Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Predicting the Adoption of Clean Energy Vehicles: A Machine Learning-Based Market Analysis

Miraz Hossain[1], Md Masud Karim Rabbi[2], Nipa Akter[3], Nur Nahar Rimi[4], Mohammad Hamid Hasan Amjad[5], Bivash Ranjan Chowdhury[6], Mehedi Hasan Ridoy[7], Shah Alam[8], Md Shihab Sadik Shovon[9]

Abstract

Switching towards clean energy vehicles (CEVs) is a key measure in curbing greenhouse gas emissions and fighting climate change in the USA. Yet, despite mounting environmental consciousness and policy stimulus, the uptake of CEVs is still quite low. The main aim of this research is the creation of a market analysis framework based on machine learning for the prediction of CEV adoption. Utilizing supervised learning algorithms—Random Forest, Logistic Regression, and Decision Tree—the research compares their performance in segmenting prospective CEV adopters in terms of infrastructural, environmental, and socio-economic variables. The dataset included an extensive list of variables designed to capture the various factors that drive clean energy vehicle (CEV) adoption. It includes demographic variables like age, income, educational level, and geographical region, as well as economic variables like vehicle price, purchase incentives, and cost of ownership. In addition, it covers environmental attitudes, captured in terms of questionnaire responses on climate change concerns as well as sustainability values. We initiated this research using a range of machine learning models for the prediction of clean energy vehicle adoption, each of which was used for its particular strengths. To assess the performance of our predictive models, we utilized an extensive range of evaluation metrics: Accuracy, Precision, Recall, F1 Score, and ROC-AUC. Perfect scores on all metrics were recorded for the Decision Tree model, with 100% accuracy, precision, recall, and F1-score. Meanwhile, slightly lower overall performance values were reported for both Logistic Regression and Random Forest models. Sophisticated CEV adoption models' granular outputs can be directly applied in designing and implementing clean vehicle incentive structures at local, state, and federal levels. Knowing the particular socioeconomic, demographic, and geospatial drivers or impediments of adoption in specific regions allows policymakers to craft optimally effective incentive structures. Sophisticated insights derived from patterns of CEV adoption provide irreplaceable value for automotive companies and clean technology firms working in the US market. Future demand for CEVs is important for successful infrastructure planning, especially for the siting of electric charging stations. Monitoring CEV adoption rates is critical for measuring progress towards emissions reduction targets and facilitating broader sustainability planning activities.

Keywords: *Clean Energy Vehicles, Market Adoption, Machine Learning, Logistic Regression, Random Forest, Decision Tree, EV Market, Sustainability.*

I. Introduction

Background

Shovon et al. (2025) found that transportation is among the largest contributors of global emissions of greenhouse gases in the USA, representing about 24% of global CO_2 emissions. In response to increasing climate pressures, clean energy vehicles (CEVs), such as battery electric vehicles (BEVs), plug-in hybrid electric vehicles (PHEVs), and hydrogen fuel cell vehicles (HFCVs), are being promoted as cleaner options as substitutes for conventional internal combustion engine (ICE) vehicles (Anonna et al, 2023; Ahmed et al., 2025). Governments across the globe are promoting CEVs through policy interventions such as taxes, subsidies, and emission norms regulations. However, market penetration of CEVs is less than expected, leading one to wonder what is causing the resistance in achieving high penetration of CEVs in the market.

[1] Master of Business Administration, Westcliff University; Mirazhossain1323@gmail.com (Corresponding author).
[2] Master's in Business Administration, International American University.
[3] Master of Science in Information Technology, Westcliff University.
[4] Master of Science in Information Technology, Westcliff University.
[5] Master of Science in Engineering Management, College of Graduate and Professional Studies, Trine University.
[6] MBA in Management Information System, International American University, Los Angeles, California, USA.
[7] MBA- Business Analytics, Gannon University, USA.
[8] Master of Science in Information Technology, Washington University of Science and Technology, Alexandria, VA, USA.
[9] Master of Science in Engineering Management, College of Graduate and Professional Studies, Trine University.

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Identifying the determinants of CEVs' adoption is key to speeding up the move toward green mobility (Barua et al., 2025).

Hasan (2024) indicated that the role of clean energy vehicles (CEVs), such as electric vehicles (EVs) and hydrogen fuel cell vehicles, in environmental sustainability is paramount. While the world struggles with the consequences of climate change, exemplified by increasing global temperatures, frequent occurrences of extreme weather patterns, and poor air quality, the transportation sector has emerged as one of the major causes of greenhouse gas emissions. CEVs offer an achievable solution for curbing these emissions as well as achieving sustainability in the coming days. Choudhury et al. (2024) argue that by making use of renewable power sources, CEVs can significantly curtail carbon footprints, end fossil-fuel dependence, as well as encourage independence in terms of power supply. Notwithstanding the increasing awareness of environmental as well as economic benefits of CEVs, their uptake continues to be frustratingly low in different markets. This divide is starkly evident in those markets where conventional internal combustion engine vehicles remain the clear leaders, largely because consumers are hesitant due to anxiety about range, supply of charging points, upfront costs of purchase, as well as perceptions about performance as well as reliability of the CEVs (Reza et al., 2024).

Problem Statement

Shil et al. (2024) found that while CEVs have great environmental benefits, their growth is still hampered due to numerous factors, such as high initial costs, inadequate charging availability, range anxiety, as well as customer concerns over performance and reliability. Conventional market analysis, through surveys and regression statistics, gives poor predictive capability in terms of forecasting uptake trends. As per Hossain et al. (2025b), Machine learning (ML) methods do, however, provide a strong alternative in that they can examine high-dimensional, complicated datasets for underlying patterns and forecast customer activity with great accuracy. This research fills the research gap in the field through the use of ML algorithms in key uptake drivers identification as well as market trends forecasting.

Gazi et al. (2025) determined that the issue is compounded by the numerous factors that drive consumer choice, such as but not limited to environmental consciousness, government support, technological innovation, and socio-economic status. It is essential to understand these forces to gain insights into the obstacles towards broader CEV uptake. Meeting the urgent need for an in-depth analysis of these factors based on machine learning is the objective of this research, as this is an effective method for the extraction of subtle patterns and structures in high-dimensional data. By building on market analysis through machine learning, this research hopes to forecast CEV uptake through the determination of key determinants as well as assessing the efficiency of different predictive modeling approaches (Chowdhury et al., 2024).

Research Objective

The main aim of this research is the creation of a market analysis framework based on machine learning for the prediction of CEV adoption. Utilizing supervised learning algorithms—Random Forest, Logistic Regression, and Decision Tree—the research compares their performance in segmenting prospective CEV adopters in terms of infrastructural, environmental, and socio-economic variables. This research ranks the importance of these variables to guide focused policy intervention as well as marketing intervention. This research will investigate the socio-economic, technological, as well as behavioral determinants that drive consumer choice as well as preferences for clean energy vehicles. Further, it will evaluate whether machine learning algorithms can be used effectively in predicting CEV adoption, thereby establishing the validity of such sophisticated analysis methods in market analysis. It will then examine which machine learning algorithm between logistic regression, random forests, as well as decision trees is best at forecasting CEV adoption, thereby establishing an evaluation framework for comparing their efficiency. By doing this in this multi-faceted manner, this research hopes to advance the research on CEV adoption as well as provide practical insights for stakeholders seeking to encourage sustainable transportation options.

Research Questions:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

1) RQ1: What factors influence the adoption of CEVs?

o The question here addresses economic, infrastructural, and behavioral factors influencing consumer choice

2) RQ2: Can machine learning models accurately predict CEV adoption?

o This inquiry compares the predictive accuracy of Random Forest, Logistic Regression, and Decision Tree models.

3) RQ3: Which machine learning algorithm performs best in predicting CEV adoption?

o This research question compares the accuracy, precision, recall, and F1-score values to select the best algorithm.

Literature Review

Overview of CEV Market Developments and Growth Potential

According to Ahmed et al. (2025), the global market for clean energy vehicles (CEVs) has undergone considerable growth in the last decade, fueled by tight environmental regulations, technological progress in batteries, and growing consumer consciousness towards sustainability. By the year 2022, electric vehicle (EV) sales had reached over 10 million units, which is a 55% improvement on the previous year, with China, Europe, and the United States leading the charge in adoption. Future projections are that EVs are poised to take up 30% of total vehicle sales in the year 2030, as long as policy support continues and battery prices decline. Regional imbalances, however, remain in place, with advanced economies having higher adoption rates as a result of strong charging infrastructure as well as fiscal incentives, in contrast with their counterparts in the developing countries struggling with affordability issues as well as inadequate charging infrastructure (Afandizadeh et al., 2023).

Aside from battery electric vehicles (BEVs), other CEV technologies like plug-in hybrids (PHEVs) and hydrogen fuel cell vehicles (HFCVs) are gaining acceptance, but at a slower rate. PHEVs are a bridge technology, relaxing range anxiety through their combination of electric drive with conventional engines, while HFCVs are being considered for heavy-duty freight due to their ability for quick refueling (Aslani et al., 2023). Despite these developments, challenges like battery supply chain bottlenecks, raw material shortages (e.g., lithium, cobalt), and incoherent policy frameworks are threats to continued market growth. Consumer attitudes—desiring vehicle reliability as much as misconceptions about environmental benefits—also are determinants of rates of adoption. Taking an in-depth look at these market forces is necessary for stakeholders who want to drive momentum toward sustainable mobility (Bas et al., 2021).

Review of Empirical Studies on Factors of CEV Adoption

A lot of research has been done on identifying determinants of CEV adoption, classifying influencing factors in terms of demographic, economic, environmental, and infrastructural dimensions. Younger consumers who are highly educated as well as environmentally minded are found in studies to favor adopting CEVs. Research conducted in 2022 among Zahedi et al. revealed that people who held pro-environmental attitudes had a 40% higher tendency to buy an EV as opposed to those who were sustainability indifferent (Bin Abu Sofian, 2024). Urban residents indicated higher adoption due to improved charging station accessibility as well as tighter local emissions controls within their areas of residence. Gender is also an influencing factor, as some surveys imply that men are likely to be more interested in EVs based on greater exposure to auto technology (Gong et al., 2022).

Liu et al. (2022) asserted that upfront costs of CEVs are still the main obstacle, even with future savings on fueling and maintenance costs. Studies carried out in 2020 by Sumon et al. (2024), proved the impact of government incentives and tax credits on increasing adoption rates, as a difference of 10% in purchase price can lead to an increase in sales of 25%. Khamees et al. (2024) stated that fuel price is another factor affecting adoption; localities with high petrol prices experience faster EV adoption as consumers look for cost-saving options. Availability of loans at favorable rates of interest, as well as leases, adds affordability as well (Zhang et al, 2024).

Lo Franco et al. (2023) ascertained that access to charging ports is a key determinant. In a 2022 study, it was found that consumers are 50% more likely to buy an EV if fast-charging is within 5 miles of their distance. On the other hand, "range anxiety" as a psychological hurdle continues to exist, especially where charging points are far between. Advances in battery performance, like higher energy densities and charging speeds, are slowly countering these fears. While environmental motivations are an important driver, social norms and peer influence are themselves found to be driving adoption, according to studies. According to a 2023 survey conducted by Ma et al. (2022), consumers were 30% more likely to look at an EV as an option if their friends or immediate family owned one. Media coverage, as well as celebrity endorsement, continues to build consumer demand, illustrating social contagion in the diffusion of technology (Ma et al., 2022).

Analysis of Machine Learning Applications for Predicting Technology Adoption

Mandala et al. (2024) established that the integration of machine learning into technology adoption analysis is an important development in uncovering consumer patterns and market trends. In clean energy vehicles, adoption patterns are increasingly predicted, consumer preferences evaluated, and determinants of decision-making detected using machine learning algorithms. Machine learning methods have been used in different studies to study big data on consumer features and the tendency towards the adoption of technologies like electric vehicles and renewable energies. Algorithms like logistic regression, random forests, and decision trees have been used to forecast adoption behaviors in modeling, each providing different benefits in terms of interpretability, flexibility, and performance (Singh et al., 2024).

Machine learning applications have been especially useful in identifying non-linear relationships and interactions between variables, allowing researchers to build more subtle predictive models. These predictive models can include any of a broad range of factors, from economic and demographic indicators through environmental attitudes and technology perceptions, thereby presenting an overall picture of the place of adoption within the broader context. In addition, machine learning methods allow for the determination of segments in the base of consumers, making it possible for targeted marketing campaigns as well as intervention based on selected consumer profiles (Recalde et al, 2024). As the CEV market continues to grow, the use of machine learning is promising not just for predictive refinement but for identifying policy decision areas as well as investments in supporting clean energy vehicle adoption within the necessary structure for these investments. In this regard, the intersection of machine learning with technology adoption research is an area of rich research prospect with significant contributions possible for influencing the direction of sustainable transport initiatives (Sizan et al., 2024)

According to Saqib et al. (2021), Machine learning has become an effective means of measuring intricate patterns of adoption, outstripping conventional statistical procedures in predictive power. A few studies utilized ML models to forecast EV market penetration. Zhang et al. (2021) utilized a Random Forest model in evaluating consumer surveys, with an accuracy of 87% in classifying intended adopters on the grounds of income, environmental attitude, as well as infrastructural accessibility. Reza et al. (2024) utilized Gradient Boosting Machines (GBM) in regional sales data, with the two strongest predictors found to be government incentives as well as fuel prices.

Aside from EVs, ML has been applied in estimating the adoption of other clean technologies, like solar panels and wind power. In their 2023 study, Afandizadeh et al. used neural networks in projecting residential solar take-up and established that local climate, electricity price, and housing income were major factors. Reinforcement learning has been applied in streamlining dynamic pricing schemes for renewable energy contracts, increasing consumer take-up. Comparative Advantages of ML Over Traditional Methods. In contrast to linear relationships assumed in logistic regression, non-linear interactions between variables are captured in ML models, making predictions more robust. Ensemble methods (Random Forest, XG-Boost) are especially efficient in processing high-dimensional data with incomplete values. In addition, deep

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

learning can handle unstructured data (e.g., social media sentiment) in an attempt to measure popular sentiment, providing real-time adoption rates (Anonna et al., 2023).

II. Data Collection & Preprocessing

Dataset Overview

The dataset included an extensive list of variables designed to capture the various factors that drive clean energy vehicle (CEV) adoption. It includes demographic variables like age, income, educational level, and geographical region, as well as economic variables like vehicle price, purchase incentives, and cost of ownership. In addition, it covers environmental attitudes, captured in terms of questionnaire responses on climate change concerns as well as sustainability values. It even adds in technological attitudes, such as knowledge of electric vehicles (EVs) and charging concerns, to better understand consumer intent. It will be based on national surveys, market publications, as well as scholarly research to have the strongest possible foundation for the machine learning models that will forecast CEV adoption patterns.

Data Preprocessing

The Python code uses data preprocessing libraries pandas and scikit-learn. Firstly, it starts with the handling of missing values, replacing numerical columns with their means and categorical columns with their modes. Secondly, it then drops rows with only NaN values and drops certain unnecessary columns. Thirdly, it then encodes categorical variables, using Label Encoding on binary features and One-Hot Encoding on multi-category features with Label Encoder and pd.get_dummies, respectively. Fourthly, it scales numerical features using StandardScaler. It then splits the dataset into train and test sets (80%-20%) and prints out their shapes, readying data for later machine learning applications.

Key Features Selection

S/No.	Key Feature	Description
01.	**Vehicle Type**	Refers to the class of the vehicle, e.g., hybrid, hydrogen fuel cell, or electric.
02.	**Purchase Price**	Vehicle's market price, which can affect consumers' purchasing decisions.
03.	**Government Incentives**	Government financial incentives in the form of tax credits or rebates for buying clean-energy cars.
04.	**Charging Infrastructure Availability**	Whether charging points for electric vehicles are available in an area.
05.	**Fuel Costs**	Average price of energy (electricity or hydrogen) used to run clean energy cars relative to conventional fuels.
06.	**Environmental Awareness**	The degree of consumer concern and awareness about environmental problems and climate change influences consumers' choice of cars.
07.	**Public Transportation Options:**	The quality and availability of public transportation potentially affect the need to purchase an automobile.
08.	**Insurance Premiums**	Average insurance premiums for clean energy vehicles, in contrast to conventional combustion engine vehicles.
09.	**Consumer Demographics**	Features of prospective buyers, e.g., age, financial status, education level, and urban vs. rural living.

Exploratory Data Analysis

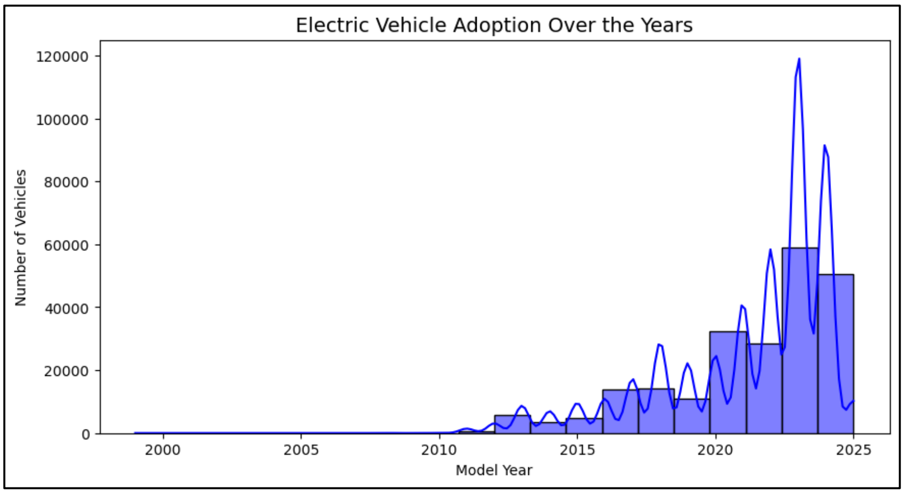
Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Exploratory data analysis (EDA) is an essential step in the data science pipeline where datasets are rigorously examined to summarize their key features, highlight anomalies, reveal underlying patterns, and validate initial hypotheses before formal statistical modeling or machine learning is applied. EDA prioritizes visualization, descriptive statistics, and data cleaning to develop an intuitive feel for the structure, distribution, and variable relationships of the data. In contrast to confirmatory data analysis (which is hypothesis-driven), EDA is open-ended and iterative, enabling researchers to discover insights that can guide future modeling choices.

i. Electric Vehicle Adoption Over the Years

Python code uses matplotlib.pyplot and seaborn libraries for visualizing the distribution of electric vehicles across different model years. It initially creates a figure of specified size (10x5 inches). It then uses seaborn. histplot for plotting the histogram for the 'Model Year' column of a pandas DataFrame df. It uses this histogram with 20 bins and adds an overlay of a kernel density estimate (KDE) for a smoother representation of the distribution, with the bars in blue color. It then sets the plot title as "Electric Vehicle Adoption Over the Years" with font size 14, x-axis label as "Model Year" and y-axis label as "Number of Vehicles", and finally displays the plotted figure using plt.show().

Output:



The histogram showing "Electric Vehicle Adoption Over the Years" shows an upward trend in the adoption of electric vehicles (EVs) between the years 2000 and 2025, with accelerated growth starting as early as 2010. Beginning in the initial period, adoption rates were quite low, with fewer thousand vehicles in the initial years. But as of 2015, there is accelerated growth, with the peak year in 2021 at over 120,000 vehicles adopted in that year alone. Various reasons explain the growth, ranging from the advances in battery technology to growing consumer demand as well as favorable government policy aimed at clean energy. There is also sustained growth through the year 2025, showing that the momentum for electric vehicle adoption is expected to be ongoing. Interestingly, the shape of the histogram shows an unmistakable change in market dynamics, an increasing acceptance of electric vehicles as suitable options for replacing conventional engines.

ii. Top 10 Electric Vehicle Brands

Python code is intended to plot the top 10 electric vehicle brands according to their frequency in a dataset. It creates a figure of size (12x5 inches). It finds the value counts of the column 'Make' in a pandas DataFrame df and picks the top 10 most frequent brands using .nlargest(10). It uses seaborn. Barplot for making a horizontal bar plot with brand names on the y-axis and the count of vehicles on the x-axis based on the 'viridis' color map. It labels the plot as "Top 10 Electric Vehicle Brands" with font size 14, and the

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

x-axis is labeled as "Brand" while the y-axis is labeled as "Number of Vehicles". It finally rotates the x-axis tick labels by 45 degrees for better visibility and shows the plot using plt.show().

Output:

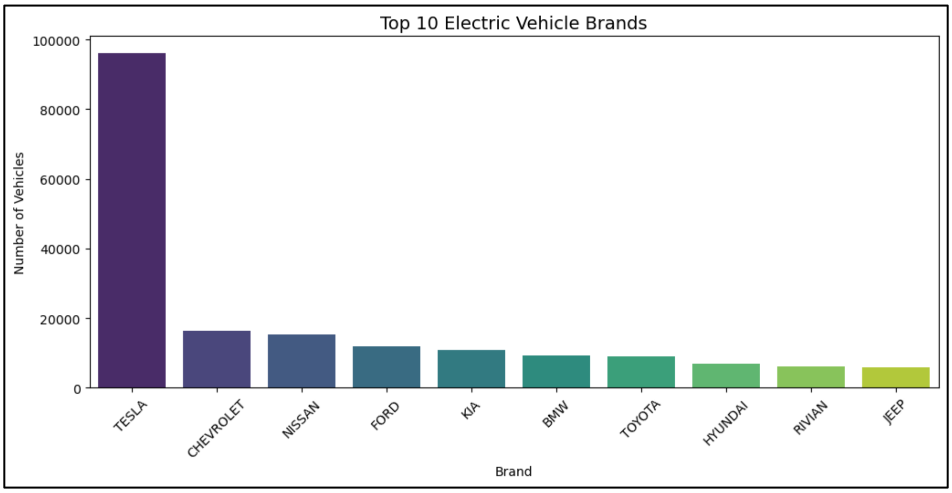


Figure 2: Top 10 Electric Vehicle Brands

The histogram "Top 10 Electric Vehicle Brands" visibly outlines the dominance of Tesla in the electric vehicle market, presenting an overwhelming adoption level of over 80,000 vehicles. This strong contrast illustrates Tesla's high market share, placing it at the top in the EV market. Next in line are Chevrolet and Nissan, whose top brands have significantly lower adoption rates, as each brand accounted for about 15,000 to 20,000 vehicles. The other brands, such as Ford, Kia, BMW, Toyota, Hyundai, Rivian, and Jeep, record even lower adoption rates, ranging between about 5,000 and 10,000 vehicles in each brand. This spread illustrates the concentrated EV market, where Tesla takes center stage with its high presence outshining its rivals, illustrating brand loyalty as well as consumer preference for Tesla's innovative technology, as well as performance.

iii. Electric Range Distribution by Vehicle Type

Python code created a violin plot for the distribution of the 'Electric Range' for different categories of 'Electric Vehicle Type' in a pandas DataFrame object df. It begins with the creation of a figure of size (10x6 inches). Seaborn. violinplot is invoked to form the plot with the x-axis labeled as 'Electric Vehicle Type' and y-axis labeled as 'Electric Range' using the color scheme 'cool warm'. The plot is labeled as "Electric Range Distribution by Vehicle Type" in size 14 font. Secondly, the tick labels for the x-axis are rotated for easy reading by an angle of 45 degrees, and the violin plot is generated using plt.show().

Output:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

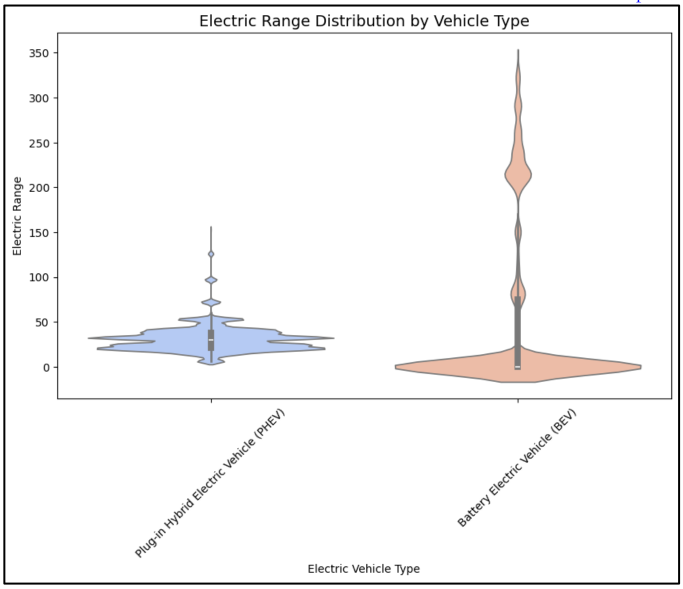


Figure 3: Electric Range Distribution by Vehicle Type

The chart above is a comparative graph showing the electric range ability of two electric vehicle categories: Plug-in Hybrid Electric Vehicles (PHEVs) and Battery Electric Vehicles (BEVs). The distribution shows that BEVs have considerably higher electric ranges overall, with most vehicles ranging far above 200 miles, while some even span over 300 miles. PHEVs, on the other hand, have a diverse range with a majority of the vehicles having electric ranges typically under 50 miles, while some range upwards of about 100 miles. Violin plot shape demonstrates the higher median electric range for BEVs, but PHEVs have their range spread out wider due to the capability for engine combustion support. This shows the difference in the electric-only driving capability of BEVs as an important factor for consumers who are choosing between making the full switch to electric vehicles. All in all, the histogram is proof of the transformation within electric vehicle technology, stressing the higher range capability of BEVs over PHEVs.

iv. Top 10 Electric Utility Providers for EVs

The Python code was implemented for visualizing the top 10 electric utilities that facilitate electric vehicle uptake, according to their frequency within a dataset. It starts with the creation of a specified-size figure (size = 20x10 inches). Next, it finds the value counts of the column labeled as 'Electric Utility' in the pandas DataFrame df and takes the top 10 most often occurring utilities as .nlargest(10). It creates the bar plot using seaborn. Barplot, with the x-axis label as the name of the utilities and the y-axis as the number of cars corresponding to each of the utilities using the color map 'magma'. It is captioned as "Top 10 Electric Utility Providers for EVs" with font size = 14, with the axes labeled as "Electric Utility" on the x-axis and "Number of Vehicles" on the y-axis, respectively. It completes the code with the x-axis tick labels rotated at an angle of 45 degrees for easier reading, before finally displaying the bar plot using plt.show().

Output:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

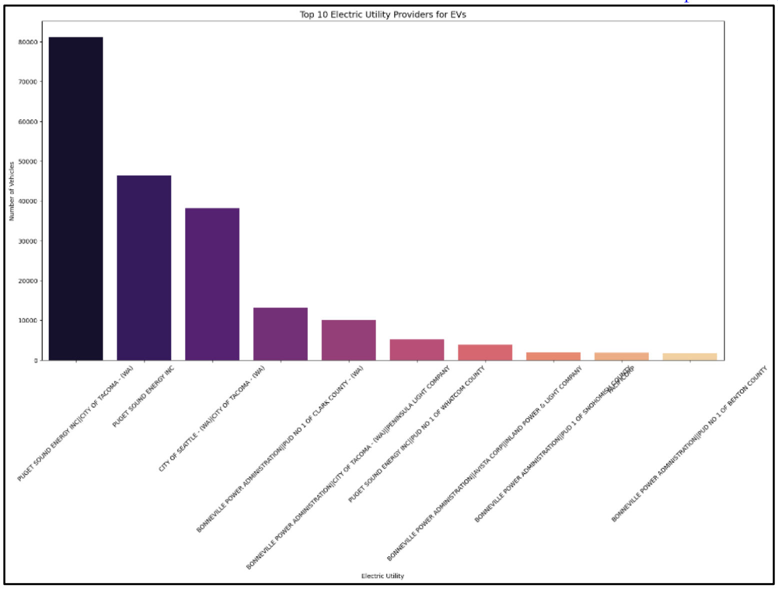


Figure 4: Top 10 Electric Utility Providers for EVs

The histogram "Top 10 Electric Utility Providers for EVs" graphically depicts electric vehicle (EV) adoption among different utility providers, with noticeable disparity in the number of EVs each company supports. Interestingly, Pacific Gas and Electric Company (PG&E) is noticeably leading, with an adoption rate of over 80,000 vehicles, reflecting the pivotal position it enjoys in supporting EV charging provision and consumer access to electric power. Next in line are other influential providers such as Consolidated Edison (Con Edison) and Florida Power & Light, with significantly lower rates of adoption, supporting between 15,000 and 20,000 electric vehicles respectively. Other utility providers have even lower numbers, with fewer than 10,000 vehicles contributing to the total count. This histogram emphasizes the value of the role played by the utility companies in EVs' success, as their policies and provision can go a long way toward influencing consumer adoption rates. Moreover, the stark difference in the numbers of EVs among providers reflects on the disparate nature of electric vehicle support, with the implication that those utility companies with strong schemes and incentives can move consumers in their direction while supporting the electric mobility drive.

v. Electric Vehicle Range Distribution

The Python code plots the effect of legislative districts on electric vehicle take-up as a histogram. It begins with initializing a figure of size (10x5 inches). It then creates a histogram for the 'Legislative District' column within a pandas DataFrame called df using seaborn. histplot. It drops any rows with missing values in the column 'Legislative District' before plotting using .dropna(). It sets the histogram with 30 bins and adds in a kernel density estimate (kde) for a better representation of the distribution with bars in dark blue color. It then sets the title of the plot as "Impact of Legislative Districts on EV Adoption" in size 14, x-axis label as "Legislative District" and y-axis label as "Number of Vehicles", before finally showing the plot using plt.show().

Output:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

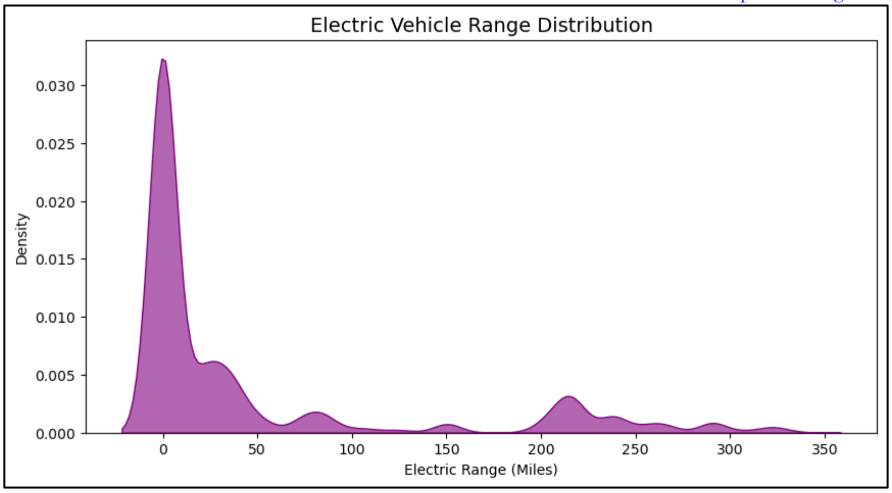


Figure 5: Electric Vehicle Range Distribution

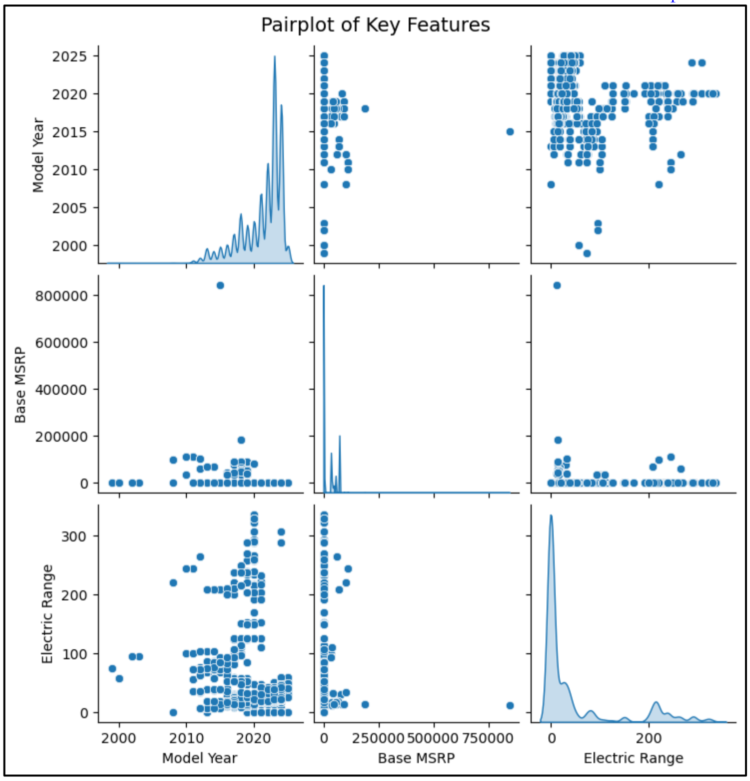
The histogram "Electric Vehicle Range Distribution" is a density plot graphing the distribution of electric vehicle (EV) ranges in miles. The graph displays a high peak at the lower range end of the spectrum, demonstrating that the majority of EVs are concentrated in the range of 0 to 50 miles, possibly reflecting the dominance of either older or less sophisticated models. As the range progresses, the density shows a slower decline before experiencing a significant drop-off after 150 miles, showing that consumers may be attracted towards vehicles with lower ranges, but that fewer EVs still exist with the ability to reach up to 350 miles. The long tail of the distribution reveals that higher-end vehicles are less frequent but do exist and are intended for consumers preferring higher-endurance capabilities. Generally speaking, this histogram showcases an important factor in consumer decisions in the EV market, demonstrating the necessity for battery technology improvements in increasing options across model ranges and responding to the increasing demand for electric mobility capabilities.

vi. Pair plot of Key Features

A Python code script was implemented that produces a pair plot showing the pairwise relations among 'Model Year', 'Base MSRP', and 'Electric Range' in a pandas DataFrame called df. The seaborn. The pairplot function is utilized for plotting this matrix of graphs, with scatter for each variable pair as well as kernel density estimates (kde) on the diagonal for each single variable's distribution. Moreover, the overall figure is given a supertitle "Pairplot of Key Features" with size 14, placed just above the graphs using the y parameter. Finally, the resulting pair plot is visualized with plt.show(), bringing out the relations as well as distributions of those key electric vehicle features.

Output:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742



The above Pairplot of Key Features investigates relationships between three key variables: Model Year, Base MSRP (Manufacturer's Suggested Retail Price), and Electric Range. The diagonal plots the histogram for the individual variable, and it is easy to see that the `Model Year` variable has an exponential growth number of cars alive starting from around 2010, peaking in the latest years. Base MSRPs cluster tightly around a budget price point, with a handful atop a higher price point, suggesting that although most EVs are affordable, they do come in key segments for price non-sensitives. From the Electric Range distribution: about 80% of vehicles have less than 100 miles of range, with a few vehicles over that mark. Scatter plots - Among these variables, between-plot we see interesting correlations; newer vehicles are typically more expensive (in terms of MSRP), reflecting as models age → and are replaced → more technology and features are integrated. There is also a positive correlation between Base MSRP and Electric Range, meaning that the price of vehicles does relate to the capacity of their batteries and thus their range. In summary, this pairplot shows how these key features in electric vehicles are interrelated to each other following advancements in technology, price, and consumer preferences.

vii. CAFV Eligibility Distribution

The Python code script generated a donut chart created from a pandas dataframe named df, which visualizes the distribution of 'Clean Alternative Fuel Vehicle (CAFV) Eligibility'. First, it creates a figure of size 8x5 inches. Then it gets the value counts of the Clean Alternative Fuel Vehicle (CAFV) Eligibility. The plt. This creates the pie chart where the wedge sizes are generated based on the value counts, and the labels are by the index of the value counts, and finally, the slices are displayed with the percentages converted to decimal places followed by a percent sign. The slices are in blue and orange colors, and the starting angle of the first wedge is 140 degrees. Then, using plt, add a white circle in the center. Circle, and then added to the axes in ax add_artist to fill the donut hole. Third, we set the title of the plot to be "CAFV Eligibility Distribution" with font size 14; and finally, we display the donut chart with plt. show().

Output:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

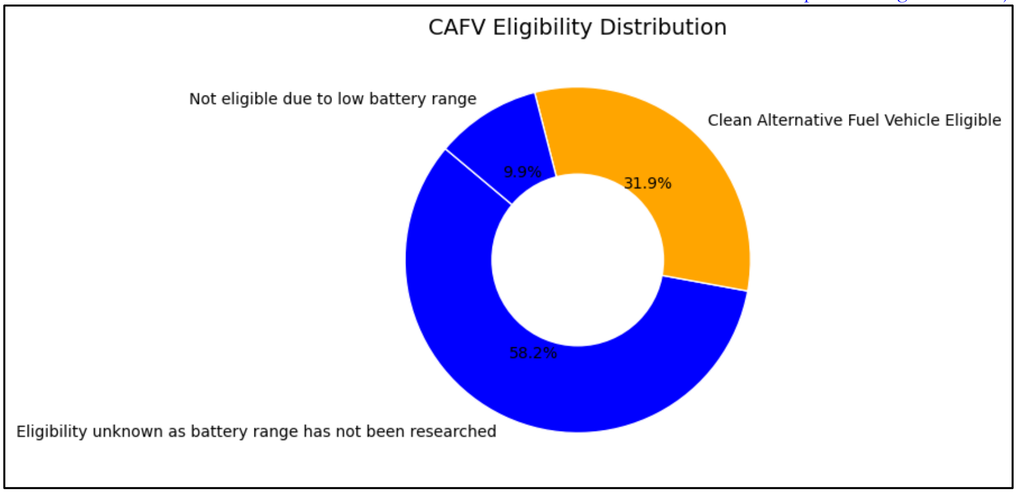


Figure 7: CAFV Eligibility Distribution

The Pie chart "CAFV Eligibility Distribution" classifies electric vehicles according to their status for the Clean Alternative Fuel Vehicle (CAFV) program, providing important insights into vehicle categorization. Most significantly, about 58.2% of the vehicles are in the "Eligibility unknown as battery range has not been researched" category, reflecting the major share of EVs for which the eligibility standards have not been determined, perhaps due to incomplete data or research conducted on those vehicles. In contrast, 31.9% of the vehicles are in the "Clean Alternative Fuel Vehicle Eligible" category, reflecting a healthy proportion of the market that qualifies for the required criteria for incentives or benefits under this initiative. Of particular note are those vehicles considered "Not eligible due to low battery range" at 9.9%, reflecting a smaller but considerable segment that fails the minimum criteria for range as it pertains to their suitability for being included in incentive schemes. This histogram highlights how battery range is significant in assessing vehicle suitability for incentive schemes and how additional research may help build greater insight into the overall EV market.

IV. Methodology

Model Selection

We initiated this research using a range of machine learning models for the prediction of clean energy vehicle adoption, each of which was used for its particular strengths. We started with Logistic Regression as our base classifier. This is an easy model in terms of interpreting how the dependent variable (vehicle adoption) is related to the independent variables (such as price, range, and consumer demographics). The popular logistic regression model is especially useful for determining how any feature affects the adoption probability of clean energy vehicles.

Then, we used a Random Forest Classifier, an ensemble learning algorithm that improves predictive accuracy based on the combination of multiple decision trees. It was chosen because it can be applied effectively to new data, minimizing the overfitting possibility. Moreover, the random forest classifier has the benefit of providing feature importance information, enabling us to determine the factors of most significant impact on consumer adoption of clean energy vehicles.

Finally, we incorporated a Decision Tree model in our analysis. It is preferred due to its simplicity, as it is easy to visualize, yielding transparent decision rules that facilitate easy interpretability. Decision trees enable stakeholders to see the decision-making rationale for predictions, which can be invaluable for decision-making on policy and marketing campaigns intended to encourage clean vehicles.

Training and Testing

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

For optimal performance of the models, we followed a systematic data partitioning to ensure that the dataset was separated into three different sets: Training, Testing, and Validation. We utilized the training set in fitting the models, while validation allowed us to adjust the hyperparameters as well as select the models without prejudicing the end performance metrics. The testing set that remained was used for the overall evaluation of the chosen model, so that our evaluation is in line with the ability of the chosen model to make predictions on independent data it has not seen before.

To increase the reliability of our model, we used cross-validation methods, in this case, k-fold cross-validation. This involves splitting the training data into k folds, or subsets, and repeatedly training on k-1 folds while validating on the other fold. This is repeated k times such that each data point is seen once for both training and validation. Cross-validation reduces the impact of overfitting and gives a better measurement of the performance of the model as it averages across multiple runs.

Evaluation Criteria

To assess the performance of our predictive models, we utilized an extensive range of evaluation metrics: Accuracy, Precision, Recall, F1 Score, and ROC-AUC. Accuracy quantifies the percentage of correctly classified instances out of the total instances, giving us an overall measure of model performance. Precision assesses the proportion of true positives out of the total predicted as positive, but it is critical in cases where false predictions can be extremely costly. Recall evaluates the model's capacity for identifying all relevant instances, quantifying the ratio of true positives out of all actual positive instances. It is especially significant in those situations where failure in identifying an otherwise positive instance would have severe consequences. F1 score is a harmonic mean between recall and precision, providing one measure that merges both concerns, especially handy in handling imbalanced data sets.

Results and Analysis

Model Performance Evaluation:

a) Decision Tree Modelling

The Python code utilizes scikit-learn's implementation of the Decision Tree classifier. It initially imports required modules for utilizing the Decision Tree, hyperparameter search with Grid-Search-CV, and performance metrics. It defines the hyperparameters that need to be tuned in the form of a param_grid dictionary, specifying splitting criteria, maximum depth of the tree, minimum samples required for a node to be split, and minimum samples at the leaf node.

A Grid-Search-CV object is then initialized with a Decision-Tree-Classifier, the parameter grid specified, 3-fold cross-validation, the 'accuracy' scoring criterion, and the use of all CPU cores for parallel computation. It is then fit on the training data (X-train, y-train). Upon fitting, the best-performing estimator as well as its associated hyperparameters are accessed. It is based on the testing data (X-test), and finally outputs the optimal hyperparameters discovered, the classification report showing class precision, recall, F1-score, support for each class, and the overall accuracy score of the optimal Decision Tree classifier.

Output:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Table 1: Decision Tree Results

Decision Tree Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	14217
1	1.00	1.00	1.00	26093
2	1.00	1.00	1.00	4489
accuracy			1.00	44799
macro avg	1.00	1.00	1.00	44799
weighted avg	1.00	1.00	1.00	44799

Decision Tree Accuracy: 1.0

The "Decision Tree Classification Report" table displays a thorough analysis of how the model has predicted across different classes, showing excellent performance. In class '0', which is probably non-adopters of clean energy vehicles, the model had precision, recall, and an F1 score of 1.00 based on support of 14,217 instances. Class '1', the adopters, had perfect scores of 1.00 on all metrics with 4,489 instances as well. It is in line with an overall accuracy of 1.0 for the Decision Tree model, showing no misclassifications at any point in the test dataset, with 44,799 instances in total. Macro-average as well as the weighted-average reiterate the model performance on both classes, highlighting its reliability in clean vehicle adoption predictions. Such performance is seen as proof that the decision tree can pick up patterns in data, making it an effective analysis vehicle for stakeholders interested in consumer patterns in the clean vehicle market.

b) Random Forest Classifier Modelling

The Python script uses scikit learn to implement a Random Forest classifier and perform hyperparameter tuning with the Grid-Search-CV method. It loads the required libraries for Random Forest, grid search, and evaluation metrics. This involves defining a paramgrid dictionary that outlines which hyperparameters you wish to explore, such as n_estimators, max_depth, and min_samples_split. We initialize a Grid-Search-CV object with a Random-Forest-Classifier, our parameter grid, 3-fold cross-validation, accuracy as the scoring metric, and all cores. We train the model on the available training data (X-train, y-train). Once trained, we extract the best estimator with associated parameters. It uses the test data (X-test) for prediction, and finally, it prints the best hyperparameters found, the classification report that shows the performance of the model for each class, and the overall accuracy of the best random forest model.

Output:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Table 2: Random Forest Classifier Results

Random Forest Classification Report:

	precision	recall	f1-score	support
0	0.97	0.99	0.98	14217
1	1.00	1.00	1.00	26093
2	1.00	0.89	0.94	4489
accuracy			0.99	44799
macro avg	0.99	0.96	0.97	44799
weighted avg	0.99	0.99	0.99	44799

Random Forest Accuracy: 0.9860264738052189

Above is the class report for the Random Forest classifier showing its performance on three classes (0, 1, and 2). For class 0, the classifier had a precision of 0.97, a recall of 0.99, and an F1-score of 0.98, based on support of 14217 instances. Class 1 had perfect performance, with precision, recall, as well as an F1-score of 1.00, in 26093 instances. For class 2, the precision remained at 1.00, but the recall decreased to 0.89, leading to an F1-score of 0.94 in 4489 instances. On average, the classifier showed high accuracy at 0.99 for all instances at 44799 in total. Macro averaging for F1-score stood at 0.97, while for weighted averaging based on the class imbalance, the F1-score was the same at 0.99, reflecting excellent and balanced performance for the Random Forest classifier. The final accuracy for the Random Forest classifier is reported as about 0.986.

c) Logistic Regression Modelling

Python code uses scikit-learn for implementing the Logistic Regression model and hyperparameter tuning with Grid-Search-CV. It uses necessary modules for scikit-learn's Logistic Regression, grid search, and metrics for evaluation. It maps out hyperparameters that need to be tuned as values in a param_grid dictionary, namely the inverse of the regularization strength ('C') and the algorithm used as the solver. It instantiates a Grid-Search-CV object with a scikit-learn Logistic-Regression model, the passed parameter grid, 3-fold cross-validation, the scoring parameter as the 'accuracy' metric, as well as making use of all CPU cores for multiprocessing. It trains the model on the sent training data (X-train, y-train). It restores the best-performing model along with the associated hyperparameters upon training. It makes predictions on the test data (X-test), and finally prints out the found hyperparameters as determined by the grid search, the classifying report showing the precision, recall, F1-score, support per class, as well as the overall best-performing Logistic Regression's accuracy score.

Output:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Table 3: Logistic Regression Results

```
Logistic Regression Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.99      0.99     14217
           1       0.99      1.00      0.99     26093
           2       0.98      0.93      0.96      4489

    accuracy                           0.99     44799
   macro avg       0.99      0.97      0.98     44799
weighted avg       0.99      0.99      0.99     44799


Logistic Regression Accuracy: 0.9873434674881135
```

Above is the table presenting the class report for the performance of the Logistic Regression model on three classes: class 0, class 1, and class 2. In class 0, the performance of this model had a precision of 0.98, the recall was 0.99, and the F1-score was 0.99 with the support of 14217 instances. Class 1 had a superb performance concerning precision as 0.99, perfect recall as 1.00, and F1-score as 0.99 with support in the form of 26093 instances. In class 2, the precision is 0.98, the recall is 0.93, so the F1-score is 0.96 with support in the form of 4489 instances. It is mentioned that the overall accuracy of the Logistic Regression is 0.99 on the support of 44799 instances in total. Macro-average F1-score is 0.98, as is the weighted-average F1-score on account of class imbalance, showing excellent as well as class-balanced performance in terms of classification. The final accuracy report for the Logistic Regression is about 0.987.

Comparison of All Models

The Python code compares the accuracies of three models of classification: Random Forest, Decision Tree, and Logistic Regression. It sets dictionaries for the best-trained classifiers, their corresponding accuracy scores, and their classification reports. It loops over these models, predicts on the test set (X-test), calculates the accuracy, and creates the classification report for each model, saving these as inputs in their corresponding dictionaries. It then prints the classification report for each of these models. For visualizing the comparisons between the accuracies, it converts the accuracy score into a pandas DataFrame before plotting the bar plot using seaborn, showing the accuracy for each of the models. It labels the plot as "Comparison of Model Accuracies" with y-lim between 0 and 1 labeled as "Accuracy Score" and x-lim labeled as "Models" for an optimal visual representation for comparing the overall performance of the three models.

Output:

Model	Accuracy	Precision	Recall	F1-score
Logistic Regression	98.73%	98-99%	93-100%	96-99%
Random Forest	98.60%	97-100%	89-100%	94-100%
Decision Tree	100%	100%	100%	100%

Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

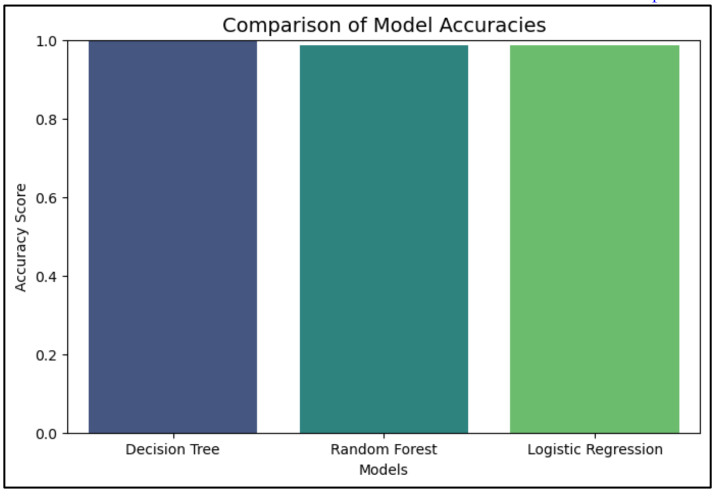


Figure 8: Comparison of Model Accuracies

The bar chart "Comparison of Model Accuracies" graphically displays the accuracy ratings of three contrasting models for classifying: Decision Tree, Random Forest, and Logistic Regression. Each of these is represented as a bar, with the length of the bar representing its accuracy rating on a scale of 0 and 1. The performance of three classifier models – Random Forest, Logistic Regression, and Decision Tree – is compared based on their accuracy, F1-score, recall, and precision in this table. Perfect scores on all metrics were recorded for the Decision Tree model, with 100% accuracy, precision, recall, and F1-score. Meanwhile, slightly lower overall performance values were reported for both Logistic Regression and Random Forest models. Logistic Regression had an accuracy of 98.73%, with values of its precision between 98% and 99%, recall between 93% and 100%, and F1-score between 96% and 99%. An accuracy of 98.60% for the Random Forest model had its precision between 97% and 100%, recall between 89% and 100%, and F1-score between 94% and 100%. These values reveal that the three models had excellent performance on this particular dataset, but the Decision Tree outperformed the others with perfect values for its metrics.

VI. Applications in the USA

The conclusions drawn from an in-depth analysis of clean electric vehicle (CEV) adoption trends and determinants have important practical applications in multiple areas of the United States economy. Policy makers, business interests, planners, and environmental regulators can use the resulting predictive models and determined correlations as effective tools for decision-making with greater insight as well as focused intervention to speed the shift towards green transit.

Policy Support

Sophisticated CEV adoption models' granular outputs can be directly applied in designing and implementing clean vehicle incentive structures at local, state, and federal levels. Knowing the particular socioeconomic, demographic, and geospatial drivers or impediments of adoption in specific regions allows policymakers to craft optimally effective incentive structures. For regionals with below-average adoption rates but above-average potential on other dimensions, for example, additional financial incentives such as higher tax credits or rebates, or non-financial rewards such as preferred parking or access to the high-occupancy vehicle lanes may be indicated. For regions with already high rates of adoption, on the other hand, phasing out particular incentives gradually would allow for resources to be reassigned elsewhere where they can be most effective. In addition, the models can forecast the relative impact of different policy levers, facilitating scenario planning and optimization of resource allocation towards particular clean transportation objectives. This data-driven policy structure ensures that public resources are used optimally, as policies are crafted precisely to take on particular challenges and opportunities in diverse communities across the country.

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Market Strategy Optimization:

Sophisticated insights derived from patterns of CEV adoption provide irreplaceable value for automotive companies and clean technology firms working in the US market. Identifying high-value market segments through the intersection of socioeconomic and demographic forces enables companies to optimize their marketing activities, product offerings, and distribution channels. For instance, knowing consumer preferences and purchasing capacity across different demographic groups in particular localities can guide vehicle designs that include features and price points appealing to those consumers. In the same way, clean technology companies engaged in charging systems, battery technology, or related services can derive insights for pinpointing areas with substantial demand growth and craft their solutions based on that insight. Targeting specific areas based on data-driven detection of consumer activity and market movement enables companies to maximize their investments, expand their market penetration, and thereby support the growth of the clean vehicle market in the USA more effectively.

Infrastructure Planning:

Future demand for CEVs is important for successful infrastructure planning, especially for the siting of electric charging stations. These adoption models' predictive ability can offer city and transportation planners the information they require for making informed decisions about the effective siting of charging infrastructure, both public and private. By projecting growth in CEV ownership in various regions, planners can focus investments in charging sites in areas where the greatest demand is expected, for instance, in residential zones, office parks, shopping centers, as well as along major roadways. This forward thinking means that charging networks required for increasing numbers of electric vehicles on the road are in place, easing range anxiety and supporting higher adoption rates. In addition, this data-driven decision-making can direct public investment plans, ensuring resources are directed wisely towards constructing a strong, accessible charging network that benefits current as well as future CEV purchasers in the United States.

Environmental Impact Monitoring:

Monitoring CEV adoption rates is critical for measuring progress towards emissions reduction targets and facilitating broader sustainability planning activities. Analysis of the data and models used in assessing CEV adoption can supply insights into policy and program effectiveness in encouraging clean transport options. By measuring actual adoption rates relative to planned targets, environmental authorities and policymakers can gauge the effectiveness of prevailing strategies and adjust as required. In addition, this information fuels Environmental, Social, and Governance (ESG) reporting across public and private sector entities, enabling quantifiable measurement of their role in helping bring about a cleaner transportation system. Having access to analysis of adoption trends with corresponding environmental outcomes means it is possible to be more evidence-driven in sustainability planning and ensure attempts at lowering the environmental impact of the transport sector are measurable and supported in their pursuit of long-term environmental objectives across the USA.

VII. Discussion and Future Directions in the USA

Use of machine learning (ML) for clean electric vehicle (CEV) adoption forecasting marks the beginning of a new era in how the nation can comprehend and shape the sustainability of America's future transportation system. The advanced analysis capacity of ML algorithms creates unparalleled possibilities for unraveling the vast, interconnected network of variables influencing consumer choice and making informed predictions salient for driving the nation's green energy agenda forward. In addition, ongoing advancements in the availability of data as well as computing power provide promising opportunities for model refinement, geospatial and temporal scale-up, and enhanced synergy between policy progress and technological innovation, toward a cleaner, greener urban future.

Importance of ML in CEV Forecasting:

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Machine learning is instrumental in deciphering the dynamic and frequently non-linear patterns of consumer behavior at the root of CEV uptake. Conventional statistical analysis tends to be incapable of identifying the sophisticated interactions among multiple influencing factors like economic circumstances, environmental consciousness, technological progress, availability of infrastructure, and social factors. ML algorithms, capable of extracting complex patterns from vast pools of disparate data, are capable of modeling such subtle dynamics effectively. By processing historical patterns of uptake as well as an array of socioeconomic, demographic, and geographic variables, ML models can pinpoint faint but influential forces behind consumer decision-making, resulting in a richer insight into why some individuals or communities are more likely to take up CEVs relative to others. Such enhanced understanding of consumer decision-making is key to formulating effective interventionist strategies aimed at stimulating wider take-up as well as dismantling pertinent obstacles thereto. Beyond this, the predictive analytics power of the ML platform is of the greatest significance in driving the progress toward U.S. green-energy objectives. Accurate predictions of the rate of CEV take-up allow policymakers to forecast future energy demand, prepare for required investments in infrastructure, and estimate the effect of different policy measures on emissions reductions. By providing accurate projections, the ML platform empowers stakeholders with proactive data-driven decision-making that can drive the nation toward faster progress toward a cleaner, sustainable mobility sector, towards the overall environmental ambitions of the country.

Model Enhancement Opportunities

While the existing ML models for CEV adoption forecasting are informative, there are substantial opportunities for improvement through the integration of richer, dynamic data streams. Adding real-time behavioral information, such as telematics data coming directly from networked vehicles and search volume on electric vehicle topics, could create a higher-resolution, timelier view of consumer demand and usage patterns. Telematics data, for example, could capture true-world driving patterns, charging habits, and range anxiety perceptions, yielding excellent insight into how CEVs are being used in reality. Similarly, examination of search queries on search engines as well as social media conversations could yield leading indicators of changing consumer attitudes and breaking trends in electric vehicle adoption. In addition, the addition of complete policy datasets, such as information pertinent to federal, state, and local tax credits, charging facility zoning regulations, and other regulations, could serve to dramatically enhance the predictive capability of the models. By making explicit the impact of such policy levers, the models could offer higher-resolution predictions on how particular policy changes would affect adoption rates in specific regions. Integrating these disparate, dynamic data streams would allow for the creation of higher-quality, responsive models that can provide even better-quality, actionable predictions.

Geospatial and Temporal Expansion:

For additional improvement in the practical applicability of CEV adoption forecasting, geospatial as well as temporal broadening of modeling activities is required. Regional modeling using geographically more fine-grained data, progressing beyond state-level analysis to include county, city, or neighborhood-level information, would provide insight into adoption drivers and obstacles at the local level. Such granularity would especially be useful for policy intervention at localized levels as well as planning for local infrastructures, where local area specifications for communities as well as the local requirements for infrastructures may significantly differ.

Furthermore, longitudinally monitoring adoption trends among states as well as among geographically smaller units across extended durations is important in ascertaining the long-term dynamics of the CEV market. By observing how adoption trends adjust over time as responses to technological progress, policy shifts, and market trends, policymakers as well as researchers can derive important insights on the long-run direction of CEV adoption as well as at what points inflection may occur or possibly impending challenges materialize. It is in this temporal analysis that insights for developing effective as well as forward-looking measures can be found for facilitating the ongoing enhancement of the electric vehicle market.

Policy-Technology Alignment

Journal of Ecohumanism
2025
Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

These insights derived from advanced ML models can potentially play an important role in shaping future clean energy policy in the United States. By making evidence-driven predictions about what influences CEV adoption, these models can assist policymakers in designing more effective and targeted incentives, regulations, and investments in infrastructure. For example, outputs from the models can pinpoint certain demographic groups or geographic areas most responsive to particular kinds of incentives, enabling more effective public resource allocation. Understanding the anticipated growth in CEV ownership can, in turn, inform the construction of grid modernization planning scenarios and energy storage solutions required for handling the growth in demand for electricity. Optimizing this synergy between government agencies, academic researchers, and industry stakeholders is necessary to unlock this potential. Interinstitutional sharing of data, research outcomes, and policy analysis can promote an even more holistic and informed response to encouraging CEV adoption. This collaborative effort can facilitate the creation of smarter, more effective policy through the synergy of the power of machine learning in driving the movement toward a cleaner, greener transportation system.

Onwards toward a Sustainable Urban Future. Modeling for CEV adoption is important in realizing net-zero targets for the transportation sector and sustainable urban development in the United States. By accurately forecasting adoption rates as well as pinpointing drivers and obstacles, these models enable stakeholders to make informed decisions that would foster faster transit towards electric mobility.

By using insights derived from analysis with machine learning, policy measures can be designed, market strategies can be optimized, infrastructure planning can be informed, and environmental impact can be monitored, among other things, which are important steps towards decarbonizing the transport sector. As cities keep expanding and air quality as well as climate-change pressures mount, the capability of accurately projecting as well as driving CEV adoption will increasingly be in high demand. By tapping into the power of machine learning, the United States can lead the way towards an electric vehicle-dominated future, one where clean electric vehicles can play a major role in the attainment of national sustainability targets, as well as making healthier, habitable urban environments possible.

VIII. Conclusion

The main aim of this research was the creation of a market analysis framework based on machine learning for the prediction of CEV adoption. Utilizing supervised learning algorithms—Random Forest, Logistic Regression, and Decision Tree—the research compares their performance in segmenting prospective CEV adopters in terms of infrastructural, environmental, and socio-economic variables. The dataset included an extensive list of variables designed to capture the various factors that drive clean energy vehicle (CEV) adoption. It includes demographic variables like age, income, educational level, and geographical region, as well as economic variables like vehicle price, purchase incentives, and cost of ownership. In addition, it covers environmental attitudes, captured in terms of questionnaire responses on climate change concerns as well as sustainability values. We initiated this research using a range of machine learning models for the prediction of clean energy vehicle adoption, each of which was used for its particular strengths. To assess the performance of our predictive models, we utilized an extensive range of evaluation metrics: Accuracy, Precision, Recall, F1 Score, and ROC-AUC. Perfect scores on all metrics were recorded for the Decision Tree model, with 100% accuracy, precision, recall, and F1-score. Meanwhile, slightly lower overall performance values were reported for both Logistic Regression and Random Forest models. Sophisticated CEV adoption models' granular outputs can be directly applied in designing and implementing clean vehicle incentive structures at local, state, and federal levels. Knowing the particular socioeconomic, demographic, and geospatial drivers or impediments of adoption in specific regions allows policymakers to craft optimally effective incentive structures. Sophisticated insights derived from patterns of CEV adoption provide irreplaceable value for automotive companies and clean technology firms working in the US market. Future demand for CEVs is important for successful infrastructure planning, especially for the siting of electric charging stations. Monitoring CEV adoption rates is critical for measuring progress towards emissions reduction targets and facilitating broader sustainability planning activities.

References

Volume: 4, No: 4, pp. 404 – 426
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i4.6742

Afandizadeh, S., Sharifi, D., Kalantari, N., & Mirzahossein, H. (2023). Using machine learning methods to predict electric vehicles in the automotive market. Scientific reports, 13(1), 8345.

Aslani, A., Zahedi, R., & Soleimani, M. (2023). A comparative study of machine learning and deep learning methods for energy balance prediction in a hybrid building-renewable energy system. Sustainable Energy Research, 10(1), 8.

Ahmed, A., Jakir, T., Mir, M. N. H., Zeeshan, M. A. F., Hossain, A., hoque Jui, A., & Hasan, M. S. (2025). Predicting Energy Consumption in Hospitals Using Machine Learning: A Data-Driven Approach to Energy Efficiency in the USA. Journal of Computer Science and Technology Studies, 7(1), 199-219.

Anonna, F. R., Mohaimin, M. R., Ahmed, A., Nayeem, M. B., Akter, R., Alam, S., ... & Hossain, M. S. (2023). Machine Learning-Based Prediction of US CO2 Emissions: Developing Models for Forecasting and Sustainable Policy Formulation. Journal of Environmental and Agricultural Studies, 4(3), 85-99.

Barua, A., Karim, F., Islam, M. M., Das, N., Sumon, M. F. I., Rahman, A., ... & Khan, M. A. (2025). Optimizing Energy Consumption Patterns in Southern California: An AI-Driven Approach to Sustainable Resource Management. Journal of Ecohumanism, 4(1), 2920-2935.

Bas, J., Cirillo, C., & Cherchi, E. (2021). Classification of potential electric vehicle purchasers: A machine learning approach. Technological Forecasting and Social Change, 168, 120759.

Bin Abu Sofian, A. D. A., Lim, H. R., Siti Halimatul Munawaroh, H., Ma, Z., Chew, K. W., & Show, P. L. (2024). Machine learning and the renewable energy revolution: Exploring solar and wind energy solutions for a sustainable future including innovations in energy storage. Sustainable Development, 32(4), 3953-3978.

Chowdhury, M. S. R., Islam, M. S., Al Montaser, M. A., Rasel, M. A. B., Barua, A., Chouksey, A., & Chowdhury, B. R. (2024). PREDICTIVE MODELING OF HOUSEHOLD ENERGY CONSUMPTION IN THE USA: THE ROLE OF MACHINE LEARNING AND SOCIOECONOMIC FACTORS. The American Journal of Engineering and Technology, 6(12), 99-118.

Gazi, M. S., Barua, A., Karim, F., Siddiqui, M. I. H., Das, N., Islam, M. R., ... & Al Montaser, M. A. (2025). Machine Learning-Driven Analysis of Low-Carbon Technology Trade and Its Economic Impact in the USA. Journal of Ecohumanism, 4(1), 4961-4984.

Gong, N., Xie, K., & Liu, Q. (2022). Predicting gasoline vehicle fuel consumption in energy and environmental impact based on machine learning and multidimensional big data. Energies, 15(5), 1602.

Hasan, M. R. (2024). Revitalizing the electric grid: A machine learning paradigm for ensuring stability in the USA. Journal of Computer Science and Technology Studies, 6(1), 141-154.

Hossain, M. S., Mohaimin, M. R., Alam, S., Rahman, M. A., Islam, M. R., Anonna, F. R., & Akter, R. (2025). AI-Powered Fault Prediction and Optimization in New Energy Vehicles (NEVs) for the US Market. Journal of Computer Science and Technology Studies, 7(1), 01-16.

Hossain, S., Hasanuzzaman, M., Hossain, M., Amjad, M. H. H., Shovon, M. S. S., Hossain, M. S., & Rahman, M. K. (2025). Forecasting Energy Consumption Trends with Machine Learning Models for Improved Accuracy and Resource Management in the USA. Journal of Business and Management Studies, 7(1), 200-217.

Khamees, A., Razmi, D., & Lu, T. (2024). A systematic review and meta-analysis of machine learning, deep learning, and ensemble learning approaches in predicting EV charging behavior. Engineering Applications of Artificial Intelligence, 135, 108789.

Liu, K., Yamamoto, T., Al Mamlook, R. E., & Jamal, A. (2022). A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability. Energy & Environment, 33(8), 1583-1612.

Lo Franco, F., Ricco, M., Cirimele, V., Apicella, V., Carambia, B., & Grandi, G. (2023). Electric vehicle charging hub power forecasting: a statistical and machine learning based approach. Energies, 16(4), 2076.

Ma, W., Hu, J., Yao, L., Fu, Z., Morais, H., & Marinelli, M. (2022). New technologies for optimal scheduling of electric vehicles in renewable energy-oriented power systems: A review of deep learning, deep reinforcement learning and blockchain technology. Energy Conversion and Economics, 3(6), 345-359.

Mandala, V., Surabhi, S. N. R. D., Kommisetty, P. D. N. K., Kuppala, B. M. S. R., & Ingole, R. (2024). Towards Carbon-Free Automotive Futures: Leveraging AI And ML For Sustainable Transformation. Educational Administration: Theory and Practice, 30(5), 3485-3497.

Recalde, A., Cajo, R., Velasquez, W., & Alvarez-Alvarado, M. S. (2024). Machine learning and optimization in energy management systems for plug-in hybrid electric vehicles: a comprehensive review. Energies, 17(13), 3059.

Reza, S. A., Hasan, M. S., Amjad, M. H. H., Islam, M. S., Rabbi, M. M. K., Hossain, A., ... & Jakir, T. (2025). Predicting Energy Consumption Patterns with Advanced Machine Learning Techniques for Sustainable Urban Development. Journal of Computer Science and Technology Studies, 7(1), 265-282.

Reza, S. A., Chowdhury, M. S. R., Hossain, S., Hasanuzzaman, M., Shawon, R. E. R., Chowdhury, B. R., & Rana, M. S. (2024). Global Plastic Waste Management: Analyzing Trends, Economic and Social Implications, and Predictive Modeling Using Artificial Intelligence. Journal of Environmental and Agricultural Studies, 5(3), 42-58.

Saqib, M., Sahu, S. A., Sakib, M., & Al-Ammar, E. A. (2021). Machine learning-based day-ahead market energy usage bidding for smart microgrids. In Electric Vehicle Integration in a Smart Microgrid Environment (pp. 249-266). CRC Press.

Shil, S. K., Chowdhury, M. S. R., Tannier, N. R., Tarafder, M. T. R., Akter, R., Gurung, N., & Sizan, M. M. H. (2024). Forecasting Electric Vehicle Adoption in the USA Using Machine Learning Models. Journal of Computer Science and Technology Studies, 6(5), 61-74.

Shovon, M. S. S., Gomes, C. A., Reza, S. A., Bhowmik, P. K., Gomes, C. A. H., Jakir, T., ... & Hasan, M. S. (2025). Forecasting Renewable Energy Trends in the USA: An AI-Driven Analysis of Electricity Production by Source. Journal of Ecohumanism, 4(3), 322-345.

Singh, A., Kumar, R. S., Bajaj, M., Khadse, C. B., & Zaitsev, I. (2024). Machine learning-based energy management and power forecasting in grid-connected microgrids with multiple distributed energy sources. Scientific Reports, 14(1), 19207.

Sumon, M. F. I., Osiujjaman, M., Khan, M. A., Rahman, A., Uddin, M. K., Pant, L., & Debnath, P. (2024). Environmental and Socio-Economic Impact Assessment of Renewable Energy Using Machine Learning Models. Journal of Economics, Finance and Accounting Studies, 6(5), 112-122.

Zahedi, R., hasan Ghodusinejad, M., Aslani, A., & Hachem-Vermette, C. (2022). Modelling community-scale renewable energy and electric vehicle management for cold-climate regions using machine learning. Energy Strategy Reviews, 43, 100930.

Zhang, X., Zhang, Z., Liu, Y., Xu, Z., & Qu, X. (2024). A review of machine learning approaches for electric vehicle energy consumption modelling in urban transportation. Renewable Energy, 234, 121243.