un lenguaje y una comprensión comunes.

En el contexto de esta investigación, es esencial comprender conceptos clave como la analítica de negocios y la inteligencia de negocios. Al integrarse en el marco conceptual, se establece una base sólida para analizar cómo las empresas pueden utilizar sus datos para mejorar su rendimiento y competitividad.

2.1. Analítica de negocios

La analítica de negocios, o BA (*business analytics*) por sus siglas en inglés, se concibe como una capacidad organizacional que convierte datos en conocimiento accionable mediante la aplicación coordinada de técnicas estadísticas, aprendizaje automático e inteligencia artificial (Liu et al., 2023). Esta perspectiva complementa la propuesta de Power et al. (2018), quienes subrayan que la verdadera ventaja competitiva emerge cuando los modelos predictivos y prescriptivos se integran en los procesos decisionales, lo que permite anticipar escenarios y recomendar la acción óptima. Desde esta óptica, la analítica de negocios trasciende la simple generación de reportes descriptivos al abarcar un ciclo completo que va desde la captura y depuración de datos hasta la automatización de recomendaciones estratégicas (Camm et al., 2020). En síntesis, la analítica de negocios actúa como el puente entre la inteligencia derivada de los datos y la obtención de

resultados tangibles, sustentando decisiones de alto impacto en un entorno empresarial cada vez más dinámico.

2.1.1. Tipos de analítica

La literatura reciente distingue cuatro niveles de analítica de negocios que se complementan de forma progresiva: descriptiva, diagnóstica, predictiva y prescriptiva (Camm et al., 2020; Sarker, 2021b). Cada uno responde a una pregunta distinta y requiere herramientas específicas:

Tabla 3. *Tipos de analítica*

Tipo	Pregunta central	Herramientas
Descriptiva	¿Qué ocurrió?	SQL, ETLs, dashboards, visualizaciones,
Diagnóstica	¿Por qué ocurrió?	Análisis <i>drill-down</i> , correlación de variables, consultas ad-hoc
Predictiva	¿Qué pasará?	Modelos de regresión, algoritmos de <i>machine learning</i> , redes neuronales, algoritmos de clasificación
Prescriptiva	¿Qué debemos hacer?	Optimización, simulación, motores de reglas, planificación de escenarios

Fuente: Adaptado de Camm (2020) y Sarker (2021b).

2.2. Inteligencia de negocios

La inteligencia de negocios, o BI (*business intelligence*) por sus siglas en inglés, se refiere al conjunto de estrategias, procesos, aplicaciones y tecnologías que permiten a las organizaciones recopilar, integrar, analizar y presentar datos empresariales relevantes. El objetivo principal de BI, según sugiere Tableau (s.f.), es respaldar la toma de decisiones informadas, y mejorar así el rendimiento y la competitividad de las empresas. Medina (2015) amplía esta definición al señalar que BI es “el conjunto de estrategias y herramientas enfocado a la administración y creación de

conocimiento mediante el análisis de datos existentes en una organización” (p. 9). Esto implica gestionar una empresa a partir de la información que genera, atendiendo las necesidades de información de ejecutivos y analistas para comprender mejor sus operaciones y definir estrategias de negocio acertadas (Medina, 2015).

2.2.1. Datos, información y conocimiento

En el contexto organizacional, es fundamental distinguir entre los conceptos de datos, información y conocimiento, ya que cada uno desempeña un papel crucial en la toma de decisiones y en la gestión empresarial. A continuación, se puntualiza cada uno de ellos.

- **Datos:** Son elementos brutos y sin procesar que representan hechos, cifras o descripciones sin un contexto específico. Estos constituyen la materia prima del conocimiento y, por sí solos, carecen de significado.
- **Información:** Se obtiene al procesar y organizar los datos, dotándolos de contexto y relevancia. La información es esencial para la toma de decisiones, ya que permite comprender situaciones específicas y facilita la identificación de patrones y tendencias.
- **Conocimiento:** Es el resultado de la interpretación y aplicación de la información en un contexto determinado. Implica la capacidad de utilizar la información para tomar decisiones estratégicas y operativas, y se considera un recurso incremental dentro de las organizaciones.

Estos tres componentes son fundamentales para la gestión efectiva del conocimiento en las organizaciones, ya que permiten transformar datos en información útil y, posteriormente, en conocimiento aplicable para la toma de decisiones estratégicas.

2.2.2. Relación y diferencias con analítica de negocios

En primer lugar, inteligencia de negocios y analítica de negocios comparten el propósito de convertir datos en conocimiento, pero se diferencian en el tipo de preguntas que responden y en el horizonte temporal que abordan. Mientras que BI se centra en describir y monitorizar el desempeño histórico—proveyendo paneles, informes y alertas sobre lo que ya ocurrió—BA avanza un paso más y, mediante modelos predictivos y prescriptivos, anticipa escenarios futuros y recomienda acciones concretas (Jiménez-Partearroyo et al., 2024).

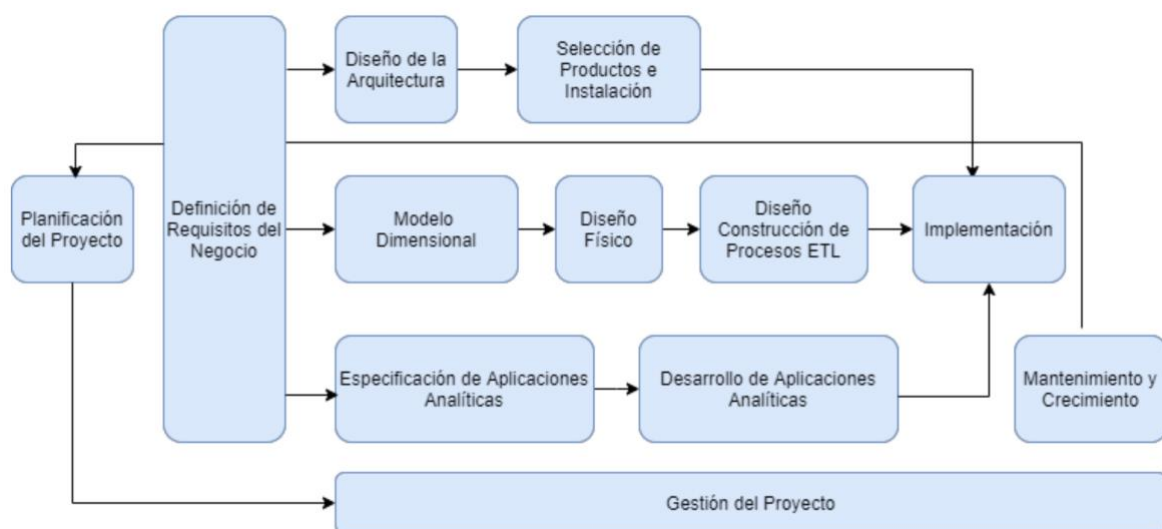
Asimismo, aunque las dos disciplinas suelen presentarse como conceptos separados, en la práctica conforman un continuo analítico. Por un lado, BI establece la infraestructura fundamental: procesos ETL, modelos de datos y almacenes corporativos que garantizan la calidad, consistencia y trazabilidad de la información. Por otro lado, BA se apoya en esa base para entrenar algoritmos, evaluar hipótesis y automatizar decisiones estratégicas (Phillips-Wren et al., 2021). Dicho de otro modo, sin un ecosistema robusto de BI, la analítica avanzada carecería de insumos confiables; sin BA, BI se limitaría a un rol meramente descriptivo y perdería la capacidad de generar ventaja competitiva futura.

En suma, BI y BA deben entenderse como capas complementarias: la primera ofrece la visión retrospectiva que asegura la gobernanza de los datos, y la segunda proyecta dicha visión hacia el futuro para apoyar decisiones de alto impacto en la organización.

2.2.2.1. Ciclo de vida de la inteligencia de negocios

En el ámbito de la inteligencia de negocios, es necesario seguir un proceso estructurado que garantice su funcionalidad y alineación con las necesidades del negocio. En este contexto, Ralph Kimball desarrolló la metodología conocida como el ciclo de vida dimensional del negocio, la cual se ha convertido en una de las más utilizadas en el diseño e implementación de soluciones de BI (Forero-Castañeda y Sánchez-García, 2021).

Figura 4. Metodología de Ralph Kimball



Fuente: Tomado de Forero-Castañeda y Sánchez-García (2021, p. 9).

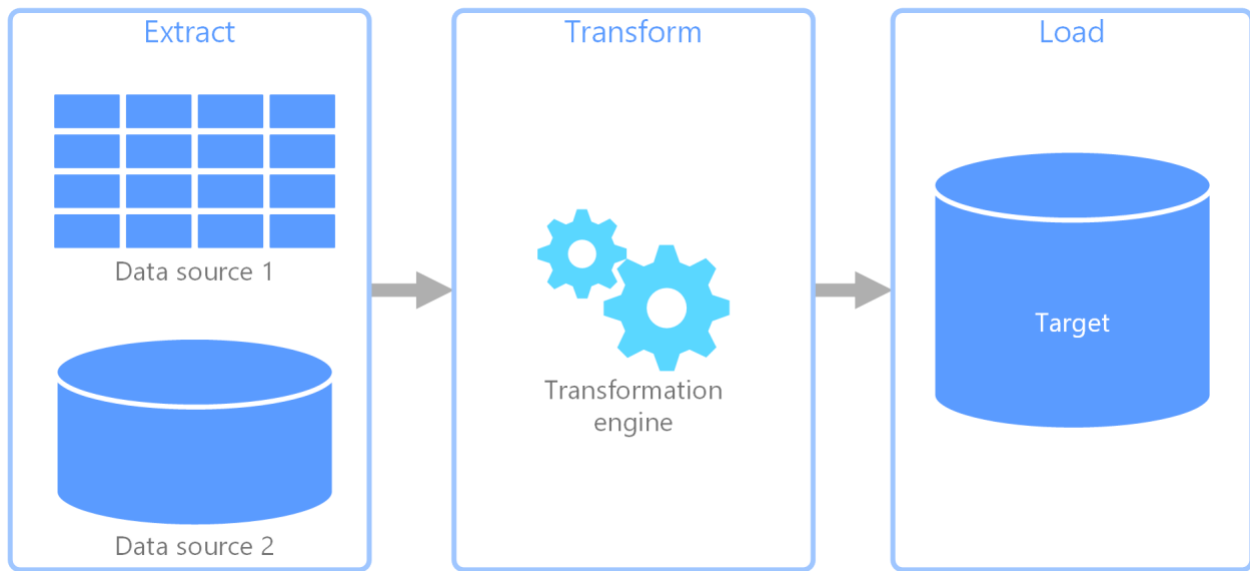
La metodología de Kimball se fundamenta en cuatro principios básicos, los cuales permiten estructurar los datos; estos principios corresponden a:

1. **Centrarse en el negocio:** Antes de proceder con el desarrollo del modelo de datos, es imprescindible contar con un profundo conocimiento del negocio. Esto permite definir un esquema de requerimientos adecuado que refleje con precisión las necesidades de la organización.
2. **Construir una infraestructura de información adecuada:** La metodología enfatiza la necesidad de realizar un análisis exhaustivo de la información para diseñar modelos eficientes, asegurando la integridad y accesibilidad de los datos.
3. **Realizar entregas en incrementos significativos:** Para evitar retrasos o entregas ineficientes, la metodología recomienda desarrollar el proyecto en fases progresivas. Estas iteraciones deben acordarse con la organización para garantizar que los plazos sean óptimos y factibles.
4. **Ofrecer una solución completa:** La metodología de Kimball exige entregar un sistema funcional que incluya herramientas de análisis, reportes visuales, capacitación para los usuarios y soporte técnico. Esto garantiza que la solución cumpla con las expectativas y requerimientos de la organización.

2.2.3. Extracción, transformación y carga (ETL)

Según Microsoft (2025), la canalización de extracción, transformación y carga —o *Extract, Transform and Load* (ETL), por sus siglas en inglés— es un proceso fundamental para el manejo de datos, ya que permite recopilar información proveniente de diversas fuentes y formatos, transformarla aplicando reglas de negocio específicas y, finalmente, cargarla en un almacén de datos u otro repositorio de destino. En la fase de extracción, se recuperan los datos originales sin alterarlos, minimizando el impacto en los sistemas de origen. Durante la transformación, se aplican diversas operaciones—como el filtrado, la ordenación, la agregación, la combinación, la limpieza y la validación—para asegurar que la información cumpla con los estándares y requerimientos del negocio. Finalmente, en la etapa de carga, los datos procesados se integran en el sistema destino, facilitando la generación de reportes, análisis y la toma de decisiones estratégicas.

Figura 5. Diagrama del proceso de extracción, transformación y carga (ETL)



Fuente: Tomado de Microsoft (2025).

2.2.3.1. Bases de datos

Marqués (2011) define una base de datos como “un conjunto de datos almacenados en memoria externa que están organizados mediante una estructura de datos. Cada base de datos ha sido diseñada para satisfacer los requisitos de información de una empresa u otro tipo de organización” (p. 2).

Siguiendo esta línea, Oracle (2020) establece que “una base de datos es una colección organizada de datos que se almacena y se puede acceder electrónicamente desde un sistema informático” (párr. 1). Esta definición resalta la importancia de contar con un sistema que no solo almacene la información de forma estructurada, sino que además permita su eficiente gestión y recuperación, lo cual es esencial para apoyar procesos analíticos y operativos en cualquier organización.

2.2.3.2. Visualización y análisis de datos

La visualización y el análisis de datos constituyen etapas fundamentales en el procesamiento de la información, ya que permiten transformar grandes volúmenes de datos complejos en representaciones gráficas y modelos analíticos de fácil interpretación. De esta manera, se pueden identificar patrones, tendencias y anomalías que son centrales para la toma de

decisiones estratégicas. En el contexto del presente proyecto, que busca desarrollar un modelo predictivo para reducir pérdidas financieras derivadas de la depreciación de vehículos eléctricos en Costa Rica, la correcta visualización de la información es crucial para detectar, de manera temprana, las variables y relaciones que influyen en la evolución del valor de los activos.

De acuerdo con Hashemi-Pour, Brush y Burns (2024), la visualización de datos “traduce la información a un contexto visual para que el cerebro humano comprenda con mayor facilidad y extraiga conclusiones, identificando patrones, tendencias y valores atípicos en grandes conjuntos de datos” (párr. 7). Así, mediante el uso de técnicas de análisis estadístico, minería de datos y herramientas de visualización, se logra no solo presentar la información de forma clara, sino también proporcionar una base sólida para validar y ajustar los modelos predictivos. En este sentido, una adecuada visualización y análisis de los datos permitirá a la entidad identificar rápidamente las tendencias del mercado y los factores críticos que afectan la depreciación de los vehículos eléctricos, optimizando así la toma de decisiones financieras.

2.3. Inteligencia de negocios en las finanzas

La inteligencia de negocios en el ámbito financiero se define como el proceso integral que combina la recolección, análisis y visualización de datos para apoyar la toma de decisiones estratégicas en las instituciones financieras. Este enfoque permite transformar grandes volúmenes de información en conocimiento útil y accionable, lo que resulta fundamental para responder a la dinámica cambiante del mercado y a los desafíos regulatorios. Según Mora (2021), las decisiones con mayor impacto financiero son aquellas que se basan en datos confiables, por lo que la adopción de sistemas de inteligencia de negocios se presenta como una herramienta competitiva esencial para mejorar la eficiencia, gestionar riesgos y fomentar la innovación en el sector.

2.3.1. Conceptos fundamentales

En esta sección se presentan los conceptos básicos que sustentan la valoración y depreciación de activos financieros, fundamentales para el análisis y la toma de decisiones en el ámbito financiero. Se exploran definiciones esenciales como la de ‘activo’, que representa los recursos económicos que una entidad posee y de los que espera obtener beneficios futuros; ‘depreciación’, que se refiere al proceso mediante el cual se distribuye el costo de un activo a lo largo de su vida útil; ‘riesgo’, que corresponde a la incertidumbre que afecta los retornos financieros; y ‘pasivo’, que se relaciona con las obligaciones que deben ser saldadas en el futuro.

Estos conceptos proporcionan la base teórica necesaria para comprender cómo se evalúan y gestionan los activos y las deudas en una organización (Damodaran, 2012).

2.3.1.1. Activo

Un activo se define como un recurso económico que una entidad posee o controla y del cual se espera que se generen beneficios económicos futuros. En el ámbito financiero, los activos incluyen no solo recursos tangibles (como edificios o vehículos), sino también intangibles (como patentes o marcas). Según Brigham y Ehrhardt (2014), los activos son fundamentales para la operación y crecimiento de una empresa, pues representan la base sobre la cual se generan ingresos.

2.3.1.2. Valoración de activo

La valoración de activo es el proceso mediante el cual se determina el valor económico actual de un recurso, basándose en la proyección de los flujos de caja futuros esperados y en el análisis de variables internas y externas que afectan su desempeño. Este proceso es fundamental para evaluar el potencial de generación de ingresos de un activo y para establecer estrategias de inversión o financiamiento. Según Damodaran (2012), “la valoración de un activo implica estimar su valor presente descontando los flujos de caja futuros a una tasa que refleje el riesgo inherente y el costo de oportunidad del capital” (p. 25). Esta definición destaca la importancia de contar con estimaciones precisas para la gestión de riesgos financieros y para la toma de decisiones estratégicas en el sector.

2.3.1.3. Depreciación de activo

La depreciación es el proceso contable mediante el cual se distribuye el costo de un activo a lo largo de su vida útil, reflejando la disminución de su valor debido al desgaste, obsolescencia o deterioro. Este proceso permite que el valor contable del activo se ajuste a la realidad del mercado y se reconozca el gasto en los periodos en que el activo contribuye a la generación de ingresos. Damodaran (2012) explica que la depreciación es crucial para determinar el valor residual y, en consecuencia, para la toma de decisiones de inversión y financiamiento.

2.3.1.4. NIC 16: Propiedad, planta y equipo en vehículos eléctricos

La NIC 16 establece los lineamientos para el reconocimiento, medición, depreciación y baja de los activos fijos, que, en el caso de los vehículos eléctricos, son fundamentales para determinar su valor contable y residual. En este contexto, la aplicación de la NIC 16 permite que

las entidades registren el costo inicial de estos vehículos y calculen su depreciación de forma sistemática a lo largo de su vida útil, considerando factores como el desgaste normal, la obsolescencia tecnológica y las condiciones de uso que afectan su rendimiento. Además, al utilizar esta norma, se logra una presentación coherente y comparable de la información financiera, lo que es crucial para evaluar el desempeño y la estabilidad de los activos en el tiempo (IFRS Foundation, 2023).

2.3.1.5. NIC 36: Deterioro del valor de activos en vehículos eléctricos

La NIC 36 se centra en la evaluación y reconocimiento de cualquier deterioro en el valor de los activos, asegurando que estos se reflejen en los estados financieros a un valor que no exceda su monto recuperable. En el ámbito de los vehículos eléctricos, dicha norma adquiere especial relevancia debido a la volatilidad del mercado, la rápida evolución tecnológica y otros factores externos que pueden afectar el valor de estos activos. Mediante la aplicación de la NIC 36, las empresas deben realizar evaluaciones periódicas del valor recuperable de sus vehículos y, en caso de identificar un deterioro, ajustar el valor contable para reflejar con precisión el valor de mercado actual. Este proceso resulta fundamental para prevenir la sobrevaloración de los activos y mitigar riesgos financieros asociados, proporcionando una visión clara y actualizada del desempeño económico del parque automotriz eléctrico (IFRS Foundation, 2023).

2.3.1.6. Riesgo

El riesgo en finanzas se refiere a la incertidumbre asociada con la posibilidad de que los resultados reales difieran de las expectativas, lo cual puede afectar la rentabilidad y la estabilidad de los flujos de caja de un activo. Hull (2018) señala que, en la gestión financiera, el riesgo se mide a través de la variabilidad de los retornos y se utiliza para evaluar la exposición de la institución a pérdidas potenciales. Esta evaluación es esencial para establecer políticas de inversión y estrategias de mitigación adecuadas.

2.3.2. Gestión de riesgos financieros

La gestión de riesgos financieros se centra en identificar, medir y mitigar las amenazas que pueden afectar la estabilidad y rentabilidad de las instituciones financieras. Como afirma Hull (2018), este proceso requiere de la integración de datos de diversas fuentes y la aplicación de técnicas analíticas avanzadas para monitorear indicadores críticos y anticipar eventos adversos en tiempo real. En este sentido, la BI se utiliza para construir modelos que permitan evaluar y gestionar dichos riesgos de forma proactiva.

2.3.2.1. Riesgo crediticio

El riesgo crediticio se refiere a la probabilidad de que un prestatario no cumpla con sus obligaciones financieras, lo cual generaría pérdidas en la cartera crediticia de la institución. Según Hull (2018), la evaluación de este riesgo se realiza a través de modelos cuantitativos que consideran variables financieras, historiales de pago y condiciones macroeconómicas, lo que posibilita la formulación de políticas de crédito orientadas a minimizar la exposición a impagos. En otras palabras, la aplicación de BI en el ámbito del riesgo crediticio permite una revisión constante y detallada de los indicadores financieros relevantes para prevenir incumplimientos.

2.3.3. Valoración predictiva de activos

La valoración predictiva de activos consiste en el uso de modelos cuantitativos y técnicas de *machine learning* para proyectar el valor futuro de los activos financieros. Como afirma Damodaran (2012), este enfoque integra datos históricos y variables del entorno para anticipar la depreciación o apreciación de los activos, facilitando la toma de decisiones en la gestión del portafolio. De esta manera, las instituciones financieras pueden optimizar la asignación de recursos y mejorar la precisión en sus decisiones de inversión.

2.3.4. Factores determinantes del valor de recuperación de activos

El valor de recuperación, o valor residual, es la estimación del monto que se espera recuperar al final de la vida útil de un activo, descontando la depreciación acumulada. Este concepto es crucial para determinar la rentabilidad de las inversiones y para desarrollar la planificación financiera. Según Pinto et al. (2020), el valor de recuperación se ve influido por diversos factores que se pueden clasificar en dos grandes grupos: factores internos y externos, además del enfoque o modelo de valoración que se emplee.

2.3.4.1. Factores internos

Los siguientes son algunas de las consideraciones para la depreciación de activos a nivel interno:

- Condición física y mantenimiento: La calidad del mantenimiento y el estado operativo del activo influyen directamente en su capacidad de retener valor.
- Antigüedad y obsolescencia: La edad del activo y la rapidez con que se vuelve obsoleto impactan negativamente en su valor residual.

2.3.4.2. Factores externos

Por su parte, estas son algunas de las consideraciones para la depreciación de activos a nivel interno:

- Condiciones de mercado: Variables como la oferta y la demanda, la competencia y las tendencias económicas afectan la valoración final del activo.
- Entorno regulatorio y económico: Las normativas, políticas gubernamentales y condiciones macroeconómicas (por ejemplo, tasas de interés e inflación) también determinan el valor de recuperación.

La consideración integral de estos factores es esencial para obtener una estimación realista y precisa del valor residual, lo que permite a las instituciones financieras formular estrategias adecuadas (Pinto et al., 2019).

2.3.5. Modelos tradicionales y predictivos

Existen dos enfoques principales que pueden ser usados para la estimación del valor de recuperación:

- **Modelos tradicionales:**
 - Se basan en métodos contables clásicos, como el descuento de flujos de caja (DCF), que utilizan datos históricos para proyectar los flujos futuros y calcular el valor presente del activo.
 - Estos métodos ofrecen una base sólida para la valoración, pero pueden ser menos sensibles a cambios repentinos en las condiciones del mercado.
- **Modelos predictivos:**
 - Incorporan técnicas de análisis de *big data* y algoritmos de *machine learning* para integrar variables en tiempo real y captar dinámicas de mercado complejas.
 - Según Hyndman y Athanasopoulos (2018), estos modelos permiten generar proyecciones más ágiles y precisas, adaptándose a la volatilidad y a las condiciones cambiantes del entorno, lo que resulta especialmente relevante en sectores de rápida evolución tecnológica.

La combinación de ambos enfoques, tal como señalan expertos como Pinto et al. (2019) y Hyndman y Athanasopoulos (2018), puede mejorar la precisión en la estimación del valor residual, proporcionando a las instituciones financieras proyecciones robustas que faciliten la toma de decisiones estratégicas.

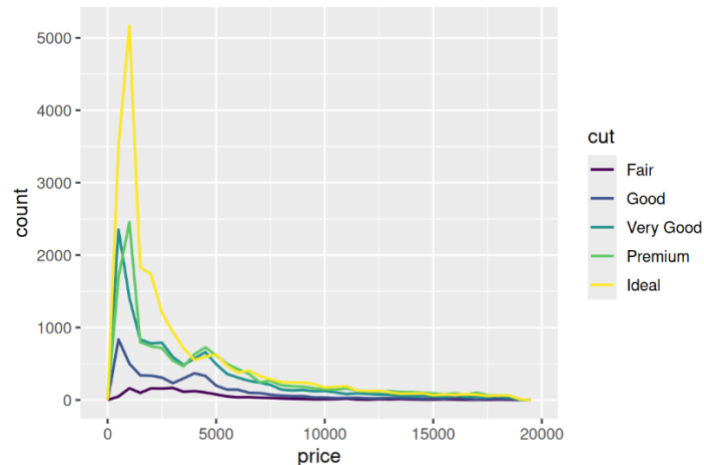
2.4. Ciencia de datos

La ciencia de datos se enfoca en analizar datos con el fin de extraer información significativa para las empresas. Se trata de un enfoque multidisciplinario que combina principios y prácticas de matemáticas, estadística, inteligencia artificial e ingeniería de computación para analizar grandes volúmenes de datos. Este análisis permite que los científicos de datos planteen y respondan preguntas fundamentales, tales como “¿qué pasó?”, “¿por qué pasó?”, “¿qué pasará?” y “¿qué se puede hacer con los resultados?”, lo que proporciona una base sólida para la toma de decisiones estratégicas y la optimización de procesos empresariales (Amazon Web Services [AWS], s.f.).

2.4.1. Metodología CRISP-DM

La metodología CRISP-DM (*Cross-Industry Standard Process for Data Mining*) es un marco estructurado y ampliamente adoptado para organizar proyectos de ciencia de datos y análisis predictivo. Según IBM (2021), CRISP-DM se compone de seis fases interrelacionadas:

1. **Comprensión del negocio:** Se define el problema, los objetivos y los criterios de éxito, asegurando que el proyecto esté alineado con las necesidades estratégicas de la organización.
2. **Comprensión de los datos:** Se explora y analiza la calidad, la estructura y la relevancia de los datos disponibles.
3. **Preparación de los datos:** Se limpian, transforman y organizan los datos para que sean aptos para el modelado.
4. **Modelado:** Se desarrollan y prueban diferentes modelos predictivos.
5. **Evaluación:** Se revisan los modelos para asegurar que cumplen con los objetivos de negocio y se selecciona el más adecuado.
6. **Despliegue:** Se implementa el modelo en un entorno operativo y se monitoriza su desempeño.



Fuente: Tomado de Wickham & Grolemund (2023).

2.4.2.1. Outliers

Los *outliers* se definen como aquellos elementos que muestran una discrepancia considerable respecto al comportamiento de la distribución de los datos restantes en una muestra estadística, al grado de alterar la representatividad del conjunto. Aunque esta definición general describe la naturaleza de los *outliers*, no proporciona criterios concretos para diferenciarlos de los valores “normales”. Por lo tanto, es necesario recurrir a definiciones adicionales que permitan categorizar y refinar la identificación de estos valores atípicos (Monhor & Takemoto, 2014).

2.4.3. Limpieza, transformación y normalización

La limpieza de datos se centra en identificar y corregir errores, valores faltantes y *outliers*, mientras que la transformación consiste en modificar el formato o estructura de los datos para adecuarlos a los requisitos del análisis. La normalización, por su parte, estandariza las variables para que tengan escalas comparables y así se facilite el procesamiento de los modelos predictivos. Wickham y Grolemund (2023) destacan que estos pasos son críticos para evitar sesgos y garantizar que los datos sean consistentes y utilizables para el modelado.

2.4.4. Técnicas de modelado predictivo

Para facilitar la comparación y selección de técnicas de modelado predictivo, se ha desarrollado el siguiente cuadro comparativo. Esta herramienta visual resume, de manera concisa, las definiciones, aplicaciones, ventajas y limitaciones de cada método, lo que permite al analista evaluar rápidamente las características de cada técnica y determinar cuál se adapta mejor a las necesidades del proyecto. Géron (2019) enfatiza la importancia de comprender en profundidad las

particularidades de cada método para mejorar la precisión de las predicciones y la eficiencia del proceso de modelado.

Tabla 4. *Cuadro comparativo de técnicas de modelado predictivo*

Técnica	Definición	Ventajas	Desventajas
Regresión lineal	Variante de la regresión lineal que incorpora un término de regularización L1, lo que fuerza a que algunos coeficientes se vuelvan exactamente cero y, por lo tanto, realiza selección automática de variables.	Al reducir el sobreajuste y simplificar el modelo mediante la eliminación de predictores irrelevantes, Lasso resulta especialmente útil cuando se dispone de un gran número de variables.	Sin embargo, puede descartar variables pertinentes si están altamente correlacionadas, es sensible al valor del hiperparámetro λ y mantiene la suposición de linealidad.
Lasso	Extensión de la regresión lineal que utiliza dos o más variables independientes para predecir la variable dependiente.	Permite modelar relaciones complejas, ajusta múltiples variables.	Puede sufrir multicolinealidad, mayor complejidad en la interpretación.
Árboles de decisión	Modelo basado en una estructura de árbol que divide el espacio de variables en regiones homogéneas, facilitando la toma de	Clasificación y regresión, segmentación de clientes.	Pueden sobreajustarse a los datos si no se podan adecuadamente.

	decisiones a través de reglas.		
<i>Random Forest</i>	Ensamble de múltiples árboles de decisión que promedia sus resultados para mejorar la precisión y robustez de la predicción.	Reduce el sobreajuste, y posee alta precisión y robustez.	Requiere mayor capacidad computacional, menor interpretabilidad que un árbol único.
<i>XGBoost</i>	Algoritmo de <i>boosting</i> que optimiza el rendimiento de los árboles de decisión de forma secuencial, ajustando errores de predicción en cada iteración.	Alta eficiencia, precisión y escalabilidad.	Parámetros complejos de ajustar, mayor riesgo de sobreajuste si no se regulariza.
<i>Gradient Boosting</i>	Método de ensamble que agrega modelos de forma secuencial (normalmente árboles de decisión), entrenando cada nuevo modelo sobre los residuos del anterior mediante descenso por gradiente.	Gracias a su capacidad para capturar relaciones no lineales complejas, manejar variables mixtas y ajustar diversas funciones de pérdida, <i>Gradient Boosting</i> ofrece un poder predictivo muy elevado.	No obstante, presenta un alto costo computacional, numerosos hiperparámetros que ajustar y cierta dificultad de interpretación, lo que incrementa el riesgo de sobreajuste si no se regula adecuadamente.

ARIMA	Modelo de series temporales que combina componentes autorregresivos (AR), de diferenciación (I) y de media móvil (MA) para capturar dependencias lineales sin estacionalidad.	Al modelar dependencias lineales en series estacionarias, ARIMA proporciona pronósticos acompañados de intervalos de confianza explícitos y mantiene una estructura relativamente interpretable.	Por otro lado, exige lograr la estacionariedad mediante diferenciación, no modela la estacionalidad ni patrones no lineales y la identificación de los órdenes (p,d,q) resulta extenuante.
SARIMA	Extensión de ARIMA que añade términos autorregresivos, de integración y de media móvil estacionales para modelar patrones periódicos.	Al incorporar términos estacionales, SARIMA se adapta a series con patrones periódicos y conserva la interpretabilidad estadística del ARIMA, ofreciendo pronósticos robustos con intervalos de confianza.	No obstante, la selección de órdenes estacionales aumenta la complejidad y el costo computacional, supone estacionariedad y linealidad tras la diferenciación e incrementa la dificultad de mantenimiento.
<i>Stacking (Stacked Generalization)</i>	Es una estrategia de ensamble que combina varios modelos base, el cual aprende a ponderar las predicciones de los primeros.	Combina sesgo y varianza de modelos heterogéneos, mejorando métricas globales.	Mayor complejidad y riesgo de fuga de datos si es mal validado

Fuente: Adaptado de Géron (2019) y Hyndman & Athanasopoulos (2021).

2.4.5. Aprendizaje automático y técnicas de ensamblaje

El aprendizaje automático (*machine learning*) es una rama de la inteligencia artificial que permite a los sistemas “aprender” patrones directamente a partir de los datos y mejorar su rendimiento de manera autónoma conforme reciben nuevos ejemplos (Sarker, 2021a). Dentro de sus enfoques supervisados, destacan los métodos basados en árboles como *Random Forest* y *Gradient Boosting*, los cuales disminuyen la varianza y corrigen secuencialmente el sesgo residual de modelos anteriores.

Un paso adicional consiste en **combinar modelos heterogéneos mediante *stacking* o *stacked generalization***. Esta estrategia, descrita por Naimi y Balzer (2018) y ampliada recientemente por Wang y Lu (2024) en un contexto industrial, se articula en dos niveles:

1. **Nivel 0 (modelos base):** Se entrenan varios estimadores (GB, RF, entre otros) mediante el empleo de validación cruzada para generar predicciones con datos que el modelo no ha visto previamente.
2. **Nivel 1 (meta-modelo):** Las predicciones anteriores se “apilan” y se utilizan como variables de entrada de un modelo que los sintetice, el cual aprende a ponderar la contribución de cada estimador.

Diversas revisiones publicadas a partir de 2018 reportan que el *stacking* suele reducir simultáneamente el sesgo y la varianza del error, superando el desempeño de los modelos individuales cuando se implementa con una correcta separación de datos (Naimi & Balzer, 2018; Liu et al., 2023).

2.4.6. Evaluación y validación de modelos predictivos

La evaluación y la validación de modelos predictivos constituyen etapas esenciales en el desarrollo de soluciones de ciencia de datos, ya que permiten determinar la precisión, robustez y capacidad de generalización de un modelo antes de su implementación en entornos operativos. En conjunto, este proceso abarca tanto la aplicación de técnicas de validación cruzada para evaluar la estabilidad del modelo en distintos subconjuntos de datos como la medición de su desempeño mediante diversas métricas. Según Géron (2019), una evaluación adecuada es fundamental para

evitar el sobreajuste y asegurar que el modelo mantenga su rendimiento al enfrentarse a nuevos datos.

2.4.6.1. Métricas de desempeño (R^2 , RMSE, MAE, MAPE)

Las métricas de desempeño son indicadores cuantitativos que permiten evaluar la precisión y la efectividad de los modelos predictivos. Entre las métricas más comunes se encuentran:

- **R^2 (coeficiente de determinación):** Mide la proporción de la variabilidad de la variable dependiente que es explicada por el modelo.
- **RMSE (error cuadrático medio):** Penaliza con severidad los errores grandes y ofrece una medida en las mismas unidades que la variable a predecir.
- **MAE (error absoluto medio):** Representa la magnitud promedio de los errores, proporcionando una medida fácil de interpretar.
- **MAPE (error porcentual absoluto medio):** Expresa el error en términos porcentuales, facilitando la comparación entre modelos de diferentes escalas.

Géron (2019) enfatiza que la selección de las métricas adecuadas es crucial para comparar distintos modelos y para afinar aquellos que mejor se adapten al problema en cuestión. Estas métricas no solo ayudan a evaluar la precisión del modelo, sino que también orientan el proceso de ajuste y optimización de hiperparámetros.

2.5. Herramientas tecnológicas

Las herramientas tecnológicas constituyen el fundamento para implementar soluciones en ciencia de datos y modelos predictivos. Estos recursos —que abarcan lenguajes de programación, plataformas y librerías especializadas— permiten manipular, analizar y visualizar grandes volúmenes de información de manera eficiente. Su constante evolución es vital para responder a las demandas derivadas del análisis de grandes cantidades de datos y del *machine learning*, facilitando la integración y procesamiento de datos en entornos empresariales complejos. Según VanderPlas (2016), disponer de un conjunto robusto de herramientas tecnológicas permite a las organizaciones transformar datos sin procesar en conocimiento valioso para la toma de decisiones estratégicas.

2.5.1.1. Python y sus librerías especializadas

Python es el lenguaje de programación elegido en la ciencia de datos por su versatilidad y facilidad de uso. Su ecosistema de librerías especializadas potencia su capacidad para procesar y analizar datos. Por ejemplo, Pandas facilita la manipulación de estructuras de datos y la realización

de operaciones de limpieza y transformación; NumPy permite realizar cálculos matemáticos y operaciones de álgebra lineal de forma eficiente; y Matplotlib es una herramienta clave para la generación de visualizaciones de alta calidad que ayudan a interpretar los resultados del análisis. Estas librerías trabajan de manera integrada para ofrecer un entorno completo que soporta tanto el análisis exploratorio como el desarrollo de modelos predictivos (McKinney, 2018).

2.6. Vehículos eléctricos

Los vehículos eléctricos son automóviles que utilizan energía eléctrica almacenada en baterías para su propulsión, lo que elimina o reduce significativamente la dependencia de combustibles fósiles y, por ende, las emisiones de gases contaminantes. La publicación de la U.S. Department of Energy's Alternative Fuels Data Center (AFDC) (2023) explica que estos vehículos incorporan componentes esenciales como el motor eléctrico, el sistema de baterías y la infraestructura de carga. Esta tecnología se destaca por su alta eficiencia energética y su menor impacto ambiental, lo que la posiciona como una alternativa clave en la transición hacia una movilidad sostenible a nivel global. Además, el continuo avance en el almacenamiento de energía y la expansión de redes de carga son factores críticos que impulsan su adopción en diferentes mercados.

2.6.1. Tipos de baterías

A continuación, se presenta una tabla comparativa que resume los principales tipos de baterías utilizadas en vehículos eléctricos junto con una breve definición de cada una. Esta fue elaborada a partir de la información disponible en la página de GreenCars (2023).

Tabla 5. Cuadro comparativo de tipos de baterías

Tipo de batería	Definición
Batería de iones de litio	Tecnología ampliamente utilizada en vehículos eléctricos modernos, caracterizada por su alta densidad energética, eficiencia y larga vida útil. Se basa en la incorporación de compuestos de litio para almacenar energía de forma efectiva.
Batería de litio-ternaria	Utiliza una combinación de níquel, manganeso y cobalto en el cátodo, lo que permite alcanzar un buen equilibrio entre densidad energética,

	durabilidad y costos. Esta tecnología es especialmente adecuada para aumentar la autonomía del vehículo.
Batería de litio-fosfato de hierro	Conocida por su gran estabilidad térmica y seguridad, aunque ofrece una densidad energética moderada. Es ideal cuando se prioriza la vida útil prolongada y la robustez, aspectos críticos en determinadas aplicaciones de movilidad eléctrica.
Batería de polímero de litio	Variante de la tecnología de iones de litio que utiliza electrolitos en forma de gel o polímero, lo que permite diseños más flexibles y una reducción de peso. Esto facilita la adaptación a diversas configuraciones y tamaños en los vehículos.
Batería de plomo-ácido	Aunque históricamente ha sido utilizada en otros tipos de vehículos, su baja densidad energética y alto peso la hacen menos competitiva para vehículos eléctricos modernos. Su uso se limita a aplicaciones específicas donde el costo inicial es una prioridad.
Batería de estado sólido	Reemplaza el electrolito líquido por uno sólido, lo que mejora significativamente la seguridad y, potencialmente, la densidad energética. Esta tecnología ofrece un mayor rendimiento en condiciones de temperatura extrema y una vida útil prolongada, aunque aún se encuentra en fases de desarrollo y comercialización limitada. Las baterías de estado sólido prometen revolucionar el sector debido a su potencial para reducir riesgos de fugas y combustión.

Fuente: Adaptado de Alkhalidi et al. (2024).

2.6.2. Diferencia entre las baterías basadas en litio y de estado sólido

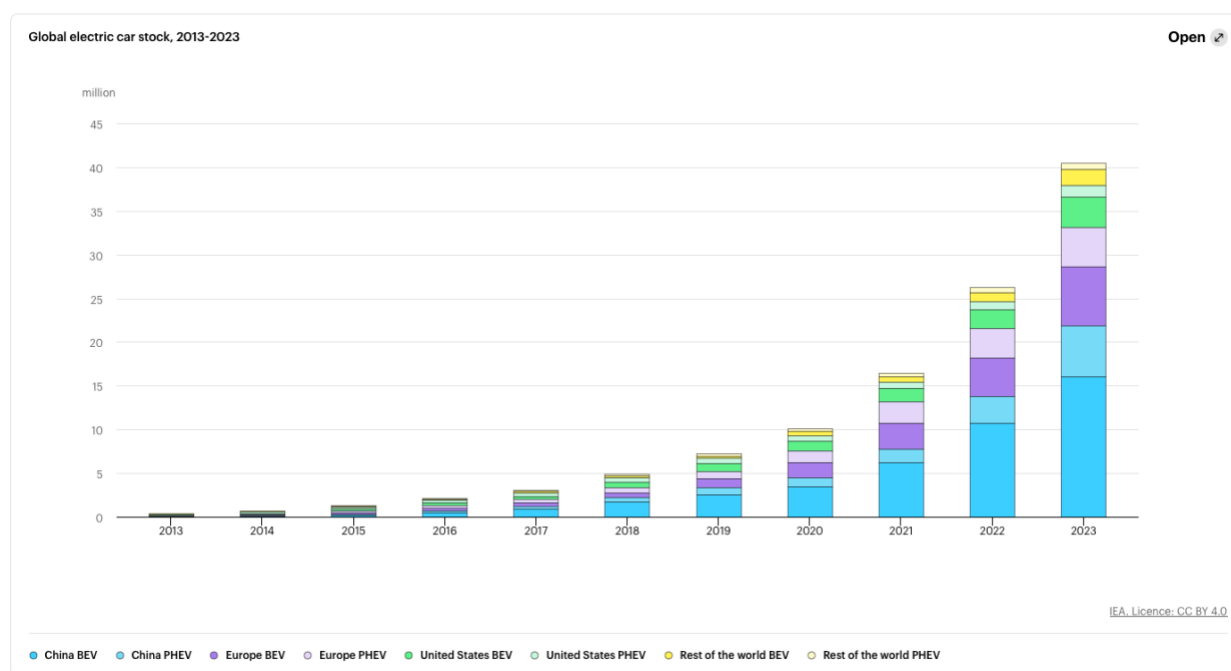
Las baterías tradicionales de iones de litio emplean electrolitos líquidos o en gel para el transporte de iones entre el ánodo y el cátodo. En contraste, las baterías de estado sólido utilizan electrolitos en forma sólida, lo que elimina el riesgo de fugas y mejora significativamente la estabilidad térmica y la densidad energética. Según Alkhalidi et al. (2024), esta innovación técnica

representa un avance crucial que, a pesar de los desafíos actuales en términos de producción a gran escala y costos, apunta a mejorar tanto la seguridad como el rendimiento de los vehículos eléctricos.

2.6.3. Contexto global

El contexto global de los vehículos eléctricos se ha acelerado en los últimos años, impulsado por políticas ambientales más estrictas, avances tecnológicos en baterías y la rápida expansión de la infraestructura de carga. El informe *Global EV Outlook 2024* de la International Energy Agency (IEA) (2024) destaca que la adopción de vehículos eléctricos ha alcanzado niveles récord a nivel mundial, evidenciando una transformación sustancial en el sector automotriz. Este crecimiento global subraya la importancia de la eficiencia energética y la reducción de emisiones, factores que están redefiniendo el panorama de la movilidad y estableciendo nuevos estándares para la industria automotriz.

Figura 8. *Inventario global de vehículos eléctricos 2013-2023*



Fuente: Tomado de International Energy Agency (2024), bajo licencia CC BY 4.0.

En la **Figura 8**, se refleja la rápida expansión de los vehículos eléctricos a nivel mundial. El eje vertical indica el número total de unidades en millones, mientras que el eje horizontal

representa los años de referencia. La barra en tonos azules corresponde a China, que ha experimentado un crecimiento significativo en los últimos años, seguida por otras regiones como Europa y Estados Unidos. De acuerdo con la IEA (2024), este incremento sostenido responde a factores como el apoyo gubernamental, los avances tecnológicos en baterías y la creciente conciencia ambiental, lo que ha impulsado la adopción de la movilidad eléctrica en diversos mercados.

2.6.4. Mercado costarricense

En Costa Rica, el mercado de vehículos eléctricos ha experimentado un crecimiento notable en respuesta a las políticas públicas de sostenibilidad y a los incentivos para la adopción de tecnologías limpias. Según el informe publicado por el Ministerio de Ambiente y Energía de Costa Rica (2024), se ha observado un aumento sostenido en el número de vehículos eléctricos importados y vendidos, lo cual refleja el creciente interés tanto de consumidores como de actores del sector automotriz. Este dinamismo se debe, en parte, a la mejora en la infraestructura de carga y a una mayor concienciación sobre la reducción de emisiones contaminantes, factores que posicionan al mercado costarricense como un entorno favorable para el desarrollo de la movilidad eléctrica.

2.6.4.1. Infraestructura y políticas de incentivos

La expansión del mercado de vehículos eléctricos en Costa Rica se ha visto favorecida no solo por la mejora en la infraestructura de carga, sino también por un sólido marco normativo que impulsa el sector. En efecto, la instalación de estaciones de carga en áreas urbanas e interurbanas ha contribuido a reducir la “ansiedad de autonomía”, lo que incrementa la confianza del consumidor en la viabilidad de estos vehículos. Además, las políticas públicas, a través de incentivos fiscales, subsidios y facilidades crediticias, han estimulado a fabricantes y concesionarios a ampliar su oferta y mejorar el acceso a esta tecnología.

Complementariamente, la normativa legal costarricense refuerza estos esfuerzos. Por ejemplo, la reforma de la Ley 9518 mediante la Ley 10209 se ha convertido en un pilar fundamental para fomentar la movilidad eléctrica. De acuerdo con el texto completo de la Normativa N° 97124, disponible en la página de la Procuraduría General de la República (2023), se establecen disposiciones que incentivan la adopción de vehículos eléctricos, promoviendo exenciones fiscales y medidas que facilitan el desarrollo de infraestructura de carga. Estas políticas

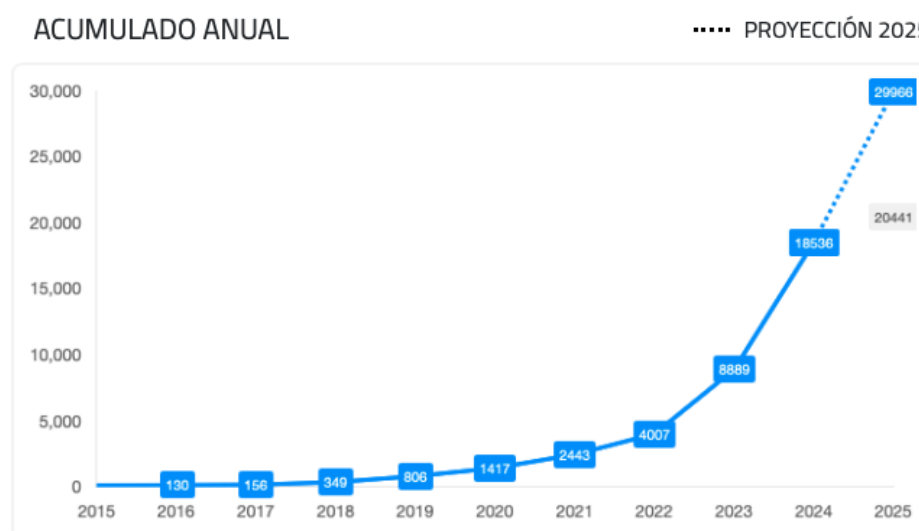
no solo facilitan la transición a una movilidad limpia, sino que también contribuyen a reducir costos operativos y a mejorar la competitividad del sector automotriz en el país.

2.6.4.2. Tendencias actuales y proyecciones futuras

Los datos publicados por la Asociación Costarricense de Movilidad Eléctrica (ASOMOVE) (2024) evidencian un crecimiento exponencial en el número de vehículos eléctricos en Costa Rica. Tal y como se ilustra en la **Figura 1**, el acumulado anual pasó de 130 unidades en 2015 a 18 536 en 2023. La marcada tendencia ascendente refleja el impacto positivo de los incentivos gubernamentales, la expansión de la infraestructura de carga y la creciente concienciación ambiental por parte de consumidores y empresas.

Además, el escenario proyectado para 2025 indica que se podrían alcanzar alrededor de 29 966 vehículos eléctricos circulando en el país, lo que consolidaría aún más la transición hacia una movilidad limpia y eficiente. Este panorama, respaldado por iniciativas de política pública y por la colaboración de diversos actores del sector, subraya el potencial de Costa Rica para posicionarse como un referente regional en la adopción de tecnologías sostenibles en el ámbito automotriz.

Figura 9. *Evolución y proyección acumulada de los automóviles eléctricos en Costa Rica*



Fuente: Adaptado de ASOMOVE (2024).

La gráfica pone de manifiesto la importancia de continuar reforzando la infraestructura de carga, las campañas de información y los incentivos financieros para sostener esta senda de

crecimiento y alcanzar el objetivo estimado. Asimismo, se enfatiza el rol de las entidades financieras y los concesionarios, quienes pueden aprovechar este entorno favorable para desarrollar soluciones que aborden la creciente demanda de vehículos eléctricos y fortalezcan la oferta de servicios de movilidad sostenible.

3. Marco metodológico

Este capítulo tiene como finalidad exponer y fundamentar los elementos metodológicos que guían el desarrollo del trabajo final de graduación. En él se presenta una descripción detallada del tipo de investigación, el enfoque y diseño adoptados, así como las fuentes de datos e información empleadas. Además, se identifican los sujetos y variables que integran el estudio, se especifican las técnicas e instrumentos de recolección de datos y se define el procedimiento metodológico en fases que estructuran el desarrollo del proyecto. Con este marco metodológico, se busca garantizar la coherencia, rigurosidad y pertinencia del estudio, asegurando que cada componente del proceso de investigación contribuya a la solución de la problemática planteada.

3.1. Tipo de investigación

La investigación puede entenderse como un “conjunto de procesos sistemáticos, críticos y empíricos que se aplican al estudio de un fenómeno o problema” (Hernández et al., 2014, p. 4, cap. 1). Este tipo de enfoque asegura que el conocimiento generado tenga rigor y relevancia, contribuyendo tanto al ámbito teórico como práctico.

La orientación de un trabajo de investigación depende del propósito y el alcance del proyecto. Por consiguiente, entender las distintas clasificaciones de las investigaciones resulta fundamental para asegurar una alineación adecuada entre el objetivo del estudio y la metodología que se empleará para alcanzar las metas establecidas. A continuación, se ofrece una descripción concisa de los diferentes tipos de investigación según su propósito, con el objetivo de identificar aquel que mejor responda a las necesidades específicas del proyecto.

3.1.1. Investigación aplicada

Su principal característica radica en que está diseñada para abordar problemas específicos, empleando el conocimiento existente para generar soluciones concretas y prácticas. De acuerdo con el Departamento Universitario Obrero y Campesino (2024), este tipo de investigación se

enfoca en adaptar teorías y conceptos a situaciones particulares, con el fin de satisfacer necesidades reales o resolver problemáticas particulares.

3.1.2. Investigación cuantitativa

El objetivo de la investigación cuantitativa es “adquirir conocimientos fundamentales y la elección del modelo más adecuado que nos permita conocer la realidad de una manera más imparcial, ya que se recogen y analizan los datos a través de los conceptos y variables medibles” (Neill et al., 2018, p. 69 / párr. 2).

Con la información previa, es posible tomar la decisión de utilizar un **enfoque aplicado** orientado a generar soluciones concretas y prácticas para la financiera Áltica, con el fin de diseñar un modelo para predecir el valor de depreciación de los vehículos eléctricos.

3.2. Enfoque y diseño de la investigación

El enfoque y diseño de la investigación definen cómo se llevará a cabo el estudio para alcanzar los objetivos planteados. En los siguientes párrafos, se procederá a plantear los diferentes enfoques y diseños de investigación que se pueden llevar a cabo.

3.2.1. Enfoque de la investigación

De acuerdo con Hernández Sampieri et al. (2014), “los enfoques cuantitativo, cualitativo y mixto constituyen posibles elecciones para enfrentar problemas de investigación y resultan igualmente valiosos. Son, hasta ahora, las mejores formas diseñadas por la humanidad para investigar y generar conocimientos” (p. 2).

Por otra parte, Hernández Sampieri et al. (2014), en su libro sobre metodología de la investigación, detallan tres enfoques investigativos, los cuales corresponden a: cuantitativo, cualitativo o mixto; estos son abordados a continuación:

Tabla 6. *Enfoques de investigación*

Enfoque	Descripción	Características
Cuantitativo	Es secuencial y probatorio. Cada etapa precede a la siguiente y no podemos “brincar” o eludir pasos. El orden es riguroso, aunque desde luego, podemos redefinir alguna fase. Parte de una idea que va acotándose y, una vez delimitada, se derivan objetivos y preguntas de investigación, se revisa la literatura y se construye un marco o una perspectiva teórica.	1. Refleja la necesidad de medir y estimar magnitudes de los fenómenos o problemas de investigación: ¿cada cuánto ocurren y con qué magnitud? 2. El investigador o investigadora plantea un problema de estudio delimitado y concreto sobre el fenómeno, aunque esté en evolución. Sus preguntas de investigación versan sobre cuestiones específicas. 3. Una vez planteado el problema de estudio, el investigador o investigadora considera lo que se ha investigado anteriormente (la revisión de la literatura) y construye un marco teórico (la teoría que habrá de guiar su estudio), del cual deriva una o varias hipótesis (cuestiones que va a examinar si son ciertas o no) y las somete a prueba mediante el empleo de los diseños de investigación apropiados.
Cualitativo	También se guía por áreas o temas significativos de investigación. Sin embargo, en lugar de que la claridad	1. El investigador o investigadora plantea un problema, pero no sigue un proceso definido

	sobre las preguntas de investigación e hipótesis preceda a la recolección y al análisis de los datos (como en la mayoría de los estudios cuantitativos), los estudios cualitativos pueden desarrollar preguntas e hipótesis antes, durante o después de la recolección y el análisis de los datos.	claramente. 2. En la búsqueda cualitativa, en lugar de iniciar con una teoría y luego “voltar” al mundo empírico para confirmar si esta es apoyada por los datos y resultados, el investigador comienza examinando los hechos en sí y, en el proceso, desarrolla una teoría coherente para representar lo que observa.
Mixto	Los métodos mixtos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos y cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada y lograr un mayor entendimiento del fenómeno bajo estudio.	1. Las etapas en las que suelen integrarse los enfoques cuantitativo y cualitativo son fundamentalmente: el planteamiento del problema, el diseño de investigación, el muestreo, la recolección de los datos, los procedimientos de análisis y/o interpretación de los datos (resultados). 2. Los métodos mixtos son multimetódicos, por lo que representan la “tercera vía”. 3. Un factor adicional que ha detonado la necesidad de utilizar los métodos mixtos es la naturaleza compleja de la gran mayoría de los fenómenos o problemas de investigación

		abordados en las distintas ciencias.
--	--	--------------------------------------

Fuente: Elaboración propia con base en la información de Hernández Sampieri et al. (2014).

Al tener claridad sobre los objetivos y el alcance de esta investigación, se concluyó que el enfoque más apropiado corresponde al **enfoque de investigación mixto**, siguiendo los conceptos de Hernández Sampieri et al. (2014). Esta elección se fundamenta en la necesidad de integrar elementos cuantitativos y cualitativos para abordar, de manera integral, la problemática planteada sobre el análisis de la información y desarrollo del modelo de cálculo de la valoración y validación para los vehículos eléctricos financiados por la entidad financiera Ática.

La investigación aplicada requiere tanto la recolección y análisis de datos cuantitativos — como ventas, depreciación y rendimiento de baterías— para garantizar la precisión en las predicciones del modelo como el análisis cualitativo de las dinámicas del entorno financiero y las percepciones de los *stakeholders* involucrados. El enfoque mixto permite combinar la fuerza del análisis cuantitativo, lo cual proporciona rigor y exactitud en las mediciones, con la perspectiva cualitativa, que ofrece un entendimiento profundo del contexto en el cual se desarrolla el proyecto, incluyendo la interpretación de las condiciones del mercado y las políticas de incentivos.

Este enfoque mixto posibilita recolectar y analizar simultáneamente los datos, lo cual es esencial para comprender cómo convergen los resultados cuantitativos y cualitativos. Por ejemplo, los datos numéricos proporcionan una base sólida para el desarrollo del modelo predictivo, mientras que el análisis cualitativo permite ajustar y validar los resultados según el contexto particular de la industria financiera en Costa Rica, incluyendo factores como la percepción del riesgo por parte de los clientes y la aceptación de los vehículos eléctricos en el mercado local.

En este contexto, el enfoque mixto resulta crucial para desarrollar un modelo predictivo robusto y, al mismo tiempo, ofrecer una visión comprensiva que abarque tanto los números como los factores subjetivos del entorno financiero. Esta combinación asegura que los modelos predictivos diseñados no solo sean precisos, sino también pertinentes y alineados con la realidad del negocio, incrementando la efectividad en la toma de decisiones y reduciendo el riesgo financiero para la entidad.

3.2.2. Diseño de la investigación

El diseño de la investigación es definido por Hernández Sampieri et al. (2014) como “plan o estrategia que se desarrolla para obtener la información que se requiere en una investigación y responder al planteamiento” (p. 128).

Tabla 7. *Diseño de la investigación*

Enfoque	Descripción
Diseño exploratorio secuencial (DEXPLOS)	El diseño implica una fase inicial de recolección y análisis de datos cualitativos seguida de otra donde se recaban y analizan datos cuantitativos.
Diseño explicativo secuencial (DEXPLIS)	El diseño se caracteriza por una primera etapa en la cual se recaban y analizan datos cuantitativos, seguida de otra donde se recogen y evalúan datos cualitativos. La mezcla mixta ocurre cuando los resultados cuantitativos iniciales informan a la recolección de los datos cualitativos.
Diseño transformativo secuencial (DITRAS)	Al igual que los diseños previos, el diseño transformativo secuencial incluye dos etapas de recolección de los datos. La prioridad y fase inicial puede ser la cuantitativa o la cualitativa, o bien, es posible otorgarle a ambas la misma importancia y comenzar por alguna de ellas.
Diseño de triangulación concurrente (DITRIAC)	Este modelo es probablemente el más popular y se utiliza cuando el investigador pretende confirmar o corroborar resultados y efectuar validación cruzada entre datos cuantitativos y cualitativos, así como aprovechar las ventajas de cada método y minimizar sus debilidades. Puede ocurrir que no se presente la confirmación o corroboración.
Diseño anidado o incrustado concurrente de modelo dominante (DIAC)	El diseño anidado concurrente colecta simultáneamente datos cuantitativos y cualitativos. Su diferencia con el diseño de triangulación concurrente reside en que un método predominante guía el proyecto (pudiendo ser este cuantitativo o cualitativo). El método que posee menor prioridad es anidado o insertado dentro del que se considera central.

Diseño anidado concurrente de varios niveles (DIACNIV)	En esta modalidad, se recolectan datos cuantitativos y cualitativos en diferentes niveles, pero los análisis pueden variar en cada uno de estos. O bien, en un nivel se recolectan y analizan datos cuantitativos; en otro, datos cualitativos y así sucesivamente.
Diseño transformativo concurrente (DISTRAC)	Este diseño conjunta varios elementos de los modelos previos: se recolectan datos cuantitativos y cualitativos en un mismo momento (concurrente) y puede darse o no mayor peso a uno u otro método, pero al igual que en el diseño transformativo secuencial, la recolección y el análisis son guiados por una teoría, visión, ideología o perspectiva, incluso un diseño cuantitativo o cualitativo (por ejemplo, un experimento o un ejercicio participativo).
Diseño de integración múltiple (DIM)	Implica la mezcla más completa entre los métodos cuantitativo y cualitativo, y es sumamente itinerante.

Fuente: Elaboración propia con base en la información de Hernández Sampieri et al. (2014).

De acuerdo con los objetivos planteados y el contexto del proyecto, en este caso, se decidió utilizar un enfoque mixto con un **diseño explicativo secuencial (DEXPLIS)**, según lo establecido por Hernández Sampieri et al. (2014). Este diseño es ideal para el desarrollo de un modelo predictivo del valor futuro de los vehículos eléctricos financiados, ya que permite un análisis estructurado y exhaustivo que integra tanto datos cuantitativos como cualitativos para comprender a fondo el problema.

El diseño DEXPLIS se caracteriza por comenzar con la recolección y análisis de datos cuantitativos, que, en el contexto del proyecto, incluyen datos sobre ventas, depreciación de los vehículos y rendimiento de las baterías. Esta primera fase cuantitativa busca establecer una base sólida para entender los patrones de comportamiento y desarrollar modelos predictivos que sean precisos y fiables. Los resultados obtenidos de esta fase cuantitativa sirven como una guía para la segunda fase, que se enfoca en la recolección y análisis cualitativo.

Este enfoque es adecuado para el proyecto, ya que facilita una comprensión multifacética del problema, combinando el rigor del análisis cuantitativo con la riqueza interpretativa del análisis

cualitativo. De este modo, el modelo predictivo desarrollado no solo tiene un sustento estadístico robusto, sino que también incorpora las perspectivas y conocimientos del entorno financiero, garantizando su aplicabilidad y relevancia dentro de la entidad financiera Áltica.

La secuencialidad del diseño DEXPLIS, donde los resultados cuantitativos informan y direccionan el análisis cualitativo, permite realizar ajustes metodológicos que aseguran una alineación continua con los objetivos del negocio y la realidad del mercado. Esto contribuye a una mejor toma de decisiones no solo al predecir el valor de los vehículos eléctricos, sino también al entender cómo y por qué estos valores cambian con el tiempo, considerando variables económicas y sociales más amplias.

En resumen, el uso del DEXPLIS proporciona una estructura sistemática para abordar tanto la dimensión cuantitativa como cualitativa del problema, asegurando que el modelo predictivo sea completo, interpretable y práctico para la entidad financiera. Esto apoya directamente la capacidad del proyecto para generar un impacto positivo al apoyar al negocio a la hora de determinar un valor de rescate de las unidades eléctricas que tenga mayor precisión y que permita mejorar la toma de decisiones financieras.

3.3. Alcance de la investigación

De acuerdo con Hernández Sampieri et al. (2014), el alcance de una investigación define los resultados esperados y orienta la metodología que se debe seguir para alcanzarlos. Es fundamental determinar el alcance antes de iniciar el estudio, ya que esto permite establecer límites conceptuales y metodológicos adecuados, así como dirigir el enfoque cuantitativo de manera precisa.

Existen diversos alcances dentro de la investigación, entre los cuales se encuentran el exploratorio, el descriptivo, el correlacional y el explicativo. Estos no son excluyentes entre sí, ya que un estudio puede abarcar más de un tipo de alcance según sus objetivos y contexto.

Tabla 8. *Alcances de la investigación*

Alcance	Descripción
---------	-------------

Alcance exploratorio	La investigación exploratoria tiene como propósito analizar un tema o problema poco estudiado, en el cual existen numerosas incógnitas o ha sido abordado de manera limitada en la literatura científica. Este tipo de estudio permite al investigador familiarizarse con fenómenos desconocidos, facilitando el desarrollo de investigaciones más profundas y estructuradas. Además, ayuda a identificar nuevas problemáticas, conceptos clave y posibles relaciones entre variables, estableciendo una base para futuras investigaciones más detalladas.
Alcance descriptivo	Tiene como objetivo detallar y especificar las propiedades, características y dimensiones de fenómenos, situaciones o contextos sometidos a análisis. Este tipo de investigación busca una mayor precisión en la representación del fenómeno estudiado, permitiendo definir con claridad qué conceptos o variables serán medidos, quiénes formarán parte del estudio y cómo se recopilarán los datos.
Alcance correlacional	Se enfoca en responder preguntas mediante la asociación entre variables, estableciendo patrones predecibles dentro de una población o grupo de estudio. Su finalidad es determinar la relación existente entre dos o más conceptos, permitiendo comprender cómo interactúan entre sí. Aunque este tipo de estudio tiene un componente explicativo, su alcance es limitado, ya que identifica relaciones sin establecer causalidad. Es común encontrar investigaciones correlacionales que exploran vínculos entre tres o más variables.
Alcance explicativo	Busca identificar las causas de eventos o fenómenos en el ámbito físico o social. Su principal objetivo es explicar por qué ocurren determinados hechos y en qué condiciones se manifiestan, además de analizar la relación causal entre dos o más variables.

Fuente: Elaboración propia con base en la información de Hernández Sampieri et al. (2014).

Para el presente estudio, se ha seleccionado un **alcance exploratorio**, ya que la problemática que se aborda aún no ha sido ampliamente investigada. La investigación busca comprender factores determinantes en la depreciación de vehículos eléctricos en Costa Rica y desarrollar un modelo predictivo basado en datos históricos para determinar el valor de recuperación de estas unidades.

3.4. Fuentes de datos e información

Conforme a lo señalado por Miranda y Acosta (2009), se consideran como fuentes todos aquellos recursos de los que se obtiene la información necesaria para cubrir las necesidades de conocimiento sobre una situación o problema planteado, y que luego se emplearán para alcanzar los objetivos previstos. En esta sección, se especifican las fuentes de datos e información que se utilizan en el proyecto.

3.4.1. Fuentes primarias

De acuerdo con Suárez (2024), las fuentes primarias se refieren a las que surgen de la experiencia directa o de la observación de eventos. Son documentos originales o las primeras versiones de un hecho o investigación. Estas fuentes incluyen, entre otras, documentos históricos, entrevistas, encuestas, diarios de campo y registros médicos. Para este proyecto, se utilizan las siguientes fuentes primarias:

Tabla 9. *Fuentes de información primarias*

Fuente de información	Importancia del documento para la investigación
Registros históricos de ventas de vehículos eléctricos en Costa Rica	Proveen datos cuantitativos cruciales para entender el comportamiento del mercado y para identificar patrones de depreciación, fundamentales para construir el modelo predictivo.
Registros de rendimiento y mantenimiento de baterías	Permiten analizar cómo el rendimiento de las baterías afecta la depreciación y valor residual de los vehículos eléctricos, proporcionando datos esenciales para la modelización del valor a futuro.

Registros de importaciones de vehículos eléctricos	Ayudan a entender la evolución del mercado de vehículos eléctricos en Costa Rica, incluyendo el volumen de importación y la aceptación del mercado, lo cual influye directamente en la valoración de activos.
Entrevistas a representantes de concesionarios de vehículos eléctricos	Recogen perspectivas clave sobre la demanda, tendencias del mercado y otros factores externos que influyen en el valor de reventa de los vehículos eléctricos, ayudando a enriquecer el análisis cualitativo.

Fuente: Elaboración propia (2025).

3.4.2. Fuentes secundarias

Siguiendo lo expresado por Suárez (2024), las fuentes secundarias son aquellas que recopilan, resumen y analizan la información proveniente de las fuentes primarias. Constituyen una recopilación de datos que han sido previamente recolectados y publicados por otros. Ejemplos de fuentes secundarias incluyen libros de texto, artículos de revistas y reportes de investigación. Para este proyecto, se utilizan las siguientes fuentes secundarias:

Tabla 10. *Fuentes de información secundarias*

Fuente de información	Importancia del documento para la investigación
Sitios <i>web</i> con información sobre tecnologías y tendencias en vehículos eléctricos	Proporcionan contexto sobre la evolución tecnológica y el mercado global y local de vehículos eléctricos, ayudando a enriquecer la comprensión del entorno donde se desenvuelve el proyecto.
Investigaciones científicas y académicas	Ofrecen antecedentes y referencias científicas sobre el uso de modelos predictivos y técnicas de aprendizaje automático en la valoración de activos, aportando un marco teórico sólido.

encontradas en bases de datos	
Repositorios de trabajos finales de instituciones académicas	Proveen estudios previos sobre la valoración de vehículos y validación de modelos predictivos, lo cual sirve como base para mejorar la metodología y justificación del proyecto.
Estudios de mercado de vehículos eléctricos publicados por consultoras	Brindan análisis detallados del mercado de vehículos eléctricos, incluyendo proyecciones y tendencias, útiles para contrastar los resultados del modelo predictivo con datos reales del mercado.
Normas y estándares de gestión de datos (ISO 8000, DAMA-DMBOK)	Proporcionan lineamientos y estándares reconocidos para la calidad y gestión de datos, que son fundamentales para garantizar la precisión y consistencia en el análisis del proyecto.
Marcos de referencia para la aplicación de citas bibliográficas como APA 7	Aseguran la correcta referenciación de fuentes de información y la calidad académica del proyecto, contribuyendo a su credibilidad y rigurosidad.

Fuente: Elaboración propia (2024).

3.5. Sujetos de investigación

Según Hernández Sampieri, et al. (2014), los sujetos de investigación son individuos o grupos cuyas características, opiniones, experiencias y condiciones resultan pertinentes para responder a las preguntas del estudio. En la **Tabla 11**, se detallan los sujetos de investigación de este proyecto.

Tabla 11. *Sujetos de investigación*

Rol	Años de experiencia	Características	Justificación
Gerente de TI	Más de 9 años	Es el encargado del departamento de TI, es el patrocinador del proyecto actual. Adicionalmente, se encarga de realizar el desarrollo de modelos de datos enfocados en la reportería de la empresa para mejorar el desempeño en el área financiera, mejorando la gestión del riesgo de la cartera.	Vela por el desarrollo del proyecto, posee un alto conocimiento del negocio, así como del tema a desarrollar para intervenir y proveer retroalimentación en caso de ser necesario para el éxito del proyecto
Analista financiero del sector automotriz	5-10 años	Especialista en valoración de activos y riesgos en el financiamiento de vehículos. Cuenta con conocimiento en depreciación de automóviles y estrategias de mitigación de riesgos.	Proporciona información sobre cómo los modelos de depreciación influyen en las decisiones de financiamiento y en la evaluación de riesgo para la entidad financiera.

Fuente: Elaboración propia (2025).

3.6. Variables de la investigación

Según Miranda y Villasís (2016), las variables en un protocolo de investigación se definen como los elementos medibles o los datos recopilados que permiten responder a las preguntas planteadas en el estudio. Estas variables se determinan en función de los objetivos del proyecto. En la **Tabla 12**, se detallan las variables consideradas para este proyecto.

Tabla 12. *Variables de investigación*

Variable	Tipo de variable	Definición conceptual	Indicador	Definición instrumento
Objetivo específico n.º1: <i>Recolectar datos históricos y relevantes sobre el mercado de vehículos eléctricos en Costa Rica, incluyendo factores como ventas, valor actual del mercado y rendimiento de baterías, a partir de una base sólida de información confiable.</i>				
Datos del mercado de vehículos eléctricos en Costa Rica	Información cuantitativa y cualitativa sobre el mercado de vehículos eléctricos en Costa Rica, incluyendo ventas, importaciones, valor del mercado actual y rendimiento de las baterías.		Lista de datos y conjuntos de datos recopilados.	Entrevista Revisión documental
Objetivo específico n.º2: <i>Analizar datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica, considerando los factores mencionados, los cuales influyen en el valor de recuperación, para su integración en el modelo del modelo.</i>				

Factores clave que influyen en el valor de recuperación	Variables específicas que impactan significativamente el valor de recuperación de vehículos eléctricos, como la antigüedad, autonomía, estado de la batería y condiciones del mercado.	Lista de factores identificados y su significancia estadística de cada factor en relación con el valor de reventa.	Revisión documental Entrevista
Calidad y consistencia de los datos	Precisión e uniformidad de los datos tras su limpieza y transformación.	Muestra/Población final de datos a utilizar.	Revisión documental
Objetivo específico n.º3: <i>Diseñar un modelo estadístico y de aprendizaje automático que integre los factores identificados, para la predicción del valor de recuperación de los vehículos eléctricos financiados por la entidad.</i>			
Modelos predictivos estadísticos	Algoritmo o conjunto de algoritmos que utilizan métodos estadísticos y de aprendizaje automático para predecir el valor de recuperación de los vehículos eléctricos.	Cantidad de modelos predictivos entrenados.	Revisión documental Observación
Objetivo específico n.º4: <i>Validar el modelo predictivo mediante pruebas con datos reales y técnicas de validación cruzada, para el aseguramiento de su exactitud y confiabilidad en la estimación del valor de recuperación de los vehículos eléctricos.</i>			
Exactitud y confiabilidad del modelo predictivo seleccionado	Capacidad del modelo para predecir con precisión y consistencia el valor de mercado de los vehículos eléctricos.	Métricas de validación del modelo.	Entrevista Revisión documental

Fuente: Elaboración propia (2025).

3.7. Técnicas e instrumentos de recolección de datos

De acuerdo con Machuca (2022), los métodos y herramientas de recolección de datos permiten obtener información relevante a partir de diversas fuentes. Antes de seleccionar una técnica específica, es fundamental tener claramente definidos los objetivos de la investigación y el tipo de información que se necesita. A continuación, se presentan los principales instrumentos de recolección de datos que se emplean en este estudio.

3.7.1. Entrevistas

Las entrevistas constituyen una técnica de recolección de datos basada en la interacción directa entre el entrevistador y uno o varios entrevistados, con el objetivo de obtener información relevante sobre un tema específico. A través del intercambio de ideas y perspectivas, se logra la construcción conjunta de significados, lo que enriquece el proceso de investigación (Hernández Sampieri et al., 2014).

Según Ryen, Grinnell y Unrau (citados por Hernández Sampieri, 2014) las entrevistas pueden clasificarse en tres categorías principales:

- **Entrevista estructurada:** Se basa en un formato previamente diseñado por el entrevistador, en el cual las preguntas son definidas de antemano y deben seguirse en un orden específico, de manera que se garantice una uniformidad en la recolección de datos.
- **Entrevista semiestructurada:** Contempla una serie de preguntas predefinidas, pero permite cierta flexibilidad en el desarrollo de la conversación. De esta manera, es posible modificar, agregar o eliminar preguntas según la dinámica del diálogo.
- **Entrevista no estructurada o abierta:** No sigue un esquema rígido de preguntas, lo que permite que el entrevistador guíe la conversación de manera espontánea, adaptándola según el contexto y las respuestas del entrevistado.

De esta forma, para la recolección de datos en el presente proyecto de investigación, se utiliza el tipo de entrevista no estructurada o abierta. En el **Apéndice D**, se puede encontrar la plantilla utilizada para capturar la información requerida para el desarrollo de este proyecto.

3.7.2. Revisión documental

La revisión documental es un método de recolección de información basado en la consulta y análisis de fuentes escritas relevantes para la investigación. De acuerdo con Machuca (2022), esta técnica permite examinar documentación existente con el propósito de extraer datos clave y establecer una base sólida de conocimiento. No obstante, su aplicación requiere de un proceso riguroso de selección y evaluación de fuentes, asegurando que la información obtenida sea veraz y pertinente para los objetivos del estudio.

Siguiendo la metodología propuesta por Machuca et al. (2022), la revisión documental en esta investigación se desarrolla en tres fases principales:

- Selección de fuentes de información: Identificación de estudios, bases de datos y documentos que proporcionen información relevante sobre la depreciación de vehículos eléctricos.
- Revisión y organización de la información: Clasificación y estructuración de los datos extraídos, asegurando su coherencia con los objetivos del estudio.
- Análisis de los datos: Evaluación de los documentos recopilados para identificar patrones, tendencias y variables clave en la predicción del valor de recuperación.

En este proyecto, la revisión documental se centró en diversas fuentes, con el propósito de garantizar un análisis integral de la problemática. Entre las fuentes consultadas, se encuentran:

- Estudios previos sobre depreciación de vehículos eléctricos: Se analizaron investigaciones científicas y reportes del sector automotriz que abordan los factores que influyen en la pérdida de valor de estos vehículos a lo largo del tiempo.
- Modelos predictivos aplicados en estudios similares: Se revisaron enfoques estadísticos y de aprendizaje automático utilizados en otras investigaciones para estimar la depreciación de activos, con el fin de seleccionar el método más adecuado para este estudio.
- Bases de datos del mercado de vehículos eléctricos en Costa Rica: Se recopilieron registros de ventas, importaciones y precios de reventa de distintas fuentes gubernamentales y comerciales.
- Documentación técnica sobre el rendimiento de baterías: Se incluyó información sobre la degradación de las baterías de los vehículos eléctricos, considerando su impacto en la depreciación del valor de los automóviles.

Para estructurar la información obtenida mediante la revisión documental, se emplea la plantilla que se presenta en el **Apéndice E**, lo que permite organizar los hallazgos de manera sistemática y facilitar su integración en el modelo predictivo propuesto.

3.7.3. Observación

La observación es una herramienta de recolección de datos que permite al investigador analizar e interpretar el entorno a través de los distintos sentidos. Según Hernández Sampieri et al. (2014), su propósito es explorar y describir los ambientes de estudio, comprendiendo la dinámica de los actores involucrados y los procesos que se desarrollan en el tiempo. Además, esta técnica facilita la identificación de patrones y relaciones entre personas, experiencias y eventos, lo que permite generar hipótesis sobre fenómenos futuros.

Al utilizar la observación como método de investigación, es importante considerar ciertos elementos clave que influyen en su aplicación:

- Entorno físico: Incluye las características del espacio donde se desarrolla la observación, así como los objetos y estructuras presentes en el ambiente.
- Contexto social y humano: Se refiere a la composición y organización de los grupos observados, tomando en cuenta aspectos como edad, nivel socioeconómico y diversidad.
- Interacciones y actividades: Engloba las acciones individuales y colectivas realizadas por los participantes, analizando qué hacen, cuándo y cómo lo hacen, y con qué propósito.
- Uso de objetos y herramientas: Examina los artefactos empleados por los participantes y la función que desempeñan dentro del contexto observado.
- Eventos significativos: Comprende la identificación de acontecimientos relevantes dentro del ambiente de estudio, estableciendo su posible impacto en la investigación.
- Descripción de los participantes: Permite caracterizar a las personas involucradas en el estudio, detallando su rol, comportamiento y características generales.

La plantilla utilizada para la documentación se puede encontrar en el **Apéndice C**.

3.8. Procedimiento metodológico de la investigación

Para desarrollar un modelo que satisfaga las necesidades de la empresa y cumpla con los objetivos establecidos, se recopiló una amplia cantidad de información. El procedimiento metodológico del proyecto se estructura en cinco fases clave que guían el desarrollo y la implementación del modelo predictivo. Primero, se realiza la recolección de los datos, recopilando

datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica para identificar factores influyentes. Seguidamente, se lleva a cabo el análisis exploratorio de datos, donde estos se limpian y preparan para asegurar su calidad y representatividad. En la tercera fase, se procede a realizar el diseño del modelo predictivo, para lo que se desarrollan y evalúan distintos algoritmos de aprendizaje automático para seleccionar el más adecuado. La cuarta fase corresponde a la validación del modelo, donde se aplican pruebas con datos de prueba y técnicas de validación cruzada, utilizando métricas específicas para establecer un punto de comparación que determine la aceptación del modelo. Finalmente, se documentan los resultados obtenidos, junto con recomendaciones para ajustes futuros o implementaciones adicionales.

A continuación, se presenta el diagrama que facilita su comprensión:

Figura 10. *Fases metodológicas de la investigación*

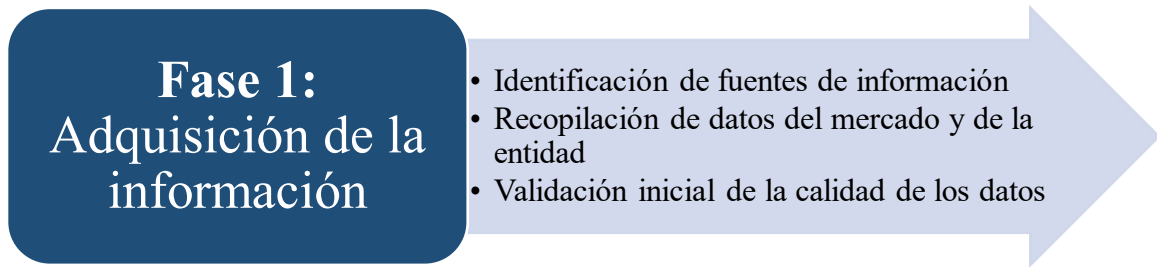


Fuente: Elaboración propia (2025).

3.8.1. Fase 1: Adquisición de la información

En primer lugar, se recopilan datos históricos y actuales sobre el mercado de vehículos eléctricos en Costa Rica, lo cual permite establecer una base sólida para el estudio. Además, se identifican las variables clave que inciden en la depreciación y el valor de recuperación de los activos. Posteriormente, se evalúa la calidad de la información obtenida, garantizando así la pertinencia y consistencia de los datos. Este proceso inicial facilita la detección de patrones y tendencias relevantes para el proyecto.

Figura 11. *Subactividades de la fase uno*

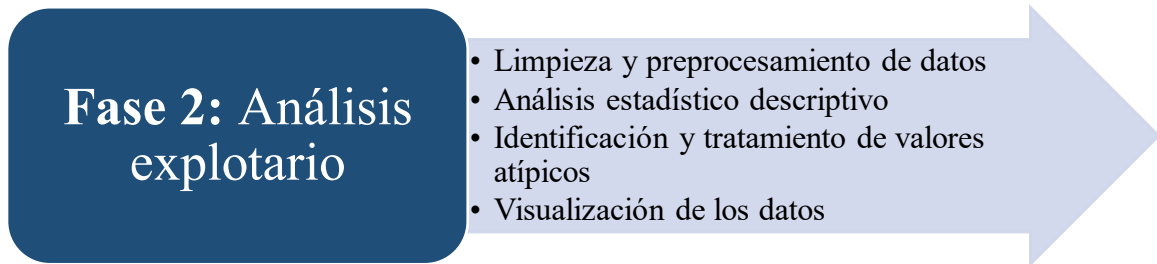


Fuente: Elaboración propia (2025).

3.8.2. Fase 2: Análisis exploratorio

Seguidamente, se procede con el análisis exploratorio de los datos recopilados, utilizando técnicas estadísticas y de visualización. Asimismo, se identifican las relaciones significativas entre las variables, lo que permite profundizar en la comprensión del fenómeno estudiado. Por consiguiente, se contrastan estos hallazgos con estudios previos para validar la información y enriquecer el marco analítico. De esta forma, se definen las bases para el desarrollo del modelo predictivo.

Figura 12. *Subactividades de la fase dos*



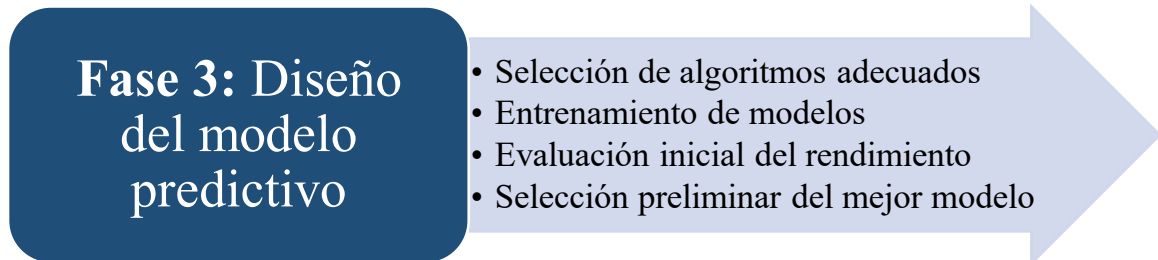
Fuente: Elaboración propia (2025).

3.8.3. Fase 3: Diseño de los modelos predictivos

Luego, se diseñan los modelos predictivos aplicando algoritmos de aprendizaje automático y técnicas estadísticas avanzadas. Para ello, se integran las variables identificadas en las fases anteriores, asegurando la robustez del modelo. Igualmente, se realizan ajustes y pruebas preliminares que orienten la selección del algoritmo más adecuado. En consecuencia, se construyen distintos modelos que compitan entre ellos, y se escoge uno que después permita estimar de manera precisa el valor futuro de los vehículos eléctricos. Adicionalmente, se plantea construir un ensamble tipo stacking que combine los dos modelos individuales de mayor

desempeño; este ensamble se evaluará frente a las demás alternativas para confirmar si proporciona la predicción más precisa.

Figura 13. *Subactividades de la fase tres*

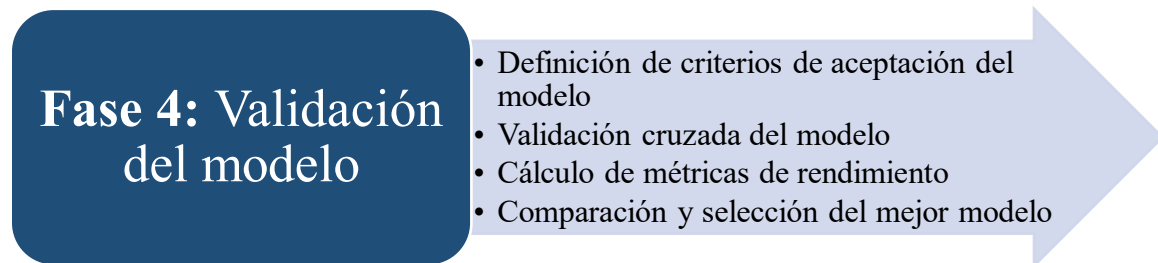


Fuente: Elaboración propia (2025).

3.8.4. Fase 4: Validación del modelo

Finalmente, se lleva a cabo la validación del modelo predictivo mediante pruebas con datos reales y técnicas de validación cruzada. En este sentido, se utilizan métricas de desempeño como el MSE y el coeficiente de determinación (R^2) para evaluar la exactitud de las predicciones. Además, se efectúan ajustes en el modelo conforme a los resultados obtenidos, garantizando su confiabilidad. Así, se culmina esta fase con un modelo optimizado y listo para su implementación en el entorno real.

Figura 14. *Subactividades de la fase cuatro*



Fuente: Elaboración propia (2025).

3.9. Operacionalización de las variables o categorías

Seguidamente, se sintetizan los fundamentos desarrollados dentro de la metodología de investigación. En la **Tabla 13**, se exponen los objetivos y su operacionalización, con el propósito de comprender su contexto en el presente proyecto.

Tabla 13. Operacionalización de las variables

Fase de investigación	Objetivo específico	Instrumentos	Variable	Sujeto
Fase 1: Adquisición de la información	Objetivo específico n.º1: <i>Recolectar datos históricos y relevantes sobre el mercado de vehículos eléctricos en Costa Rica, incluyendo factores como ventas, valor actual del mercado y rendimiento de baterías, a partir de una base sólida de información confiable.</i>	Entrevista Revisión documental	Datos del mercado de vehículos eléctricos en Costa Rica	Gerente de TI Analista financiero del sector automotriz

Fase 2: Análisis exploratorio	Objetivo específico n.º2: <i>Analizar datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica, considerando los factores mencionados, los cuales influyen en el valor de recuperación, para su integración del modelo.</i>	Revisión documental	Factores clave que influyen en el valor de recuperación	Analista financiero del sector automotriz
		Revisión documental	Calidad y consistencia de los datos	Gerente de TI

Fase 3: Diseño del modelo predictivo	Objetivo específico n.º3: <i>Diseñar un modelo estadístico y de aprendizaje automático que integre los factores identificados, para la predicción del valor de recuperación de los vehículos eléctricos financiados por la entidad.</i>	Revisión documental Observación	Modelos predictivos estadístico	Gerente de TI
--	--	--	--	----------------------

Fase 4: Validación del modelo	Objetivo específico n.º4: <i>Validar el modelo predictivo mediante pruebas con datos reales y técnicas de validación cruzada, para el aseguramiento de su exactitud y confiabilidad en la estimación del valor de recuperación de los vehículos eléctricos.</i>	Entrevista Revisión documental	Exactitud y confiabilidad del modelo predictivo seleccionado	Gerente de TI Analista financiero del sector automotriz
---	---	-----------------------------------	--	--

Fuente: Elaboración propia (2025).

3.10. Tabla resumen del procedimiento metodológico de la Investigación

En la siguiente sección, se puede apreciar el resumen del procedimiento metodológico de la investigación por medio de la matriz de trazabilidad de los objetivos y su relación entre las fases del procedimiento metodológico, a continuación en la **Tabla 14**.

Tabla 14. Resumen del procedimiento metodológico

Objetivo específico	Marco conceptual	Marco metodológico	Resultados o propuesta	Conclusiones	Recomendaciones
---------------------	------------------	--------------------	------------------------	--------------	-----------------

Objetivo específico n.º1: <i>Recolectar datos históricos y relevantes sobre el mercado de vehículos eléctricos en Costa Rica, incluyendo factores como ventas, valor actual del mercado y rendimiento de baterías, a partir de una base sólida de información confiable.</i>	2.1 Analítica de negocios La analítica de negocios, o BA (<i>business analytics</i>) por sus siglas en inglés, se concibe como una capacidad organizacional que convierte datos en conocimiento accionable mediante la aplicación coordinada de técnicas estadísticas, aprendizaje automático e inteligencia artificial (Liu et al., 2023). Esta perspectiva complementa la propuesta de Power et al. (2018), quienes subrayan que la verdadera ventaja competitiva emerge cuando los modelos predictivos y prescriptivos se integran en los procesos decisionales, lo que permite anticipar escenarios y recomendar la acción óptima. Desde esta óptica, la analítica de negocios trasciende la simple generación de reportes descriptivos al abarcar un ciclo completo que va desde la captura y depuración de datos hasta la automatización de recomendaciones estratégicas (Camm et al., 2020). En síntesis, la analítica de negocios actúa como el puente entre la inteligencia derivada de los datos y la obtención de resultados tangibles, sustentando decisiones de alto impacto en un entorno empresarial cada vez más dinámico.	Fase 1: Adquisición de la información	Resultados de la Fase 1: Adquisición de la información	Conclusión: Objetivo específico uno	Recomendaciones
---	--	--	---	--	------------------------

3.10.1. Tipos de analítica

La literatura reciente distingue cuatro niveles de analítica de negocios que se complementan de forma progresiva: descriptiva, diagnóstica, predictiva y prescriptiva (Camm et al., 2020; Sarker, 2021b). Cada uno responde a una pregunta distinta y requiere herramientas específicas:

Tabla 3. *Tipos de analítica*

Tipo	Pregunta central	Herramientas
Descriptiva	¿Qué ocurrió?	SQL, ETLs, dashboards, visualizaciones,
Diagnóstica	¿Por qué ocurrió?	Análisis <i>drill-down</i> , correlación de variables, consultas ad-hoc
Predictiva	¿Qué pasará?	Modelos de regresión, algoritmos de <i>machine learning</i> , redes neuronales,

		algoritmos de clasificación
Prescriptiva	¿Qué debemos hacer?	Optimización, simulación, motores de reglas, planificación de escenarios

Fuente: Adaptado de Camm (2020) y Sarker (2021b).

Inteligencia de negocios

2.3 Inteligencia de negocios en las finanzas

2.5 Herramientas tecnológicas

Objetivo específico n.º2: <i>Analizar datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica, considerando los factores mencionados, los cuales influyen en el valor de recuperación, para su integración del modelo.</i>	2.3 Inteligencia de negocios en las finanzas 2.4 Ciencia de datos 2.6 Vehículos eléctricos	Fase 2: Análisis exploratorio	Resultados de la Fase 2: Análisis exploratorio (EDA)	Conclusión: Objetivo específico dos	Recomendaciones
---	--	----------------------------------	--	--	-----------------

Objetivo específico n.º3: <i>Diseñar un modelo estadístico y de aprendizaje automático que integre los factores identificados, para la predicción del valor de recuperación de los vehículos eléctricos financiados por la entidad.</i>	2.4 Ciencia de datos	Fase 3: Diseño de los modelos predictivo	Resultados de la Fase 3: Diseño de los modelos predictivos	Conclusión: Objetivo específico tres	Recomendaciones
--	----------------------	---	--	---	-----------------

Objetivo específico n.º4: <i>Validar el modelo predictivo mediante pruebas con datos reales y técnicas de validación cruzada, para el aseguramiento de su exactitud y confiabilidad en la estimación del valor de recuperación de los vehículos eléctricos.</i>	2.4 Ciencia de datos	Fase 4: Validación del modelo	Resultados de la Fase 4: Validación del modelo Propuesta de solución	Conclusión: Objetivo específico cuatro	Recomendaciones
--	----------------------	----------------------------------	--	---	-----------------

Fuente: Elaboración propia (2025).

4. Análisis de resultados

Los datos generados al aplicar las actividades e instrumentos descritos en el marco metodológico se han estudiado, comparado y contrastado para extraer hallazgos relevantes. Esta sección incluye el núcleo crítico del Trabajo Final de Graduación, donde se discuten en profundidad los resultados obtenidos, se justifica las alternativas de solución propuestas y se evidencian las ventajas de su implementación.

4.1. Resultados de la Fase 1: Adquisición de la información

A continuación, se describen los datos recopilados de las fuentes primarias y secundarias: registros de ventas, rendimiento de baterías, importaciones y entrevistas a concesionarios. Se evaluó la calidad, cobertura y representatividad de cada conjunto de datos, identificando brechas y confirmando la pertinencia de las variables clave para el modelo.

En el marco del primer objetivo específico, se buscó construir un repositorio fiable sobre el mercado costarricense de vehículos eléctricos (VE), por lo que la Fase 1 se desarrolló siguiendo la lógica de **CRISP-DM**, cuyo primer paso consiste en comprender y depurar los datos antes de cualquier análisis exploratorio. En los siguientes párrafos, se describen, sin extrapolaciones, los insumos efectivamente disponibles después del proceso de limpieza y filtrado.

4.1.1. Fuentes consolidadas

Seguidamente, se presenta una tabla resumen con las fuentes de datos recopiladas que fueron depuradas, enriquecidas y utilizadas para la construcción del modelo final. Para un detalle del proceso de depuración y la descripción de las fuentes, véase el **Apéndice O**. Revisión documental #1.

Tabla 15. *Tabla de fuentes consolidadas*

Origen	Registros iniciales	Rango temporal	Registros retenidos
Áltica (Histórico de ventas)	179	2017-2025	177
Red Motors (Avalúos)	472	2019-2025	207
Web Scraping de CRAutos (Áltica)	606	2020-2023	575

<i>Web Scraping</i> de CRAutos (propio)	360	2023-2025	359
--	-----	-----------	-----

Fuente: Elaboración propia (2025).

En síntesis, el conjunto final de datos integra 1318 registros procedentes de cuatro fuentes complementarias, de modo que se garantiza una cobertura temporal continua entre 2017 y 2025.

4.1.2. Rendimiento de las baterías

No se dispuso de mediciones directas del estado de las baterías de los vehículos estudiados. No obstante, se incorporaron dos factores que, según la literatura, capturan buena parte de la variabilidad del desempeño de estas:

- **Tipo de batería:** A partir de fichas técnicas de páginas *web* como EV-Database, se asignó a cada modelo la tecnología predominante (Lithium-ion, LFP, NMC, Blade, entre otras), variable que influye en la autonomía y la tasa de degradación.
- **Condición cualitativa del vehículo:** Las fuentes internas ya clasificaban cada unidad como excelente, buena o regular, y también utilizaban la categoría de ‘desconocido’. Este indicador se mantuvo, ya que, aunque no es específico de la batería, aporta contexto sobre el mantenimiento y uso real de la unidad.

4.1.3. Criterios de depuración y enriquecimiento

En primer lugar, a cada conjunto de datos se le aplicó un filtro estricto para conservar únicamente vehículos estrictamente eléctricos, es decir, que se excluyeron los vehículos híbridos, así como aquellos vehículos de combustión clasificados erróneamente en los conjuntos de datos, los cuales fueron removidos de forma manual. Adicionalmente, se removieron registros duplicados en los datos provenientes del *web-scraping* realizado por Áltica, para evitar sesgos de representación en el modelo.

4.1.4. Validación experta

Para contextualizar las variables económicas, se llevó a cabo una entrevista abierta con el analista financiero de Red Motors, lo cual permitió profundizar en los criterios usados por la empresa para estimar el valor de reventa, apreciar el impacto del estado de la batería y entender la lógica de compra-venta de vehículos usados eléctricos. Para la transcripción resumida y los detalles discutidos, véase el **Apéndice N**.

4.1.5. Resultado de la fase y alienación con los objetivos

La fase concluye con un conjunto de datos limpio y enriquecido que respalda la representatividad del mercado local: abarca ocho años de transacciones, incluye cuatro fuentes complementarias y contempla variables tanto mecánicas como económicas y dimensionales. Este repositorio se convierte en la materia prima para la Fase 2: Análisis exploratorio de datos, donde se investigarán patrones, *outliers* y correlaciones que expliquen la depreciación de los VE en Costa Rica.

4.2. Resultados de la Fase 2: Análisis exploratorio (EDA)

Aquí se presentan los principales hallazgos del EDA: resumen estadístico de variables, detección y tratamiento de *outliers*, y patrones o correlaciones más significativas (p. ej., relación entre antigüedad del vehículo y su depreciación). Se acompaña de visualizaciones que sustentan la selección de factores determinantes. Una vez consolidado el conjunto de datos de 1 320 registros depurados, se aplicó un EDA con el fin de identificar los predictores que explican con mayor fuerza el valor de recuperación y, simultáneamente, diagnosticar posibles problemas de colinealidad antes de entrenar los modelos.

4.2.1. Introducción y delimitación

Conforme al segundo objetivo específico, esta sección examina las evidencias empíricas que vinculan características técnicas y de mercado con el valor de recuperación (VR) de los vehículos eléctricos vendidos en Costa Rica. Se parte de la premisa de que un VR estimado con precisión minimiza la exposición al riesgo crediticio de las entidades financieras y refuerza la planeación de su cartera (Damodaran, 2012). A lo largo de los apartados subsiguientes, se describen, primero, la fiabilidad del conjunto de datos; después, la influencia de los factores identificados sobre los precios de transacción y, por extensión, sobre el VR proyectado.

4.2.2. Calidad y consistencia de los datos

El proceso de limpieza partió de 1 617 registros brutos y concluyó con 1 318 registros válidos, excluyendo aquellos atribuibles a duplicados o a vehículos erróneamente clasificados. La

Figura 15. Descripción del conjunto de datos presenta los 24 campos finales, distinguiendo variables continuas, discretas y categóricas. Estas tareas de limpieza y estandarización fueron documentadas y pueden apreciarse en el **Apéndice P**.

Figura 15. Descripción del conjunto de datos

Descripción del Dataset			
Tipos de Columnas y Características			
Nombre de la Columna	Tipo de Dato	Tipo de Característica	Descripción
ValorVentaDolares	flotante	Númerica (Continua)	Precio de venta en dólares
Kilometraje	flotante	Númerica (Continua)	Kilometraje del vehículo
Estilo	string	Categórica	Tipo de carrocería (sedán, SUV, etc.)
Estado	string	Ordinal	Estado/condición del vehículo
FechaVenta	datetime	Temporal	Fecha en la que se vendió el vehículo
Antigüedad_Vehiculo	entero	Númerica (Discreta)	Antigüedad del vehículo en años
Año_Venta	entero	Númerica (Discreta)	Año en que se vendió el vehículo
Mes_Venta	entero	Númerica (Discreta)	Mes en que se vendió el vehículo
Estado_filled	string	Categórica	Estado con valores nulos imputados
make	string	Categórica	Marca estandarizada (minúsculas)
model	string	Categórica	Modelo estandarizado (minúsculas)
year	entero	Númerica (Discreta)	Año de fabricación (verificado)
style	string	Categórica	Carrocería estandarizada (minúsculas)
range_km	flotante	Númerica (Continua)	Autonomía eléctrica en kilómetros
usable_battery_kwh	flotante	Númerica (Continua)	Capacidad útil de la batería en kWh
rsp_dollars	flotante	Númerica (Continua)	Precio de venta recomendado en dólares
real_battery_kwh	flotante	Númerica (Continua)	Capacidad real de la batería en kWh
battery_type	string	Categórica	Tipo de tecnología de batería utilizada
number_of_cells	flotante	Númerica (Discreta)	Número de celdas de la batería
weight	flotante	Númerica (Continua)	Peso del vehículo en kg
length	flotante	Númerica (Continua)	Longitud del vehículo en mm
width	flotante	Númerica (Continua)	Ancho del vehículo en mm
height	flotante	Númerica (Continua)	Altura del vehículo en mm
efficiency	flotante	Númerica (Continua)	Métrica de eficiencia energética

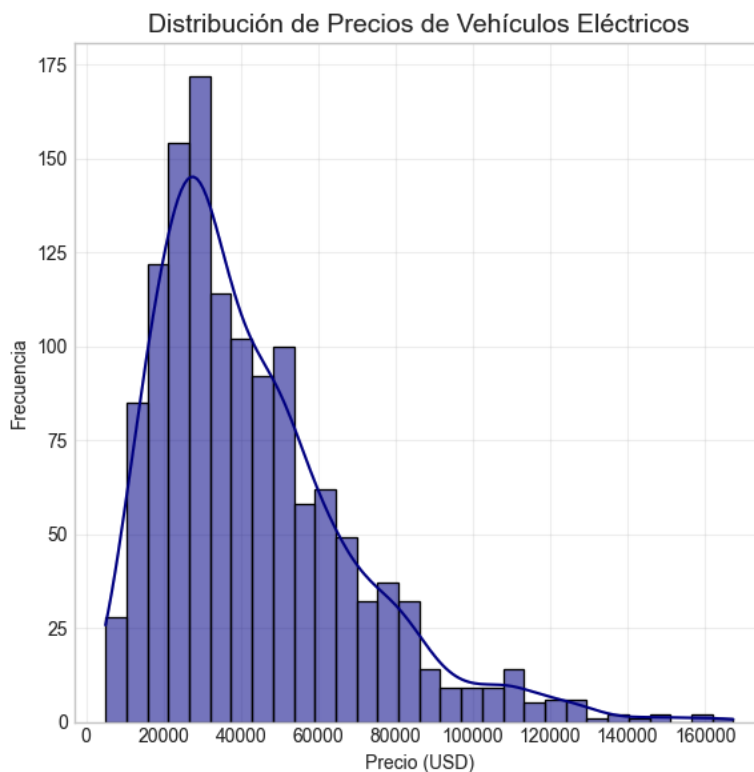
Fuente: Elaboración propia (2025).

4.2.3. Panorama del mercado

La **Figura 16** muestra la distribución de precios. La mayoría de los autos se vende por menos de \$60 000, pero existe un pequeño segmento de lujo que supera los \$120 000. Esta diferencia sugiere tratar los precios con una escala logarítmica en el modelo, de manera que los valores extremos no dominen el resultado.

También se observó que los vehículos con baterías mayores de 75 kWh suelen ubicarse en la parte alta de la curva, lo que confirma la importancia del rendimiento energético al momento de fijar precios.

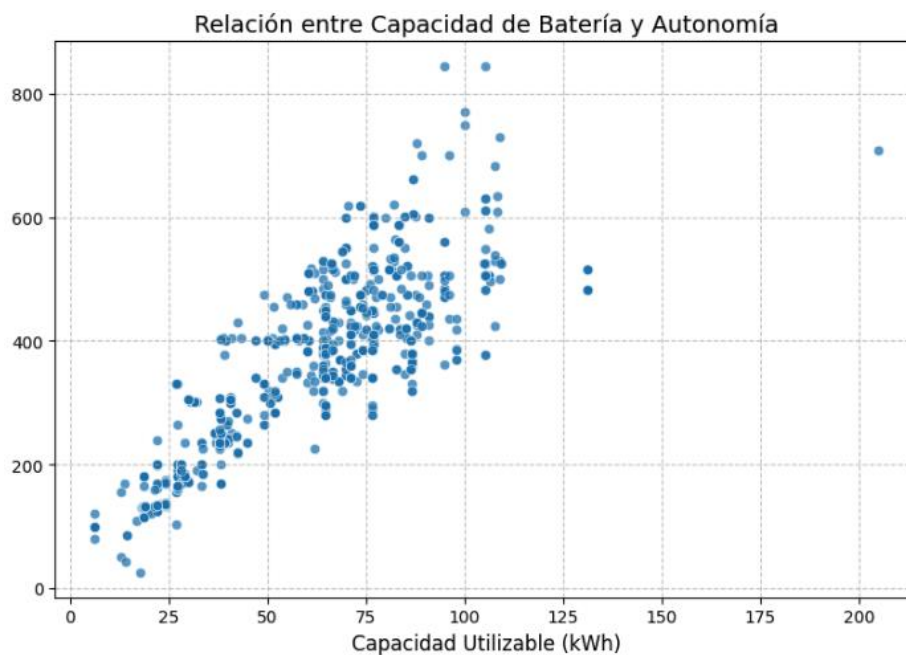
Figura 16. *Distribución de precios de vehículos eléctricos en Costa Rica*



Fuente: Elaboración propia (2025).

4.2.4. Relación entre la capacidad útil de la batería y la autonomía

Figura 17. *Relación entre capacidad de la batería y la autonomía*



Fuente: Elaboración propia (2025).

El diagrama de dispersión (**Figura 18**) pone de manifiesto una asociación positiva clara entre la energía almacenada en la batería y la distancia que el vehículo puede recorrer con una carga completa. A medida que la capacidad utilizable supera los 75 kWh, los puntos tienden a agruparse por encima de los 400 km, mientras que las unidades con menos de 25 kWh raramente superan los 200 km. Se observan, además, algunos casos atípicos (p. ej., un modelo con ~200 kWh y >700 km), lo que sugiere innovaciones de alta densidad energética.

La tendencia respalda los resultados estadísticos reportados en la sección **4.2.2**, donde la correlación de Pearson ($> 0,60$) entre capacidad y valor de recuperación (VR) se explicaba en parte por la autonomía como variable intermedia. Este patrón confirma que un mayor alcance operacional constituye un atributo percibido como valioso por el mercado, lo que refuerza la inclusión de ambas variables en los modelos de regresión. Asimismo, la dispersión creciente en el rango superior indica rendimientos decrecientes: cada kWh adicional aporta menos kilómetros conforme se incrementa la capacidad, fenómeno que deberá considerarse al interpretar los coeficientes en modelos lineales y en la calibración de algoritmos de ensamble.

Con esta evidencia gráfica, se completa el análisis exploratorio de las variables técnicas clave, el cual sirve como puente entre la caracterización descriptiva **4.2.3** y la selección definitiva de predictores **4.2.6**.

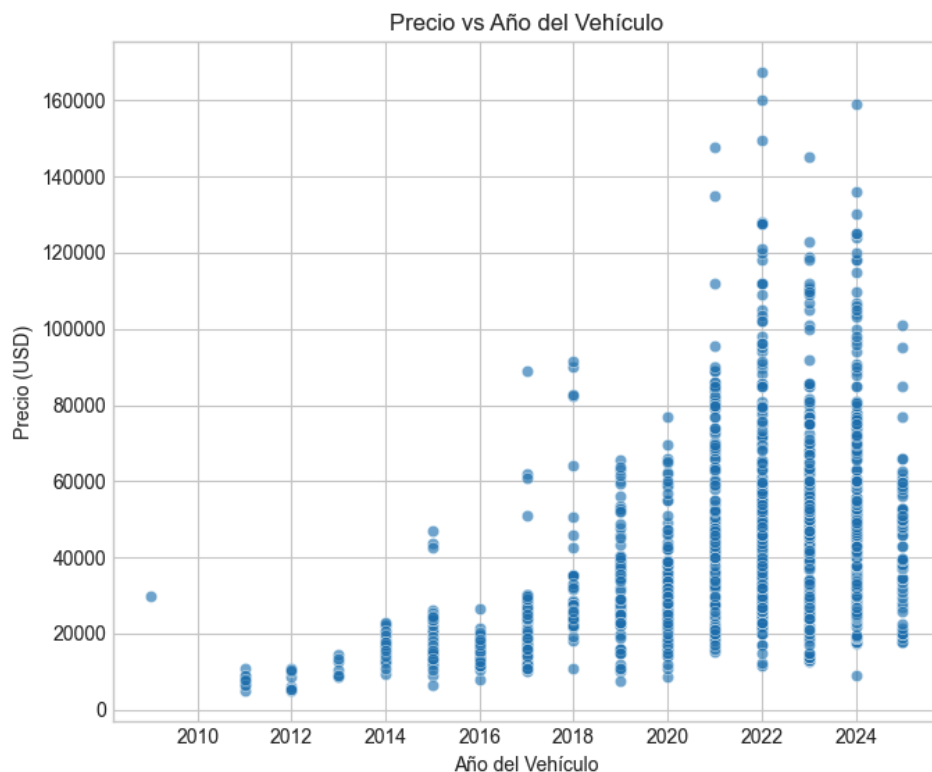
4.2.5. Antigüedad del vehículo y su impacto

La relación entre el año modelo del vehículo y su precio puede verse en la **Figura 18**, donde los modelos anteriores a 2018 raras veces superan los 40 000 USD, mientras que las unidades de 2021 a 2025 alcanzan picos que triplican esa cifra.

Cuando se agrupan los vehículos por rangos de edad (**Figura 19**), es posible notar tres patrones claros:

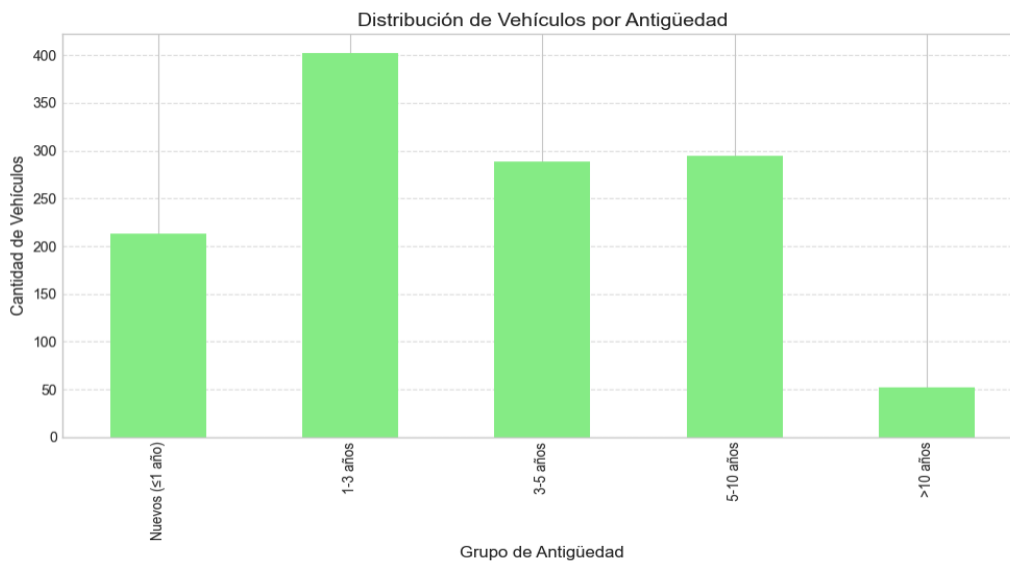
- **Nuevos (≤ 1 año):** retienen alrededor del **78 %** de su valor de lista.
- **1-3 años:** mantienen un **68 %**.
- **5 -10 años:** bajan a un **45 %**, y la caída se acrecienta a partir del quinto año.

Figura 18. *Precio vs año de fabricación*



Fuente: Elaboración propia (2025).

Figura 19. *Distribución de vehículos por antigüedad*



Fuente: Elaboración propia (2025).

Estos resultados confirman que la edad del auto es un predictor principal del VR y que su efecto no es lineal, ya que la depreciación se acelera después de la primera mitad de vida útil.

4.2.6. Factores técnicos y estructurales

La matriz de correlaciones encontrada en el **Anexo V** resume, en un solo plano de color, la relación lineal entre cada par de variables. Los tonos rojizos indican asociaciones positivas, los azulados negativas y la intensidad muestra la fuerza del vínculo. Un valor cercano a +1 significa que dos variables crecen al mismo tiempo; un valor cercano a -1 señala que, cuando una sube, la otra baja; cifras próximas a 0 evidencian relación débil o nula. Al analizarla detalladamente, destacan tres grupos de resultados:

1. **Batería y autonomía como impulsores del precio:** La capacidad útil de la batería (*usable_battery_kWh*) y la autonomía real (*range_km*) se asocian de forma clara con el valor de venta (coeficientes > 0.60). En la práctica, los vehículos que ofrecen mayor alcance por carga tienden a retener mejor su valor de mercado, un indicador clave para proyectar su valor de recuperación.
2. **Dimensiones físicas con efecto moderado:** Los valores del peso, el largo y el ancho muestran correlaciones positivas intermedias (≈ 0.55). Aunque no son tan influyentes como la batería, los autos más robustos y espaciosos se perciben como de mayor calidad, lo que añade algunos puntos al valor final.
3. **La antigüedad resta valor de forma consistente:** La variable *Antigüedad_Vehículo* presenta un coeficiente negativo cercano a -0.60 . Esto confirma la tendencia observada en la sección 4.2.4: a mayor edad, menor precio de mercado y, por ende, menor valor de recuperación esperado.

Un detalle adicional que se debe mencionar es que las variables *usable_battery_kWh* y *real_battery_kWh* comparten casi la misma información (correlación ≈ 0.99). Para evitar que el modelo “cuente dos veces” el mismo efecto, se conservará solo la primera. Este filtro, junto con la futura transformación logarítmica del precio, prepara un conjunto de variables compacto y sin redundancias.

4.2.7. Síntesis e integración

En conjunto, el EDA confirma que los aspectos de edad, kilometraje, capacidad útil/real de batería, autonomía nominal y segmento de mercado son los predictores más influyentes sobre *ValorVentaDolares*. Para evitar multicolinealidad, se retendrá un solo indicador de capacidad y se agruparán las variables dimensionales. Además, la distribución sesgada de precios aconseja probar transformaciones logarítmicas en los modelos lineales y pérdidas robustas en los algoritmos de ensamble.

Estos hallazgos nutren la siguiente sección, en la que se explica cómo cada técnica (Regresión múltiple, Lasso, *Gradient Boosting*, SARIMA) incorpora las variables seleccionadas y gestiona los patrones descubiertos.

4.3. Resultados de la Fase 3: Diseño de los modelos predictivos

En esta tercera fase del ciclo CRISP-DM, se materializa el tercer objetivo específico, el cual consiste en *diseñar un modelo estadístico y de aprendizaje automático capaz de integrar los factores identificados durante el análisis exploratorio*. De manera progresiva, el apartado detalla la lógica seguida para transformar las hipótesis previas en arquitecturas concretas: se define la canalización de preparación de datos, se justifican los algoritmos comparados —incluidos los modelos temporales ARIMA y SARIMA contemplados inicialmente— y se describe la estrategia de búsqueda de hiperparámetros que soporta la selección del ensamblaje final. Dicho recorrido, apoyado en las subactividades previstas para la Fase 3, permite demostrar la coherencia metodológica entre los hallazgos del EDA y el diseño de los modelos predictivos estadísticos que sirven de base a la validación posterior.

4.3.1. Delimitación del problema y criterios de modelado

El propósito es estimar el VR de vehículos eléctricos al término del contrato de financiamiento. Este VR se modela como función de variables técnicas (capacidad útil de batería, autonomía, potencia) y comerciales (edad, kilometraje, antigüedad). Para que el modelo sea operativo en un flujo de tasación mensual se establecieron tres criterios:

1. Interpretabilidad parcial para respaldar decisiones de riesgo.
2. Actualización semestral sin requerir intervención manual extensa.
3. Desempeño mínimo esperado: $R^2 \geq 0.8$ y $MAE \leq 8\,000$ USD.

4.3.2. Arquitectura de la canalización de datos

Para estandarizar el *pre-processing* y mantener un flujo reproducible, se construyó un **ColumnTransformer** de *scikit-learn*, el cual orquesta dos flujos o *sub-pipelines* especializados: uno para variables numéricas y otro para variables categóricas. Puede consultarse el **Apéndice P** y **Apéndice Q** para obtener mayor detalle sobre este punto.

- **Sub-pipeline numérico**

1. **Imputación por mediana:** Se sustituye cualquier valor faltante por la mediana de la columna; esta medida es robusta a valores extremos y mantiene la distribución original mejor que la media.

2. **Escalado robusto (RobustScaler):** El escalado se basa en el rango intercuartílico (IQR), de modo que los *outliers* influyen mínimamente en el ajuste de los parámetros de escala. Así, los modelos sensibles a la magnitud de las variables—por ejemplo, los basados en regularización L1/L2—reciben datos en un rango comparable.
- **Sub-pipeline categórico**
 1. **Imputación por etiqueta “other”:** Los valores nulos se etiquetan como “*other*”, evitando la pérdida de registros y capturando grupos poco frecuentes en una clase controlada.
 2. **One-Hot Encoding** con salida densa: Cada categoría se transforma en una columna binaria, lo que habilita a los algoritmos lineales y de árbol a interpretar la información nominal sin introducir orden artificial. La salida se mantiene en formato denso para facilitar la integración con los ensambles que no aceptan matrices dispersas nativamente.

El ***ColumnTransformer*** actúa como *envoltorio* único que encapsula todas las transformaciones y las aplica de forma coherente durante el ajuste, la validación cruzada y el despliegue. Gracias a esta abstracción, cualquier estimador (Lasso, *Random Forest*, *Gradient Boosting*, etc.) puede “conectarse” en la etapa posterior sin reescribir código o correr riesgo de *data leakage*. Además, al quedar serializado dentro de un objeto Pipeline, se garantiza que los pasos de imputación y escalado se entrenan exclusivamente con los datos de entrenamiento en cada iteración o *fold*, de manera que se preserve la integridad estadística del proceso.

4.3.3. Selección de algoritmos y justificación

El diseño contempló tres familias de algoritmos, representados en la siguiente:

Tabla 16. *Familias de algoritmos*

Familia	Algoritmo	Rol dentro del diseño	Motivo de selección/Descarte
Lineales regularizados	Lasso	Línea base interpretable	Reduce multicolinealidad y revela la importancia de cada

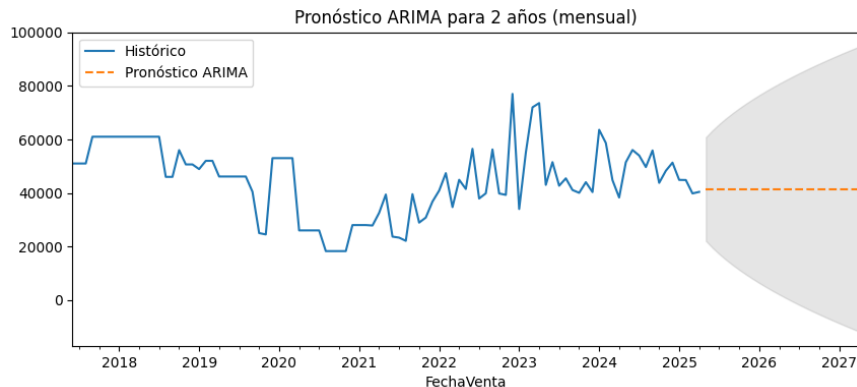
			variable
Ensamblados no lineales	Árbol de decisión, <i>Random Forest</i> , <i>Gradient Boosting</i> , <i>XGBoost</i>	Capturar interacciones complejas	Robustos a outliers, alto desempeño con poco ajuste y buena generalización
Serie temporales	ARIMA y SARIMA	Explorar componente temporal	Modelan tendencia estacional; descartados por bajo desempeño

Fuente: Elaboración propia (2025).

4.3.4. Modelos ARIMA y SARIMA

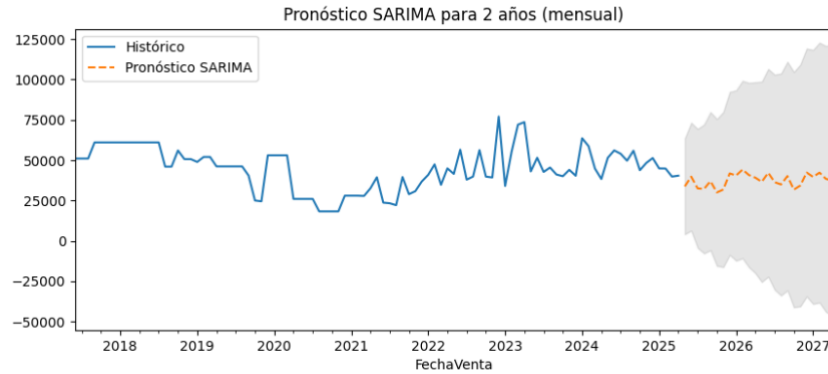
A continuación se muestran los pronósticos generados con los modelos ARIMA y SARIMA para los próximos 24 meses:

Figura 20. *Modelo ARIMA predicción*



Fuente: Elaboración propia (2025).

Figura 21. *Modelo SARIMA predicción*



Fuente: Elaboración propia (2025).

En ambos gráficos se aprecia la serie histórica (línea azul sólida) y el pronóstico (línea naranja discontinua) con su banda de incertidumbre (área gris). El modelo ARIMA (0, 1, 1) (Figura 20) produce un pronóstico prácticamente plano y bandas muy anchas, lo que refleja poca capacidad para explicar la variación futura del precio; sus errores de prueba superan los \$20 000. El modelo SARIMA (0, 1, 1)×(1, 1, 1){12} (Figura 21) ajusta mejor la estacionalidad anual y reduce moderadamente la dispersión, pero todavía mantiene un MAE cercano a \$18 000. Dado que ambos resultados están significativamente por encima del umbral operativo ($\leq 10\,000$ USD), los enfoques temporales se descartan para producción y se conservan solo como referencia metodológica.

4.3.5. Diseño del modelo ensamblado (*Stacking*)

Los mejores dos modelos *Random Forest* y *Gradient Boosting* (según R^2 -CV) se integraron en un *StackingRegressor* con una regresión Ridge como metamodelo. El ensamblaje combina la baja varianza del bosque con el bajo sesgo del *boosting*, logrando R^2 -CV $\approx 0,83$ y cumpliendo los criterios fijados. Cabe destacar que el XGBoost no fue contemplado para ser utilizado en el *stacking* debido a que corresponde a una variación del *Gradient Boosting*, por lo que se optó por utilizar el modelo fuente o base. La elección puede consultarse con mayor detalle en el **Apéndice R**.

4.4. Resultados de la Fase 4: Validación del modelo

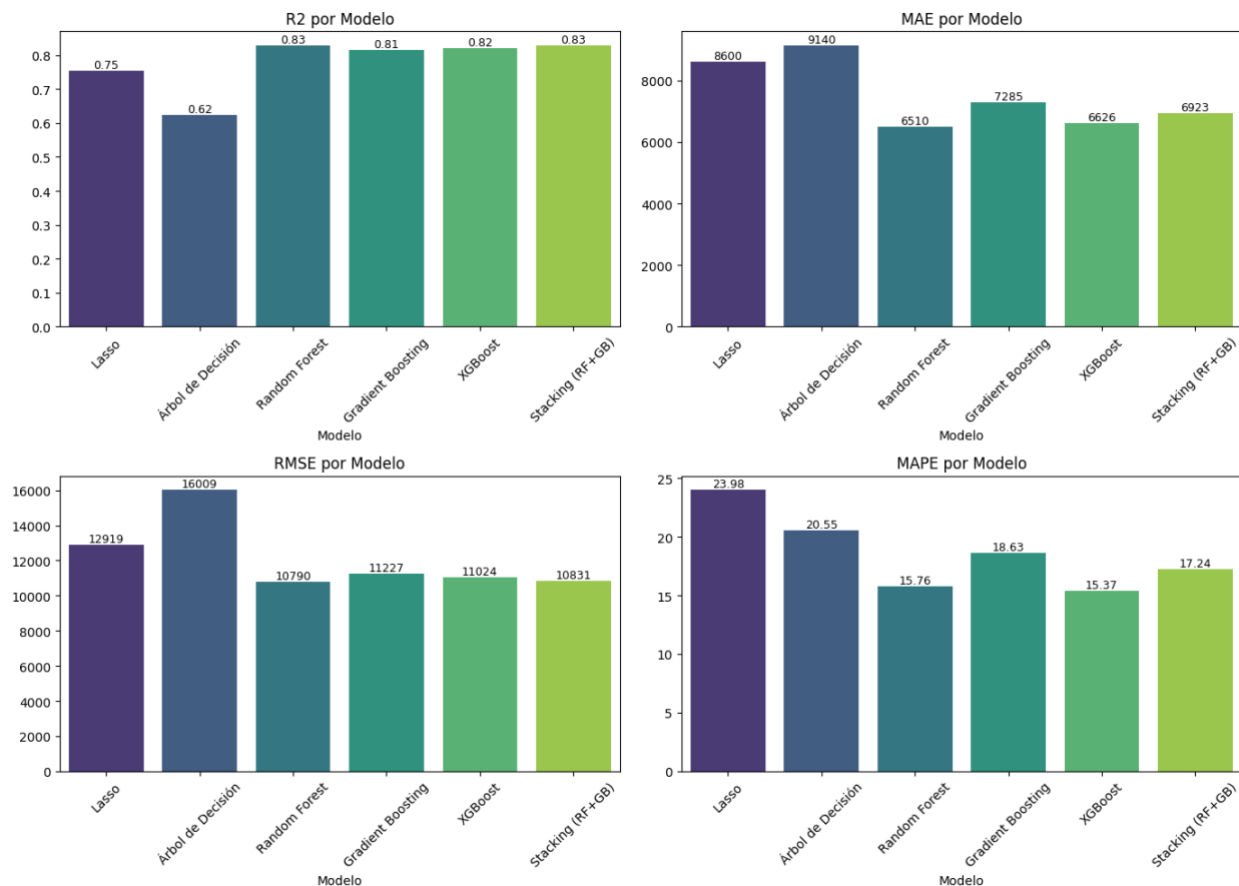
Una vez definido el mejor candidato de modelo, la Fase 4 aborda la *evaluación rigurosa de su exactitud y confiabilidad*, con el fin de alcanzar el cuarto objetivo específico. Siguiendo las pautas metodológicas descritas para la validación del modelo, esta sección expone el esquema de particionamiento de datos, las métricas empleadas y los resultados obtenidos en el conjunto de prueba, contrastándolos con los umbrales de aceptación establecidos por la entidad. Asimismo, se

interpretan los hallazgos clave, errores medios, estabilidad en validación cruzada y análisis de residuales que evidencian la robustez del modelo seleccionado y justifican su idoneidad para apoyar la estimación operativa del valor de recuperación de los vehículos eléctricos.

4.4.1. Exactitud y confiabilidad

Para determinar la confiabilidad, se aplicó un esquema entrenamiento/prueba (80-20 %) con validación cruzada estratificada. Las métricas clave fueron R^2 , MAE, RMSE y MAPE. La **Figura 22** resume el desempeño en el conjunto de prueba:

Figura 22. Desempeño de los modelos evaluados



Fuente: Elaboración propia (2025).

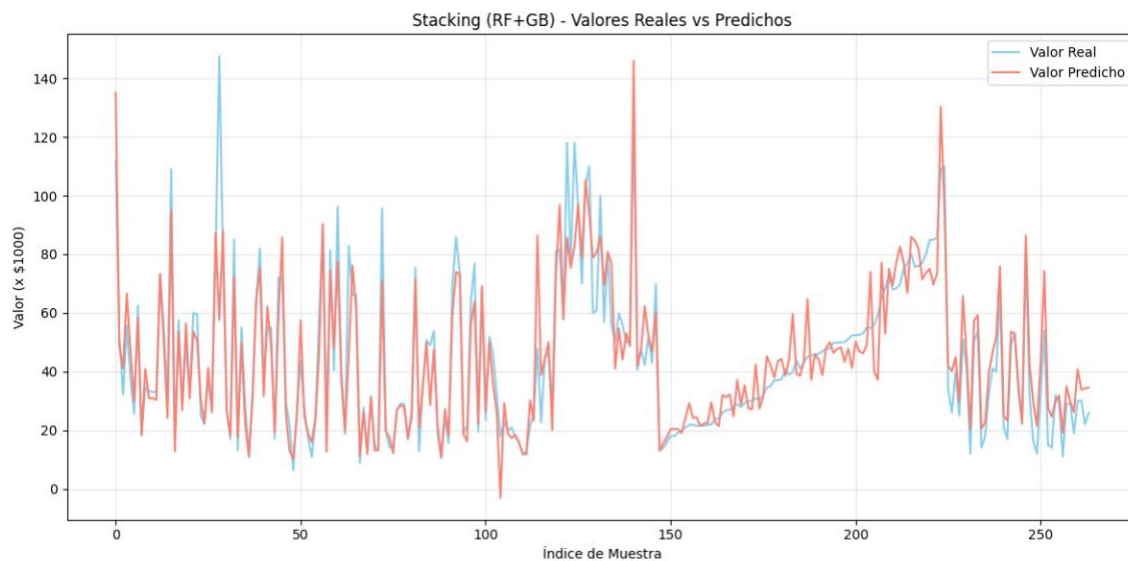
A partir de la figura anterior, se puede interpretar lo siguiente:

- El **Stacking Regressor** superó a los modelos individuales, explicando **83 %** de la variabilidad de precios y manteniendo errores absolutos medios alrededor de \$6900.
- El **Random Forest** presentó el MAE y MAPE más bajos, lo que lo convierte en un excelente *benchmark* para escenarios donde se prioriza el error promedio sobre la capacidad explicativa global.

- El modelo lineal (Lasso) evidencia limitaciones frente a relaciones no lineales —aunque es útil como línea base y aporta interpretabilidad—.

Para garantizar la confiabilidad, se verificó la estabilidad de las métricas. La desviación estándar del R^2 en CV para el *stacking* fue ≤ 0.02 , y la **Figura 23** muestra ausencia de patrones sistemáticos en los residuales. La opinión del gerente de TI y del analista financiero sobre los resultados de estas métricas se detallan en el **Apéndice S**.

Figura 23. *Valor real vs. valor predicho*



Fuente: Elaboración propia (2025).

Con estos resultados, se valida el cuarto objetivo específico. El modelo seleccionado exhibe exactitud y consistencia suficientes para ser desplegado como herramienta de apoyo en la valoración de recuperación, reduciendo la exposición al riesgo financiero de la entidad.

En suma, el *Stacking Regressor* compuesto por *Random Forest* (reducción de varianza) y *Gradient Boosting* (reducción de sesgo) ofrece el mejor compromiso global entre potencia predictiva y estabilidad. Su $R^2 \approx 0,83$ cumple holgadamente el umbral de explicabilidad ($\geq 0,75$), mientras que su MAE $\approx \$6900$ se mantiene dentro del rango operativo ($\leq \$10\,000$). Además, la desviación estándar de las métricas en validación cruzada ($\sigma \leq 0,02$) confirma una varianza muy baja entre pliegues, lo que mitiga el riesgo de sobreajuste y asegura un desempeño consistente en producción. Al combinar los puntos fuertes de ambos algoritmos, el ensamblaje captura relaciones no lineales complejas sin sacrificar la robustez ni la interpretabilidad relativa. Por estas razones, el

ensamblaje *Stacking* Regressor se selecciona como modelo definitivo por su plena alineación con los requisitos técnicos y de negocio para la tasación del valor de recuperación, así como por su capacidad para proporcionar a la entidad la mayor reducción plausible del riesgo financiero.

5. Propuesta de solución

Sobre la base de los resultados expuestos en el capítulo anterior, donde el ensamblaje *Stacking Regressor* explicó 83 % de la variabilidad del precio y mantuvo un MAE aproximado de \$6 800, este capítulo presenta la ruta para transformar el modelo analítico en una herramienta operativa dentro del proceso de tasación de Áltica. Primero se describe la solución —tanto su arquitectura técnica como su integración en el flujo de avalúos—; luego se detalla un plan de implementación que abarca piloto, despliegue y gobernanza; y, finalmente, se estiman los costos de elaboración y mantenimiento del modelo. Con ello, se cierra el ciclo CRISP-DM al pasar de la validación al despliegue, garantizando que los beneficios observados en la fase de prueba se materialicen en la práctica corporativa.

5.1. Descripción de la solución

La solución propuesta convierte el modelo *Stacking Regressor* en un servicio operable dentro del valor de recuperación de los vehículos eléctricos en Áltica. Para ello se encapsula la transformación y preparación de datos (ColumnTransformer con subpipelines numérico y categórico) y el algoritmo de ensamble en un objeto *pipeline* de *scikit-learn*, cumpliendo los criterios de interpretabilidad y desempeño mínimo definidos en la sección 4.3.1. El servicio resultante se expone mediante una API interna que recibe las características del vehículo, ejecuta el *pipeline* serializado y devuelve la estimación del valor de recuperación (VR), integrándose de forma transparente en el proceso de avalúos para vehículos eléctricos.

5.1.1. Arquitectura analítica del modelo

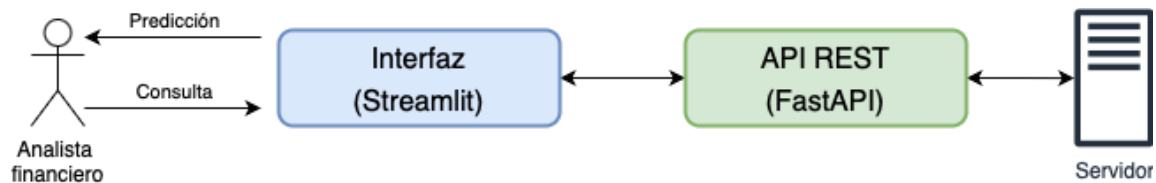
En primer lugar, la propuesta descansa en dos artefactos independientes creados con **joblib**: *preprocesador.joblib* y *modelo.joblib*. El *preprocesador.joblib* contiene el **ColumnTransformer**, que, tal y como se definió en la sección 4.3.2, imputa los valores faltantes con la mediana y aplica un escalado robusto a las variables numéricas, mientras que a las variables categóricas les asigna la etiqueta *other* y realiza *one-hot encoding*; de este modo, se garantiza que los datos enviados a producción reciban exactamente el mismo tratamiento que durante el entrenamiento. Por su parte, *modelo.joblib* almacena el *Stacking* modelo en la validación final.

Asimismo, ambos archivos se alojan en el servidor analítico en Áltica empleado para procesamiento de datos. Allí se ejecutan dos capas ligeras en Python:

1. **API REST (FastAPI + Uvicorn):** Al iniciarse, carga los dos artefactos; cada petición convierte los atributos del vehículo en un DataFrame, aplica el preprocesador y devuelve la estimación del valor de recuperación (VR) en milisegundos.
2. **Interfaz (Streamlit):** Mediante un formulario *web*, el analista introduce año, capacidad útil, autonomía, kilometraje y tipo de batería, y recibe al instante la predicción proveniente de la API.

Por consiguiente, esta arquitectura mantiene la coherencia entre entrenamiento y uso operativo, evita fugas de información y permite sustituir, de forma independiente, el modelo o el preprocesador cuando proceda un nuevo entrenamiento, al tener disponibilidad a más información.

Figura 24. *Flujo de predicción del valor de recuperación*



Fuente: Elaboración propia (2025).

El diagrama anterior muestra, de izquierda a derecha, el recorrido de los datos: la persona usuaria accede al interfaz Streamlit, que envía las variables del vehículo a la API FastAPI. La API, alojada en el servidor analítico, aplica primero el preprocesador (*joblib*) para transformar la entrada y, a continuación, pasa el resultado al modelo (*joblib*), el cual devuelve la estimación del valor de recuperación (VR). Finalmente, la respuesta se presenta de nuevo en la interfaz.

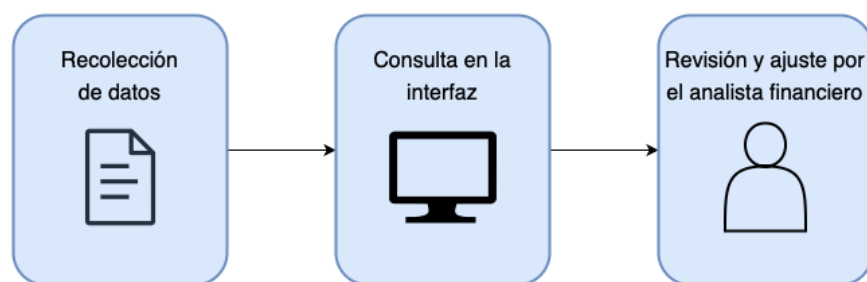
5.1.2. Flujo de datos y proceso de tasación

En lugar de integrarse a un sistema transaccional, el modelo se utiliza como herramienta de apoyo directo para el perito valuador. De esta manera, se preserva la intervención experta, al tiempo que se aporta una referencia cuantitativa coherente con los análisis de la sección 4. El procedimiento se desarrolla en tres hitos sencillos, los cuales se ilustran en la **Figura 25**.

1. **Captura de atributos del vehículo:** El valuador reúne los datos técnicos y comerciales necesarios (año-modelo, capacidad utilizable de la batería, autonomía real, kilometraje, tipo de batería, precio de compra, entre otros). Estos campos coinciden con los predictores finales definidos durante la construcción del modelo, por lo que no se requiere información adicional.

2. **Consulta al modelo mediante el formulario *web*:** Con los datos a la mano, la persona encargada abre la interfaz Streamlit alojada en el servidor de la empresa, introduce la información y hace clic en *Predecir*. El formulario envía la petición a la **API FastAPI**, la cual aplica *preprocesador.joblib*, ejecuta *modelo.joblib* y devuelve la estimación del valor de recuperación (VR) de forma inmediata.
3. **Revisión y ajuste con juicio experto:** El sistema presenta el VR predicho junto con el rango indicativo del error medio absoluto ($\approx$ \$6 900). A partir de dicha cifra, el valuador decide —con base en factores cualitativos que el modelo no captura, como el estado estético o la inminente salida de un modelo sustitutivo— si mantiene el valor sugerido o lo calibra antes de emitir el informe final.

Figura 25. *Flujo de trabajo con el modelo de predicción*



Fuente: Elaboración propia (2025).

5.1.3. Infraestructura tecnológica

Esta subsección detalla el entorno donde se alojarán el API y la interfaz, asegurando que el modelo opere con tiempos de respuesta ágiles y con el menor costo posible. Dado que Áltica ya dispone de un servidor para tareas analíticas, no se requiere de la adquisición de ningún servicio adicional.

Tabla 17. *Infraestructura tecnológica del servidor*

Aspecto	Especificación recomendada
Sistema operativo	Windows Server 2022 Datacenter Version 21H2
Recursos	Procesador: Intel(R) Xeon(R) CPU E5-2673 v4 @ 2.30GHz 2.29 GHz

	RAM: 32.0 GB Tipo de sistema operativo: 64-bit
Entorno de Python	Python 3.10 + pip
Librerías clave	Scikit-learn, joblib, fastapi, uvicorn, pandas, numpy streamlit
Orquestación	Servicio en el TaskScheduler, el cual ejecuta un script que es encarga de ejecutar el API y la interfaz

Fuente: Elaboración propia (2025).

5.1.4. Gobierno y ciclo de vida del modelo

El modelo se entrega como un par de artefactos *preprocesador.joblib* y *modelo.joblib*, depositados en el servidor privado usado por el equipo analítico, y los archivos se incluyen en la rutina de respaldo nocturno ya configurada para el servidor.

Aunque no se prevé reentrenamiento dentro de este proyecto, el gerente de TI tendrá a su disposición los *notebooks* originales con el código fuente y la configuración de entrenamiento. Si en el futuro se recopilan datos más recientes o nuevas variables, podrá ejecutar de nuevo el flujo de entrenamiento, evaluar los resultados y sustituir los archivos *joblib* por la versión mejorada, siguiendo el mismo esquema de protección y respaldo. De este modo, se preserva la estabilidad operativa actual sin cerrar la puerta a mejoras posteriores.

5.2. Costos de elaboración del modelo

Con el fin de determinar la inversión requerida para desarrollar y poner en marcha la solución, se identificaron los egresos vinculados al recurso humano y a gastos externos puntuales. Es necesario documentar dichos costos de la elaboración del presente modelo.

5.2.1. Costos salarios de los involucrados

Algunos datos utilizados en este cálculo son confidenciales; por consiguiente, se muestran únicamente los resultados agregados, evitando la divulgación de información sensible conforme a la solicitud de la organización.

A continuación, se describen los principales costos identificados para las tres personas que participaron en la formulación del proyecto:

- Estudiante investigador:

Este rubro corresponde al estipendio mensual asignado al estudiante responsable de la recopilación de datos, su análisis posterior, el diseño y entrenamiento del modelo, la validación correspondiente y la redacción del documento. En este caso, para obtener el salario de este puesto, se utiliza lo acordado al inicio del desarrollo del Trabajo Final de Graduación.

Tabla 18. *Costo por salario de Estudiante investigador*

Total de salario mensual	Cantidad de meses	Costo total
\$800	4	\$3200

Fuente: Elaboración propia (2025).

- Gerente de TI:

En relación con los costos dedicados al salario de este involucrado, estos se calculan según las horas dedicadas a la participación en los instrumentos de recolección de datos, así como en las reuniones de progreso del Trabajo Final de Graduación. Para ello, se consultó la página *web* de Glassdoor para obtener el salario de un puesto similar correspondiente a un *manager* de TI, el cual cuenta con una remuneración promedio de \$5 700 mensuales, incluyendo el monto correspondiente de las cargas sociales.

Tabla 19. *Costo por salario de Gerente de TI*

Total de salario por hora	Cantidad de horas	Costo total
\$33	15	\$495

Fuente: Elaboración propia (2025).

- Analista financiero del sector automotriz:

En lo que se refiere a los costos dedicados al salario de este involucrado, estos se calculan según las horas dedicadas a la participación en los instrumentos de recolección de datos, así como en las reuniones de progreso del Trabajo Final de Graduación. Para ello, se consultó la página *web* de Glassdoor para obtener el salario de un puesto similar correspondiente a un *Analista financiero*,

el cual cuenta con una remuneración promedio de \$4 000 mensuales, incluyendo el monto correspondiente de las cargas sociales.

Tabla 20. *Costo por salario de Analista financiero del sector automotriz*

Total de salario por hora	Cantidad de horas	Costo total
\$24	4	\$96

Fuente: Elaboración propia (2025).

5.2.2. Costos herramienta utilizada

También fue necesario considerar otros costos directos, los cuales tuvieron un impacto durante el desarrollo del Trabajo Final de Graduación. Estos costos corresponden al enriquecimiento de la información para su uso en la construcción del modelo, y provienen de las peticiones realizadas al API de ChatGPT para completar la información de los registros depurados.

Tabla 21. *Costo por herramienta*

Total de costo por consulta	Cantidad de consultas	Costo total
\$0.03	1824	\$57

Fuente: Elaboración propia (2025).

No se incluyen costos de infraestructura adicionales, dado que el servidor ya forma parte de uso operativo de Áltica, y no se requirió la adquisición adicional de ninguna licencia, espacio de almacenamiento o poder de procesamiento.

5.2.3. Costos totales

Una vez obtenida una clara percepción de todos los costos relacionados con los salarios de los involucrados, así como el costo de la herramienta utilizada, se obtienen los costos totales, tal como se muestra en la siguiente tabla.

Tabla 22. *Costos totales*

Concepto	Monto en dólares
----------	------------------

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

Total de salario pagado al estudiante	\$3200
Total de salario pagado al Gerente de TI	\$495
Total de salario pagado al Analista del sector automotriz	\$96
Costo herramienta utilizada	\$57
Total	\$3848

Fuente: Elaboración propia (2025).

6. Conclusiones

Este capítulo sintetiza los hallazgos más relevantes del análisis de resultados y verifica el logro de cada objetivo propuesto. Primero, se exponen las conclusiones relativas a los cuatro objetivos específicos, los cuales siguen el flujo metodológico de adquisición, análisis, diseño y validación. Finalmente, se presenta la conclusión asociada al objetivo general. El propósito es demostrar, de forma concisa y con evidencia cuantitativa, cómo el modelo predictivo diseñado responde a la problemática de la variabilidad en el valor de recuperación (VR) de los vehículos eléctricos financiados por la entidad.

6.1. Conclusión: Objetivo específico uno

El primer objetivo se define como: *Recolectar datos históricos y relevantes sobre el mercado de vehículos eléctricos en Costa Rica, incluyendo factores como ventas, valor actual del mercado y rendimiento de las baterías, a partir de una base sólida de información confiable*. Las conclusiones definidas fueron las siguientes:

- Se consolida un repositorio de 1 318 registros válidos, procedentes de cuatro fuentes (ventas, avalúos y dos *web-scrapings*), tras depurar 1 617 registros brutos. La **Tabla 15** y la sección **4.1.1** evidencian que se cumplen los criterios de integridad definidos.
- La validación experta en la sección **4.1.4** confirmó la representatividad de las variables técnicas y comerciales, garantizando que los ocho años de transacciones reflejan fielmente el mercado costarricense.

6.2. Conclusión: Objetivo específico dos

El segundo objetivo se define como: *Analizar los datos históricos y actuales del mercado de vehículos eléctricos en Costa Rica, considerando los factores mencionados, los cuales influyen en el valor de recuperación, para su integración en el modelo*. Las conclusiones alcanzadas se detallan a continuación:

- La antigüedad del vehículo reduce sistemáticamente el valor de recuperación. La variable *Antigüedad_Vehículo* muestra correlación negativa cercana a $-0,60$ y la curva de depreciación se acelera a partir del quinto año; consultar gráficos de la sección **4.2.4**.

- El kilometraje actúa como refuerzo del efecto edad. Aunque su aporte es más moderado (correlación alrededor de $-0,45$), el análisis confirma que los kilómetros recorridos se sitúan entre los cinco predictores más influyentes; revisar la sección 4.2.5.
- La relación entre capacidad y autonomía es aproximadamente proporcional dentro del rango observado. Aunque existe dispersión creciente por encima de 100 kWh, la tendencia global se mantiene positiva y sin evidencia clara de rendimientos decrecientes, por lo que captura adecuadamente el patrón principal; véase **Figura 17**.
- Se evitó la multicolinealidad y se mitigaron sesgos de escala. Para controlar redundancias, se retuvo una sola medida de capacidad de batería y se aplicaron transformaciones logarítmicas al precio antes de alimentar algoritmos lineales; los detalles se encuentran en la sección 4.2.7.

6.3. Conclusión: Objetivo específico tres

El tercer objetivo se define como: *Diseñar un modelo estadístico y de aprendizaje automático que integre los factores identificados, para la predicción del valor de recuperación de los vehículos eléctricos financiados por la entidad*. Las conclusiones se definen a continuación:

- Siguiendo la metodología detallada en las secciones 4.3.3 y 4.3.5, se seleccionó un *Stacking Regressor* (*Random Forest + Gradient Boosting*) que alcanzó $R^2 \approx 0,83$ en la validación cruzada y que cumplió los criterios operativos.
- La arquitectura elegida combina baja varianza y bajo sesgo, proporcionando un equilibrio entre potencia explicativa e interpretabilidad, tal como se justifica en la comparación de familias de algoritmos de la **Tabla 16**.

6.4. Conclusión: Objetivo específico cuatro

El cuarto objetivo se define como: *Validar el modelo predictivo mediante pruebas con datos reales y técnicas de validación cruzada, para el aseguramiento de su exactitud y confiabilidad en la estimación del valor de recuperación de los vehículos eléctricos*. Las conclusiones se definen a continuación:

- La sección 4.4.1 reporta que el modelo de ensamble o *Stacking* obtuvo $R^2 = 0,83$, $MAE = \$6\,900$ y una desviación estándar de $R^2 \leq 0,02$, demostrando estabilidad en los pliegues y ausencia de sobreajuste; véase **Figura 22** y **Figura 23**.

- Los modelos temporales ARIMA y SARIMA son descartados debido a su rendimiento, ya que sus respectivos MAE superan en \$18 000 el criterio definido, muy por encima del rango aceptable, tal como se observa en la sección 4.3.4.

6.5. Conclusión: Objetivo general

El objetivo general se define como: *Desarrollar un modelo funcional de un modelo predictivo analítico del valor de recuperación de los vehículos eléctricos financiados por la entidad asegurando su exactitud y confiabilidad mediante pruebas y validación técnica, durante el primer semestre de 2025.* La conclusión alcanzada es la siguiente:

- Se ha desarrollado un modelo predictivo operativo que estima con alta precisión el valor de recuperación (VR) de los vehículos eléctricos financiados por Áltica. El ensamblaje *Stacking Regressor*, obtiene un $R^2 = 0,83 \pm 0,02$ y un MAE $\approx \$6\,900$, lo que disminuye la incertidumbre sobre la depreciación y ofrece a la gerencia de riesgos una herramienta objetiva para ajustar sus políticas crediticias, como puede comprobarse en la sección 4.4.1 y en la **Figura 23**. *Valor real vs. valor predicho.*

7. Recomendaciones

Con base en los resultados obtenidos y las conclusiones expuestas en el capítulo anterior, a continuación se presentan las acciones prioritarias para mejorar la exactitud del modelo, reforzar su uso operativo y asegurar la sostenibilidad del proceso de tasación a largo plazo. Estas recomendaciones combinan ajustes en la captura de datos, alineación con las políticas de riesgo y posibles extensiones del conjunto de predictores, de modo que la herramienta continúe aportando valor conforme evolucione el mercado de vehículos eléctricos y la disponibilidad de información. En consecuencia, se recomienda:

- Mantener un registro histórico anual de precios de lista y de transacción de cada modelo y versión de vehículo eléctrico; la variable resultante (“PrecioNuevo”) permitirá al algoritmo aprender tasas de depreciación normalizadas por cohorte de mercado.
- Incorporar, tan pronto como sea posible, la medición del estado de salud (SoH) de la batería como predictor adicional, dado que refleja directamente la capacidad útil remanente y, por ende, el valor de recuperación.
- Programar un reentrenamiento integral dentro de seis meses, empleando todos los registros capturados desde la puesta en producción y verificando que el MAE no empeore respecto de la versión actual.
- Definir una política de excepción que autorice al valuador a apartarse de la recomendación del modelo solo cuando existan factores externos excepcionales debidamente documentados (p. ej., siniestros graves o modificaciones irreversibles).
- Ampliar la base de datos mediante convenios con concesionarias adicionales para alcanzar entre 2 000 y 5 000 registros, incrementando la variabilidad del conjunto de entrenamiento y fortaleciendo la robustez del modelo ante nuevos segmentos y marcas.
- Explorar variables macroeconómicas y de mercado, de manera que se incluyan aspectos como tipo de cambio, incentivos fiscales vigentes y precio de combustibles. Esto puede aumentar la robustez del modelo frente a fluctuaciones externas, especialmente en los segmentos premium.

8. Referencias

- Alkhalidi, A., Khawaja, M. K., & Ismail, S. M. (2024). Solid-state batteries, their future in the energy storage and electric vehicles market. *Science Talks*, 11, 100382. <https://doi.org/10.1016/j.sctalk.2024.100382>
- Áltica. (s.f.). *Altica – Servicios*. Áltica. <https://alticacr.com/>
- Amazon Web Services. (s.f.). *What is data science?* Amazon. <https://aws.amazon.com/es/what-is/data-science/>
- Asociación Costarricense de Movilidad Eléctrica [ASOMOVE]. (2024). *Indicadores – Autos*. ASOMOVE. <https://asomove.org/indicadores/autos>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf>
- Brigham, E. F., & Ehrhardt, M. C. (2013). *Financial management: Theory & practice* (14th ed.). Cengage Learning. <https://www.scribd.com/document/807881221/PDF-Financial-Management-Theory-and-Practice-14th-Edition-Eugene-F-Brigham-Michael-C-Ehrhardt-download>
- Camm, J. D., Cochran, J. J., Fry, M. J., & Ohlmann, J. W. (2020). *Business analytics: Descriptive, predictive, prescriptive* (4th ed.). Cengage Learning. https://www.researchgate.net/profile/Gerhard-Wilhelm-Weber/publication/361492487_G-W_Weber_and_J_Parikh_Business_Analytics_-_Descriptive_Predictive_Prescriptive_by_Jeffrey_D_Camm_James_J_Cochran_Michael_J_Fry_Jeffrey_W_Ohlmann_Cengage_Learning_Boston_MA_02210_USA_ISBN-978-0-357-131/links/655432823fa26f66f4006fb0/G-W-Weber-and-J-Parikh-Business-Analytics-Descriptive-Predictive-Prescriptive-by-Jeffrey-D-Camm-James-J-Cochran-Michael-J-Fry-Jeffrey-W-Ohlmann-Cengage-Learning-Boston-MA-02210-USA-ISBN-97.pdf
- Damodaran, A. (2012). *Investment valuation: Tools and techniques for determining the value of any asset* (3^a ed.). Wiley. <https://suhaconsulting.com/wp-content/uploads/2018/09/investment-valuation-3rd-edition.pdf>

- Departamento Universitario Obrero Campesino [DUOC UC]. (2024). Diferencia entre la investigación básica y la aplicada. <https://bibliotecas.duoc.cl/investigacion-aplicada/Investigacion-basica-y-aplicada>
- Domínguez, S., Pedreros, S., Delgadillo, D., & Anzola, J. (2024). *A depreciation method based on perceived information asymmetry in the market for electric vehicles in Colombia*. Scilit. <https://www.scilit.com/publications/3d39f0589990adade172c1bff28464b9>
- Esteban Nieto, N. (2018). *TIPOS DE INVESTIGACIÓN*. Universidad Santo Domingo de Guzmán. <http://repositorio.usdg.edu.pe/bitstream/USDG/34/1/Tiposde-Investigacion.pdf>
- Forero-Castañeda, D. A., & Sánchez-García, J. A. (2021). Introducción a la inteligencia de negocios basada en la metodología Kimball. *Revista Tecnología, Investigación y Academia*, 9(1), 5-17. <https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://revistas.udistrital.edu.co/index.php/tia/article/download/18082/17993/112771&ved=2ahUKEwjTk8Tc5J6MAxWLRDABHSu1BZoQFnoECA4QAAQ&usg=AOvVaw3IM7qpmD8SfLaPSBWlrsxw>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. <https://jerryfriedman.su.domains/ftp/trebst.pdf>
- Gautam, P., Pode, G., Pode, R., Ayetor, G., & Diouf, B. (2024). Depreciation in the Electric Vehicle Transition Sustainability of the Second-Hand Electric Vehicle Market. *Vehicles*, 6(4), 2044-2074. https://www.researchgate.net/publication/386474117_Depreciation_in_the_Electric_Vehicle_Transition_Sustainability_of_the_Second-Hand_Electric_Vehicle_Market
- Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (2ª ed.). O'Reilly Media. <https://github.com/yanshengjia/ml-road/blob/master/resources/Hands%20On%20Machine%20Learning%20with%20Scikit%20Learn%20and%20TensorFlow.pdf>
- GreenCars. (2023). *Different types of EV batteries*. GreenCars. <https://www.greencars.com/greencars-101/different-types-of-ev-batteries>

- Hashemi-Pour, C., Brush, K., & Burns, E. (2024, 16 agosto). What is data visualization? SearchBusinessAnalytics, TechTarget.
<https://www.techtarget.com/searchbusinessanalytics/definition/data-visualization>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. del P. (2014). *Metodología de la investigación* (Sexta ed.). McGraw Hill. <https://www.esup.edu.pe/wp-content/uploads/2020/12/2.%20Hernandez,%20Fernandez%20y%20Baptista-Metodolog%C3%ADa%20Investigacion%20Cientifica%206ta%20ed.pdf>
- Hull, J. (2018). *Risk management and financial institutions* (5^a ed.). Wiley. https://dn720002.ca.archive.org/0/items/quant_books/Risk%20Management%20_%20Financial%20Institutions%20-%20J.%20C.%20Hull.pdf
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2^a ed.). OTexts. <https://otexts.com/fpp2/>
- IBM. (2021). CRISP-DM help overview. IBM Documentation. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=dm-crisp-help-overview>
- IFRS Foundation. (2023). IAS 16 Property, plant and equipment. <https://www.ifrs.org/issued-standards/list-of-standards/ias-16-property-plant-and-equipment/>
- IFRS Foundation. (2023). IAS 36 Impairment of assets. <https://www.ifrs.org/issued-standards/list-of-standards/ias-36-impairment-of-assets/>
- International Energy Agency [IEA]. (2024). Global EV Outlook 2024. IEA. <https://www.iea.org/reports/global-ev-outlook-2024>
- Jiménez-Partearroyo, M., Medina-López, A., & Rana, S. (2024). Business intelligence and business analytics in tourism: Insights through Gioia methodology. *International Entrepreneurship and Management Journal*, 20(3), 2287-2321. <https://doi.org/10.1007/s11365-024-00973-7>
- Liu, S., Chen, J., & Ouyang, L. (2023). A review on business analytics: Definitions, techniques, applications and challenges. *Mathematics*, 11(4), 899. https://www.researchgate.net/publication/368440278_A_Review_on_Business_Analytics_Definitions_Techniques_Applications_and_Challenges
- Liu, S., Feng, Z., Zhao, X., & Liang, Y. (2023). Hybrid stacking ensemble learning for predicting crop yield with multisource data. *Computers and Electronics in Agriculture*, 208, 107821. <https://doi.org/10.1016/j.compag.2023.107821>

- Machuca, F. (2022). 8 técnicas de recolección de datos: descubre un mundo más allá de la encuesta. Crehana Blog. <https://www.crehana.com/blog/transformacion-digital/tecnicas-recoleccion-de-datos/>
- Marqués, M. (2011). *Bases de datos* (Primera ed.). Publicacions de la Universitat Jaume I. <https://bdigital.uvhm.edu.mx/wp-content/uploads/2020/05/Bases-de-Datos.pdf>
- McKinney, W. (2018). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython* (2^a ed.). O'Reilly Media. [https://github.com/Jffrank/Books/blob/master/Python%20for%20Data%20Analysis.%20Data%20Wrangling%20with%20Pandas%2C%20NumPy%2C%20and%20IPython%20\(2017%2C%20O'Reilly\).pdf](https://github.com/Jffrank/Books/blob/master/Python%20for%20Data%20Analysis.%20Data%20Wrangling%20with%20Pandas%2C%20NumPy%2C%20and%20IPython%20(2017%2C%20O'Reilly).pdf)
- Medina La Plata, E. (2015). *Business Intelligence: Una guía práctica*. Universidad Peruana de Ciencias Aplicadas (UPC). <https://repositorioacademico.upc.edu.pe/handle/10757/338680>
- Microsoft. (2025). *Extracción, transformación y carga de datos (ETL)*. Microsoft Learn. <https://learn.microsoft.com/es-es/azure/architecture/data-guide/relational-data/etl>
- Ministerio de Ambiente y Energía [MINAE]. (2024). *Movilidad sostenible*. MINAE. <https://minae.go.cr/energia/Movilidad%20Sostenible.aspx>
- Miranda, M., & Villasís, M. (2016). El protocolo de investigación IV: las variables de estudio. *Revista Alergia México*, 63(3), 303-310. <https://revistaalergia.mx/ojs/index.php/ram/article/view/199/350>
- Miranda, U., & Acosta, Z. (2009). *Fuentes de información para la recolección de información cuantitativa y cualitativa* [Documento PDF]. <https://docs.bvsalud.org/biblioref/2018/06/885032/texto-no-2-fuentes-de-informacion.pdf>
- Monhor, D., & Takemoto, S. (2014). Understanding the concept of outlier and its relevance to the assessment of data quality: Probabilistic background theory. *Earth, Planets and Space*, 57(11), 1009. <https://doi.org/10.1186/bf0335188>
- Mora, G. (2021). Inteligencia de negocios como una herramienta competitiva dentro de las instituciones financieras. *Revista Académica ECO*, 24, 1-20. <https://doi.org/10.36631/ECO.2021.24.01>
- Naimi, A. I., & Balzer, L. B. (2018). Stacked generalization: An introduction to super learning. *European Journal of Epidemiology*, 33(5), 459-464. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6089257/>
-

- Neill, D. A., Quezada Abad, C., & Arce Rodríguez, J. (2018). Investigación cuantitativa y cualitativa. En D. A. Neill & L. Cortez Suárez (Eds.), *Procesos y fundamentos de la investigación científica* (pp. 68-87). Universidad Técnica de Machala. <https://repositorio.utmachala.edu.ec/bitstream/48000/14232/1/Cap.4-Investigaci%C3%B3n%20cuantitativa%20y%20cualitativa.pdf>
- Oracle. (2020). *¿Qué es una base de datos?* Oracle. <https://www.oracle.com/database/what-is-database/>
- Phillips-Wren, G., Daly, M., & Burstein, F. (2021). Reconciling business intelligence, analytics and decision support systems: More data, deeper insight. *Decision Support Systems*, 146, 113560. <https://doi.org/10.1016/j.dss.2021.113560>
- Pinto, J. E., Henry, E., Robinson, T. R., Stowe, J. D., & Wilcox, S. E. (2020). Equity asset valuation. (Fourth edition ed.). Wiley. http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=032078684&sequence=000003&line_number=0002&func_code=DB_RECORDS&service_type=MEDIA
- Prada Madrid, E. (2008). Los insumos invisibles de decisión: datos, información y conocimiento. *Anales de Documentación*, (11), 183-196. <https://www.redalyc.org/pdf/635/63501110.pdf>
- Procuraduría General de la República. (2023). Normativa N.º 97124: Disposiciones para la aplicación de incentivos a la movilidad eléctrica. Sistema Costarricense de Información Jurídica (SCIJ). https://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?nValor1=1&nValor2=97124
- Power, D. J., Sharda, R., & Burstein, F. (2018). Defining business analytics: An empirical approach. https://www.researchgate.net/publication/327204216_Defining_business_analytics_an_empirical_approach
- Roberson, L. (2024). *Accelerating electric vehicle adoption in the United States: The impact of experience, incentives, and resale value*. The George Washington University. <https://www.proquest.com/dissertations-theses/accelerating-electric-vehicle-adoption-united/docview/3014023436/se-2?accountid=27651>

- Sarker, I. H. (2021a). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3). https://www.researchgate.net/publication/350297716_Machine_Learning_Algorithms_Real-World_Applications_and_Research_Directions
- Sarker, I. H. (2021b). Data science and analytics: An overview from data-driven smart computing, decision-making and applications perspective. *SN Computer Science*, 2(5), 377. https://pmc.ncbi.nlm.nih.gov/articles/PMC8274472/pdf/42979_2021_Article_765.pdf
- Suárez, E. (2024). *Guía práctica sobre los distintos tipos de fuentes de información*. Experto Universitario. <https://expertouniversitario.es/blog/fuentes-de-informacion/>
- Tableau. (s.f.). *Inteligencia de negocios: qué es y por qué es importante*. Tableau. <https://www.tableau.com/es-mx/learn/articles/business-intelligence>
- U.S. Department of Energy, Alternative Fuels Data Center. (2023). All-electric vehicles. <https://afdc.energy.gov/vehicles/electric-basics-ev>
- VanderPlas, J. (2016). *Python data science handbook: Essential tools for working with data*. O'Reilly Media. https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://avys.omu.edu.tr/storage/app/public/mustafa.kaya/135286/Python_Datascience.pdf&ved=2ahUKewjEzr-CtcmNAxUISDABHZzAGJUQFnoECBwQAQ&usg=AOvVaw36oePAjFhJ1vwmsqSTIxUQ
- Wang, Q., & Lu, H. (2024). A novel stacking ensemble learner for predicting residual strength of corroded pipelines. *npj Materials Degradation*, 8(87), 1-10. https://www.researchgate.net/publication/383268239_A_novel_stacking_ensemble_learner_for_predicting_residual_strength_of_corroded_pipelines
- Wickham, H., & Grolemund, G. (2023). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data* (2nd ed.). O'Reilly Media. <https://r4ds.hadley.nz/EDA.html>

9. Apéndices

Apéndice A. Cronograma de actividades

Cronograma																
Actividad	Semanas															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Reunión con profesor tutor																
Reunión con empresa Áltica																
Revisión y ajustes del anteproyecto																
Revisión de la documentación sobre el mercado de los VEs en Costa Rica																
Recopilación de los datos iniciales, así como extracción de CRAutos																
Limpieza y procesamiento de los datos recopilados																
Análisis exploratorio de los datos (EDA) para identificar patrones y correlación de los atributos																
Investigación sobre modelos estadísticos y técnicas de aprendizaje automático adecuados																
Desarrollo de los modelos																
Ajuste y optimización del modelo seleccionado																
Cálculo de métricas de desempeño																
Validación del modelo predictivo																
Elaboración de análisis y resultados																
Desarrollo de la interfaz de consulta																
Entrega del trabajo final de graduación																

Fuente: Elaboración propia (2025)

Apéndice B. Plantilla de minuta de reuniones

Reunión No.		Fecha	
Lugar:		Hora de inicio y finalización	
Objetivo:			
Participantes:			
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable
Firmas			
Próxima reunión			
Temas por tratar		Fecha	Convocados

Fuente: Adoptado de Machote de minuta-ejemplo, por S. Mora, 2017, Google Sites (https://drive.google.com/file/d/1b65nGntkFzFPO4xQ_10QyIoL90sTFsrM/view)

Apéndice C. Plantilla hoja de registro observacional

Hoja de observación	
Fecha	
Elemento observado	
Rol del observador	
Principales hallazgos	
-	
-	

Fuente: Elaboración propia (2025).

Apéndice D. Plantilla de entrevista abierta

Entrevista

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

No. Entrevista	
Fecha	
Objetivo	
Participantes	
Rol dentro del proyecto	Nombre
Preguntas	
Pregunta	Respuesta

Fuente: Elaboración propia (2025).

Apéndice E. Plantilla de revisión documental

Revisión documental	
No. Revisión	
Objetivo	
Fuente	
Información obtenida	

Fuente: Elaboración propia (2025).

Apéndice F. Minuta de Reunión #1

Reunión No.	1	Fecha	11/03/2025
Lugar:	Google Meet	Hora de inicio y finalización	3:30 PM – 4:30 PM
Objetivo:	Primera Reunión TFG - Coordinación		
Participantes:	Hernán Rojas, Federico Torres y Mauricio Loría		
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable
1	Coordinación Profesor Tutor y contraparte de la empresa	Se dejó claro los aspectos a desarrollar	Todos acordamos Responsable: Mauricio Loría

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

		planteados en los objetivos del proyecto	
2	Consulta de apoyo, para obtener información para facilitar el desarrollo del proyecto	Hernán se iba a encargar de conversar para hacer lo posible en obtener la información	Todos acordamos Responsable: Hernán Rojas
Firmas			
Hernán Rojas Federico Torres Mauricio Loría			
Próxima reunión			
Temas por tratar		Fecha	Convocados
Corrección secciones variables, inclusión marco teórico NICS, corrección herramientas		28/03/2025	Mauricio Loría Federico Torres

Fuente: Elaboración propia (2025).

Apéndice G. Minuta de Reunión #2

Reunión No.	2	Fecha	28/03/2025
Lugar:	Google Meet	Hora de inicio y finalización	10:00 AM – 11:00 AM
Objetivo:	Reunión Seguimiento #2		
Participantes:	Federico Torres y Mauricio Loría		
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable
1	Corrección variables de investigación	Volver a plantearlas para que se adecuen a los objetivos específicos	Responsable: Mauricio Loría

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

2	Reordenar marco teórico	Replantear la agrupación y el orden de las secciones del marco teórico	Responsable: Mauricio Loría
Firmas			
Federico Torres Mauricio Loría			
Próxima reunión			
Temas por tratar		Fecha	Convocados
Consulta información de registros para el proyecto y entrega del segundo avance		04/04/2025	Mauricio Loría Federico Torres

Fuente: Elaboración propia (2025).

Apéndice H. Minuta de Reunión #3

Reunión No.	3	Fecha	04/04/2025
Lugar:	Google Meet	Hora de inicio y finalización	12:30 PM – 1:00 PM
Objetivo:	Reunión Seguimiento #3		
Participantes:	Federico Torres y Mauricio Loría		
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable
1	Información registros de otras concesionarias	Hernán reporta que todavía no existe avance, la persona responsable estaba fuera del país	Responsable: Hernán Rojas
2	Entrega segundo avance	Se realiza la entrega del segundo avance con las correcciones	Responsable: Mauricio Loría

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

		solicitadas, a la espera de retroalimentación	
3	Cambio de nombre del proyecto	Eliminar la palabra prototipo y reemplazar por modelo	Acuerdo profesor tutor Responsable: Mauricio Loría
Firmas			
Federico Torres Mauricio Loría			
Próxima reunión			
Temas por tratar		Fecha	Convocados
Consultas fuentes de datos y análisis exploratorio		14/04/2025	Mauricio Loría Federico Torres

Fuente: Elaboración propia (2025).

Apéndice I. Minuta de Reunión #4

Reunión No.	4	Fecha	14/04/2025
Lugar:	Google Meet	Hora de inicio y finalización	10:00 AM – 11:00 AM
Objetivo:	Reunión Seguimiento #4		
Participantes:	Federico Torres y Mauricio Loría		
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable
1	Fuentes de datos adicionales	Hernán confirma que la persona responsable, intentó contactar	Responsable: Mauricio Loría
2	Análisis exploratorio	Revisar la segmentación de los	Responsable: Mauricio Loría

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

		registros de las 4 fuentes recopiladas para asegurar la distribución de la información dentro de los años recopilados	
3	Cambio de nombre del proyecto	Eliminar la palabra prototipo y reemplazar por modelo	Acuerdo profesor tutor Responsable: Mauricio Loría
4	Aceptación segundo avance	Se acepta el segundo avance del proyecto	
Firmas			
Federico Torres Mauricio Loría			
Próxima reunión			
Temas por tratar		Fecha	Convocados
Consulta información de registros para el proyecto y elaboración de los modelos		25/04/2025	Mauricio Loría Federico Torres

Fuente: Elaboración propia (2025).

Apéndice J. Minuta de Reunión #5

Reunión No.	5	Fecha	25/04/2025
Lugar:	Google Meet	Hora de inicio y finalización	1:00 PM – 1:45 PM
Objetivo:	Reunión Seguimiento #5		
Participantes:	Federico Torres y Mauricio Loría		
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

1	Fuentes de datos adicionales	Hernán confirma que la persona responsable, intentó contactar, pero no fue posible que la concesionaria que se planteó compartiera los datos requeridos	Continuamos con las fuentes disponibles Responsable: Mauricio Loría
2	Más tiempo para resolver desarrollar el análisis de resultados	Se solicita más tiempo al profesor, debido a temas personales	Responsable: Mauricio Loría
3	Enriquecimiento de la información por medio del uso de la herramienta del API de ChatGPT	Se explica la metodología para enriquecer la información	Aprueba Federico Torres Responsable: Mauricio Loría
4	Inicio de diseño y evaluación de los modelos, así como los criterios de aceptación de las métricas	Se escogen los modelos a desarrollar, y se define las métricas aceptables	Aprueba Federico Torres Responsable: Mauricio Loría
Firmas			
Federico Torres Mauricio Loría			
Próxima reunión			
Temas por tratar		Fecha	Convocados
Entrega del Avance III		09/05/2025	Mauricio Loría Federico Torres

Fuente: Elaboración propia (2025).

Apéndice K. Minuta de Reunión #6

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

Reunión No.	6	Fecha	09/05/2025
Lugar:	Google Meet	Hora de inicio y finalización	1:00 PM – 1:45 PM
Objetivo:	Reunión Seguimiento #6		
Participantes:	Federico Torres y Mauricio Loría		
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable
1	Entrega III Avance	Se entrega el avance III se espera por retroalimentación	Responsable: Federico Torres
2	Avance IV	Se procede a avanzar mientras tanto en las secciones restantes del proyecto	Responsable: Mauricio Loría
3	Desarrollo propuesta de la solución	Interfaz para consultar por el valor de recuperación de un vehículo eléctrico	De acuerdo a mi profesor tutor y lo solicitado por la contraparte de la empresa Responsable: Mauricio Loría
Firmas			
Federico Torres Mauricio Loría			
Próxima reunión			
Temas por tratar		Fecha	Convocados
Retroalimentación Avance III y desarrollo Avance IV		28/05/2025	Mauricio Loría Federico Torres

Fuente: Elaboración propia (2025).

Apéndice L. Minuta de Reunión #7

Reunión No.	7	Fecha	28/05/2025
Lugar:	Google Meet	Hora de inicio y finalización	11:30 AM – 12:00 AM
Objetivo:	Reunión Seguimiento #7		
Participantes:	Federico Torres y Mauricio Loría		
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable
1	Retroalimentación Avance III	Agregar sección referente al modelo <i>stacking</i> utilizado para su mejor comprensión en el proyecto	Responsable: Mauricio Loría
Firmas			
Federico Torres Mauricio Loría			
Próxima reunión			
Temas por tratar		Fecha	Convocados
Entrega Avance IV		30/05/2025	Mauricio Loría Federico Torres Hernán Rojas

Fuente: Elaboración propia (2025).

Apéndice M. Minuta de Reunión #8

Reunión No.	8	Fecha	30/05/2025
Lugar:	Google Meet	Hora de inicio y finalización	11:30 AM – 12:00 AM
Objetivo:	Reunión Seguimiento #8		

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

Participantes:	Federico Torres, Hernán Rojas y Mauricio Loría		
Temas tratados			
No.	Asunto	Comentarios	Acuerdo y responsable
1	Corrección Avance III	Se aplican los cambios solicitados del Avance III	Responsable: Mauricio Loría
2	Entrega Avance IV	Se entrega el Avance IV con la propuesta de solución, conclusiones y recomendaciones	Responsable: Mauricio Loría
3	Coordinación avances desarrollo del proyecto	Se indica que se encuentra satisfecho con el trabajo realizado, y espera utilizarlo pronto	Acordada la aprobación por parte del profesor tutor y el Gerente de TI Responsable: Mauricio Loría
Firmas			
Federico Torres Hernán Rojas Mauricio Loría			
Próxima reunión			
Temas por tratar		Fecha	Convocados
-		-	-

Fuente: Elaboración propia (2025)

Apéndice N. Entrevista #1

Entrevista	
No. Entrevista	1
Fecha	11/03/2025

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

Objetivo	Comprender el mercado de usados de vehículos eléctricos	
Participantes		
Rol dentro del proyecto	Nombre	
Analista financiero del sector automotriz	Gustavo Vargas	
Gerente de TI	Hernán Rojas	
Preguntas		
Pregunta	Respuesta	
¿Cómo describiría la evolución del mercado de VE en Costa Rica en los últimos años?	Crecimiento sostenido >40 % anual impulsado por incentivos fiscales y aumento de la red de carga.	
¿De dónde provienen los datos usados para fijar precios?	Históricos de avalúos realizados, listados de otras concesionarias, publicaciones en Facebook Marketplace y CRAutos, así como ajustes basados en experiencia previa; no hay base de datos unificada.	
¿Cuáles son los factores que más inciden en la degradación de la batería?	Ciclos de carga rápidos, temperaturas altas, profundidad de descarga frecuente, así también el no uso de una unidad, a concepción contraria no utilizar regularmente una unidad puede deteriorar significativamente una batería. También las llantas, debido a que se gastan más rápido al tener mayor peso las unidades por las baterías.	
¿Observa diferencias de depreciación según marca o química de batería?	Sí; marcas con mejor garantía de batería y químicas NMC muestran menor depreciación que LFP en el mercado local. Además, toda esta información del estado y condición de las unidades es almacenada en nuestro sistema. Uno de los retos mencionados es la rápida mejora tecnológica de los modelos nuevos, que dejan obsoletos a los anteriores sin que sus	

	precios cambien mucho, lo que ocasiona una preferencia de compra por los modelos recientes debido a su mejor rendimiento y autonomía.
¿Qué valor ve en un modelo predictivo automatizado?	Podría sin dudas ayudar en la valoración de las unidades, como te digo el juicio de experto es clave, pero podría servir de referencia para futuras valoraciones.

Fuente: Elaboración propia (2025)

Apéndice O. Revisión documental #1

Revisión documental	
No. Revisión	1
Objetivo	Conocer, consolidar y evaluar las fuentes de datos históricas y actuales sobre el mercado costarricense de vehículos eléctricos (VE)
Fuente	Documentación interna y externa: Registros históricos de ventas de Áltica Avalúos de Red Motors Web-scraping de unidades en CRAutos (propio y de Áltica) Fichas técnicas de EV-Database
Información obtenida	- Se identificaron cuatro fuentes principales de datos estructurados: – Históricos de ventas de Áltica (178 registros, 2017–2025) – Avalúos de Red Motors (207 registros, 2019–2025) – Web-scraping de CRAutos por Áltica (576 registros, 2020–2023) – Web-scraping independiente de CRAutos (359 registros, 2023–2025) - Tras la depuración (eliminación de duplicados y vehículos no eléctricos) se consolidó un total de 1 318 registros válidos a partir de 1 617 registros brutos.

	• Variables capturadas: autonomía real, antigüedad, kilometraje, precio de venta y condición cualitativa del vehículo. • Se detectó la ausencia de mediciones directas del estado de salud (SoH) de la batería; se usaron variables sustitutas como el tipo de la batería. • La pertinencia y representatividad de las fuentes se validó mediante entrevista abierta con el analista financiero del sector.
--	---

Fuente: Elaboración propia (2025).

Apéndice P. Revisión documental #2

Revisión documental	
No. Revisión	2
Objetivo	Conocer, consolidar y evaluar las fuentes de datos históricas y actuales sobre el mercado costarricense de vehículos eléctricos (VE)
Fuente	Bitácoras de limpieza en notebooks de Python (Pandas)
Información obtenida	• Identificados valores nulos en 7 de 13 variables (0,5 % a 6,2 %). • Imputación: mediana para numéricas y modo para categóricas. • Unificación de unidades: kilometraje y autonomía en km; precio en dólares. • Se creó un pipeline reproducible de pre-procesamiento (ColumnTransformer + RobustScaler) para evitar fugas de datos.

Fuente: Elaboración propia (2025)

Apéndice Q. Revisión documental #3

Revisión documental	
No. Revisión	3

Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica

Objetivo	Analizar las relaciones entre variables, identificar dependencias lineales y reducir la multicolinealidad para optimizar la selección de predictores.
Fuente	Matriz de correlación de Pearson y visualizaciones heatmap generadas en Jupyter.
Información obtenida	• Identificados valores nulos en 7 de 13 variables (0,5 % a 6,2 %). • Edad y kilometraje correlacionan negativamente con el precio ($\rho \approx -0,60$ y $\rho \approx -0,45$). • Se detectó colinealidad moderada entre capacidad usable y capacidad nominal de batería. • Transformación del estado a valor ordinal para su mejor interpretación en el modelo.

Fuente: Elaboración propia (2025).

Apéndice R. Revisión documental #4

Revisión documental	
No. Revisión	4
Objetivo	Comparar los principales algoritmos y justificar la elección del modelo final que estima el valor residual de los vehículos eléctricos.
Fuente	<i>Notebooks</i> de experimentación en Python y registro de métricas en hoja de cálculo.
Información obtenida	• Se evaluaron tres grupos de modelos: lineales (Lasso), ensambles de árboles (<i>Random Forest</i> y <i>Gradient Boosting</i>) y series temporales de referencia (ARIMA/SARIMA). • Hiperparámetros ajustados con <i>RandomizedSearch</i> y validación cruzada de 5 pliegues. • El ensamble <i>Stacking Regressor (RF + GB)</i>, alcanzó el mejor desempeño ($R^2 \approx 0,83$; $MAE \approx \\$6\,900$) y cumplió los umbrales fijados ($R^2 \geq 0,80$; $MAE \leq \\$8\,000$). • Modelos lineales y temporales no superaron dichos umbrales y fueron descartados.

--	--

Fuente: Elaboración propia (2025).

Apéndice S. Entrevista #2

Entrevista	
No. Entrevista	2
Fecha	11/05/2025
Objetivo	Comprender el mercado de usados de vehículos eléctricos
Participantes	
Rol dentro del proyecto	Nombre
Analista financiero del sector automotriz	Gustavo Vargas
Gerente de TI	Hernán Rojas
Preguntas	
Pregunta	Respuesta
¿Cómo califican la exactitud del modelo a la luz de las métricas finales ($R^2 \approx 0,83$; MAE $\approx \$6\,900$)?	Ambos coinciden en que el desempeño es “alto” y suficiente para decisiones preliminares; sin embargo, el analista señala que un MAE de \$6 900 aún requiere ajuste fino mediante juicio experto.
¿Qué opinan sobre la disminución de la varianza con respecto a otros modelos ?	Se valora positivamente porque aporta estabilidad; el gerente de TI destaca que reduce el riesgo de resultados impredecibles entre consultas.
¿Cómo planean integrar el modelo en el flujo de valoración?	Ejecutar el modelo como primer paso de valuación; el analista revisará las predicciones y aplicará correcciones, de acuerdo a lo que crea justo en el momento, además será necesario comparar las primeras ejecuciones con el juicio de experto y el modelo.

¿Cómo se desempeñaron los modelos de ARIMA y SARIMA?	Estos modelos tuvieron un resultado muy por debajo de lo esperado lo cual, me lleva a descartarlos de las opciones debido a que su variabilidad es muy alta y de esta formas los intervalos de confianza son pocos significativos, llevando a una predicción muy poco acertada.
¿Dónde se podrá consultar el modelo?	Podemos ejecutarlo en nuestro ambiente de servidor, y posteriormente valorar su integración con el CRM para facilitar su uso, por ahora necesitaríamos algo sencillo.

Fuente: Elaboración propia (2025)

10. Anexos

Anexo I. Hoja de control de cambios

Nombre Proyecto



Hoja de Control de Cambios			
Datos Generales del Cambio			
Nº Cambio	1		
Solicitante		Fecha de solicitud del cambio	DD/MM/AAAA
Responsable de la implementación		Fecha de realización del cambio	DD/MM/AAAA
Estado	☐ Aprobado ☐ En Revisión ☐ Rechazado		
Detalles del Cambio			
Categoría	Introducción / Alcance / Marco Teórico / Metodología / ...		
Descripción detallada	•		
Justificación			
Implicaciones de realizar el cambio			
Impacto	Especificar si el cambio genera impacto en otras áreas del proyecto, tales como recursos, cronogramas, otros proyectos, entre otros.		
Comentarios/ Observaciones			

Revisado por:

Elaborado por:

(Prof. tutor)

(Estudiante)

Revisado por:

Aprobado por:

(Empresa)

(Coordinadora de TFG)

Fuente: Adoptado de HojaControlCambio, por S. Mora, 2017, Google Sites
(<https://drive.google.com/file/d/1jfnTi5TB8DXpnHBpWKe9Mfxqf3dNDny0/view>)

Anexo II. Cambio #1 – Título del proyecto

Modelo predictivo para reducir pérdidas financieras
por la variabilidad en el valor de recuperación de los vehículos eléctricos
en Costa Rica



Hoja de Control de Cambios			
Datos Generales del Cambio			
N° Cambio	1		
Solicitante	Prof. Tutor	Fecha de solicitud del cambio	08/04/2025
Responsable de la implementación	Estudiante	Fecha de realización del cambio	08/04/2025
Estado	☒ Aprobado ☐ En Revisión ☐ Rechazado		
Detalles del Cambio			
Categoría	Título		
Descripción detallada	Cambio del título del proyecto, para realizar la sustitución de la palabra Prototipo por Modelo. Es decir: • Prototipo Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica		
Justificación	Comprende de una forma más acertada el enfoque del proyecto		
Implicaciones de realizar el cambio	Cambio en el título del proyecto únicamente		
Impacto	No genera impacto en otras áreas del proyecto, tales como recursos, cronogramas, otros proyectos, entre otros.		
Comentarios/ Observaciones			

Revisado por:

Dr. Federico Torres Carballo

FEDERICO TORRES
CARBALLO (FIRMA)
(Prof. tutor)

Firmado digitalmente por
FEDERICO TORRES
CARBALLO (FIRMA)
Fecha: 2025.04.10
20:20:45 -06'00'

Elaborado por:

Mauricio Loria Rodríguez

MAURICIO LORIA RODRIGUEZ
(Estudiante)

Firmado digitalmente por
MAURICIO LORIA RODRIGUEZ
(FIRMA)
Fecha: 2025.04.10
19:20:45 -06'00'

Revisado por:

Hernán Rojas Blanco

HERNAN ROJAS BLANCO
(FIRMA)

(Empresa)

HERNAN
PABLO ROJAS
BLANCO (FIRMA)
Firmado digitalmente por
HERNAN PABLO
ROJAS BLANCO
(FIRMA)
Fecha: 2025.04.08
13:25:31 -06'00'

Aprobado por:

Máster Yarima Sandoval Sánchez

YARIMA SANDOVAL SANCHEZ
(FIRMA)

(Coordinadora de T.C.)

Firmado digitalmente por
YARIMA TATIANA
SANDOVAL SANCHEZ
(FIRMA)
Fecha: 2025.05.21 11:29:39
-06'00'

Fuente: Elaboración propia (2025).

Anexo III. Carta de Filóloga

San José, Costa Rica, 08 de junio de 2025

Tecnológico de Costa Rica

Escuela de Administración de Tecnologías de Información

Estimados señores:

Me pongo en contacto con ustedes para hacer constar que yo, Eunice Arias Corrales, número de identificación 2-0727-0426, filóloga, número de carné 86928 del Colegio de Licenciados y Profesores en Letras, Filosofía, Ciencias y Artes, efectué la revisión filológica del trabajo titulado *Modelo predictivo para reducir pérdidas financieras por la variabilidad en el valor de recuperación de los vehículos eléctricos en Costa Rica*, realizado por Mauricio Loria Rodríguez, número de identificación 1-1742-0846, como proyecto de graduación para optar por el grado de Licenciatura en Administración de Tecnología de Información.

Como parte de esta revisión, contemplé la corrección de ortografía, estilo, coherencia y cohesión, así como la intención comunicativa, según la norma culta de la lengua española y los principios del lenguaje académico. Asimismo, respeté el estilo personal del autor y el lenguaje técnico del área de conocimiento específica a lo largo del texto. A partir de lo anterior, considero que el trabajo se encuentra listo para lo correspondiente.

Atentamente,



Bach. Eunice Arias Corrales

Filóloga, Universidad de Costa Rica

Anexo IV. Resumen de minutas aprobado

Próxima reunión		
Temas por tratar	Fecha	Convocados
-	-	-

Revisado por:

Dr. Federico Torres Carballo

FEDERICO TORRES
CARBALLO
(FIRMA)

Firmado digitalmente por
FEDERICO TORRES
CARBALLO (FIRMA)
Fecha: 2025.06.06 19:21:18
-06'00'

(Prof. tutor)

Revisado por:

Hernán Rojas Blanco

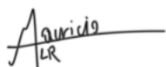
HERNAN
PABLO ROJAS
BLANCO
(FIRMA)

Firmado digitalmente
por HERNAN PABLO
ROJAS BLANCO
(FIRMA)
Fecha: 2025.06.09
12:59:45 -06'00'

(Empresa)

Elaborado por:

Mauricio Loría Rodríguez


LR

(Estudiante)

Anexo V. Matriz de correlación entre variables numéricas

