Predicting Electric Vehicle (EV) Buyers in India: A Machine Learning Approach

Sushil Kumar Dixit1 · Ashirwad Kumar Singh2

Received: 26 October 2021 / Accepted: 21 April 2022 / Published online: 18 May 2022
© Springer Japan KK, part of Springer Nature 2022

Abstract
Electric mobility has been around for a long time. In recent years, with advancements in technology, electric vehicles (EVs) have shown a new potential to meet many of the challenges being faced by humanity. These challenges include increasing dependence on fossil fuels, environmental concerns, challenges posed by rapid urbanization, urban mobility, and employment. However, the adoption of electric vehicles has remained challenging despite consumers having a positive attitude toward EVs and big policy pushes by governments in many countries. Marketers from the electric vehicle (EV) industry are finding it difficult to identify genuine buyers for their products. In this context, the present study attempts to develop a machine learning model to predict whether a person would “Buy” or “Won’t Buy” an electric vehicle in India. To develop the model, an exploration of EV context was done first by conducting a text analysis of online content relating to electric vehicles. The objective was to find frequently occurring words to gain a meaningful understanding of the consumer’s interests and concerns relating to electric vehicles. The machine learning model indicates that age, gender, income, level of environmental concerns, vehicle cost, running cost, vehicle performance, driving range, and mass behavior are significant predictors of electrical vehicle purchase in India. The level of education, employment, and government subsidy are not significant predictors of EV uptake.

Keywords Electric vehicles · EV buyers in India · EV adoption · Factors affecting EV purchase · Machine learning model to predict EV buyers

Sushil Kumar Dixit
sushil.dixit@yahoo.com
Ashirwad Kumar Singh
ashirwad_kumarsingh@lbsim.ac.in

1 Lal Bahadur Shastri Institute of Management, New Delhi, India
2 Denave India Pvt Ltd, Noida, Uttar Pradesh, India
1 Introduction

In business and economics literature, the term “industry” has been defined in varied context with different meanings for different purposes. Researchers and policy makers define industries depending on their objectives [1–4]. In general, “industry” refers to a set of business activities that is slightly domain specific. With this view, the authors define “electric vehicles industry” as a subset of automobile industry, including all businesses involved in the manufacture, trade, and service of all types of vehicles powered by electrical energy, and associated and ancillary businesses. Electric vehicles are considered one of the most important means by which some of the serious challenges being faced by modern societies are met, e.g., energy security, environmental deterioration, and urban mobility.

All industries evolve in time and space. New industries emerge and old ones vanish with changes in technology or consumer preferences. For the last few years, the electric vehicle industry has evolved in differing contexts. In case of electric vehicles (EVs), the government plays a significant role in shaping not only the perception, but also adoption of EVs by the masses. Governments across the world are coming up with electric vehicle policy, focusing on reducing dependence on fossil fuels, meeting environmental concerns and challenges posed by rapid urbanization, enhancing employment, among others. Governments in many countries have facilitated the adoption of EVs by policy interventions such as supporting research and development, infrastructure development, and financial incentives to industry and consumers.

Over the last few years, the Indian government has started focusing on electric vehicles [5, 6]. Recently, the Indian government declared that it aims to have EV sales account for 30% of private cars, 70% of commercial vehicles, and 80% of two and three wheelers by 2030 as there is an immediate need to de-carbonize the transport sector [7]. The central and state governments have both initiated policy measures to promote manufacture and adoption of EVs. To date, 15 state governments have announced EV policy for their states. Key components of the Indian government’s EV policy are making electric vehicles economically viable, developing charging/swapping infrastructure, technology advancement, and focusing on small and public vehicles to make an early impact. The EVs are also seen to contribute to economic development and employment in India. Many automobile manufacturers have recently launched EV models in two-wheeler and four-wheeler segments.

At the time of conducting this study, the Indian automobile industry was undergoing a severe slowdown, affecting overall consumer perception and sentiments. Retail prices of automobile fuel were at an all-time high, and economically, India’s GDP growth rate had slowed down. Steps like demonetization and Goods and Services Tax (GST) implementation has also negatively affected many small- and medium-size industries. The Reserve Bank of India (RBI) has been continuously reducing the policy rate, and the COVID-19 pandemic has also impacted all industries across the word, including the automobile industry.

The EV industry is moving at a fast pace in most of the countries, not only in terms of evolution of technology, but also in terms of government policy and
consumer expectations. Studies conducted in the EV domain in the past will become less relevant soon due to the fast-evolving nature of the industry. Many studies have been conducted in the past on the Indian automobile industry or its traditional segments, but the literature on the EV industry in India is limited and fragmented. Most of the earlier studies attempted to understand consumer sentiments toward electrical vehicles in western developed countries and China. Few studies focusing on understanding electric vehicle uptake were also conducted in the context of developed nations. In the past, only a few studies were conducted explaining consumer understanding and expectations toward EVs in the Indian context [8, 9]. A study focusing on consumer concerns for electric vehicles and understanding the factors affecting electric vehicle uptake was missing in the Indian context. Indian policy makers and industry professionals lacked much needed insight into the EV domain. 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 help them design better policies and strategies to give a big boost to EV adoption by Indian consumers.

This research paper is divided into two sections. Section 1 focuses on text analysis of Indian consumers’ posts on social media platforms to understand the user concerns toward electric vehicles. The text analysis of social media posts along with the review of existing literature provides insights into the context of the Indian electric vehicle industry, as well as the possible factors affecting EV uptake in India. Section 2 presents a machine learning model to predict who will buy an electric vehicle in India. It considers demographic, social, contextual, level of environmental awareness, and other relevant considerations of Indian consumers to develop a machine learning classification model that could predict whether a person would “Buy” or “Won’t Buy” an electric vehicle.

1.1 Objectives of Research

The two specific objectives of this study are:

1. Understanding Indian consumers’ concerns with electrical vehicles.
2. Developing a Predictive Machine Learning Model that can classify whether an Indian consumer will “Buy” or “Won’t Buy” an electric vehicle.

1.2 Literature Review

As previously mentioned, most of the existing studies examining consumer sentiments and factors relating to electrical vehicles have been conducted in either western or Chinese context. Shepherd et al. [10] developed a system dynamics model using factors such as subsidies, vehicle driving range, and availability of charging points, and concluded that subsidies have little impact, except in conditional marketing scenarios. Coffman et al. [11] concluded that despite significant performance improvements, most governments’ goals for EV uptake could not be met. Mixed
evidence was found for the role of government incentives in EV uptake; however, public charging infrastructure availability was found to have a significant impact. The authors also noted the presence of an “attitude-action” gap, indicating a significant gap between having a positive attitude toward electric vehicles and actually buying one. In their study, Christidis and Focas [12] identified that income, educational attainment, and urbanization level had a significant impact on EV uptake in the European Union (EU). The study also found that the local conditions and regional variations have a major, if not deciding, effect on EV purchase. Kim et al. [13] in their study of 31 countries found that the share of electric vehicles in different markets was correlated with their relative price as compared to internal combustion engine vehicle, number of models available, and vehicle driving range. However, they observed that the relationship between electric vehicle market share and availability of charging infrastructure was insignificant. While studying electric vehicle adoption in the USA, Soltani-Sobh et al. [14] found that electricity price, use of urban roads, and government incentives play a significant role in EV adoption.

Wang et al. [15] studied factors affecting public acceptance of electric vehicles in Shanghai, China, and concluded that the level of available technology, marketing efforts, perceived risks, and the level of environmental awareness have significant effects on electric vehicle acceptance. In their study in Thailand, Thananusak et al. [16] found that performance factors like driving range, speed, and safety were more important than the availability of charging infrastructure, financial considerations like purchase and resale price, and operating and maintenance costs. One important finding was that an individual’s environmental concerns affect their decision to purchase electric vehicles, and they were also willing to pay a higher price premium for electric vehicles due to their positive impact on the environment. However, the price premium factor was found to have a negative moderating effect on the relationship between their intention to buy an electric vehicle and environmental concern. Tu and Yang [17] in their study on Taiwanese consumers found that resource availability and opinions from consumers surroundings, along with their environmental awareness, influence consumer EV purchase intentions. Li et al. [18] in their systematic study of 1846 papers to understand the factors affecting EV purchase found that all factors can be categorized into demographic, situational, and psychological factors. In their study on electric vehicle usage, Sang and Bekhet [19] found that for Malaysian consumers, the electric vehicle acceptance was significantly related to demographics, financial benefits, performance attributes, environmental concerns, social influences, infrastructure availability, and government interventions.

Kim et al. [20] in their study in Korea examining consumer intentions for purchasing an electric vehicle found that prior experience in driving electric vehicles, along with factors like number of vehicles in the household, educational achievement, availability of parking, and perception of government incentives significantly affect consumers’ intentions for purchasing electric vehicles. Sierzchula et al. [21] in their study in 30 countries found that the electric vehicle market share in different countries was positively correlated with financial incentives, availability of charging infrastructure, and local production. It was further observed that availability of charging infrastructure had the strongest correlation with electric vehicle adoption. Verma et al. [9] in their study on identifying factors affecting electric vehicle
adoption in Bangalore, India, noted that the key motivators in electric vehicle adoption were perceived environmental benefits and financial incentives. Kumar et al. [8] studied challenges to the adoption of electric vehicles and concluded that sharing economy and public utilities will play a critical role in EV purchase, considering the high cost coupled with consumers’ low purchasing power and lack of availability of charging infrastructure in India. The study also recognized the role of government in terms of interventions at different levels to meaningfully enhance EV adoption in India. A 2020 study by Castrol in India noted that consumers generally have a positive attitude toward electric vehicles and estimated the Indian EV market would reach $2 billion by 2025. The study identified vehicle price, charging time, and driving range as the most important challenges in EV adoption in India. The average price point of $3100, charging time of 35 min, and vehicle driving range of 401 KM were identified as the tipping points to achieve mainstream EV adoption in India [22]. Higueras-Castillo et al. [23] in their study to find factors that affect electric vehicle purchase intentions in Spain conclude that driving range, financial incentives, and vehicle reliability are the most important predictors of the purchase intention. Bennett and Vijaygopal [24] found that a weak link exists between attitude and willingness to purchase an electric vehicle. Lin and Wu [25] examined the reasons for electric vehicle purchase by Chinese consumers and concluded that demographic characteristics such as gender, age, and marital status, along with attitude-related factors such as network externality, environmental awareness and concerns, price acceptability, government incentives, and vehicle performance have a significant impact on consumers’ willingness to purchase electric vehicles.

Therefore, it can be concluded that a majority of the earlier studies are context specific. The factors identified and studied also vary from one context to other. The studies do not converge in terms of identifying and listing factors affecting electric vehicle uptake in different contexts. There is no comprehensive list of the factors affecting electric vehicle uptake; thus, the present study aims to not only list the factors that are relevant in the current Indian context, but also develop a model to predict who will buy an electric vehicle in India.

1.3 Research Design

The present study was conducted in two stages. The first stage focused on identifying consumer interests and concerns related to electric vehicles. This stage is similar to opinion extraction or sentiment classification [26, 27], and involves gathering and analyzing individuals’ opinions about some issue, event, product, etc. [28, 29]. In the research context, opinion extraction can be understood as exploring and understanding issues that matter to Indian consumers in the electric vehicle context. Understanding public opinion helps in making better decisions. Presently, social media has become an important tool to express opinions on the issues that really matter to the masses [30, 31]. Content created and shared in the EV context on Twitter by Indian people during January and February 2021 was collected by using N-Capture, i.e., a web browser extension that allows quick access and capture of web and social media content based on the keywords of interest. As N-Capture accesses all the data
available on selected keywords, it can be treated like a census rather than a sampling study. To understand the opinions and concerns of Indian people for electric vehicles, a text analysis of the collected content was performed. Therefore, this stage was exploratory in nature with a focus on understanding Indian consumers’ opinions and concerns for electric vehicles. Similar approaches were also adopted in earlier studies [32–34]. The accessed text was analyzed with the help of SPSS and R. The outcome was a frequency count for the most often occurring words and word combinations. The findings can be presented either as a frequency distribution table or a word cloud. A word cloud presents high-frequency words in a visualization with different sizes. The size of different words indicates the frequency of its occurrence [35]. The word cloud provided much needed understanding about EV concerns of Indian consumers. Understanding consumer concerns and opinions relating to electric vehicles helped the researchers in designing a questionnaire for the second section of this paper.

Section 2 of the paper develops a predictive machine learning model to predict who will or will not buy EVs. This may be treated as descriptive research using survey methodology. A questionnaire was developed for the purpose of data collection. The questionnaire collected data on the parameters of concerns identified from the literature review and text analysis findings, and was tested for content/face validity by taking the opinion of 8 experts [36]. The experts were consistent in their opinions about the relevance of the parameters included in the questionnaire. Some changes related to the wording of statements were made. The changes suggested by the experts were incorporated into the questionnaire, which was sent to the respondents as a Google document to collect the data. The sampling methodology is convenience sampling, the reason for which is its efficiency in terms of time and money. In addition, due to electric vehicles being a subject of common concern, the sampling method provided a readily available sample in the given research context. A total of 245 respondents returned usable questionnaires. The collected data was analyzed with the help of SPSS and R to develop a predictive machine learning model.

2 Analysis and Findings

2.1 Understanding Stakeholders’ Concerns Using Text Analysis

Social media, e-newspapers, and review websites have now become an active tool for everyone to express their views and opinions on specific issues. Analyzing these helps in gaining insights and a general overview of public opinions or concerns. India is a highly populated country with a huge population falling under the group of “working-age population,” and, according to the Ministry of Statistics and Programme Implementation, the government of India is 15–64 years of age. The digital penetration in India has increased significantly over the last few years. As per the Telecom Regulator Authority in India on June 30th, 2019, the tele-density, i.e., the number of telephonic connections for every 100 individuals living in the area, in Indian urban areas was 160. Further availability of 4G and LTE services across most of India has made the digital space more vibrant than ever.
Social media has now become a platform for exchanging ideas, concerns, and opinions by the masses. Twitter is the most widely used platform for this purpose by all groups of people. According to Statista, the number of Twitter users in India was estimated to be 24.45 million in October 2021. Almost all major influencers such as corporate leaders, policymakers, policy advocates, journalists, and media houses have a Twitter account; therefore, the platform is one among many social media platforms one can use to gain an insight into what consumers are discussing. For this study, tweets were extracted, and text analysis was carried out to understand what online users are talking about regarding electric vehicles in India. The tweets were collected by using hashtags #EVIndia and #EVIssues. The extracted tweets were then analyzed for the most frequently occurring words in the dataset. The sole purpose of this was to understand what the most frequently discussed topics are related to electric vehicles. Then, a word cloud was formed for better visualizations, which also became the basis for the questionnaire that was later floated to make a predictive model that could classify whether a consumer will “Buy” or “Won’t Buy” an electric vehicle.

### 2.2 Hashtag: #EVIndia

The first hashtag we used to collect content was #EVIndia. To analyze the collected content, a frequency bar plot and word cloud were formed, which are presented in (Figs. 1, 2). It can clearly be seen that the most frequently used word was “electric,” followed by “charging,” “vehicle,” “battery,” and so on. It can be concluded from these high-frequency words that the Indian consumers are talking about electric vehicles. This simply means that electric vehicles have attracted the attention of Indian users and they are discussing and sharing their concerns. When it comes to concerns, the words “charging” and “battery” are used most frequently, indicating
that consumers are most concerned about battery-charging issues. The word “experience” is another high-frequency word, but surprisingly, they are not talking much about vehicle price, government incentives, or the maintenance costs.

### 2.3 Hashtag #EVIssues #EVIndia

These two hashtags together were chosen deliberately to capture public opinion and concerns related to EVs in India. The frequency bar plot and word cloud obtained are presented in (Figs. 3, 4). From the bar plot, it can be observed that “battery” and “products” are the words most often used. Other high-frequency words are “experience,” “solution,” “showcased,” “innovators,” and “exhibitors.” This also indicates that consumers are most concerned about the battery and related issues. Frequent use of the word “innovators” may indicate a discussion about the firms innovating in the domain of electric vehicles. They also discuss showcasing and exhibition of electrical vehicles and related technologies. In this case, other words such as “price,” “operating cost,” and “subsidies” did not turn up as words occurring with high frequency.

In summary, from the above text analysis, it may be concluded that Indian consumers are aware, concerned, and talking about electric vehicles and related issues. However, the amount and deepness of discussion relating to electric vehicles is quite limited. Consumers are most concerned about battery and battery-charging issues. Surprisingly, vehicle price and maintenance costs were not something they were concerned or talking much about.
2.4 Model to Predict Who Will Buy

A predictive machine learning model was developed to predict whether a consumer in India will or will not buy an electric vehicle depending on the input variables. There are many popular machine learning classification models that can predict with decent accuracy, but in this research paper, the logistic regression algorithm was used for forecasting. The advantages of using a logistic regression are that the normalization of data is not required, scaling of the data is not required, and missing data does not impact model building. However, one of the disadvantages of using
logistic regression models is that training analysts to use the model takes longer, which at times becomes a concern [37–39].

Surveys are commonly used to understand consumer behavior. A primary survey involves first-hand data collection by the researcher, and the data obtained is further analyzed to gain insights. To collect data for the model, a questionnaire was floated among the respondents. The variables chosen were based on the literature review and the outcomes obtained from the text analysis of the tweets. The variables used in the model are a combination of demographic, social, contextual, level of environmental awareness, and other relevant considerations of Indian consumers. The predictive machine learning model developed can classify whether an Indian consumer would “Buy” or “Won’t Buy” an electric vehicle. A total of 245 respondents replied to the google form questionnaire sent to them. The (Table 1) presents the data dictionary for the variables chosen for the questionnaire.

2.5 Sample Description

The survey questionnaire was floated among respondents and received a total of 245 usable responses. The sample description is presented in (Table 2). The respondents consisted of 140 males and 105 females. In terms of their marital status, 59% were married and 41% were unmarried. Regarding their age, the majority of respondents were less than 25 years of age representing 44% of the total and were closely followed by the 26- to 35-year age group representing 38%. People above these two age groups represented only 22%; thus, the majority consisted of younger generation respondents. The reason for this could likely be that the younger respondents are more aware and concerned about EVs in India. In terms of educational attainment, a majority were post-graduate and represented 44% of respondents. This was followed by graduates with 33% representation. People who had not graduated represented 22% of the respondents. In terms of income, the majority of the respondents were from below 5 Lakh income group representing 45%, followed by 5–8 Lakh income group representing 31%. This is in line with the large proportion of respondents who were from the student community or early on in their employment. In terms of employment, the majority of 63% was in service, followed by 24% in business, and 12% in the unemployed category. In terms of their geographical location, east, west, north, south, and central India were represented with 25, 24, 29, 12, and 10% respondents in each group, respectively. From the age distribution, it can be observed that the sample has a slight bias toward the younger population. This seems to be normal because it is these younger generation people who are typically more concerned about new technological developments, including electric vehicles. This group is also more active in voicing their concerns relating to a subject of interest on internet and social platforms. Therefore, the sample was a fair representation of all groups from all geographical regions in India in the context of the present study. Similarly, the results of present research may be generalizable over the larger Indian population.
Table 1  Data dictionary for variables taken in the survey

| Factor            | Variable                        | Options                                                                 |
|-------------------|---------------------------------|-------------------------------------------------------------------------|
| Demographic       | Gender                          | Male or female                                                          |
|                   | Educational qualification       | postgraduate and above, graduate, up to higher secondary                |
|                   | Marital status                  | Married or unmarried                                                    |
|                   | Annual household income         | 4 categories were defined: ‘below INR 5 lakh’, ‘5–8 lakh’, ‘8–12 lakh’ and ‘above 12 lakh’ |
|                   | Employment                      | Business, service, and not employed                                     |
|                   | Geographical location           | North, East, West, South and Central                                    |
| Social/ attitude  | Network externality/ mass behavior | People around you are buying EVs. (Rating 1–7) ≤ strongly disagree and ≥ strongly agree |
|                   | Price acceptability             | Currently electrical vehicle prices in Indian market are reasonable as compared to internal combustion engines. (Rating 1–7) 1–> strongly disagree and 7–> strongly agree |
|                   | Running cost                    | Operating/running cost of electrical vehicle is reasonable. (Rating 1–7) 1–> strongly disagree and 7–> strongly agree |
|                   | Government subsidy sufficiency  | Government subsidies/incentives to buy EVs are reasonable. (Rating 1–7) 1–> strongly disagree and 7–> strongly agree |
|                   | Charging infrastructure sufficiency | Currently the EV charging infrastructure in India is reasonable. (Rating 1–7) 1–> strongly disagree and 7–> strongly agree |
|                   | Level of environmental concern  | I am concerned about environmental degradation by the automobile industry. (Rating 1 –7) 1–> strongly disagree and 7- > strongly agree |
|                   | Driving range of EV             | The driving range of currently available EVs is reasonable. (Rating 1–7) 1–> strongly disagree and 7–> strongly Agree |
| Knowledge         | Vehicle performance             | The performance of currently available EVs is reasonable. (Rating 1–7) 1–> strongly disagree and 7–> strongly agree |
| Outcome           | Buying intention of EVs         | 1 = Will buy, 0 = Will not buy                                          |
Before proceeding to the logistic regression model building, correlation among the continuous variables were explored. No variable was found to have a strong correlation with other variables, which indicates that the issue of multicollinearity was not present. Multicollinearity is a statistical phenomenon in which predictor variables in a logistic regression model are highly correlated. The existence of collinearity inflates the variances of the parameter estimates, resulting in incorrect inferences about relationships between predictor and outcome variables [40].
2.6 Machine Learning Classification Model

A logistic regression machine learning model was developed on SPSS for the data collected and outputs have been summarized in (Tables 3, 4, 5, 6, 7, 8). Logistic regression is a classification technique that predicts whether something is true or false. The output of the logistic regression model is presented below. As previously mentioned, a total of 245 cases were included in the model. No instance of a missing case was reported. The dependent variable was coded as 1 to indicate an intention to buy, and 0 for no intention to buy an electrical vehicle. (Table 5) presents the omnibus tests of the model coefficients. The value given in the significance column is the probability of getting a Chi Square statistic (114.525) when the null hypothesis is true. The model is statistically significant as the *p* values reported are less than 0.05. The degree of freedom (df) column is an indication

| Table 3 | Case processing summary |
|---|---|
| Unweighted cases | N | Percent |
| Selected cases | | |
| Included in analysis | 245 | 100.0 |
| Missing cases | 0 | .0 |
| Total | 245 | 100.0 |
| Unselected cases | 0 | .0 |
| Total | 245 | 100.0 |

| Table 4 | Dependent variable encoding |
|---|---|
| Dependent variable encoding | | |
| Original value | Internal value |
| Not buy | 0 |
| Buy | 1 |

| Table 5 | Omnibus tests of model coefficients |
|---|---|
| Omnibus tests of model coefficients | | |
| | Chi-square | df | Sig |
| Step 1 | | | |
| Step | 114.525 | 13 | .000 |
| Block | 114.525 | 13 | .000 |
| Model | 114.525 | 13 | .000 |

| Table 6 | Model summary |
|---|---|
| | Step | −2 Log likelihood | Cox an snell R square | Nagelkerke R square |
| 1 | 208.734a | .373 | .510 |

*a*Estimation terminated at iteration number 6 because parameter estimates changed by less than .001
of the number of predictors in the model. For the current model, df is 13, which indicates that the model uses 13 predictors.

(Table 6) presents the model summary. The logistics regression does not have an $R^2$ to predict the variation in the outcome variable that can be explained by the model predictors as the case in OLS regression. In place of $R^2$, a large variety of pseudo-$R^2$ statistics has been developed. Two such statistics developed by Cox and Snell and Nagelkerke are presented in (Table 6). However, these statistics are not good equivalents to $R^2$ statistics on OLS. These should be treated as supplementary to other evaluative indices such as overall evaluation of the model, test of individual regression coefficients, and goodness-of-fit test statistics.

The classification (Table 7) is used to calculate model accuracy, precision, and recall. The “observed” column presents the number of “Buy” and “Won’t

### Table 7  Classification table

| Observed          | Predicted          | Percentage correct |
|-------------------|--------------------|--------------------|
| Buy or not buy    | Not buy            | 64                 |
| Not buy           | Buy                | 27                 |
|                  |                    | 70.3               |
| Overall percentage |                   | 81.2               |

*The cutoff value is .500

### Table 8  Variables in the equation

| Variable(s) entered on step 1: gender, age, education, income, Employment, level of environmental concern, vehicle cost, running cost, charging infrastructure, government subsidy, vehicle performance, the driving range of EVs is reasonable, people around are buying EVs
| B     | S.E  | Wald | df | Sig   | Exp(B) |
|-------|------|------|----|-------|--------|
| Gender | .981 | .448 | 4.794 | 1 | .029 | 2.666 |
| Age    | .723 | .293 | 6.109 | 1 | .013 | 2.062 |
| Education | .059 | .443 | .018 | 1 | .894 | 1.061 |
| Income | 1.612 | .314 | 26.394 | 1 | .000 | 5.015 |
| Employment | − .400 | .365 | 1.203 | 1 | .273 | .670 |
| Level of environmental concern | .564 | .220 | 6.589 | 1 | .010 | 1.757 |
| Vehicle cost | .095 | .198 | .232 | 1 | .030 | 1.100 |
| Running cost | − .913 | .232 | 15.443 | 1 | .000 | .401 |
| Charging infrastructure | .743 | .196 | 14.379 | 1 | .000 | 2.102 |
| Government subsidy | − .320 | .170 | 3.536 | 1 | .060 | .726 |
| Vehicle performance | 1.075 | .227 | 22.513 | 1 | .000 | 2.930 |
| Driving range of EVs | − .955 | .206 | 21.581 | 1 | .000 | .385 |
| Mass behavior | − .448 | .136 | 10.886 | 1 | .001 | .639 |
| Constant | − 3.631 | 1.903 | 3.642 | 1 | .056 | .026 |

aVariable(s) entered on step 1: gender, age, education, income, Employment, level of environmental concern, vehicle cost, running cost, charging infrastructure, government subsidy, vehicle performance, the driving range of EVs is reasonable, people around are buying EVs
Buy” as the dependent variables. The “predicted” column presents predicted values of the dependent variables based on the full logistic model. 64 cases are observed to be “Won’t Buy” and are correctly predicted to be “Won’t Buy;” 135 cases are observed to be “Buy” and are correctly predicted to be “Buy.” On the other hand, 27 cases are observed to be “Won’t Buy,” but are predicted to be “Buy;” 19 cases are observed to be “Buy” but are predicted to be “Not Buy.”

MODEL ACCURACY = (TN + TP) / (TN + TP + FP + FN) = 199/245 = 81.22%

PRECISION = TP / (TP + FP) = 135/162 = 83.33%

RECALL = TP / (TP + FN) = 135/154 = 87.66%

Table 8 presents the variables in the logistic regression equation. Different columns of this table present significant information to predict the outcome variable. Column B indicates log-odds units to predict dependent variables from independent variables. The prediction equation is:

\[
\log(\frac{p}{1-p}) = b_0 + b_1 \times x_1 + b_2 \times x_2 + b_3 \times x_3 + b_4 \times x_4 \ldots
\]

The estimates for different variables indicate a predicted increase or decrease in the predicted log-odds to buy an electric vehicle that would be indicated by one unit increase or decrease holding all other predictors constant. For all predictor variables that are not significant, the coefficients are not significantly different from 0. To identify coefficients that are significant values in the column labeled Wald are used. These columns provide the Wald chi-square value and two-tailed p value used in testing the null hypothesis that the coefficient (parameter) is 0. Each p value is compared with the preselected value of alpha to determine whether it is a statistically significant predictor. Coefficients having p values less than alpha are statistically significant. To predict the purchase of an electric vehicle, the authors have chosen alpha to be 0.05, which means that in all cases where alpha is reported to be less than 0.05, the null hypothesis can be rejected, and it can be concluded that the coefficient is significantly different from 0.

From the Sig, it can be observed that age, gender, income, level of environmental concerns, vehicle cost, running cost, vehicle performance, driving range, and mass behavior are significant predictors of electric vehicle purchase. On the other hand, level of education, employment, and government subsidy were not found to be significant predictors of e-vehicle purchase or uptake. The constant of the model was also reported to not be a statistically significant predictor of electrical vehicle purchase. Exp(B) are the odds ratios for the predictor e-vehicle “Buy” or “Won’t Buy.” These are exponentials of the coefficients. The odds ratio indicates the change in the odds of the outcome variable given a unit change in any predictor variable.
2.7 Conclusions and Suggestions

After carrying out the literature review, it can be concluded that the electric vehicle industry in different regions and countries is at different levels of evolution. Many countries are making serious efforts in the development of electric vehicles and related infrastructure. Governments are also providing right policy, environment, and financial support to increase electric vehicle uptake. A wide variety of factors has been studied to influence the uptake of electric vehicles in different contexts. From the text analysis, it can be concluded that people are talking about electric vehicles in India, but the issues they are discussing on the internet and social media platforms are not very serious or deep. The reason for this lower interest is likely that the availability of electric vehicles is still limited for common consumers in India. The most used words in Indian electric vehicle context are “battery,” “charging,” and “experience.” Government subsidies or incentives for increasing electric vehicle uptake were not significantly discussed. It can be concluded from the text analysis that charging stations and batteries continue to be the most discussed words in the EV context on social media.

A Logistic Regression model with 81.22% accuracy was developed, which classifies consumers into either the “Buy” or “Won’t Buy” category. The model included age, gender, income, level of environmental concerns, vehicle cost, running cost, vehicle performance, driving range, and mass behavior as significant predictors to classify a consumer in either category. However, level of education, employment, and government subsidy were not found to be significant predictors in e-vehicle purchase or uptake. The findings of the model can be used by electric vehicle marketers to enhance design, delivery, and marketing for better uptake of electric vehicles in India’s personal passenger vehicle segment.

2.8 Scope for Future Research

For the present study, the text analysis carried out was based on random tweets obtained from N-Capture software and thus future research may explore and capture conversations from a greater number of social media platforms. The wider content is expected to provide a better understanding of people’s opinions and concerns relating to electric vehicles. Similarly, a predictive model may be developed with better indicators that may also evolve with time. Finally, for data collection, a wider sample covering respondents with a wide demographic and regional representation may be surveyed to obtain results with better generalization potential.

Declarations

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.
References

1. Abell, D. F. (1980). *Defining the business: the starting point of strategic planning*. Englewood Cliffs.
2. Nightingale, J. (1978). On the definition of ‘industry’ and ‘market’. *The Journal of Industrial Economics.*, 27(1), 31–40. https://doi.org/10.2307/2098116
3. Porter, M. E. (1979). How competitive forces shape strategy. *Harvard Business Review.*, 57, 137–145.
4. Wadegaonkar, D. W. (1981). Concept of industry. *Journal Indian Law Institute.*, 23(3), 400–420.
5. Juyal, S., & Chander, S. (2018) Zero Emission Vehicles (ZEVs): Towards a Policy Framework. New Delhi: NITI Aayog and World Energy Council
6. KPMG. (2020) Shifting gears: the evolving electric vehicle landscape in India. New Delhi
7. Kumar, R., Jha, A., Damodaran, A., Bangwal, D., & Dwivedi, A. (2020). Addressing the challenges to electric vehicle adoption via sharing economy: An Indian perspective. *Management of Environmental Quality: An International Journal.*, 32(1), 82–99. https://doi.org/10.1108/MEQ-03-2020-0058
8. Verma, M., Verma, A., & Khan, M. (2020). Factors Influencing the adoption of electric vehicles in Bengaluru. *Transportation Developing Economies.*, 6(2), 17. https://doi.org/10.1007/s40890-020-0100-x
9. Shepherd, S., Bonsall, P., & Harrison, G. (2012). Factors affecting future demand for electric vehicles: A model based study. *Transport Policy.*, 20, 62–74. https://doi.org/10.1016/j.trpol.2011.12.006
10. Christidis, P., & Focas, C. (2019). Factors affecting the uptake of hybrid and electric vehicles in the European Union. *Energies*, 12(18), 3414.
11. Tu, J. C., & Yang, C. (2019). Key factors influencing consumers’ purchase of electric vehicles. *Sustainability*, 11(14), 3863. https://doi.org/10.3390/su11143863
12. Sierzchula, W., Bakker, S., Maat, K., & van Wee, B. (2014). The influence of financial incentives and other socio-economic factors on electric vehicle adoption. *Energy Policy*, 68, 183–194. https://doi.org/10.1016/j.enpol.2014.01.043
13. CASTROL (2020): Electric vehicle adoption in India: Study reveals three tipping points. Retrieved January 2, 2020, from https://timesofindia.indiatimes.com/business/india-business/electric-vehicle-adoption-in-india-study-reveals-three-tipping-points/articleshow/77917482.cms
23. Higueras-Castillo, E., Guillén, A., Herrera, L.-J., & Liébana-Cabanillas, F. (2020). Adoption of electric vehicles: Which factors are really important? *International Journal Sustainable Transportation*. https://doi.org/10.1080/15568318.2020.1818330

24. Bennett, R., & Vijaygopal, R. (2018). Consumer attitudes towards electric vehicles: Effects of product user stereotypes and self-image congruence. *European Journal Marketing.*, 52(3/4), 499–527. https://doi.org/10.1108/EJM-09-2016-0538

25. Lin, B., & Wu, W. (2018). Why people want to buy electric vehicle: An empirical study in first-tier cities of China. *Energy Policy.*, 112, 233–241. https://doi.org/10.1016/j.enpol.2017.10.026

26. Abbasi, A., Chen, H., & Salem, A. (2008). Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums. *ACM Transactions Information Systems (TOIS)*., 26(3), 1–34.

27. Pang, B., & Lee, L. (2009). *Sentimental extraction and opinion analysis*. Springer Press.

28. Rout, J. K., Choo, K.-K.R., Dash, A. K., Bakshi, S., Jena, S. K., & Williams, K. L. (2018). A model for sentiment and emotion analysis of unstructured social media text. *Electronic Commerce Research, 18*(1), 181–199. https://doi.org/10.1007/s10663-017-9257-8

29. Abirami, A., & Gayathri, V. (2017) A survey on sentiment analysis methods and approach. 2016 Eighth International Conference on Advanced Computing (ICoAC): IEEE 72–6

30. Bradley, P. (2010). Be where the conversations are: The critical importance of social media. *Business Information Review.*, 27(4), 248–252. https://doi.org/10.1177/026638210390976

31. Nielsen, R. K., & Schrøder, K. C. (2014). The relative importance of social media for accessing, finding, and engaging with news. *Digital Journalism.*, 2(4), 472–489. https://doi.org/10.1080/21670811.2013.872420

32. Capelillo, N. D., & Kurani, K. S. (2012). Households’ stories of their encounters with a plug-in hybrid electric vehicle. *Environment and behavior.*, 44(4), 493–508. https://doi.org/10.1177/0013916511402057

33. Habich-Sobiegalla, S., Kostka, G., & Anzinger, N. (2018). Electric vehicle purchase intentions of Chinese, Russian and Brazilian citizens: An international comparative study. *Journal Cleaner Production.*, 205, 188–200. https://doi.org/10.1016/j.jclepro.2018.08.318

34. Junquera, B., Moreno, B., & Álvarez, R. (2016). Analyzing consumer attitudes towards electric vehicle purchasing intentions in Spain: Technological limitations and vehicle confidence. *Technological Forecasting and Social Change.*, 109, 6–14. https://doi.org/10.1016/j.tfs.2016.05.006

35. Dyson, P. (2012). Slum Tourism: Representing and Interpreting ‘Reality’ in Dharavi. *Mumbai Tourism Geographies.*, 14(2), 254–274. https://doi.org/10.1080/14616688.2011.609900

36. Hardesty, D. M., & Bearden, W. O. (2004). The use of expert judges in scale development: Implications for improving face validity of measures of unobservable constructs. *Journal Business Research.*, 57(2), 98–107.

37. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley and Sons.

38. Nunan, D., Birks, D. F., & Malhotra, N. K. (2020). *Marketing research: an applied approach*. London: Pearson.

39. Peng, C.-Y.J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal Educational Research.*, 96(1), 3–14. https://doi.org/10.1080/00220670209598786

40. Midi, H., Sarkar, S. K., & Rana, S. (2010). Collinearity diagnostics of binary logistic regression model. *Journal Interdisciplinary Mathematics.*, 13(3), 253–267. https://doi.org/10.1080/09720502.2010.10700699

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.