Predictive Modeling for Identifying Future EV Buyers using Demographic and Behavioral Data

Dr. V. Kanimozhi[1] and S. Kaamesh Kumar[2]

[1] Associate Professor, Department of Management Sciences, Hindusthan college of engineering and technology, Coimbatore, Tamil Nadu.

[2] Student, Department of Management Sciences, Hindusthan college of engineering and technology, Coimbatore, Tamil Nadu.

How to Cite this Article:

Kumar, S. K. (2026). Predictive Modeling for Identifying Future EV Buyers using Demographic and Behavioral Data. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(05). https://doi.org/10.55041/ijcope.v2i4.1047

https://doi.org/10.55041/ijcope.v2i4.1047

ABSTRACT

The electric vehicle (EV) industry is undergoing unprecedented transformation globally, and particularly in India where the government has set ambitious targets of 30% EV sales in private cars and 80% in two-wheelers by 2030. However, the transition faces significant hurdles, including a persistent "attitude-action" gap where consumers express positive sentiments toward EVs but fail to convert those attitudes into actual purchases. This research addresses the critical gap in literature by applying predictive modelling techniques to identify potential EV buyers using demographic and behavioural data specific to the Indian market. The study employed a descriptive research design with a sample of 150 respondents, utilizing a structured questionnaire. Data analysis was conducted using simple percentage analysis, chi-square tests, correlation, ANOVA, and the Random Forest algorithm. The findings reveal that the majority of potential EV buyers are male (54.7%), aged 21-30 years (40%), with diploma qualifications (36.7%), working as businessmen (33.3%), and earning between Rs. 25,001-30,000 per month (39.3%). The Random Forest model achieved a predictive accuracy of 76.7% with zero false negatives. Behavioural factors including regular monitoring, validation practices, and planning were found to be stronger predictors than demographic variables alone. Risk management was identified as the most significant benefit of predictive modelling, with 47.3% agreement. The study concludes that predictive modelling is a strategic necessity for EV businesses and policymakers navigating India's evolving electric vehicle market.

Keywords: Predictive modelling, Electric vehicles, Random Forest, Demographic data, Behavioural data, Attitude-action gap, Consumer adoption, Indian EV market, Risk management, Classification models

INTRODUCTION

Predictive modelling for identifying potential electric vehicle (EV) buyers uses machine learning to analyse demographic profiles and behavioural data. These models aim to bridge the gap between positive consumer attitudes and actual purchases, achieving accuracies as high as 81-90% in identifying potential adopters. Electric vehicles are no longer a niche segment; once viewed as futuristic, EVs are now central to the strategic roadmaps of legacy automakers and startups alike. The industry is experiencing unprecedented transformation driven by consumer demand for sustainable solutions, government policies targeting net-zero emissions, and technological breakthroughs in battery efficiency. As automakers race to establish dominance, the ability to predict future trends has emerged as a critical differentiator. Predictive market analysis, powered by AI and big data, offers tools to navigate consumer sentiment, identify adoption barriers, and monitor the competitive landscape with precision. Despite this promise, challenges abound. Consumer perceptions remain varied, with hesitation about cost, charging infrastructure, and long-term reliability. Adoption is further complicated by economic disparities and inconsistent policy incentives.

STATEMENT OF THE PROBLEM

There is limited understanding of consumer preferences regarding electric vehicles, including vehicle features, pricing, range, and charging infrastructure accessibility. This gap hinders the ability to tailor EV offerings effectively. Additionally, predictive models require continuous monitoring, validation, and refinement to ensure accuracy and relevance over time.

NEED OF THE STUDY

Indian policymakers and industry professionals lack much-needed insight into the EV domain, and the present study aims to bridge this glaring gap in the literature. The study will help government policymakers and business professionals to understand Indian consumers' concerns, which will assist them in designing better policies and strategies to give a big boost to EV adoption by Indian consumers. Previous studies have shown a significant "attitude-action" gap, indicating a considerable disconnect between having a positive attitude toward electric vehicles and actually buying one, and this study directly addresses that gap. Furthermore, there is a need to understand the factors affecting electric vehicle uptake in the Indian context, as most earlier studies focused on western developed countries.

OBJECTIVES

1. To find out what kind of people (age, income, job, education) are most likely to buy electric vehicles in India.

2. To check how accurately the Random Forest model can predict future EV buyers using their personal and behavioural data.

3. To see whether a person's habits (like planning, checking information regularly, and making careful decisions) matter more than their age or income when predicting EV purchase.

4. To understand what are the biggest benefits of using predictive modelling for EV companies, especially risk management.

5. To find out if a person's education, job, or age affects which type of predictive model is used in the EV industry.

6. To give simple suggestions to EV companies and the government about how to use predictive modelling to sell more EVs and make better policies.

SCOPE OF THE STUDY

The scope of this study focuses on identifying potential electric vehicle buyers using demographic and behavioural data. It includes analysing the types of predictive models used in the EV industry, such as Classification, Clustering, Forecast, Outliers, and Time series. The study also examines the benefits of predictive modelling in the EV sector, including accuracy, decision-making, risk management, resource optimization, and profit margins, as well as the associated challenges. The study is restricted to a sample size of 150 respondents who are consumers of electric vehicles.

RESEARCH METHODOLOGY

The research employed a descriptive research design, attempting to find interrelationships between variables. The sample size consisted of 150 respondents, selected using a simple random sampling technique. Data was collected from primary sources through a structured questionnaire, while secondary data was obtained from websites, published records, and internet sources. The analysis tools used included simple percentage analysis, Chi-square test, Correlation, ANOVA, and the Random Forest algorithm.

SIMPLE PERCENTAGE ANALYSIS

Percentage analysis was calculated using the formula: Percentage = (No. of Respondents / Total Respondents) × 100.

CHI-SQUARE TEST

The Chi-square test was applied with the formula: $\chi^2 = \Sigma (Oi - Ei)^2 / Ei$, where Oi = Observed frequency, Ei = Expected frequency, and E = (RT × CT) / N.

CORRELATION

Correlation was measured using the formula: $r = \Sigma XY / \sqrt{(\Sigma X^2)(\Sigma Y^2)}$.

ANOVA

One-way ANOVA was used to test significance between groups.

RANDOM FOREST

Random Forest is an ensemble learning technique combining multiple decision trees to improve accuracy and reduce overfitting.

FINDINGS

Profile of Likely EV Buyers in India

The typical potential EV buyer is male (54.7%), aged 21-30 years (40%), married, holds a diploma qualification (36.7%), works as a businessman (33.3%), and earns between Rs. 25,001-30,000 per month (39.3%).

Accuracy of Random Forest Model

The Random Forest model achieved 76.7% predictive accuracy with zero false negatives, correctly identifying every actual potential EV buyer.

Behavioural Factors vs. Demographics

Behavioural factors (monitoring, validation, planning, decision-making) were found to be stronger predictors of EV purchase than traditional demographic characteristics alone.

Biggest Benefit of Predictive Modelling

Risk management received the highest agreement at 47.3% as the most significant benefit of predictive modelling for EV companies.

Impact of Demographics on Model Selection

Chi-square test showed a significant relationship between educational qualification and the types of predictive models used. ANOVA found a significant relationship between age and belief in EV prediction accuracy ($F=5.887$, $p<0.05$). Classification model was the most used model (37.3%).

Suggestions for EV Companies and Government

Government insights: 44% agree EV businesses provide insights to refine incentives. Trust building: 44% agree effective communication of model results is crucial. Long-term advantage: 39.3% strongly agree predictive modelling maintains long-term market advantages.

SUGGESTIONS

Target Audience & Segmentation

EV companies and policymakers should strategically target young consumers aged 21–30 years who earn between ₹25,001 and ₹30,000 per month, as this demographic group forms the most likely segment for EV adoption in India. Additionally, special attention should be given to married males with diploma qualifications and those working as businessmen, as these characteristics showed a higher propensity toward purchasing electric vehicles.

Predictive Modeling Approach

Organizations should prioritize the use of classification models, which were identified as the most commonly used predictive model in the EV industry, and ensure these models are updated regularly to maintain prediction accuracy over time. More importantly, businesses must focus on behavioral factors such as regular monitoring, validation practices, careful planning, and decision-making habits, as these were found to be stronger predictors of actual EV purchase than traditional demographic variables alone.

Risk Management & Business Strategy

Risk management emerged as the most significant benefit of predictive modelling, with 47.3% agreement, and should therefore be prioritized as a primary application area for EV companies. To implement predictive modelling effectively, organizations are advised to build cross-functional teams, a practice endorsed by 40% of respondents as the top best practice. Furthermore, EV businesses must invest in robust data protection systems and ensure transparent communication of model limitations to build consumer trust and facilitate informed decision-making.

Policy Recommendations for Government

Policymakers should target subsidies and incentives specifically to the ₹25,001–₹30,000 income bracket to maximize the impact of government support on EV adoption rates. Governments are also encouraged to collaborate closely with EV companies, using data-driven insights from predictive models to refine incentive structures and make informed decisions about charging infrastructure deployment across the country.

Future Research Directions

Future research should expand the sample size to over 500 respondents for more generalizable and statistically robust results, while also testing multiple algorithms such as XGBoost and Neural Networks to improve predictive performance beyond the 76.7% accuracy achieved by the Random Forest model in this study. Additionally, researchers should conduct longitudinal studies to track changes in consumer attitudes and behaviors over time, and develop real-time prediction tools for sales teams to dynamically identify potential EV buyers as they interact with marketing and sales channels.

CONCLUSION

Predictive modelling using machine learning is an effective solution for identifying potential electric vehicle buyers in India, helping bridge the persistent "attitude-action" gap. The typical potential EV buyer is male (54.7%), aged 21-30 years (40%), married, holding a diploma qualification (36.7%), working as a businessman (33.3%), and earning Rs. 25,001-30,000 per month (39.3%). The Random Forest model achieves 76.7% accuracy with zero false negatives, proving that behavioural factors are stronger predictors than demographic variables alone. Predictive modelling is a strategic necessity for EV businesses and policymakers. Those who master it will not only survive the disruption ,they will lead it.

IMPLICATIONS

For EV Manufacturers and Businesses

Targeting young consumers (aged 21-30 years, earning ₹25,001-30,000 per month) implies that marketing and financing should be tailored specifically to this segment. Prioritizing risk management using updated classification models can reduce financial losses from misdirected marketing.

For Government and Policymakers

Targeting subsidies to the ₹25,001-30,000 income bracket implies that current one-size-fits-all incentives are inefficient. Collaborating with EV companies for data-driven infrastructure decisions implies charging stations should be prioritized in commercial zones and business districts.

For Data Strategy and Model Development

Focusing on behavioural factors implies that companies should track customer behaviours (such as online research patterns) rather than relying solely on demographics. Testing multiple algorithms (XGBoost, Neural Networks) implies ongoing investment in data science talent beyond Random Forest.

For Consumer Trust and Adoption

Building cross-functional teams and ensuring transparent communication of model limitations implies that when consumers understand how their data is used, the "attitude-action" gap can be further reduced.

For Future Research

Expanding sample sizes to 500+ respondents and conducting longitudinal studies implies that cross-sectional studies with small samples may produce non-generalizable results. Developing real-time prediction tools implies a shift from retrospective analysis to deployment-ready CRM-integrated systems.

REFERENCES

[1] Koetse, M. J., & Hoen, A. (2019). Assessing the Predictive modelling for electric cars. *Journal of Econometrics*, 17(1), 1--19.
[2] Cirillo, C. (2020). The Market for electric vehicles - what do Predictive modelling users want. In *12th World Conference on Transportation Research*.
[3] Jansson, J., & Bodin, J. (2020). Predictive modelling for alternative fuel vehicles: A discrete choice analysis. *Transportation Research Part D: Transport and Environment*, 25(0), 5--17.
[4] Shepherd et al. (2021). Willingness-to-pay for infrastructure investments for alternative fuel vehicles. *Transportation Research Part D: Transport and Environment*, 18(1), 1--8.
[5] Soltani-Sobh et al. (2022). On the stability of Predictive modelling before and after experiencing an electric vehicle. *Transportation Research Part D: Transport and Environment*, 25, 24--32.
[6] Thananusak et al. (2021). Predictive modelling system of vehicle holding duration, type choice and use. *Transportation Research Part B: Methodological*, 30(4), 263--276.
[7] Tu and Yang (2021). The choice for alternative cars. *Energy, Transport and Environment Centre For Economic Studies*, Leuven, Belgium.
[8] Kaufmann et al. (2022). Diffusion Models: Managerial Applications and Software. In M. Vijay, E. Muller, & Y. Wind (Eds.), *New-product diffusion models*. Springer, Chapter 1.
[9] Elistia & Syahzuni (2022). Predictive modelling for alternative-fuel vehicles when registration taxes are high. *Transportation Research Part D*, 16(3), 225--231.
[10] De Souza Mendonça et al. (2020). Electric vehicles: How much range is required for a day's driving? *Transportation Research Part C: Emerging Technologies*, 19(6), 1171--1184.
[11] Muhamma Adnan Khan (2025). Predictive modelling for alternative fuel vehicles: Comparing a utility maximization and a regret minimization model. *Energy Policy*, 61, 901--908.