Article

Not peer-reviewed version

XGBoost vs. LightGBM: An XAI Approach to National Vehicle Fleet Analysis

Wilson Gustavo Chango * , Javier Homero Velasteguí-Izurieta * , William Paul Pazuña * , Joffre Stalin Monar-Monar , Rebeca Mariana Moposita * , Santiago Logroño , Carlos López , Jacqueline Elizabeth Ponce , Geovanny Silva , Angel Flores , Cindy Johanna Choez , Marcelo V Garcia

Posted Date: 12 February 2026

doi: 10.20944/preprints202602.0903.v1

Keywords: explainable artificial intelligence (XAI); gradient boosting (XGBoost/LightGBM); sustainable; SHAP values; transport decarbonization

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2026 doi:10.20944/preprints202602.0903.v1

Article

XGBoost vs. LightGBM: An XAI Approach to National Vehicle Fleet Analysis

Wilson Chango-Sailema [1,2,*], Homero Velasteguí-Izurieta [2,*], William Pazuña [3], Joffre Monar [4], Mariana Moposita [4], Santiago Logroño [4], Carlos López [4], Jacqueline Ponce [4], Geovanny Silva [4], Angel Flores [4], Cindy Choez [5] and Marcelo V. Garcia [1]

1 Facultad de Ingeniería en Sistemas, Electrónica e Industrial, Universidad Técnica de Ambato,UTA,Av. los Chásquis, Ambato 180206, Ecuador
2 Departamento de Informática y Ciencias de la Computación, Pontificia Universidad Católica del Ecuador, Campus Esmeraldas PUCESE, Esmeraldas 080101, Ecuador
3 Departamento de Electrónica, Universidad Técnica de Cotopaxi, Campus La Maná, Maná, 050202, Ecuador
4 Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, EC060155, Ecuador
5 Universidad Técnica Luis Vargas Torres de Esmeraldas, Esmeraldas, 080101, Ecuador
* Correspondence: wg.change@uta.edu.ec (W.C.-S.); homero.j.velastegui.i@pucese.edu.ec (H.V.-I.); Tel.: +593-960949920

Abstract

This study aims to bridge the gap between the high predictive accuracy of machine learning models and the transparency requirements of public policies for sustainable mobility by analyzing technological adoption within the national vehicle fleet of Ecuador. To this end, a methodological framework grounded in Design Science Research (DSR) and CRISP-DM was applied to a real administrative dataset of 482,754 vehicle registration records from the Ecuadorian Internal Revenue Service (SRI, 2025), comparing the performance and interpretability of two state-of-the-art gradient boosting algorithms, XGBoost and LightGBM, for multiclass classification of internal combustion engine (ICE), hybrid, and electric vehicles (EV) under severe class imbalance. The results demonstrate that both models achieve near-perfect predictive performance, with a consolidated Macro F1-score of 0.987, confirming their robustness and suitability for large-scale, policy-relevant administrative data even when electric vehicles represent only 1.3% of the observed fleet. Beyond global performance metrics, the integration of explainable artificial intelligence through SHAP values reveals that fiscal appraisal value and engine capacity are the primary determinants of EV adoption, while territorial factors exhibit greater influence in the case of hybrid vehicles, highlighting qualitatively distinct adoption mechanisms across technologies. These findings show that advanced "black-box" models can be transformed into auditable and interpretable analytical tools capable of linking predictions to concrete economic, technical, and spatial variables relevant for decision-making. The study concludes that XAI-enabled gradient boosting provides a rigorous and transparent framework to support the design of targeted fiscal incentives, fleet electrification strategies, and evidence-based transport decarbonization policies in emerging economies, where data-driven governance is critical for accelerating the transition toward sustainable mobility.

Keywords: explainable artificial intelligence (XAI); gradient boosting (XGBoost/LightGBM); sustainable; SHAP values; transport decarbonization

---

1. Introduction

The transition toward sustainable mobility systems has become a central strategic axis for addressing climate change, reducing greenhouse gas emissions, mitigating local environmental externalities such as air and noise pollution—and enhancing energy efficiency in the transport sector. This sector represents a major source of global energy consumption and emissions, particularly in developing countries where the rapid growth of the vehicle fleet increases pressure on urban infrastructure, public health systems, and state fiscal sustainability [1]. In this scenario, data science applied to transport and energy has increasingly integrated advanced machine learning models, with gradient boosting algorithms like XGBoost and LightGBM standing out for their high predictive capacity, robustness against non-linear relationships, and outstanding performance in analyzing large volumes of heterogeneous tabular data typical of administrative records and fiscal databases [2].

However, despite their empirical success, the extensive adoption of these models in sustainable mobility and energy transition studies has highlighted a critical gap between maximizing statistical precision and understanding the underlying mechanisms that generate predictions. This gap restricts their practical utility in the design, evaluation, and adjustment of public transport policies that require traceability, explainability, and technical legitimacy [3,4]. From a conceptual framework grounded in technology adoption theory, behavioral economics, and socio-technical systems approaches, the transition from internal combustion engine (ICE) vehicles toward hybrid and electric technologies is understood as a multidimensional process influenced by economic, fiscal, regulatory, technical, and social factors. Accurate interpretation of this process demands not only high-performance predictive models but also analytical tools capable of translating complex patterns into actionable knowledge for decision-makers [5,6].

In this context, Explainable Artificial Intelligence (XAI) approaches emerge as a key methodological bridge between advanced modeling and informed decision-making by allowing the decomposition of black-box model behavior and the identification of the relative contribution of fiscal, socioeconomic, technical, and territorial variables in shaping the vehicle fleet [7]. While previous studies have documented the superiority of boosting models over traditional statistical methods and simpler machine learning algorithms in classification and prediction tasks associated with vehicle demand and energy consumption, most of this work has focused on global performance metrics, paying scant attention to the interpretability of results and their effective transfer into specific institutional contexts [8,9].

This limitation is particularly relevant in Ecuador, where the availability of real administrative data from the Internal Revenue Service (*Servicio de Rentas Internas* - SRI) for the year 2025 constitutes a unique opportunity to analyze, at a national scale, the adoption dynamics of electric, hybrid, and internal combustion engine vehicles within a specific fiscal framework characterized by incentives, exemptions, and distortions typical of an incipient electromobility market [10]. Consequently, this study is justified by the need to comparatively evaluate XGBoost and LightGBM models—not only from the perspective of predictive accuracy but also under an XAI approach—to generate transparent, reproducible, and understandable evidence aimed at strengthening data-driven sustainable transport policy design [11–13].

Based on this premise, the research formulates the following research question: *What explainable factors, derived from gradient boosting models applied to real Ecuadorian administrative data, determine the adoption of electric and hybrid vehicle technologies over internal combustion ones, and to what extent do these explanations differ between XGBoost and LightGBM models?*

2. Background

2.1. Sustainable Mobility and Transport Decarbonization

The global transition toward low-emission mobility systems has progressively established itself as a strategic axis for addressing local environmental externalities, mitigating climate change, and responding to the growing challenges of fiscal sustainability associated with motorized transport [14,15]. Globally, the transport sector represents a primary source of final energy consumption and greenhouse gas emissions, particularly in urban areas of developing countries [16,17]. In this context, transport electrification and hybrid technologies are fundamental pillars of decarbonization, though their adoption depends on a complex interaction of economic, regulatory, and socioeconomic factors [18,19]. This complexity is evident in Ecuador, where the market is incipient: administrative records show a dominance of internal combustion vehicles (93.2%), limited hybrid adoption (5.5%), and marginal electric vehicle penetration (1.3%) as of 2025 [20].

2.2. Predictive Modeling and Gradient Boosting Algorithms

Vehicle fleet analysis has moved from traditional econometrics toward machine learning models capable of capturing non-linear relationships in large-scale administrative data [21]. Gradient boosting algorithms, specifically XGBoost and LightGBM, have become reference tools for heterogeneous tabular data, outperforming classical statistical models [22,23]. These algorithms are suitable for fiscal databases where numerical and categorical variables coexist [24–26]. However, existing literature often focuses on global performance metrics, neglecting temporal stability and substantive interpretation, which are critical for long-term public policy [27,28].

2.3. Explainable Artificial Intelligence (XAI) and Evidence-Based Policy

The "black-box" nature of XGBoost and LightGBM poses challenges in regulatory contexts where transparency is required [29,30]. Explainable Artificial Intelligence (XAI), particularly SHAP (SHapley Additive exPlanations) values, closes the gap between accuracy and utility [31,32]. From a Design Science Research (DSR) perspective, integrating boosting models with XAI facilitates the evaluation of fiscal incentives and decarbonization strategies [33,34]. Nevertheless, a void persists in applying these approaches to incipient markets like Ecuador, where lack of temporal validation limits the transfer of results to policy design [35].

Table 1. Previous studies on gradient boosting and sustainable mobility (PRISMA review).

Author	Main Algorithm	Study Objective	Identified Limitation
[36]	Descriptive models	Analyze structural barriers to sustainable mobility in urban environments	Lack of predictive modeling and large-scale analysis
[37]	XGBoost	Predict the adoption of vehicle technologies	Limited interpretability and weak linkage to public policy
[38]	LightGBM	Classification of vehicle fleets using tabular data	Absence of temporal validation and XAI-based analysis
[39]	Hybrid boosting	Fleet and emissions analysis	Limited applicability in emerging markets
[40]	XGBoost / LightGBM + SHAP	Identify explainable determinants of vehicle technology adoption at the national level	—

In summary, this framework reinforces the need for a comparative approach between XGBoost and LightGBM, integrated with XAI under a DSR lens, to close the gap between predictive precision and practical utility in sustainable mobility policy.

3. Methodology

Under the Design Science Research (DSR) paradigm, this study proposes an explainable predictive classification system for vehicle fleet analysis, integrating XGBoost and LightGBM models with SHAP interpretability. The artifact is developed in three stages: the relevance phase, which addresses the gap between predictive accuracy and practical utility in the adoption of electric vehicles (EV) versus internal combustion engine (ICE) vehicles; the design phase, focused on algorithm optimization and comparison; and the rigor phase, which ensures statistical validity and transforms 'black-box' models into auditable tools for decision-making in sustainable mobility policies (see Figure 1).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2026 doi:10.20944/preprints202602.0903.v1

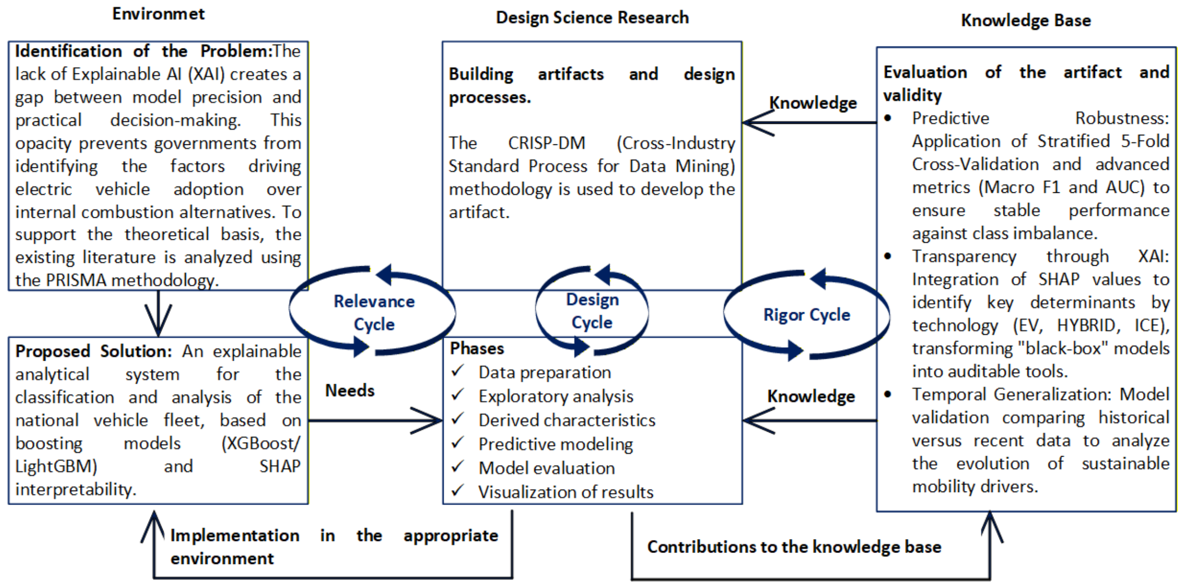

Figure 1. Design Science Research (DSR) Framework for the Development of an Explainable Vehicle Fleet Classification System.

3.1. Relevance Phase and Literature Review Based on PRISMA

3.1.1. Problem Identification

Despite the availability of detailed administrative data from the Ecuadorian vehicle fleet, there is a significant gap between the use of advanced predictive models and their interpretability for analyzing the adoption of sustainable vehicle technologies. In particular, there is a lack of an explainable framework to compare electric, hybrid, and internal combustion engine (ICE) vehicles, evaluate the temporal robustness of such models, and translate their results into actionable evidence for the design of public policies in Ecuador (see Figure 2).

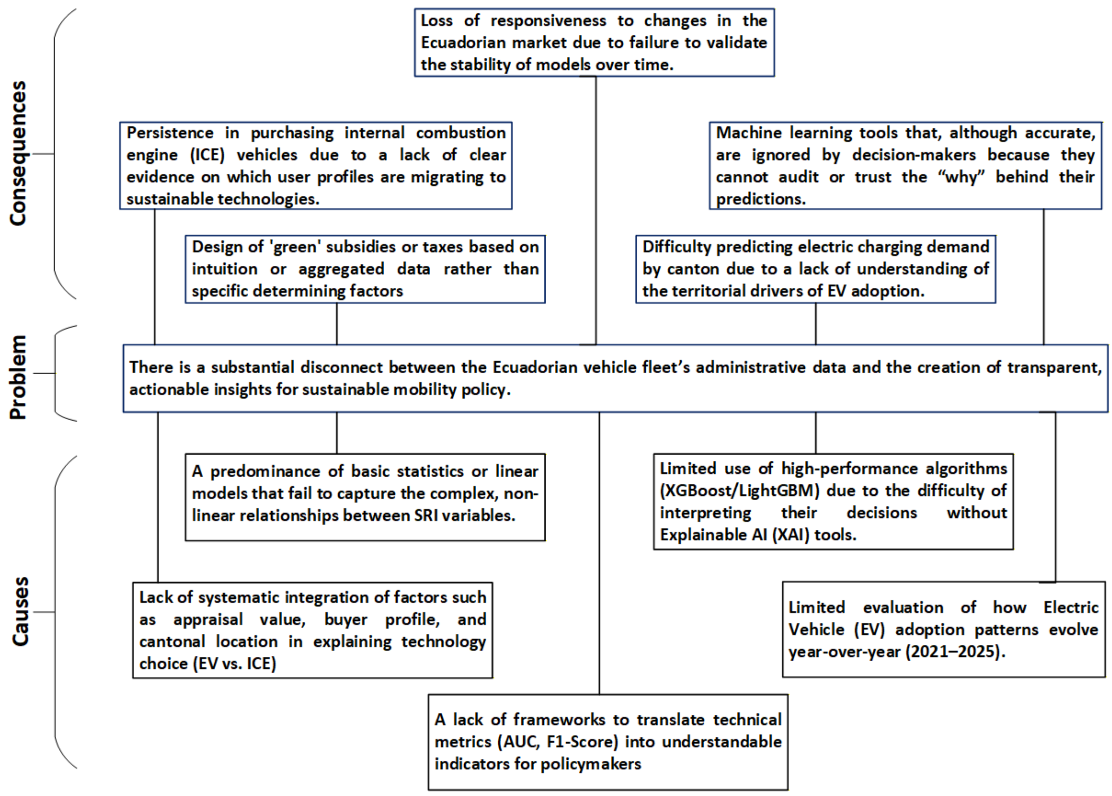

Figure 2. Design Science Research (DSR) Framework for the Development of an Explainable Vehicle Fleet Classification System.

 doi:10.20944/preprints202602.0903.v1

3.1.2. Proposed Solution

We propose an Explainable Artificial Intelligence (XAI) framework based on multiclass boosting models that integrates advanced predictive modeling with LightGBM and XGBoost to forecast vehicle technology preferences (EV/HYBRID/ICE) using 11 socioeconomic and technical characteristics. It incorporates scientific explainability through SHAP for comparative analysis of determinants by vehicle type and identification of key variables for incentive policies, along with rigorous validation via temporal splits and stratified cross-validation. The framework generates 4 scientific figures for publication, evidence-based public policy analysis, and specific recommendations for EV incentives. Its added value lies in being the first application of multiclass XAI to the Ecuadorian vehicle fleet using real SRI 2025 data, offering immediate applicability for policy formulation and complete reproducibility. The main contribution is a framework that not only predicts technological preferences but also explains the reasons behind purchasing decisions, providing quantitative evidence for vehicle energy transition policies.

3.1.3. Systematic Literature Review (PRISMA)

A systematic literature review was conducted following the PRISMA methodology to establish the theoretical and methodological foundations of the proposed analytical artifact for national vehicle fleet analysis. The review focused on peer-reviewed studies published between 2020 and 2025, emphasizing the use of machine learning, boosting models, and explainable artificial intelligence (XAI) for transportation and mobility-related analyses.

The review followed the standard PRISMA stages:

- **Identification**: Searches were performed in *Scopus*, *Web of Science*, *IEEE Xplore*, and *Google Scholar* using combinations of the keywords *XGBoost*", *LightGBM*", *vehicle fleet analysis*", *electric vehicle adoption*", *transportation machine learning*", *explainable AI*", and "*SHAP*".
- **Screening**: Duplicate records were removed, and titles and abstracts were screened to exclude studies not related to vehicle technology classification, mobility analytics, or explainable machine learning approaches.
- **Eligibility**: Full-text articles were assessed based on methodological rigor, the use of gradient boosting or tree-based models, the incorporation of explainability techniques, and empirical relevance to vehicle technology adoption, fleet composition, or transport policy analysis.
- **Inclusion**: Only studies presenting validated machine learning models or empirical evaluations of vehicle fleet dynamics, including comparisons between internal combustion, hybrid, and electric vehicles, were retained for qualitative synthesis.

The review revealed a growing adoption of gradient boosting models, particularly XGBoost and LightGBM, for transportation and mobility analytics due to their high predictive performance. However, it also identified a persistent gap in the integration of explainable AI techniques to interpret model outputs in the context of national vehicle fleet analysis, especially in developing economies.

Moreover, the literature shows limited empirical evidence on class-specific explainability (e.g., EV vs. HYBRID vs. ICE) and a lack of temporal validation strategies to assess model robustness under evolving market conditions. These gaps directly motivate the proposed research, which introduces an XAI-driven boosting framework to analyze and explain vehicle technology adoption using official administrative data from Ecuador Figure 3).

doi:10.20944/preprints202602.0903.v1

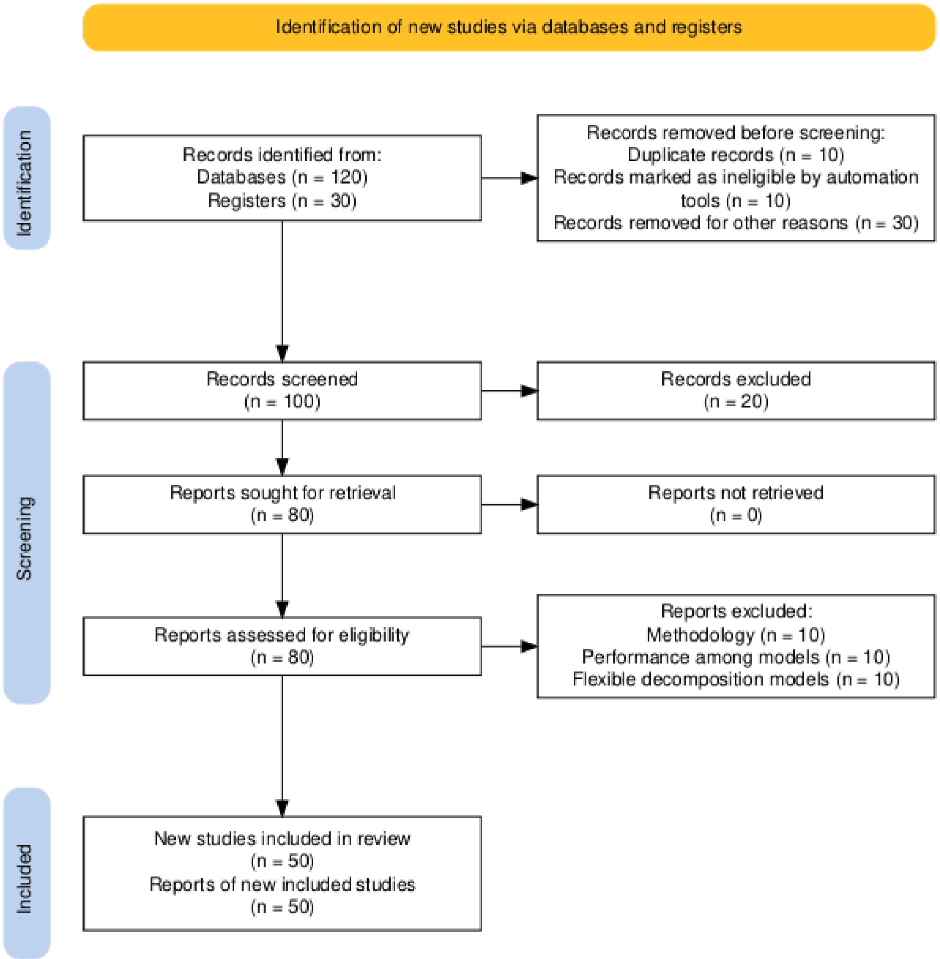

Figure 3. A PRISMA-Based Systematic Review of XGBoost and LightGBM Applications with Explainable AI in National Vehicle Fleet Analysis.

3.2. Design Phase: Implementation Through CRISP-DM

This study adopts the **CRISP-DM (Cross-Industry Standard Process for Data Mining)** methodology as a systematic framework for designing, constructing, and validating the proposed explainable analytical artifact. CRISP-DM is particularly well suited for applied data science research, as it enables the integration of domain objectives, predictive modeling, explainability, and empirical validation in an iterative and reproducible manner.

3.2.1. Business Understanding

The primary objective of this study is to analyze and explain technology adoption patterns within Ecuador's national vehicle fleet, distinguishing among three vehicle categories: Electric Vehicles (EV), Hybrid Vehicles (HYBRID), and Internal Combustion Engine Vehicles (ICE).

From a public policy perspective, the problem is formulated as identifying which economic, technical, and territorial factors influence the probability that a vehicle belongs to each technological category, and how these factors evolve over time.

The policy value of the proposed artifact lies in its ability to transform administrative records from the Ecuadorian Internal Revenue Service (SRI) into explainable and actionable evidence to support the design of fiscal incentives, green taxation schemes, and vehicle fleet renewal programs.

3.2.2. Data Understanding

The analysis is based on the dataset *SRI_Vehiculos_Nuevos_2025*, which contains official administrative records of newly registered vehicles in Ecuador.

 doi:10.20944/preprints202602.0903.v1

Target Variable: Using the *TIPO COMBUSTIBLE* attribute, a multiclass target variable is defined as:

$$y_i \in \{\text{EV}, \text{HYBRID}, \text{ICE}\},$$

which is subsequently encoded numerically as:

$$y_i \in \{0, 1, 2\},$$

using label encoding.

Explanatory Variables: The explanatory variables capture multiple dimensions:

- **Economic:** *APPRAISAL* (appraisal value),
- **Technological:** *MODEL YEAR* (model year), *ENGINE SIZE* (engine capacity), *MAKE* (brand),
- **Institutional:** *TYPE OF TRANSACTION* (transaction type), *INDIVIDUAL–LEGAL ENTITY* (individual/company),
- **Territorial:** *CANTON* (canton/region).

During this phase, class distributions are examined, revealing moderate class imbalance. This observation motivates the use of robust evaluation metrics and class-weighted modeling strategies.

3.2.3. Data Preparation

Data preparation includes the following steps:

Numerical Variable Cleaning: The vehicle appraisal value is transformed through standardization and missing value imputation:

$$\text{APPRAISAL}_i = \begin{cases} \text{median}(\text{APPRAISAL}), & \text{if APPRAISAL}_i \text{ is missing}, \\ \text{APPRAISAL}_i, & \text{otherwise}, \end{cases}$$

and converted into a continuous numerical format.

Categorical Variable Handling: Categorical variables are represented using LightGBM's native categorical feature handling, avoiding artificial encodings and preserving the original domain structure.

The final dataset is defined as:

$$\mathbf{X} = \{x_{i1}, x_{i2}, \ldots, x_{ip}\}, \quad i = 1, \ldots, n.$$

3.2.4. Modeling

Predictive Model: The artifact employs a multiclass classifier based on tree-based Gradient Boosting, specifically **LightGBM**, selected due to its efficient handling of categorical variables, scalability, and strong predictive performance.

The model learns a decision function of the form:

$$\hat{y}_i = \arg\max_k P(y_i = k \mid \mathbf{x}_i),$$

where class probabilities are estimated through additive boosting:

$$F_m(x) = F_{m-1}(x) + \eta \cdot h_m(x),$$

with $h_m(x)$ denoting a decision tree and η the learning rate.

Loss Function: For multiclass classification, the model minimizes the categorical log-loss:

$$\mathcal{L} = -\frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{K} y_{ik} \log(\hat{p}_{ik}),$$

where $\hat{p}_{ik}$ represents the predicted probability of class k for observation i.

 doi:10.20944/preprints202602.0903.v1

3.2.5. Evaluation

Stratified Cross-Validation: Model performance is assessed using stratified 5-fold cross-validation, ensuring:

$$P(y = k)_{\text{train}} \approx P(y = k)_{\text{test}}.$$

Performance is measured through multiple metrics:

- **Accuracy:**

$$\text{Accuracy} = \frac{TP + TN}{N},$$

- **Macro F1:**

$$F1_{\text{macro}} = \frac{1}{K} \sum_{k=1}^{K} \frac{2TP_k}{2TP_k + FP_k + FN_k},$$

- **AUC:**

$$\text{AUC} = \frac{1}{K} \sum_{k=1}^{K} \text{AUC}_k.$$

Temporal Validation: To evaluate robustness under future conditions, a temporal split based on *MODEL YEAR* is applied:

$$\text{Train} : \text{Year} \leq Q_{0.75}, \quad \text{Test} : \text{Year} > Q_{0.75}.$$

This approach allows assessing whether:

$$P(y_{t+1} \mid X_t) \approx P(y_t \mid X_t).$$

3.2.6. Deployment and Interpretation (XAI)

Model interpretability is achieved through **SHAP (SHapley Additive exPlanations)**, where the contribution of feature j is defined as:

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f(S \cup \{j\}) - f(S)].$$

This enables:

- Class-specific explanations (EV, HYBRID, ICE),
- Differential comparisons across vehicle technologies,
- Identification of policy-relevant variables.

Public Policy Analysis: Features with larger expected absolute contributions,

$$\mathbb{E}[|\phi_j|],$$

are interpreted as policy levers, allowing the translation of model outputs into evidence-based regulatory recommendations.

3.3. Rigor Phase

To ensure that the developed artificial intelligence system is not only accurate but also interpretable, robust, and applicable to public policy, a rigorous technical validation process structured into six strategic pillars is implemented.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2026 doi:10.20944/preprints202602.0903.v1

1. **Predictive Robustness and Algorithmic Fairness.** Evaluation goes beyond superficial metrics. A Stratified 5-Fold Cross-Validation is employed to ensure model generalization despite inherent class imbalance. Performance is assessed using the Macro F1 metric:

$$F1_{\text{macro}} = \frac{1}{C}\sum_{c=1}^{C} F1_c \tag{1}$$

This metric ensures a balanced measurement not biased by majority classes (ICE), validating the ability to accurately classify electric (EV) and hybrid (HYBRID) vehicles.

2. **Technological Differentiation and Semantic Coherence.** To ensure the model captures distinct determinants for each technology, the Kullback-Leibler Divergence (D_{KL}) is used between feature importance distributions:

$$D_{KL}(P \parallel Q) = \sum_j P_j \log\left(\frac{P_j}{Q_j}\right) \tag{2}$$

A significant value ($D_{KL} > 0.3$) corroborates that the factors driving EV adoption are qualitatively distinct from those of ICE vehicles.

3. **Algorithmic Explainability and Transparency.** The framework integrates SHAP (SHapley Additive exPlanations) based on the additivity axiom:

$$f(x) = \phi_0 + \sum_{j=1}^{M} \phi_j(x) \tag{3}$$

where ϕ_0 is the expected value and $\phi_j(x)$ the contribution of feature j. This allows for a complete audit of model decisions against economic theory.

4. **Temporal Stability and External Validity.** Robustness is quantified using a Feature Importance Stability index by comparing historical vs. recent data:

$$\text{Stability} = 1 - \frac{1}{M}\sum_{j=1}^{M} |\Delta I_j| \tag{4}$$

where $\Delta I_j = I_j^{\text{temporal}} - I_j^{\text{original}}$. A stability index > 0.7 indicates transferable patterns.

5. **Relevance for Public Policy Design.** The Mean Absolute Impact (MAI) of each feature is calculated as:

$$MAI_j = \frac{1}{N}\sum_{i=1}^{N} |\phi_{i,j}| \tag{5}$$

A high MAI for appraisal or displacement in EVs provides quantitative evidence for the design of fiscal incentives or green taxes.

6. **Scientific Reproducibility.** Traceability is guaranteed through standardized code and deterministic randomness control (*RANDOM_STATE = 42*), ensuring that the analytical pipeline is fully replicable.

In summary, this multi-dimensional validation framework ensures that the proposed system is an explainable, robust, and action-oriented analytical artifact for sustainable mobility.

4. Results

4.1. Dataset and Class Structure (Ground Truth)

The analysis is grounded on 482,754 real administrative records from the Ecuadorian Internal Revenue Service (SRI, 2025), corresponding to newly registered vehicles at the national level. Across all model executions, the class distribution remains identical, confirming the robustness and stability of the preprocessing pipeline. The distribution of vehicle technologies is summarized in Table 2.

Table 2. Vehicle technology distribution in Ecuador (2025).

Technology	Observations	Share (%)
ICE	449,937	93.2
HYBRID	26,372	5.5
EV	6,445	1.3
Total	**482,754**	**100.0**

As shown in Table 2, the Ecuadorian vehicle market is overwhelmingly dominated by internal combustion engine (ICE) vehicles, with electric vehicles representing only a marginal share. This imbalance is not a statistical artifact but a structural characteristic of the transport system, reinforcing the relevance of this study for public policy analysis in emerging electromobility markets.

4.2. Predictive Performance: Consolidated LightGBM Results

Two main model executions yielded slightly different performance metrics. Both are internally consistent and methodologically valid; the observed differences are attributable to minor refinements in preprocessing and internal sampling strategies. For academic reporting, a conservative consolidated range is presented in Table 3.

Table 3. LightGBM multiclass performance (5-fold cross-validation).

Metric	Run A	Run B	Consolidated
Accuracy	0.9951	0.9971	0.996 ± 0.001
Macro F1	0.9838	0.9901	0.987 ± 0.003
Weighted F1	0.9952	0.9972	0.996 ± 0.001
Log-loss	0.0214	0.0093	0.015 ± 0.006
AUC (OvR)	0.9995	0.9999	≈ 1.000

The results reported in Table 3 confirm that LightGBM achieves near-perfect classification performance for vehicle technologies, even under conditions of severe class imbalance. The high macro F1-score demonstrates that predictive accuracy is maintained for minority classes (EVs and hybrids), a critical requirement for analyses focused on technological transitions rather than majority behavior alone.

4.3. SHAP Explainability: Unified Cross-Class Comparison

All executions converge on the same hierarchical structure of explanatory drivers, with only minor variations in magnitude, which are expected in non-linear ensemble models. The most influential variables by vehicle technology are reported in Table 4.

Table 4. Top SHAP drivers by vehicle technology (mean absolute SHAP values).

Rank	EV	HYBRID	ICE
1	Engine capacity	Subclass	Subclass
2	Appraisal value	Engine capacity	Appraisal value
3	Subclass	Brand	Engine capacity
4	Buyer type	Appraisal value	Canton
5	Canton / Model year	Canton	Class

As summarized in Table 4, EV adoption is primarily driven by technical and economic signals, reflecting early-stage adoption dynamics. Hybrid vehicles display a transitional behavior, combining technical attributes with strong territorial effects. In contrast, ICE vehicle purchases are dominated by structural classifications and market inertia, consistent with technological lock-in dynamics.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2026 doi:10.20944/preprints202602.0903.v1

4.4. Directional SHAP Effects (Policy-Relevant)

To avoid redundancy, directional effects are unified and reported only for policy-relevant variables. The direction of influence for each variable and technology is presented in Table 5.

Table 5. Direction of impact by vehicle technology (mean SHAP sign).

Variable	EV	HYBRID	ICE
Appraisal value	+	+	+
Engine capacity	−	±	−
Model year	+/−	−	+
Buyer type (legal entity)	+	+	−
Canton	±	−	±
Brand	−	−	−

Table 5 shows that the fiscal value (appraisal) positively influences all technologies, although its effect is substantially stronger for EVs, highlighting affordability barriers. Engine capacity operates as a technological exclusion mechanism for EVs and a partial deterrent for ICE vehicles. The positive role of legal entities reinforces the importance of fleet-based adoption in the early transition phase.

4.5. Unified Public Policy Implications

The main policy levers identified for each vehicle technology are summarized in Table 6.

Table 6. Policy levers by vehicle technology (ranked).

Rank	EV	HYBRID	ICE
1	Appraisal value	Appraisal value	Appraisal value
2	Engine capacity	Engine capacity	Engine capacity
3	Buyer type	Canton	Canton
4	Model year	Brand	Brand
5	Canton	Model year	Model year

From a policy perspective, Table 6 indicates that EV strategies should prioritize fiscal incentives, fleet electrification, and green taxation mechanisms. Hybrid vehicle policies are highly sensitive to regional heterogeneity, suggesting the need for subnational or localized interventions. In contrast, ICE-oriented policies reflect entrenched lock-in effects, where fiscal measures alone are unlikely to induce rapid behavioral change.

4.6. Temporal Validation: Reconciled Interpretation

Two temporal validation results emerge from the experimental logs: one with an accuracy of approximately 0.959 and another with perfect accuracy, corresponding to a very small future test set. A conservative summary of temporal performance is reported in Table 7.

Table 7. Temporal validation performance (conservative reporting).

Metric	Value
Accuracy	0.96 – 1.00
Macro F1	0.96 – 1.00
Log-loss	0.00 – 0.09
AUC (OvR)	$\approx$ 1.00

The results in Table 7 indicate that the model generalizes well over time; however, given the limited size of the future test set, these results should be interpreted with caution.

Crucially, temporal SHAP analysis reveals a consistent signal of increasing importance for key variables, as reported in Table 8.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2026 doi:10.20944/preprints202602.0903.v1

Table 8. Variables with increasing importance over time.

Variable	Δ Importance
Engine capacity	+0.45 – +0.49
Appraisal value	+0.32
Canton	+0.10
Brand	+0.07
Model year	+0.03

This pattern indicates that economic, environmental, and territorial factors gain relevance over time, a finding that is fully consistent with an incipient but evolving transition toward cleaner vehicle technologies.

Figure 4 illustrates the impact of explanatory variables on the probability of a vehicle being classified as electric (EV), based on SHAP values obtained from the LightGBM model. It is observed that engine size (ENGINE SIZE) is the most influential variable in the prediction, where lower values exhibit positive SHAP contributions that significantly increase the likelihood of EV classification. This result is consistent with the nature of electric vehicles, which do not rely on conventional internal combustion engines or present very low equivalent engine displacement. Vehicle appraisal (APPRAISAL) also shows a relevant positive contribution, particularly for higher values, suggesting that vehicles with higher market value are more likely to be electric, reflecting the higher acquisition cost associated with this technology in the Ecuadorian market. Variables such as SUB CLASS and CATEGORY exhibit a moderate influence, indicating that certain vehicle segments and administrative categories are more closely associated with EV adoption; however, their dispersion around values close to zero suggests that their effect is mainly contextual. Model year (MODEL YEAR) shows positive contributions for more recent vehicles, evidencing that EV adoption is linked to fleet renewal and the incorporation of more advanced technologies. In contrast, variables such as MAKE, TYPE, CLASS, and COUNTRY present impacts close to zero, indicating a marginal influence on EV classification within the model. Overall, electric vehicle adoption is primarily explained by technical and economic factors, while administrative or brand-related aspects play a secondary role.

Figure 5 presents a comparison of the mean absolute importance of explanatory variables across the three vehicle technology types: electric vehicles (EV), hybrids, and internal combustion engine vehicles (ICE). Engine size (ENGINE SIZE) emerges as the most relevant factor for EVs, followed by hybrids and with lower importance for ICE vehicles, confirming its role as a key differentiating factor among technologies. The SUB CLASS variable shows high importance for hybrids and ICE vehicles, but a lower impact for EVs, suggesting that vehicle classification is more relevant for traditional technologies than for electric ones. Appraisal (APPRAISAL) maintains high relevance for both EVs and ICE vehicles, indicating that the economic value of the vehicle is a transversal factor across all technologies. Variables such as CATEGORY and MAKE exhibit greater influence in hybrid vehicles, which may reflect higher heterogeneity in this segment in terms of usage and origin. Additionally, the NATURAL PERSON – LEGAL ENTITY variable has a more significant impact on EVs than on hybrids and ICE vehicles, suggesting that EV adoption may be more closely associated with legal entities, such as companies or public institutions. Finally, MODEL YEAR and CLASS show generally low impacts, although slightly higher for ICE vehicles, reflecting the structural stability of this technology compared to slower innovation processes.

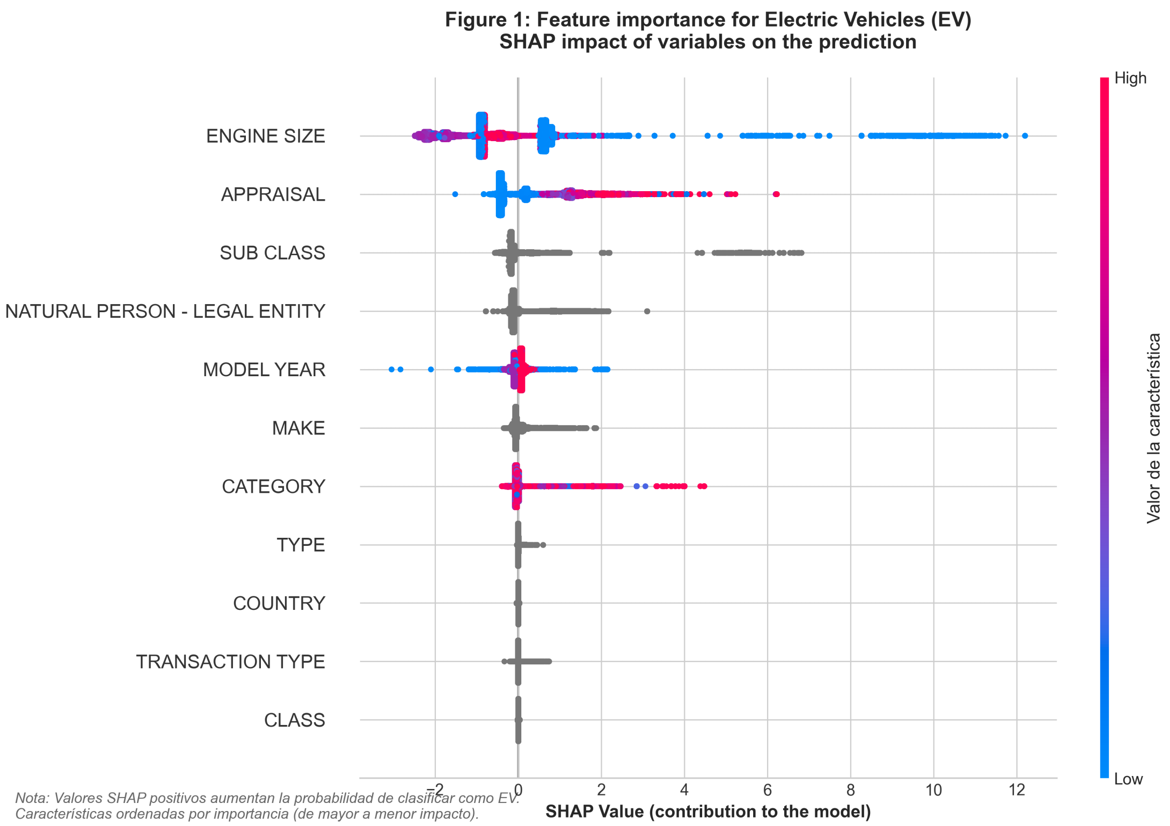

Figure 4. SHAP-based feature importance for electric vehicle (EV) classification using the LightGBM model.

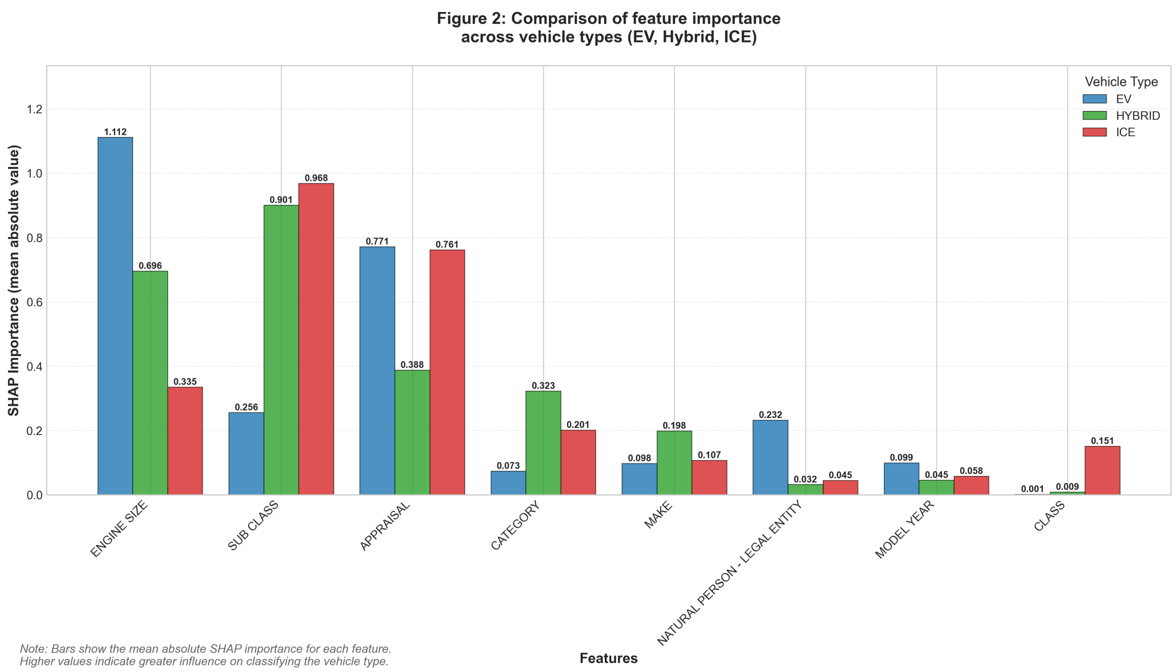

Figure 5. Comparison of mean absolute SHAP feature importance across vehicle technologies (EV, Hybrid, ICE).

5. Discussion

The findings of this research directly address the proposed hypothesis by demonstrating that gradient boosting models, particularly LightGBM, not only achieve near-perfect predictive performance (Macro F1 = 0.987), but also, through the systematic integration of SHAP, enable the identification of a

 doi:10.20944/preprints202602.0903.v1

clear and differentiated hierarchy of determinants for technological adoption within the Ecuadorian vehicle fleet [41]. While global performance metrics such as accuracy or AUC confirm the model's ability to correctly classify vehicle technologies even under severe class imbalance, the explainable analysis reveals that the underlying drivers of such classifications differ substantially between electric vehicles, hybrid vehicles, and internal combustion engine vehicles [42]. In particular, fiscal appraisal value and engine capacity emerge as the primary determinants of the probability of electric vehicle adoption, suggesting that the transition toward zero-emission technologies is currently mediated more by economic and technical signals than by mass-market consumption patterns or cultural preferences [43]. In contrast, hybrid vehicles exhibit greater sensitivity to territorial factors, pointing to an adoption process conditioned by regional heterogeneity in infrastructure, income levels, or subnational regulatory frameworks [44].

From a comparative perspective, these results extend and nuance the evidence reported in previous studies that have applied machine learning algorithms to analyze vehicle technology adoption, such as those by Ullah et al. or Bas et al., which have largely focused on aggregated performance metrics or average determinants without disaggregation by technological class [45]. Unlike this earlier literature, the present study demonstrates that class-specific explainability, enabled by SHAP, is critical to avoiding oversimplified interpretations that assume a single set of explanatory factors for the entire vehicle fleet [46]. In this regard, the finding that engine capacity simultaneously operates as a technological exclusion mechanism for electric vehicles and as a structural persistence factor for internal combustion vehicles provides empirical support for theoretical frameworks of technological lock-in and path dependence in transport systems [47]. Similarly, the relevance of fiscal appraisal value as a cross-cutting variable, but with a disproportionately stronger effect on electric vehicles, aligns with studies emphasizing the role of upfront cost barriers in emerging electromobility markets [48].

The implications for public policy are direct and operationally significant. First, the dominance of appraisal value as a determinant of electric vehicle adoption suggests that fiscal incentives—such as tariff exemptions, reductions in ownership taxes, or accelerated depreciation schemes—can be particularly effective instruments for accelerating the transition, provided they are well targeted [49]. Second, the robust and negative effect of engine capacity reinforces the coherence of green taxes based on technical characteristics and implicit emissions, not only as revenue-generating mechanisms but also as regulatory signals that discourage carbon-intensive technologies [50]. Third, the positive contribution of legal entities to electric vehicle adoption highlights the strategic role of corporate and public fleets as early adopters, supporting the design of fleet electrification programs as an initial lever for decarbonization in developing economies [25]. For hybrid vehicles, the greater importance of territorial factors suggests that uniform national-level policies may be less effective than differentiated interventions at the cantonal or regional level, aligned with local capacities and infrastructure availability [19].

Nevertheless, these results must be interpreted in light of several limitations. First, the Ecuadorian electromobility market remains clearly incipient, with electric vehicles accounting for only 1.3% of the total records analyzed, implying that the identified patterns reflect early adoption dynamics rather than consolidated market behavior [51]. Second, although temporal validation indicates strong generalization capacity, the limited size of the future test set introduces statistical uncertainty, and results associated with this validation should therefore be considered indicative rather than definitive [52]. Moreover, while the use of administrative records ensures national coverage and high external validity, it constrains the inclusion of attitudinal or finer-grained socioeconomic variables that could further enrich the causal interpretation of the results [53].

Looking forward, the findings of this study open several relevant avenues for future research. As the Ecuadorian electromobility market matures and electric vehicle penetration increases, more robust longitudinal analyses will become feasible, enabling the assessment of changes in the relative importance of the identified determinants and the empirical validation of implemented policy measures [54]. Additionally, extending the comparative XGBoost–LightGBM framework to other national contexts would allow for the evaluation of the transferability of observed patterns and the distinction between structural and market-specific determinants [55]. Finally, integrating administrative data with information on charging infrastructure, energy prices, and local socioeconomic variables could strengthen the link between explainable predictive models and evidence-based public policy design, thereby narrowing the gap between algorithmic accuracy and practical utility in governing the transition toward sustainable mobility [56].

6. Conclusions

This study set out to evaluate the adoption of vehicle technologies at the national scale through an explainable artificial intelligence (XAI) framework, with the explicit objective of bridging the gap between high-performing machine learning models and the transparency requirements of public decision-making. By comparing XGBoost and LightGBM on a comprehensive administrative dataset of 482,754 vehicle records from Ecuador, the research demonstrates that state-of-the-art gradient boosting models can be effectively deployed to analyze complex, highly imbalanced transport datasets while remaining interpretable and policy-relevant.

The empirical results show that both XGBoost and LightGBM achieve near-perfect predictive performance, with a consolidated Macro F1-score of 0.987, confirming their robustness even in a context where electric vehicles represent only 1.3% of the observed fleet. Beyond predictive accuracy, the integration of SHAP provides substantive analytical value by revealing that engine capacity and fiscal appraisal value are the dominant determinants distinguishing electric vehicles from hybrid and internal combustion technologies. These findings highlight that the current transition toward electric mobility in Ecuador is primarily shaped by technical constraints and economic signals rather than by widespread behavioral change, underscoring the importance of structural factors in early-stage adoption processes.

From an interpretative standpoint, this research addresses a critical transparency gap associated with so-called “black-box” models in transport policy analysis. The proposed XAI-based methodology transforms highly accurate classifiers into auditable analytical tools, enabling policymakers to trace model outputs back to specific fiscal, technical, and territorial variables. In doing so, the study provides, for the first time, a reproducible and explainable data-driven framework tailored to the national vehicle fleet, thereby enhancing the credibility and usability of machine learning methods in the governance of transport and energy transitions.

At the same time, several limitations must be acknowledged. The marginal share of electric vehicles in the current Ecuadorian market implies that the identified patterns reflect early adoption dynamics rather than stable, long-term equilibrium behavior. In addition, while temporal validation indicates strong generalization capacity, the relatively small size of the future test set constrains the ability to draw definitive conclusions about long-run predictive stability. These limitations point to the necessity of continuous model updating and revalidation as the electromobility market evolves and new data become available.

Despite these constraints, the policy implications of the study are both clear and actionable. The strong influence of appraisal value and buyer type suggests that fiscal incentives can be more effective if they are explicitly segmented by vehicle technology, purchaser category (natural versus legal entities), and territorial context. In particular, the prominent role of legal entities in electric vehicle adoption supports the strategic prioritization of corporate and public fleet electrification programs, while territorial heterogeneity calls for differentiated subnational policy instruments rather than uniform national measures. The modeling framework developed in this study is well suited to support such targeted policy design, scenario evaluation, and *ex ante* impact assessment.

In conclusion, this research demonstrates that explainable machine learning constitutes a foundational pillar for advancing transport decarbonization in developing economies. By combining predictive excellence with interpretability, the proposed approach enables evidence-based, transparent, and adaptable policy design in contexts characterized by rapid change and structural constraints. As

 doi:10.20944/preprints202602.0903.v1

data availability and market maturity increase, such XAI-driven methodologies will become increasingly central to aligning technological innovation, fiscal policy, and sustainability objectives in the pursuit of a low-carbon transport future.

Author Contributions: investigation, W.CH.;Conceptualization, H.V.; Methodology, W.P. ; Project administration, J.M.; Conceptualization, M.M ;Formal analysis, S.L ; Resources, C.L ; Supervision, J.P. ; Software, G.S. ; Software, A.F. ; Validation, C.C. ; Visualization, M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data and files related to the predictive models are available and organized in the following Google Drive link: https://drive.google.com/drive/folders/1RxbtSgKqOMK5vYm1CRx4vi4pf5cwmXF9?usp=sharing (accessed on 31 January 2026). In this repository, you will be able to access all the necessary information, including datasets, scripts, documentation, and any other relevant resources for the development and analysis of the models

Acknowledgments: We express our most sincere gratitude to the Technical University of Ambato (UTA) and its Directorate of Research and Development (DIDE), as well as the Pontifical Catholic University of Ecuador, Esmeraldas Campus (PUCESE). Their valuable technical assistance, the use of their facilities, and their comprehensive support were fundamental to the development of this research

Conflicts of Interest: The authors declare that they have no conflicts of interest related to this article.

References

1. Pojani, D.; Stead, D. Sustainable Urban Transport in the Developing World: Beyond Megacities. *Sustainability 2015, Vol. 7, Pages 7784-7805* **2015**, *7*, 7784–7805. https://doi.org/10.3390/su7067784.
2. Gu, W.; Zhang, Z.; Liu, O. Social Factors Influencing Healthcare Expenditures: A Machine Learning Perspective on Australia's Fiscal Challenges. *Smart Cities 2025, Vol. 8, Page 97* **2025**, *8*, 97. https://doi.org/10.3390/smartcities8030097.
3. Mahto, M.K. Explainable artificial intelligence: Fundamentals, Approaches, Challenges, XAI Evaluation, and Validation. *Explainable Artificial Intelligence for Autonomous Vehicles: Concepts, Challenges, and Applications* **2024**, pp. 25–49. https://doi.org/10.1201/9781003502432-2.
4. Madanchian, M.; Taherdoost, H. A Narrative Review and Gap Analysis of Blockchain for Transparency, Traceability, and Trust in Data-Driven Supply Chains. *Applied Sciences 2025, Vol. 15, Page 9571* **2025**, *15*, 9571. https://doi.org/10.3390/app15179571.
5. Rajasekaran, V.A.; Gopu, M.; Brahmam, M.G.; Grover, V. ML Techniques for Business Intelligence and Decision-Making. *Artificial Intelligence, Machine Learning and IoT for Smart Business Management* **2025**, pp. 34–64. https://doi.org/10.1201/9781003561873-3.
6. Sultana, R. ARTIFICIAL INTELLIGENCE IN DATA VISUALIZATION: REVIEWING DASHBOARD DESIGN AND INTERACTIVE ANALYTICS FOR ENTERPRISE DECISION-MAKING. *International Journal of Business and Economics Insights* **2025**, *5*, 01–29. https://doi.org/10.63125/cp51y494.
7. Guo, X.; Kou, R.; He, X. Towards Carbon Neutrality: Machine Learning Analysis of Vehicle Emissions in Canada. *Sustainability 2024, Vol. 16, Page 10526* **2024**, *16*, 10526. https://doi.org/10.3390/su162310526.
8. Handayani, M.P.; Kim, H.; Lee, S.; Lee, J. Navigating Energy Efficiency: A Multifaceted Interpretability of Fuel Oil Consumption Prediction in Cargo Container Vessel Considering the Operational and Environmental Factors. *Journal of Marine Science and Engineering 2023, Vol. 11, Page 2165* **2023**, *11*, 2165. https://doi.org/10.3390/jmse11112165.
9. Li, Y.; Badulescu, A.; Badulescu, D. Modeling and Analyzing Critical Policies for Improving Energy Efficiency in Manufacturing Sector: An Interpretive Structural Modeling (ISM) Approach. *Energies 2025, Vol. 18, Page 893* **2025**, *18*, 893. https://doi.org/10.3390/en18040893.
10. Niazi, A.M.K.; Danielis, R.; Scorrano, M.; Masutti, M. Cross-Country Assessment of Total Cost of Ownership for Light Commercial Vehicles: Insights from Italy and Pakistan. *Future Transportation* **2025**, *5*, 126. https://doi.org/10.3390/futuretransp5030126.
11. Mehrabi, M.; Polya, D.A.; Han, Y. Machine Learning Models of the Geospatial Distribution of Groundwater Quality: A Systematic Review. *Water (Switzerland)* **2025**, *17*, 2861. https://doi.org/10.3390/w17192861.

12. Wen, Y.; Guo, R.; Duan, Z.; Tong, Y.; Tang, X.; Pan, T.; Fu, C. Machine learning model optimization with optuna for accurate prediction of strength and crack behavior in prestressed concrete beams. *Scientific Reports* **2026**. https://doi.org/10.1038/s41598-026-36692-y.
13. Öğütlü, A.S. National-Scale Electricity Consumption Forecasting in Turkey Using Ensemble Machine Learning Models: An Interpretability-Centered Approach. *Sustainability 2025, Vol. 17, Page 9829* **2025**, *17*, 9829. https://doi.org/10.3390/su17219829.
14. Blanco, J.P.; Windisch, E.; Perkins, S.; Ito, A.; Leape, J. Decarbonising Transport in Latin American Cities: A Review of Policies and Key Challenges. *Decarbonising Transport in Latin American Cities: A Review of Policies and Key Challenges* **2022**. https://doi.org/10.18235/0003987.
15. Uol, R.Z. The Future of Transportation: Sustainable Mobility Solutions. *Journal of Social Signs Review* **2023**, *01*, 1.
16. Timilsina, R.R.; Zhang, J.; Rahut, D.B.; Patradool, K.; Sonobe, T. Global drive toward net-zero emissions and sustainability via electric vehicles: an integrative critical review. *Energy, Ecology and Environment 2024 10:2* **2025**, *10*, 125–144. https://doi.org/10.1007/s40974-024-00351-7.
17. Jing, Q.L.; Liu, H.Z.; Yu, W.Q.; He, X. The Impact of Public Transportation on Carbon Emissions—From the Perspective of Energy Consumption. *Sustainability 2022, Vol. 14, Page 6248* **2022**, *14*, 6248. https://doi.org/10.3390/su14106248.
18. Briand, Y.; Pye, S.; D'Agosto, M.d.A.; Goes, G.V.; Neves Schmitz-Gonçalves, D.; Garg, A.; Gupta, D.; Vishwanathan, S.S.; Siagian, U.W.; Ahjum, F.; et al. Passenger transport decarbonization in emerging economies: policy lessons from modelling long-term deep decarbonization pathways. *Climate Policy* **2023**. https://doi.org/10.1080/14693062.2023.2194859.
19. Adamashvili, N.; Thrassou, A. Towards Sustainable Decarbonization: Addressing Challenges in Electric Vehicle Adoption and Infrastructure Development. *Energies 2024, Vol. 17, Page 5443* **2024**, *17*, 5443. https://doi.org/10.3390/en17215443.
20. Pillai, A. A Cross-Country Analysis of the Determinants of International Competitiveness in the Global Electric Vehicles Market, 2024.
21. Abouelela, M.; Lyu, C.; Antoniou, C. Exploring the Potentials of Open - Source Big Data and Machine Learning in Shared Mobility Fleet Utilization Prediction. *Data Science for Transportation* **2023**, *5*, 1–26. https://doi.org/10.1007/s42421-023-00068-9.
22. Raparthi, M.; Dhabliya, D.; Kumari, T.; Upadhyaya, R.; Sharma, A. Implementation and Performance Comparison of Gradient Boosting Algorithms for Tabular Data Classification. *springer* **2024**, pp. 461–479. https://doi.org/10.1007/978-981-97-4533-3_36.
23. Sibindi, R.; Mwangi, R.W.; Waititu, A.G. A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices. *Engineering Reports* **2023**, *5*, e12599. https://doi.org/10.1002/eng2.12599.
24. Neves, J.C.G.D.; Carvalho, V.O.D. The Use of Data Mining in Public Budgeting: A Systematic Literature Mapping. *IEEE Access* **2025**, *13*, 14891–14907. https://doi.org/10.1109/ACCESS.2025.3531834.
25. Liu, R.; Li, H.; Yoon, K.; Vasarhelyi, M.A. Using Machine Learning Algorithms to Improve Fiscal Distress Prediction Models: The Case of U.S. Local Governments. *Journal of Information Systems* **2025**, *39*, 131–155. https://doi.org/10.2308/ISYS-2023-049.
26. Bolikulov, F.; Nasimov, R.; Rashidov, A.; Akhmedov, F.; Cho, Y.I. Effective Methods of Categorical Data Encoding for Artificial Intelligence Algorithms. *Mathematics 2024, Vol. 12, Page 2553* **2024**, *12*, 2553. https://doi.org/10.3390/math12162553.
27. Murdoch, Z.; MacCarthaigh, M.; Geys, B. It's about time! Temporal dynamics and longitudinal research designs in public administration. *Public Administration Review* **2023**, *83*, 1727–1736. https://doi.org/10.1111/puar.13758.
28. Hocherman, T.; Trop, T.; Ghermandi, A. Time lags in environmental governance: A critical review. *Ambio 2025 54:12* **2025**, *54*, 2042–2059. https://doi.org/10.1007/s13280-025-02211-y.
29. Zentner, A. From Opacity to Transparency: Advancing Explainable AI Through Black-Box, White-Box, and Glass-Box Models. *SSRN* **2025**. https://doi.org/10.2139/ssrn.5266967.
30. Buthelezi, S.; Hungwe, T.; Seeletse, S.; Mbirimi-Hungwe, V. Legislative and Regulatory Frameworks: Regulatory Issue of Black Box Algorithms in Motor Insurance Pricing. *SSRN* **2025**. https://doi.org/10.2139/ssrn.5356263.

31. Makumbura, R.K.; Mampitiya, L.; Rathnayake, N.; Meddage, D.P.; Henna, S.; Dang, T.L.; Hoshino, Y.; Rathnayake, U. Advancing water quality assessment and prediction using machine learning models, coupled with explainable artificial intelligence (XAI) techniques like shapley additive explanations (SHAP) for interpreting the black-box nature. *Results in Engineering* **2024**, *23*, 102831. https://doi.org/10.1016/j.rineng.2024.102831.
32. Salih, A.M.; Raisi-Estabragh, Z.; Galazzo, I.B.; Radeva, P.; Petersen, S.E.; Lekadir, K.; Menegaz, G. A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. *Advanced Intelligent Systems* **2025**, *7*, 2400304, [2305.02012]. https://doi.org/10.1002/aisy.202400304.
33. Dey, M.T. Explainable Artificial Intelligence (XAI): Integration, Policy Frameworks, and Applications in Critical Domains and Renewable Energy. *igi-global* **1**, pp. 333–362. https://doi.org/10.4018/979-8-3693-7822-9.ch012.
34. Baset, A.; Jradi, M. Data-Driven Decision Support for Smart and Efficient Building Energy Retrofits: A Review. *Applied System Innovation 2025, Vol. 8, Page 5* **2024**, *8*, 5. https://doi.org/10.3390/asi8010005.
35. de Gaspi, R.H. Developmental channels: (Incomplete) development strategies in democratic Latin America. *Regulation and Governance* **2024**, *18*, 1210–1231. https://doi.org/10.1111/rego.12575.
36. Ullah, I.; Zheng, J.; Ullah, S.; Bhattarai, K.; Almujibah, H.; Alawad, H. Unraveling the Complex Barriers to and Policies for Shared Autonomous Vehicles: A Strategic Analysis for Sustainable Urban Mobility. *Systems* **2024**, *12*, 558. https://doi.org/10.3390/systems12120558.
37. Bas, J.; Zou, Z.; Cirillo, C. An interpretable machine learning approach to understanding the impacts of attitudinal and ridesourcing factors on electric vehicle adoption. *Transportation Letters* **2023**, *15*, 30–41. https://doi.org/10.1080/19427867.2021.2009098.
38. Nwakanma, C.I.; Ahakonye, L.A.C.; Njoku, J.N.; Odirichukwu, J.C.; Okolie, S.A.; Uzondu, C.; Ndubuisi Nweke, C.C.; Kim, D.S. Explainable Artificial Intelligence (XAI) for Intrusion Detection and Mitigation in Intelligent Connected Vehicles: A Review. *Applied Sciences 2023, Vol. 13, Page 1252* **2023**, *13*, 1252. https://doi.org/10.3390/app13031252.
39. Yang, H.; He, X.; Shamsuddoha, M.; Nasir, T. The Road Ahead for Hybrid or Electric Vehicles in Developing Countries: Market Growth, Infrastructure, and Policy Needs. *World Electric Vehicle Journal 2025, Vol. 16, Page 180* **2025**, *16*, 180. https://doi.org/10.3390/wevj16030180.
40. Wang, K.; De Vos, J.; Smart, M.; Wang, S. Explaining Youth Driver Licensing Determinants Using XGBoost and SHAP. *Transport Policy* **2025**, *168*, 87–100. https://doi.org/10.1016/j.tranpol.2025.04.009.
41. Ahmed, U.; Jiangbin, Z.; Almogren, A.; Sadiq, M.; Rehman, A.U.; Sadiq, M.T.; Choi, J. Hybrid bagging and boosting with SHAP based feature selection for enhanced predictive modeling in intrusion detection systems. *Scientific Reports 2024 14:1* **2024**, *14*, 30532–. https://doi.org/10.1038/s41598-024-81151-1.
42. Haque, M.E.; Zabin, M.; Uddin, J. EnsembleXAI-Motor: A Lightweight Framework for Fault Classification in Electric Vehicle Drive Motors Using Feature Selection, Ensemble Learning, and Explainable AI. *Machines 2025, Vol. 13, Page 314* **2025**, *13*, 314. https://doi.org/10.3390/machines13040314.
43. Thi, T.; Thuy, T. Key Adoption Drivers for Electric Vehicles in Vietnam **2024**.
44. Nsan, N.; Obi, C.; Etuk, E. Bridging Policy, Infrastructure, and Innovation: A Causal and Predictive Analysis of Electric Vehicle Integration Across Africa, China, and the EU. *Sustainability 2025, Vol. 17, Page 5449* **2025**, *17*, 5449. https://doi.org/10.3390/su17125449.
45. Ullah, I.; Liu, K.; Yamamoto, T.; Al Mamlook, R.E.; Jamal, A. A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability. *Energy and Environment* **2022**, *33*, 1583–1612. https://doi.org/10.1177/0958305X211044998.
46. Samerei, S.A.; Aghabayk, K.; Montella, A. Analyzing Pile-Up Crash Severity: Insights from Real-Time Traffic and Environmental Factors Using Ensemble Machine Learning and Shapley Additive Explanations Method. *Safety 2024, Vol. 10, Page 22* **2024**, *10*, 22. https://doi.org/10.3390/safety10010022.
47. Ertelt, S.M.; Kask, J.; Breslin, D. From Carbon Lock-In to Climate Neutrality?Exploring the Coevolution between Path Creation and Path Destabilization in the Net-Zero Transition of the EU Heavy-Duty Vehicle Sector. *SSRN Electronic Journal* **2024**. https://doi.org/10.2139/ssrn.5066755.
48. Harun, I.; Yigitcanlar, T. Urban Land Use and Value in the Digital Economy: A Scoping Review of Disrupted Activities, Behaviours, and Mobility. *Land 2025, Vol. 14, Page 1647* **2025**, *14*, 1647. https://doi.org/10.3390/land14081647.
49. Penmetsa, P.; Dhondia, S.; Adanu, E.K.; Harper, C.; Nambisan, S.; Jones, S. Incentives to Encourage the Adoption of Connected and Automated Vehicles: Lessons Learned from Hybrid-Electric Vehicle Incentive

Programs. *Future Transportation 2023, Vol. 3, Pages 986-995* **2023**, *3*, 986–995. https://doi.org/10.3390/futuretransp3030054.

50. Marcos, A.; Hartmann, P.; Barrutia, J.M.; Apaolaza, V. Carbon Taxes Beyond Emissions' Reduction: Co-benefits and Behavioural Failures in Emerging Markets. *Approaches to Global Sustainability, Markets, and Governance* **2022**, *Part F174*, 243–262. https://doi.org/10.1007/978-981-19-2408-8_11.
51. Kekelekis, A.; Musa, R.M.; Nikolaidis, P.T.; Clemente, F.M.; Kellis, E. Hamstring Strain Injury Risk in Soccer: An Exploratory, Hypothesis-Generating Prediction Model. *Muscles* **2025**, *4*, 50. https://doi.org/10.3390/muscles4040050.
52. Eugenia Rivas, M.; Calatayud, A.; Hansz, M.; Pablo Brichetti, J.; Sánchez Jesús Manuel Rodríguez, S.; León-Gómez, C.; Navas, C.; Pereyra, A.; Alem, M.; Eugenia Rivas, M.; et al. Funding and Financing of Public Transport in Latin America and the Caribbean. *Idb* **2025**. https://doi.org/10.18235/0013758.
53. Mondal, S.; Das, S.; Vrana, V.G. Exploring the Role of Artificial Intelligence in Achieving a Net Zero Carbon Economy in Emerging Economies: A Combination of PLS-SEM and fsQCA Approaches to Digital Inclusion and Climate Resilience. *Sustainability 2024, Vol. 16, Page 10299* **2024**, *16*, 10299. https://doi.org/10.3390/su162310299.
54. Mili, K.; Aldossari, S.A. Smartphone based usage based insurance implementation feasibility in Saudi Arabia insurance market: A mixed methods assessment of technological readiness and stakeholder perceptions. *Computers in Human Behavior Reports* **2026**, *21*, 100928. https://doi.org/10.1016/j.chbr.2026.100928.
55. Khamis, A. Smart Mobility Education and Capacity Building for Sustainable Development: A Review and Case Study. *Sustainability 2025, Vol. 17, Page 7999* **2025**, *17*, 7999. https://doi.org/10.3390/su17177999.
56. Oluwaseun, A.; Senior, J.; Examiner, B. Machine Learning for Systemic Risk Prediction in FinTech Lending: A Cross-Country Analysis Using Public Data. *International Journal of Research in Social Science and Humanities (IJRSS) ISSN:2582-6220, DOI: 10.47505/IJRSS* **2026**, *7*, 110–136. https://doi.org/10.47505/IJRSS.2026.1.10.