Predicting Electric Vehicle Adoption in India: A Machine Learning Approach and Text Analysis of Consumer Sentiment for OLA Electric Vehicles

Nabanita Das

Department of Information and Technology, Raghunath Jew Junior College, Cuttack, Odisha, India
Email Id: das.navnita@gmail.com

Amaresh Sahu

Master in Computer Application, Ajay Binay Institute of Technology, Cuttack, Odisha, India
Email Id: amaresh_sahu@yahoo.com

Hemanta Kumar Bhuyan

Department of Information Technology & Computer Applications,
Vignan's Foundation for Science, Technology & Research (Deemed to be University) Guntur, Andhra Pradesh, India
Email Id: hmb.bhuyan@gmail.com

Anurag Nanda

Department of Information Technology and Management, Ajay Binay Degree College, Cuttack Odisha, India
Email Id: anuragnada001@gmail.com

Abstract: *The globe is currently extremely concerned about the impending threat of climate change, which is being exacerbated by rising global temperatures. Regrettably, India's greenhouse gas emissions are rising substantially, placing it in the top ten global emitters. Air pollution plays a crucial part in this environmental issue, acting as a significant accelerator for the greenhouse effect. Specifically, 10% of India's air pollution is caused by the transportation industry. Recognising the gravity of the problem, the Indian government has moved decisively to reduce air pollution, putting particular focus on promoting the use of electric vehicles. Nevertheless, customer attitudes, views, and knowledge about electric vehicles determine how successful these programmes are. In a diverse market, finding the proper customers is a challenging issue for marketers in the electric vehicle space. Recent research endeavours to develop a machine learning model that can forecast sales of electric vehicles (EVs) in India, taking into account the country's dynamic topography. In order to construct the model, OLA conducted a thorough investigation of the EV landscape, which included a textual analysis of social media comments pertaining to EVs. The primary objective is to identify frequently used terms that reveal important details about consumer concerns and interests for OLA electric cars. In this research with the help of web scraping and natural language processing we found the sentiments of EV users. Also linear regression was used for determining the EV future demands.*

Keywords: *Linear Regression, Sentiment Analysis using NLTK, Natural Language processing (NLP) and Web scrapping*

Introduction

Within the domain of business and economic literature, the term "industry" takes on various meanings contingent upon the context and objectives outlined by Abell, D. F. (1980), Nightingale, J. (1978), Porter, M. E. (1979), and Wadegaonkar, D. W. (1981)[21]. In a broader sense, "industry" denotes a compilation of business activities with a subtle domain specificity. In line with this perspective, authors delineate the "electric vehicles industry" as a subset of the automobile industry, encompassing all entities involved in the production, trade, and servicing of vehicles powered by electrical energy, along with associated and ancillary businesses. Electric vehicles stand out as pivotal contributors to addressing contemporary societal challenges such as energy security, environmental degradation, and urban mobility.

Many industries, by their inherent nature, undergo evolutionary changes over time and space, with emerging sectors displacing older ones due to technological advancements and shifts in consumer preferences. In recent years, the EV industry has undergone substantial transformations in diverse contexts. In the realm of electric vehicles (EVs), government influence plays a decisive role in shaping public perception and fostering mass adoption. Governments across the globe are formulating electric vehicle policies with a focus on diminishing reliance on fossil fuels, addressing environmental concerns, managing challenges posed by rapid urbanization, and augmenting employment opportunities. Numerous countries have implemented policy interventions to support research and development, infrastructure development, and provide financial incentives to both the electric vehicle industry and consumers.

In the Indian context, the government's commitment to electric vehicles has witnessed a surge in recent years, as discussed by Juyal, S., & Chander, S. (2018), and KPMG. (2020) in their work titled "Shifting Gears." A notable development is the government's ambitious target of achieving EV sales comprising 30% of private cars, 70% of commercial vehicles, and 80% of two and three-wheelers by 2030, driven by the imperative to decarbonize the transport sector (Khan, S., 2021)[22]. Both central and state governments have implemented policy measures to incentivize the manufacturing and adoption of EVs, with 15 state governments already announcing EV policies. Key components of the Indian government's EV policy include making electric vehicles economically viable, developing charging/swapping infrastructure, advancing technology, and prioritizing small and public vehicles for an immediate impact. EVs are also perceived as contributors to economic development and employment in India, with several automobile manufacturers recently launching EV models in both two-wheeler and four-wheeler segments.

This research paper is divided into two distinct sections. The first section delves into the text analysis of Indian consumers' social media posts, aiming to comprehend user concerns regarding electric vehicles. The integration of social media text analysis with a comprehensive review of existing literature provides valuable insights into the context of the Indian electric vehicle industry and potential factors influencing EV adoption. The second section introduces a linear regression machine learning model designed to predict OLA vehicle sales based on feedback obtained from social media, offering a quantitative perspective on consumer sentiments and market trends.

This research paper has integrated 3 separate techniques as web scrapping, NLP, machine learning. As per the paragraph above, the 1st section consists of NLP using VADER sentiment analyser of NLTK library with web scrapping which will collect review of every people who have shared their comments in that social media site. Those comments will be evaluated as per the VADER sentiment analyser of NLTK library.

In the 2nd section the machine learning part comes into action. After analysing the sentiments of every comment, prediction of OLA EV is done using linear regression algorithm, which predicts the future sales of OLA EV by previous year's data.

Literature Review

In the prior discussions, a substantial body of research on consumer sentiments and determinants related to electric vehicles has primarily focused on Western and Chinese contexts. Shepherd Et Al crafted a system dynamic model that incorporated reasons such as subsidies, milage, and the availability of charging infrastructure[1]. Their findings suggested that subsidies had limited impact, particularly in dependent marketing scenarios. Cofman et al. noted significant performance improvements; however, government targets for electric vehicle (EV) adoption remained elusive. Divergent evidence surfaced concerning the influence of government incentives on EV adoption, with the availability of public charging infrastructure emerging as a noteworthy factor[2].

Christidis and Focas pinpointed income, educational attainment, and urbanization level as pivotal factors influencing EV adoption within the European Union[3]. They underscored the impact of local conditions and regional variations on EV purchasing decisions[4]. Kim et al. conducted a study across 31 countries, drawing correlations between the market share of electric vehicles and factors such as relative pricing, available models, and driving range[5]. Soltani-Sobh et al. directed their focus to the USA, revealing that electricity prices, urban road usage, and government incentives played pivotal roles in shaping EV adoption[6].

Wang and his research team delved into an exploration of the factors influencing the public's acceptance of electric vehicles (EVs) in Shanghai, with a dedicated focus on technological dimensions, marketing strategies, perceived risks, and environmental consciousness[7]. Their research brought to light the intricate dynamics at play among these factors, providing valuable insights into how they collectively shape the attitudes and preferences of the public towards electric vehicles. Simultaneously, Thananusak and his research team conducted an extensive study in Thailand, emphasizing the supremacy of performance-related factors over considerations such as the access of charging infrastructure and financial aspects[8]. The study underscored the pivotal role of performance attributes in steering the consumer's decision-making process concerning electric vehicles. Significantly, both studies converged on the central significance of environmental concerns as a driving force behind the public's willingness to incur a premium for electric vehicles. The positive environmental impact associated with EVs emerged as a compelling factor, inspiring individuals to adopt these vehicles despite potential cost implications.[9]

Tu and Yang's examination of Taiwanese consumers unveiled the considerable impact of factors like resource availability, social opinions, and environmental awareness on their inclination to adopt electric vehicles (EVs). In a parallel investigation[10], Li et al. conducted a thorough analysis, categorizing the determinants of EV purchases into demographic, situational, and psychological dimensions. Simultaneously, Sang and Bekhet identified critical factors influencing Malaysian consumers, including demographics, financial benefits, performance attributes, environmental concerns, social influences, infrastructure availability, and government interventions, all significantly contributing to the acceptance of EVs[11].

Kim et al. explored consumer intentions to acquire EVs in Korea, pinpointing key determinants such as prior EV driving experience, household vehicle count, educational attainment, parking availability, and the perception of government incentives[12]. Sierzchula et al. broadened their research to encompass 30 countries, establishing positive correlations between EV market share and financial incentives, charging infrastructure, and local production.[13] In the Indian context, Verma et al. delved into factors impacting EV adoption in Bangalore, emphasizing perceived environmental benefits and financial incentives as pivotal drivers[14].Kumar et al. identified challenges to EV adoption in India, emphasizing the significance of the sharing economy and public utilities due to high costs and low purchasing power[15]. Government interventions

were underscored as essential for meaningfully enhancing EV adoption[16]. A 2020 study by Castrol in India projected a positive attitude toward EVs, estimating a $2 billion EV market by 2025, with challenges identified including price of vehicle, time to charge & range of driving. Shifting focus to Spain, Higueras-Castillo et al[17]. emphasized range of driving, financial advantage and reliability of vehicle as key predictors of EV purchase intentions[18]. Bennett and Vijaygopal identified a relation between attitude and willingness to purchase an EV[19]. Lin and Wu scrutinized factors influencing EV purchases in China, highlighting the significance of demographic characteristics, attitude-related factors, and vehicle performance[20].

Methodology and model specifications

The results of performance of OLA electric vehicle have been evaluated by applying the methods:

I. Linear Regression

II. Sentiment Analysis using NLTK

III. Natural Language processing (NLP)

IV. Web scrapping

We have adopted following steps for detecting the sentiments of OLA users:

a) Step-1

First, we collect people's reviews as data from various websites regarding OLA Electric using web scrapping.

b) Step-2

Data collection is done using Web scrapping as we scrape the data embedded in <p> in the HTML file of the website using the Beautiful Soup class of bs4 library.

c) Step-3

Then we analyze the sentiment score of each review using NLP and NLTK library. The positive sentiment is the sentiment whose score is between 0 to 1, the negative sentiment is the sentiment that has a score in between 0 to -1 and the neutral sentiment is the sentiment that has a score 0. Then we plot the sentiment intensity of positive, negative, and neutral using Matplotlib. After that, we calculate the number of positive, negative, and neutral sentiments and plot them in a bar graph using Matplotlib.

d) Step-4

At last, we predict the sales of OLA electric vehicles using Linear regression by giving it the last 5 years of sales of OLA electric vehicles.

Details of the Methodologies

Linear regression

It is one of the machine algorithms used for prediction-related works. It is of 2 types as Linear regression using single variable and Linear regression using multiple variables.

Linear regression using single variable

Linear regression is a ML algorithm that calculates a linear relationship between dependent variable and independent characteristics. When there is only one independent characteristic, it is called simple linear regression. more features, this is multiple linear regression. Linear regression helps predict continuous output variables based on input characteristics. Its simplicity and interpretability make it the basis for more complex algorithms. It is not only predictive but also the basis of advanced models.

The equation of this type of linear regression is : $y = mx + C$……..(1)

Here, y = dependent variable

x = independent variable

m = slope of the equation

C = y-intercept

Linear regression using multiple variable

Multiple linear regression, a powerful ML algorithm that extends linear regression of single var to model the relationship between one dependent continuous vars and multiple independent variables. Let's dive into the details:

Model representation:

Multiple linear regressions tries to find the linear equation that best fits the observed data points.

The equation is: $y = m_1x_1 + m_2x_2 + \ldots + m_nx_n + C$ ……(2)

here:

(Y) represents dependent variable (output).

$(x_1, x_2, \ldots, x_n)$ are independent variables (traits).

$(m_1, m_2, m_3, \ldots, m_n)$ are the model coefficients.

C is the y-intercept

Hyperplane:

Unlike simple linear regression (which deals with a single feature), multivariate regression works in a multidimensional space.The equation represents the hyperplane, which is essentially a line of best fit with three or more dimensions.

The regressor tries different equations to find the hyperplane that best represents the training data.

In this case, I have used linear regression using single variable which is also known as time – series analysis.

Sentiment Analysis Using NLTK

Sentiment analysis is an exciting area of natural language processing (NLP) that aims to check the emotional tone or emotion of the text.

Sentiment analysis involves classifying text into positive, negative, or neutral categories. It is commonly used for social media analysis, posts, reviews, comments and other text data.

NLTK provides a built-in sentiment analyzer, VADER (Valence Aware Dictionary and Sentiment Reasoner).

VADER is vocabulary-based and works well on social media text, with Vader, you can quickly analyze sentiment.

Web Scrapping

Beautiful Soup is a Python library for parsing HTML and XML documents.

Accessing HTML content:

Send an HTTP request to the web page you want to capture using the request library.

Get the raw HTML content from the server response.

Creating a parse tree:

Create a parse tree (soup object) from the HTML content using Beautiful Soup.

A parse tree represents the nested structure of HTML.

Data extraction:

Trace the parse tree to find relevant elements (tags) using methods such as find, find all and attribute-based searches. Extract data such as links, text or specific attributes.

Empirical results

Section 1 delves into the textual exploration of the opinion of Indian consumers expressed on various social media, aiming to discern prevalent sentiments &concerns related to electric vehicles. By combining the findings from social media text analysis with a comprehensive review of existing literature, this section endeavors to offer a nuanced understanding of the Indian electric vehicle landscape. Additionally, it seeks to identify potential factors influencing the adoption of electric vehicles in the Indian market. Figure 1. Depicts the sentiment analysis of OLA vehicle customers through bar graph. Figure 2. Shows sentiment score through heatmap.

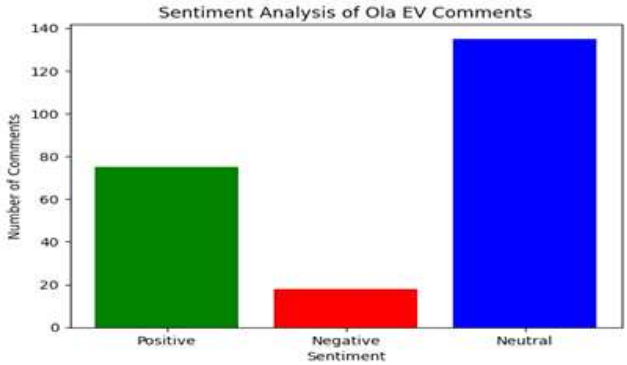

Figure 1: Sentiment analysis of OLA EV comments

Source: Author

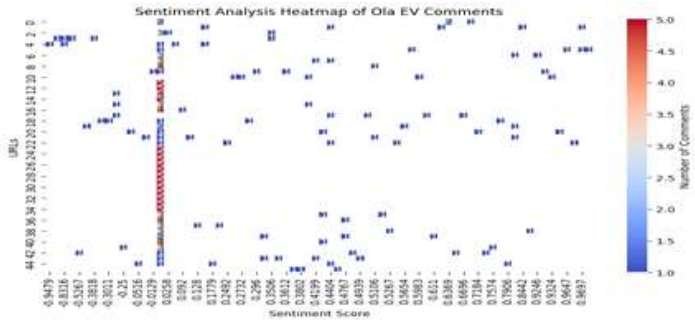

Figure 2: Sentiment analysis Heatmap of OLA EV comments

Source: Primary data

Section 2 introduces a machine learning framework centered around linear regression for forecasting the sales performance of OLA vehicles. The model leverages insights extracted from social media feedback, contributing to a data-driven approach in predicting sales outcomes for OLA electric vehicles. Figure 3. shows the confusion matrix for classification of customer sentiments and Figure 4. depicts the sales prediction of OLA vehicle in 2023.

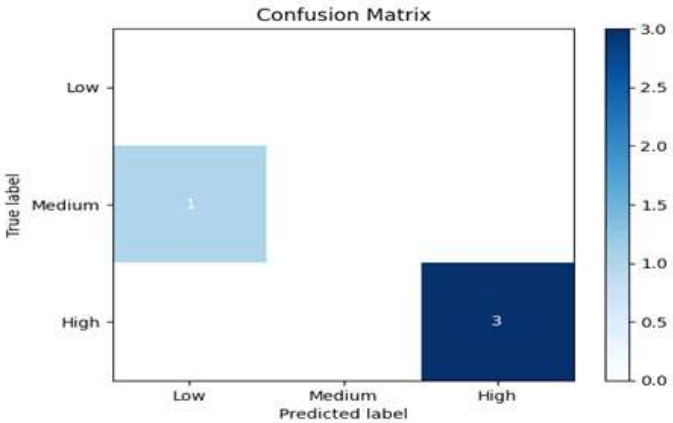

Figure 3: Confusion Matrix

Source: Primary data

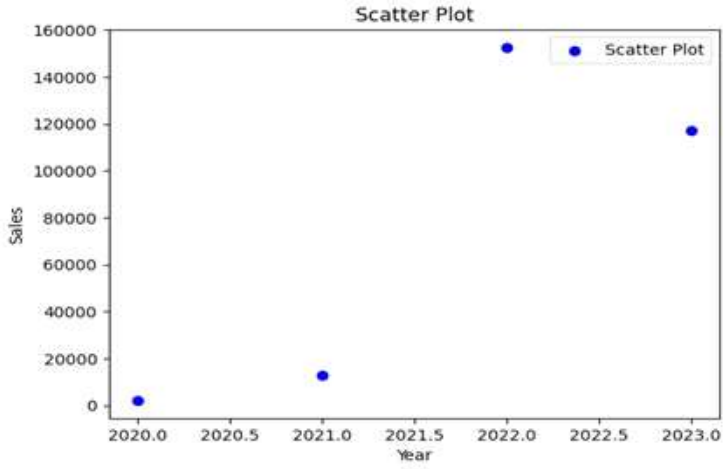

Figure 4: Scattor Plot

Source: Primary Data

Future Scope

Advanced Predictive Models: Further development of machine learning algorithms could lead to more accurate predictive models for electric vehicle (EV) adoption in India. Researchers can explore advanced techniques such as deep learning, ensemble methods, and reinforcement learning to improve forecast accuracy. The integration of big data analytics and IoT sensors can provide real-time insights into consumer behavior leading to more accurate predictions.

Incorporation of Socio-Economic Factors: Future research can delve into the socioeconomic factors that influence EV adoption worldwide. India. Factors such as income level, government policies, infrastructure development, and cultural perceptions significantly influence consumer behavior towards electric cars. Incorporating these factors into predictive models can provide more insight into adoption patterns.

Geospatial Analysis: Geospatial analysis can be used to identify regional variations in EV adoption across states and cities in India. By analyzing spatial data such as population density, income distribution, availability of charging infrastructure and environmental policies, researchers can identify hotspots for electric vehicles and design targeted strategies to promote adoption in specific areas.

Longitudinal studies: Conducting longitudinal studies can. to give insights. On long-term trends and dynamics of EV in India. By tracking consumer opinions and adoption patterns over time, researchers can identify emerging trends, barriers and opportunities for growth. Longitudinal studies can also help evaluate the effectiveness of government policies and marketing campaigns to promote EV adoption.

Integration of alternative data sources: In addition to textual analysis of consumer opinions, future studies can examine alternative data sources, such as social media activities, online forums, and customer reviews, to gain a vast understanding of consumer choices &understanding of OLA electric vehicles. NLP techniques can be used to extract valuable insights from non structured data sources that improve the predictive ability of models.

Collaboration with industry stakeholders: Collaboration with industry stakeholders, including OLA Electric and other electric vehicles, can facilitate proprietary components. knowledge and domain knowledge. Collaborative projects can help researchers develop customized predictive models and actionable insights that

address the specific needs and challenges of the Indian EV industry.

Policy Recommendations: Research findings can be used to formulate policy recommendations to accelerate deployment. of electric cars in India. Insights from predictive models and consumer attitude analysis can help policymakers design effective incentives, infrastructure investments, and regulatory frameworks to overcome barriers to adoption of sustainable mobility solutions and promote their use.

Impact evaluation: Future research can focus on environmental evaluation , economic and social impacts of EV adoption in India. By quantifying the benefits of the transition to electric vehicles, researchers can provide stakeholders with valuable information to support investment decisions and policy-making aimed at achieving long-term sustainability goals.

Cross-cultural analysis: cross-country and regional studies can provide valuable information on the factors that influence electric vehicle adoption vehicles and consumer preferences. By analyzing cross-cultural differences in attitudes, perceptions and adoption, researchers can identify both universal trends and region-specific factors that influence the use of electric cars. It informs global strategies to promote sustainable transport.

Interdisciplinary research: collaboration between researchers from different disciplines such as engineering, economics, sociology and environmental sciences can promote interdisciplinary research methods to address the complex challenges of mainstreaming. of electric vehicles. Combining knowledge from multiple disciplines can lead to more comprehensive understanding and innovative solutions to promote sustainability in India and beyond.

Using this future research, researchers can contribute to knowledge development and support efforts to accelerate electric cars. in India thereby reducing environmental pollution, reducing dependence on fossil fuels and promoting sustainable economic development.

Conclusion

A multitude of prior investigations exhibit context-specific findings, wherein identified factors fluctuate across diverse contexts. A notable absence of consensus prevails regarding a comprehensive compilation of factors that significantly impact electric vehicle (EV) adoption. In light of these gaps, this study strives not only to enumerate pertinent factors within the present Indian context but also endeavors to construct a predictive model aimed at discerning potential EV adopters in India.

References

1. Abell, D. F. (1980). Defining the business: the starting point of strategic planning. Englewood Cliffs. https://www.scirp.org/reference/referencespapers?referenceid=1723060
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.https://doi.org/10.1007/978-1-349-20317-8_10
4. Wadegaonkar, D. W. (1981). Concept of industry. Journal Indian Law Institute, 23(3), 400–420.https://www.jstor.org/stable/43950760?ab_segments=&searchKey=
5. Juyal, S., & Chander, S. (2018). Zero Emission Vehicles (ZEVs): Towards a Policy Framework. New Delhi: NITI Aayog and World Energy Council. https://e-amrit.niti.gov.in/assets/admin/dist/img/new-fronend-img/report-pdf/EV_report.pdf
6. KPMG. (2020). Shifting gears: the evolving electric vehicle landscape in India. New Delhi.https://assets.kpmg.com/content/dam/kpmg/in/pdf/2020/10/electric-vehicle-mobility-ev-adoption.pdf
7. 7Shepherd, 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.tranpol.2011.12.006

8. Cofman, M., Bernstein, P., & Wee, S. (2017). Electric vehicles revisited: A review of factors that affect adoption. Transport Reviews, 37(1), 79–93.

https://doi.org/10.1080/01441647.2016.1217282

9. Christidis, P., & Focas, C. (2019). Factors affecting the uptake of hybrid and electric vehicles in the European Union. Energies, 12(18), 3414. http:/doi.org//10.3390/en12183414

10. Kim, J. H., Lee, G., Park, J. Y., Hong, J., & Park, J. (2019). Consumer intentions to purchase battery electric vehicles in Korea. Energy Policy, 132, 736–743. https://doi.org/10.1016/j.enpol.2019.06.028

11 Soltani-Sobh, A., Heaslip, K., Stevanovic, A., Bosworth, R., &Radivojevic, D. (2017). Analysis of the electric vehicles adoption over the United States. Transportation Research Procedia, 203–212.https://doi.org/10.1016/j.trpro.2017.03.027

12. Wang, N., Tang, L., & Pan, H. (2018). Analysis of public acceptance of electric vehicles: An empirical study in Shanghai. Technological Forecasting and Social Change, 126, 284–291. https://doi.org/10.1016/j.techfore.2017.09.011

13. Thananusak, T., Rakthin, S., Tavewatanaphan, T., &Punnakitikashem, P. (2017). Factors affecting the intention to buy electric vehicles: Empirical evidence from Thailand. Journal of Electric and Hybrid Vehicles, 9(4), 361–381. https://doi.org/10.1504/ijehv.2017.089875

14. Tu, J. C., & Yang, C. (2019). Key factors influencing consumers' purchase of electric vehicles. Sustainability, 11(14), 3863. https://doi.org/10.3390/su11143863

15. Li, W., Long, R., Chen, H., & Geng, J. (2017). A review of factors influencing consumer intentions to adopt battery electric vehicles. Renewable and Sustainable Energy Reviews, 78, 318–328. https://doi.org/10.1016/j.rser.2017.04.076

16. Sang, Y.-N., &Bekhet, H. A. (2015). Modelling electric vehicle usage intentions: An empirical study in Malaysia. Journal of Cleaner Production, 92, 75–83. https://doi.org/10.1016/j.jclepro.2014.12.045

17. Kim, S., Lee, J., & Lee, C. (2017). Does driving range of electric vehicles influence electric vehicle adoption? Sustainability, 9(10), 1783. https://doi.org/10.3390/su9101783

18. 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

19. 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

20. Higueras-Castillo, E., Guillén, A., Herrera, L.-J., & Liébana-Cabanillas, F. (2020). Adoption of electric vehicles: Which factors are really important? International Journal of Sustainable Transportation. https://doi.org/10.1080/15568318.2020.1818330

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

22. 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