Exploratory Data Analysis for electric vehicle (EV) Driving range prediction: A Comparative Study

Prince Kumar “Department of *Electrical and Electronics* International Institute of Information Technology Bhubaneswar, India” [B320033@iiit-bh.ac.in](mailto:B320033@iiit-bh.ac.in)

Priyanka Rai “Department of *Electrical and Electronics* International Institute of Information Technology Bhubaneswar, India” [B320034@iiit-bh.ac.in](mailto:B320034@iiit-bh.ac.in)

Ayush Kumar “Department of *Electrical and Electronics* International Institute of Information Technology Bhubaneswar, India” [B320014@iiit-bh.ac.in](mailto:B320014@iiit-bh.ac.in)

Dr. Debanai Prasad Mishra “Head of Department of *Electrical and Electronics* International Institute of Information Technology Bhubaneswar, India” [debani@iiit-bh.ac.in](mailto:debani@iiit-bh.ac.in)

***Abstract*—** As a green and ecologically beneficial form of transportation, electric cars (EVs) are gaining popularity. One of the biggest challenges EV users have been being able to predict the amount of driving their vehicles can do on a single charge**.** Electric vehicles (EVs) are becoming increasingly popular as a sustainable and environmentally friendly mode of transportation One of the biggest challenges EV users have is predicting the amount of driving time their vehicles will have on one battery charge. Planning a trip and reducing range anxiety depends on an accurate range estimate. The goal of this study is to anticipate the EV driving range using machine learning methods. In this research, several regression models for predicting EV driving range will be developed and compared. A real-world dataset comprising various factors affecting EV range, such as power, trip distance, energy consumption, driving style, and environmental factors, is used for analysis. To manage missing values, outliers, and categorical variables, the dataset is preprocessed using exploratory data analysis methods. The findings of this study contribute to the expanding area of EV range prediction and provide EV buyers, producers, and regulators insightful information. The user experience can be improved, EV adoption can be boosted, and effective design of the charging infrastructure is made possible with accurate range prediction. The study also highlights the importance of model selection and data pretreatment in making accurate predictions.

*Keywords— Electric vehicle (EV), Data preprocessing, Exploratory Data Analysis (EDA), Machine learning, Regression models, Deep Multilayer Perceptron (MLP), Range prediction*

# Introduction

A carbon-neutral economy and a sustainable environment are made possible by electric vehicles (EVs). Power usage models that can accurately and consistently anticipate the consumption of energy are essential for the deployment of EVs to be successful. Driver anxiety will be greatly reduced by increasing the energy efficiency of EVs, which will also offer a crucial foundation for the management, operation, and development of the infrastructure for charging.

With better battery technology and the demand for minimal or zero-emission vehicles, electric vehicles are a strong contender to take the place of combustion engine-powered engine. Despite these vehicles' advantages, the general public has not given them much popularity. ‘Due to the limited infrastructure for charging and consequently shorter driving range, BEV drivers may have range anxiety, or worry that the battery capacity may deplete before reaching their destination.’ [1]. To minimise range anxiety and increase the usability of EVs, applications are needed that help drivers reach their destinations safely without wasting a great deal of time or money. These applications' primary goals are to improve and accurately predict an efficient driving range. Drivers often save as much as twenty percent of the charge in their batteries as a precautionary measure [2], which has a negative impact on how efficiently the battery uses energy.

This research paper is organised as stated: An overview of the literature on the consumption of energy of EVs and the variables that influence it is given in the Background section. The Dataset section outlines the steps taken to prepare the data for analysis. Subsequently, the algorithm selection and model preparation/examination processes are discussed in the Algorithm section. The Results section presents a comparison of the effectiveness of various algorithms. The key results of the research and their significance for predicting the range of electric vehicles are finally summarised in the Conclusion section.

# BACKGROUND AND RELATED WORK

Based on trajectory information of electric vehicles (EVs), researchers have carried out a number of studies on the user's travel and charging behavioural patterns.

The study of how to increase the capacity of batteries or driving range for electric cars (EVs) is based on the driving habits of EV users.

To be able to optimally utilize battery capacity, Li et al. [3] presented an integrated distribution model that described the

daily trip miles. The outcomes of the tests demonstrated the way the mixed distribution model was capable of meeting various drivers' demands. Furthermore, Dong and Lin [5] created “the concept of BEV viability by employing a stochastic modelling approach to characterise the behaviours of BEV drivers.” In order to find ways to lessen range anxiety, the comfort levels of drivers with various driving traits were examined. However, the researchers discovered that the factors are linked even if the driving behaviour that distinguish BEVs are stochastic. Brady and O'Mahony [6] used “A stochastic modelling approach after studying the dependency structure between the six variables using a nonparametric copula function.” The result was a daily trip itinerary and billing profile.

The most thorough approach to reducing air pollution is to deploy electric cars (EVs). Governments are thus promoting the purchase and usage of these vehicles in place of cars with internal combustion engines [7]. EV sales reportedly increased 72% globally in 2018 in contrast to 2017, and they saw a 2.1% rise in market share. [9]. The small market share of electric vehicles may seem odd given the benefits listed above and the presence of large companies in the sector, but it is due to a number of factors, the most prominent of which are their high purchase costs, prolonged charging times when compared to cars powered by fossil fuels, and their limited range per charge [12].

For data-driven predictions, like those generated by ML systems, a large training dataset is preferred. [13]. A few papers have suggested data sharing between cars and the cloud so that users might gain from the knowledge of other consumers, ultimately producing forecasts that are more correct. By gathering data on BEVs' energy usage while navigating a road stretch, Grubwinkler et al. presented an energetic route map built through crowdsourcing [14]. To collect data from the general public for the forecasting of vehicle energy consumption, Tseng and Chau used the participatory sensing approach [15]. Straub et al. [16] proposed an alternative approach to developing an energetic roadmap by collecting driving profiles from the crowd and using machine learning techniques to fill in gaps in the information. This method effectively removed any potential limitations in the coverage of data, resulting in a more accurate and reliable energy roadmap.

In recent years, data-driven methods have become more widely used as an effective way of estimating consumption and gauging driving range. The rationale is that, when compared to more traditional ways, they are more reliable and cost-effective, and this is because internet of things innovations has reduced the costs associated with deployment. “To reduce the expenses associated with installing sensors and transferring data from cars, a considerable amount of information is extracted from the vehicle's network and transmitted to the cloud. This data may then be processed by machine learning algorithms to offer a variety of helpful services” [11].

One of the main problems with ML is the uneven distribution of the training dataset. In general, machine learning models' ability to accurately predict outcomes on testing data would suffer if the distributions of the training and testing sets are different. However, it should be noted that the ML methodologies discussed above do not specifically tackle this challenge. Instead, they mostly use the driving range as the ML algorithms' regression model objective.

# DatASET DESCRIPTION AND PREPROCESSING

We conducted our research using a publicly available dataset called SpritMonitor [19]. Due to its broad collection of data on vehicle fuelling, SpritMonitor came out as the most appropriate alternative among the many datasets we took into consideration. On the crowdsourcing website SpritMonitor, users may provide details on the makes, models, features, and fuel use of their vehicles. Each record includes details such as the distance travelled after the last fill-up, the amount of petrol used, the kind of tire and petrol, and other relevant statistics. The dataset, which is a useful resource for our investigation, mostly includes information from well-known and commonly used car types. Table 1 lists a few electric vehicles that have received the most fuelling records from various consumers, representing various cars.

|  |  |  |
| --- | --- | --- |
| MODEL | NUMBER OF RECORDS | NUMBER OF USERS |
| Mitsubishi i-MiEV | 6555 | 21 |
| Nissan leaf | 6871 | 102 |
| Tesla Model S | 9079 | 114 |
| BMW i3 | 5396 | 84 |
| Renault Zoe | 13619 | 225 |
| Volkswagen e-golf | 7950 | 71 |
| Kia Soul | 3714 | 24 |
| Opel Ampera | 7634 | 24 |
| Citroen Saxo | 2859 | 9 |

1. ELECTRIC VEHICLES WITH THE MOST RECORDED FUEL DATA LIST. [19]

Similar to earlier studies, it is essential to train distinct machine learning models for each type of electric car owing to the substantial variances among them. We decided to use the Volkswagen e-Golf for our subsequent investigation. This decision was based on the records-to-users ratio since a larger ratio produces a dataset that is more evenly distributed. Furthermore, the frequency of outliers was lower in data obtained by people driving the Volkswagen e-Golf than in records from other types of automobiles.

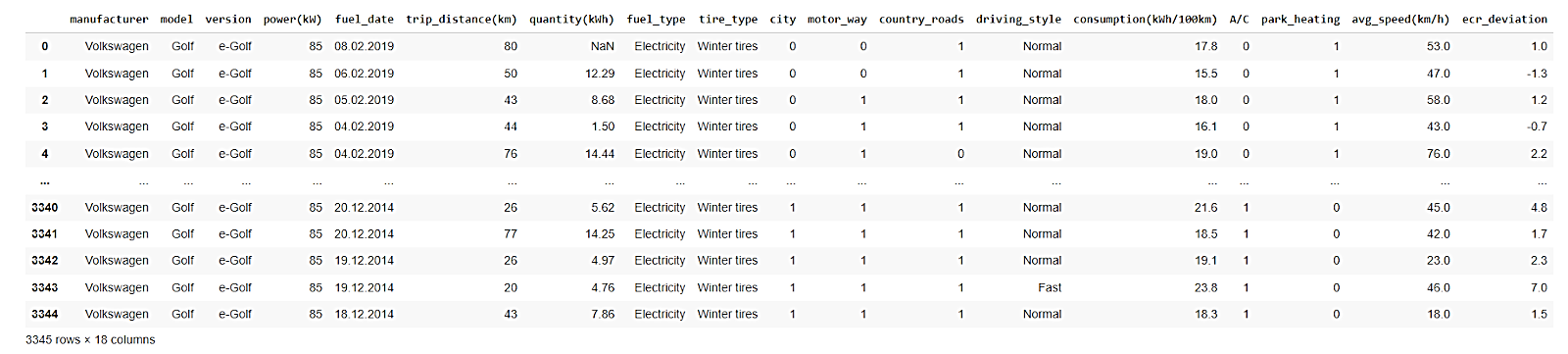
## **Data Collection and variables**

The dataset for this research was gathered using crawlers, which took data from numerous sources on electric automobiles. Fig 1. displays a representative row of data. The dataset is stored in a CSV file and contains a range of variables affecting EV driving range. The following variables are included:

|  |  |
| --- | --- |
| INPUT VARIABLES | avg\_speed(km/h) |
| tyre\_type |
| city |
| motor\_way |
| A/C |
| driving\_style |
| ecr\_deviation |
| country\_roads |
| park\_heating |
| quantity(kWh) |
| consumption(kWh/100km) |
| TARGET OUTPUT | trip\_distance(km) |

1. INPUT/OUTPUT DATA FORMAT

* **Manufacturer:** The manufacturer of the electric vehicle.
* **Model:** The particular electric car model.
* **Version:** The version or edition of the electric vehicle.
* **Power (kW):** The electric vehicle's kilowatt-rated power.
* **Fuel Date:** Date of the journey.
* **Trip Distance (km):** the km that were covered throughout the excursion.
* **Quantity (kWh):** The quantity of electricity consumed during the trip in kilowatt-hours.
* **Fuel Type:** The type of fuel used, which in this case is electricity.
* **Tire Type:** The type of tires used during the trip.
* **City:** A driving indicator variable for cities.
* **Motor Way:** Indicator variable for driving on motorways.
* **Country Roads:** Indicator variable for driving on country roads.
* **Driving Style:** The driving style during the trip.
* **Consumption (kWh/100km):** The energy consumption rate in kilowatt-hours per 100 kilometres.
* **A/C:** Indicator variable for the use of air conditioning.



1. IMPLEMENTED DATASET

* **Park Heating:** Indicator variable for the use of the car's heating system.
* **Average Speed (km/h):** The journey's average speed in kilometres per hour.
* **ECR Deviation:** The deviation of the energy consumption rate from the manufacturer's declared figure, which is stated as 16.8 kilowatts per hundred kilometres.

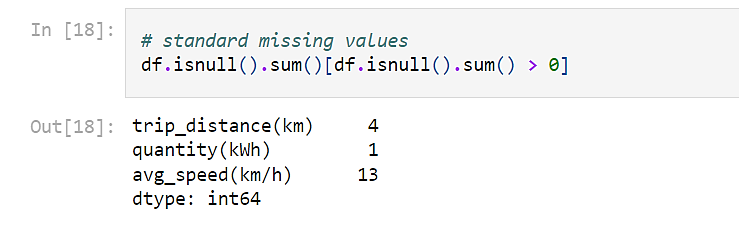
For determining the remaining driving range of EVs, the eleven criteria listed in Table 2 were used as the input variables. Since we are using supervised learning in this case, the model output should be as near to the target output as possible.

## **Data Preprocessing Techniques**

Due to abnormal data points brought on by device failures, the raw dataset cannot be utilized directly. Before doing exploratory data analysis (EDA), the dataset has to be preprocessed to preserve data consistency and quality. Thus, the raw data was processed as follows:

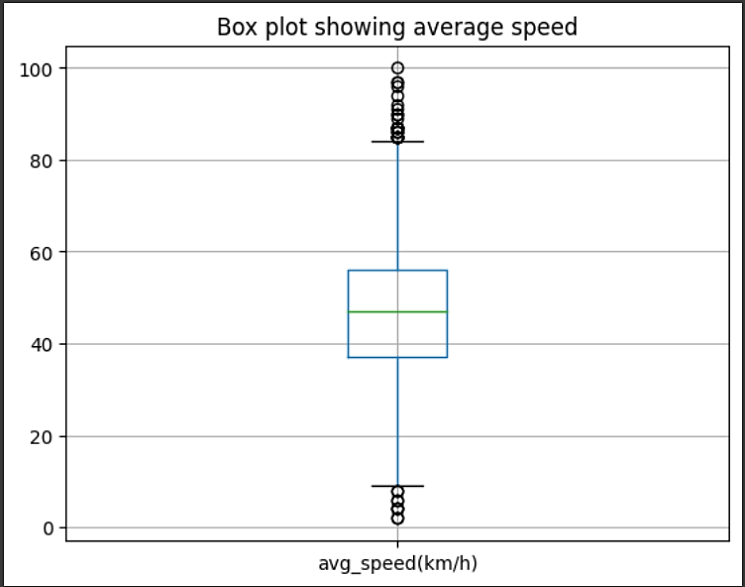
1. Handling Missing Values
2. Removing Useless Columns
3. Labeling the Target Variable

The figure 2, below displays the Standard missing values. As quantity (kWh) has only one missing value. As a

result, we made the decision to drop the one record of quantity(kWh) and also the records for trip\_distance which can be used as test data.

1. COUNT OF MISSING VALUES

Given that ‘avg\_speed(km/h)’ had numerical missing values, we filled them using mean and median imputation based on whether the data distribution is symmetric or skewed using the box plot in Fig. 3. We chose the mean value to impute missing values because the data is symmetric within the range of 40-60 km/h. For Outliers data points we

used median imputation.

1. BOX PLOT SHOWING AVERAGE SPEED

We eliminated the unnecessary columns such as "manufacturer," "model," "version," "power(kW)," and "fuel\_type" since each had just the same value. Fuel\_date is likewise a pointless feature since it represents the data date tests that the respective users have run., 'motor\_way', 'country\_roads', park\_heating', 'A/C, 'city' are binomial data types that were transformed to string types by encoding. The desired variable for determining driving range was identified as the 'trip\_distance(km)' column. The training set is one half of the data sets, while the test set is the other half, in the same ratio as of Total Data Set with random = 2, and then they are compared by determining their R2 scores.

* 1. ***Exploratory Data Analysis (EDA) Approaches***

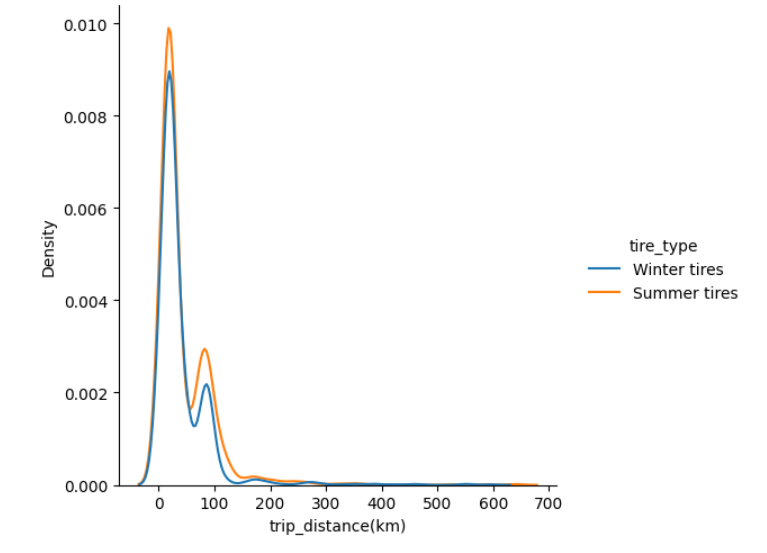
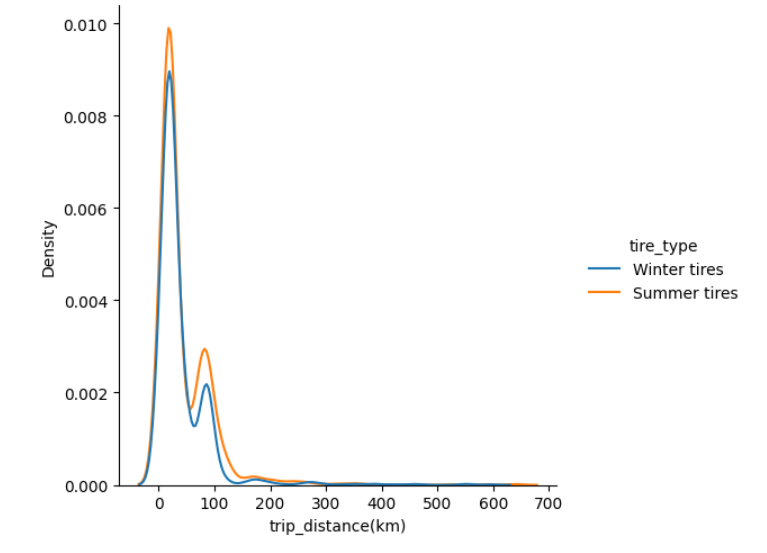
In order to learn more about the dataset, comprehend the connections between variables, and spot anomalies or patterns that can have an impact on the range prediction models, EDA approaches are used. The EDA approaches used in this study include:

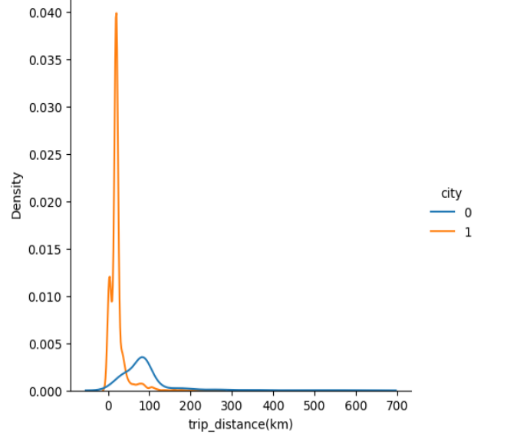
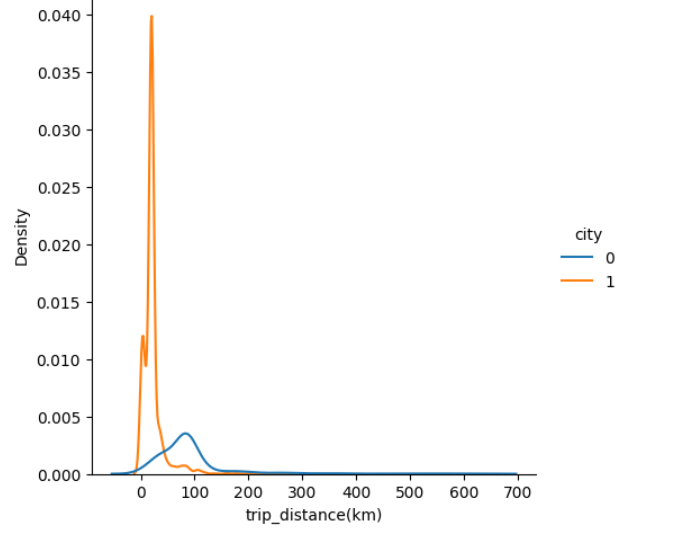
#### **UNIVARIATE**

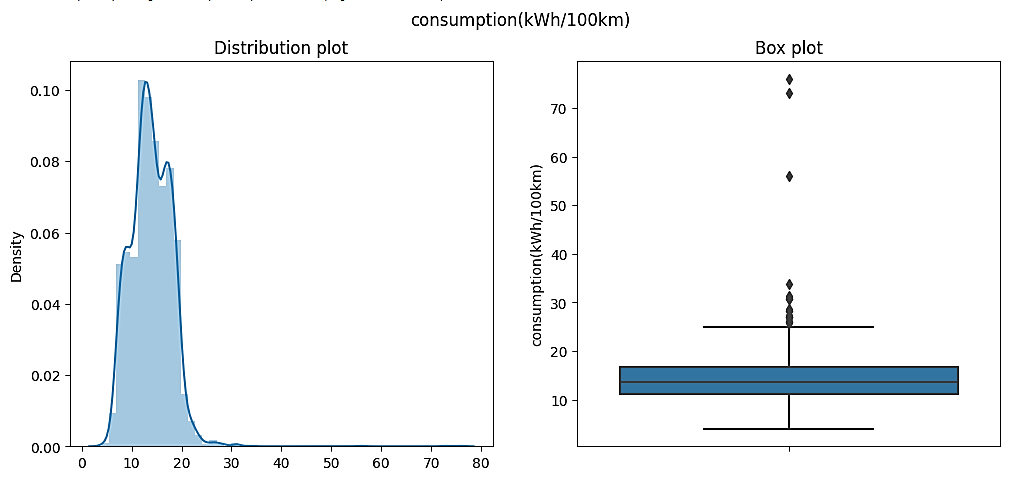
During the univariate analysis, we evaluated individual variables in the dataset to determine their distributions and features. As an example, we estimated the mean, median, and standard deviation of descriptive statistics for variables like average speed, quantity, and energy consumption rate. We also visualized the distributions using histograms, box plots, and density plots to identify any outliers or skewness in the data as shown in Fig. 6.

#### **BIVARIATE**

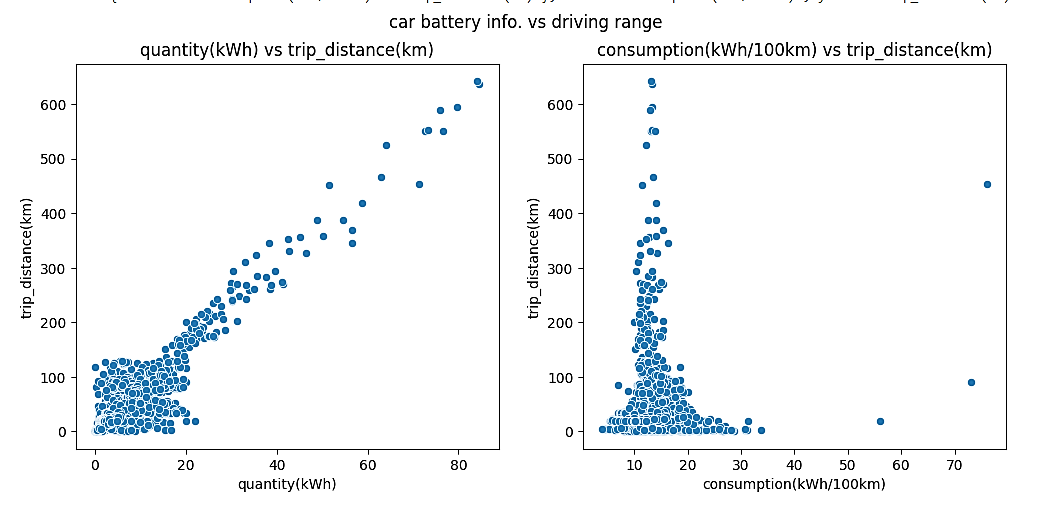
In the bivariate analysis, we concentrated on analyzing the interactions between pairs of variables to find linkages and dependencies. For instance, we examined if certain parameters had an effect on the range of electric cars by using scatter plots to visualize the link between trip distance and other characteristics.



1. DRIVING RANGE VS. TIRE TYPE
2. DRIVING RANGE VS. CITY

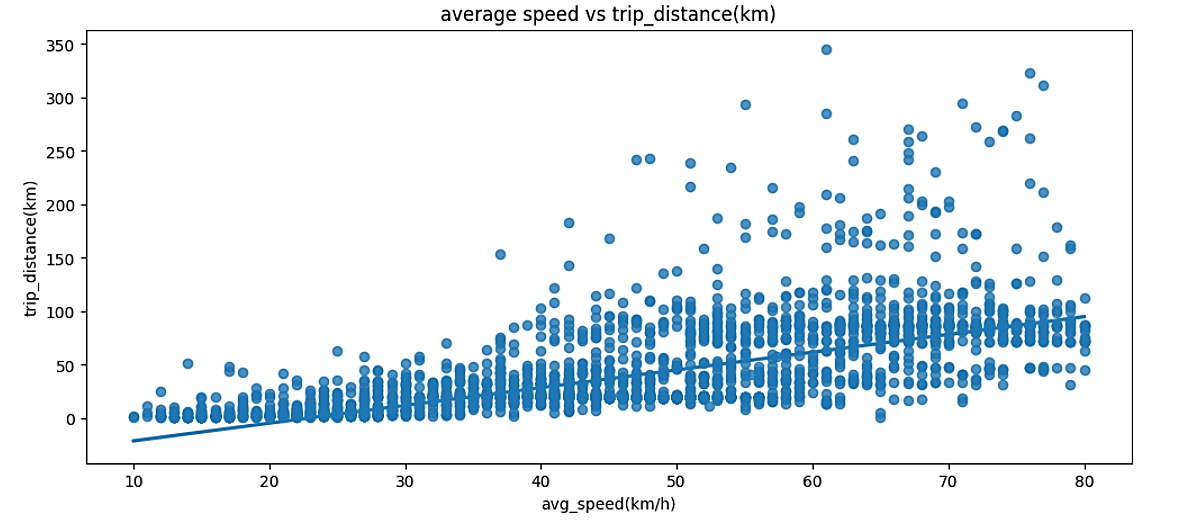


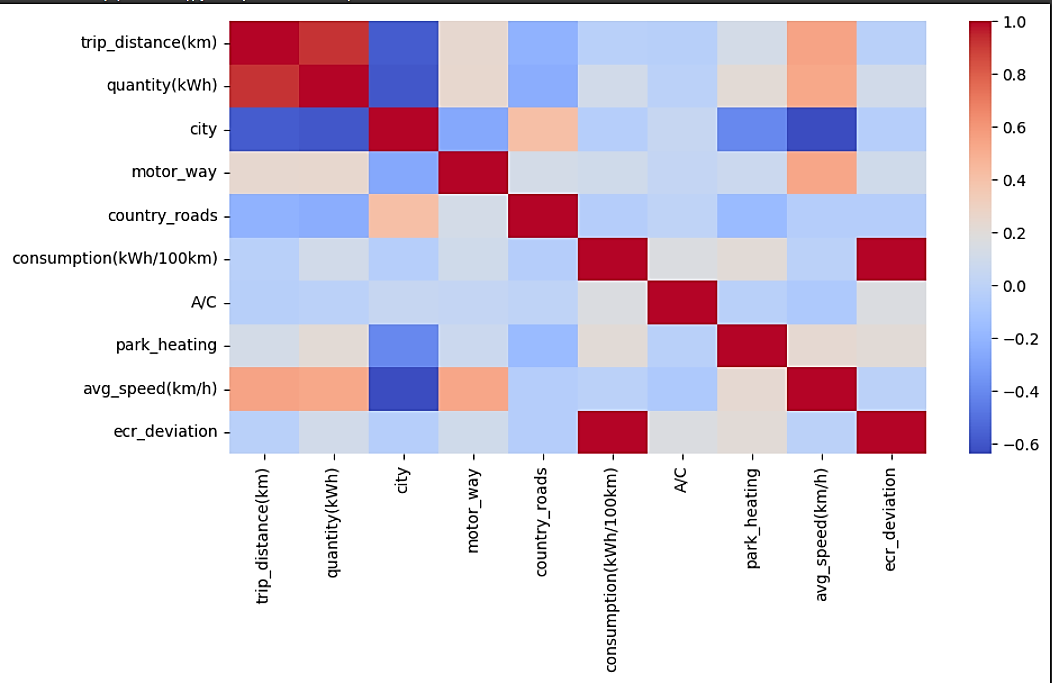
DENSITY AND BOX PLOT OF CONSUMPTION(KWH/100KM)

1. 

SCATTERED PLOT OF CAR BATTERY INFO VS. DRIVING RANGE

SCATTERED PLOT OF AVERAGE SPEED VS. DRIVING RANGE

1. 



1. CORRELATION MATRIX

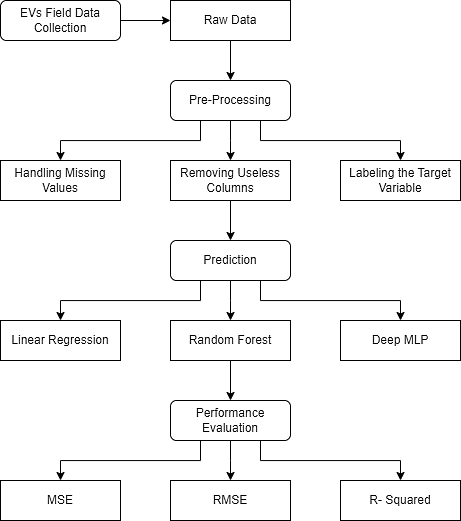
The data points exhibit an increasing trend up to 80 km/hr., as seen in the plot’s Fig 8, therefore the average speed will be a characteristic to take into account for distance range, with trip\_distance being shorter in cities = 1 in Fig 5. For all tire types, the distribution of travel mileage remained the same as shown in Fig 4. As a result, it is not a helpful attribute for estimating the range of distance. The amount beyond 20 kWh is directly proportional to the travel distance. In Fig 7, Quantity in the 0 to 20 range may not be able to calculate the journey distance by itself, thus we incorporated some other characteristics to do so. In EVs with an energy consumption range of 10 to 20, the traveled distance is greater. When the air conditioner is running and the park heating is not turned on, the energy usage is greater.

#### **MULTIVARIATE**

We studied interactions between three or more variables in the multivariate analysis to comprehend intricate patterns and relationships. For instance, we visualized the correlation matrix between variables like trip distance with other variables and auxiliary loads using heatmaps in Fig 9.

As a result, we were able to pinpoint the factors that were strongly connected and may be significantly affecting the EV driving range. We noticed Very few outliers over 50 kwh/100km in energy consumption and the outliers in quantity are in the range >= 40. Few outliers are below 10 and over 80 km/h in average speed.

Fig. 10 depicts the EV range prediction function's flow pattern. We generated a clean and relevant dataset for further research and model building by using these data pre-processing approaches and conducting thorough exploratory data analysis.



1. PROPOSED PROCESSING STRUCTURE FOR PREDICTING ENERGY CONSUMPTION IN ELECTRIC VEHICLES

# REGRESSION MODELS

When predicting a target variable that is continuous based on a number of input variables, regression models are often employed in the analysis of data and ML. In this work, we investigated a number of well-liked regression models for estimating the motor range of electric cars using a variety of input characteristics. To formulate the motor range of EV’s, some of the machine learning algorithms incorporated are - Linear Regression, Random Forest and Deep Multi-Layer Perceptron (MLP). The last two of which are wholesome techniques.

Linear Regression algorithm is a ML method that aims to apply relationships to illustrate the outcome of an event on the basis of data for the independent variables. The observed fitted line is a straight line that closely approximates the individual data points. The aim of the algorithm is to reduce the mathematical disparity between the actual values provided by the manufacturer, and predicted values. The formula for the equation is as given below:

(1)

Where:

Here, the dependent parameter Y stands in for the electric vehicle's driving range. , , ..., are the independent variables which affect driving range. , , , ..., are the coefficients of the independent variables. ε is the disturbance term or error variable in the data. The coefficients β0, β1, β2, ..., βn are computed to reduce the total squared deviations between the actual and predicted values.

MLP (Multi-Layer Perceptron) is a neural network made up of several connected layers which changes the input dimension into the desired dimension. Neurons (or nodes) are conjoined to form a neural network in such a manner that some of the outputs are also feeded as their input. One node serves as an input, one node serves as an output, and there may be any number of hidden layers, each with any number of nodes.

Deep MLP can capture complex correlations between elements like power, trip distance, energy consumption, and driving style, leading to more accurate range estimates. By using its deep architecture and non-linear activation functions, the method offers the potential to uncover odd patterns and correlations in the data that may be difficult to capture with linear models like linear regression. Four hidden layers, each with 64 neurons, were used to apply the Rectified Linear Unit (ReLU) activation function. The benefit of ReLU over other activation functions, such as the sigmoid or hyperbolic tangent, is that it enables the network to learn more rapidly and avoids the saturation issue. The ReLU function is defined as max(0.0, x), where x provides the input to the activation function. It returns the input value if it is positive, otherwise zero. The research shows that mini-batches may be handled well during training thanks to the usage of the Adam optimization algorithm with a 32-batch size. The size of a batch is a reflection of the quantity of samples used in each iteration.

Next algorithm which is used in this report for the comparative study for the driving range of EV’s is Random Forest (RF). Random Forest Regression is a supervised learning approach that uses collective learning, which

integrates predictions from different machine learning models to enhance the accuracy of predictions in regression situations.

To create Random Forest Regression, we imported the RandomForestRegressor class from the sklearn package, made an instance of it, and assigned it to a variable. In this scenario, we put the n\_estimators argument to 50, which indicates our random forest would consist of 50 trees.

Using the. fit() method, we train the model by modifying the weights depending on the data values to boost accuracy. Once the training is complete, our model is ready to generate predictions based on the learnt patterns from the training data.

# Result and discussions

To evaluate the performance of different regression models for EV range prediction, we compared linear regression, random forest (RF) and deep multilayer perceptron (Deep MLP) algorithms. The efficacy of the models is evaluated using regression measures. These metrics compute the prediction error, or the difference between actual and predicted values. Since our focus is on minimizing significant outlier errors, mean squared error (MSE) is a preferable option over mean absolute error (MAE). The common regression evaluation metrics are listed below.

A quantity known as the MSE measures the average of the squared discrepancies between the output that was anticipated and the output that was actually produced. The squared error is preferred because it doesn't differentiate between overestimations or underestimations, but simply indicates that the prediction was inaccurate.

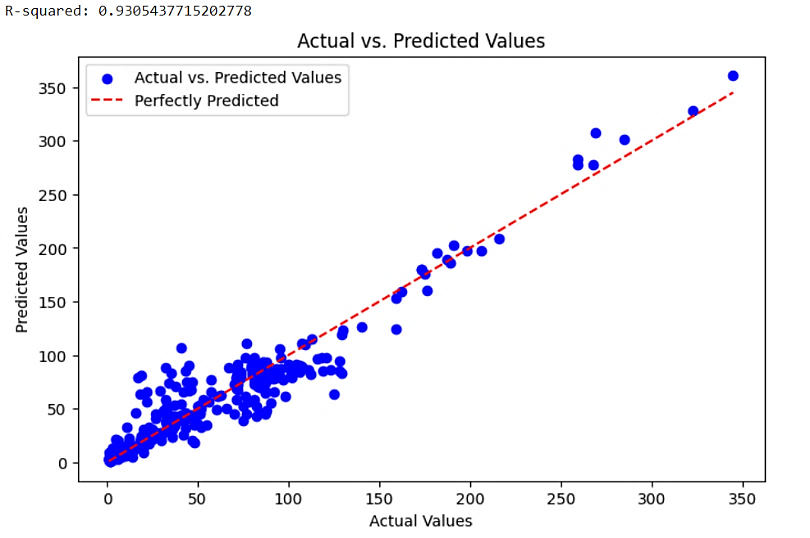
(2)

Here denotes the true value target variable and is the predicted value or the output. Lower the MSE value, closer is the predicted value to actual result.

The score is the next evaluation criteria, which measures how much of the target variable's fluctuation can be accounted for by the model's characteristics. It provides an indication of how well the model performs in explaining the variability of the outcome variable and is formulated as:

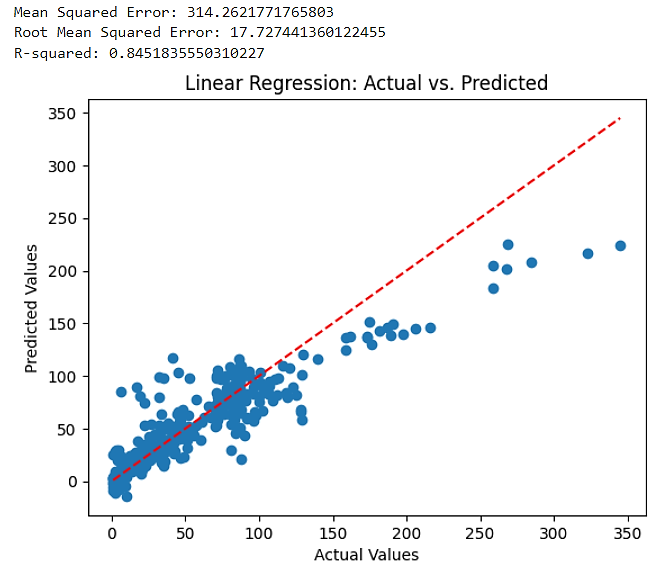
(3)

(4)

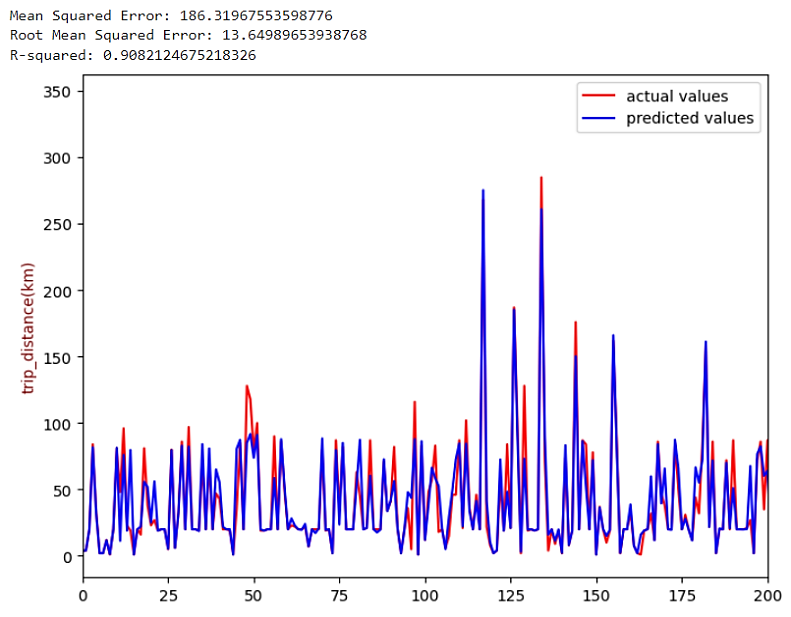
The performance goes on increasing as this score reaches 1. Table 3 displays the results of our comparative analysis of multiple regression models for EV range prediction. It is evident that all of the models scored rather well for accuracy, with R-squared values ranging from 0.84 to 0.93.The deep MLP model obtained the highest accuracy rating with an R-squared of 0.93. With R-squared values of 0.90, the random forest model likewise scored well in terms of accuracy. Despite being the simplest model, the linear regression model's accuracy score, which was 0.84, was comparatively lower than other models.

|  |  |  |  |
| --- | --- | --- | --- |
| REGRESSION MODELS | Root Mean Squared Error (RMSE) | Mean Squared Error (MSE) | R- Squared |
| Linear Regression | 17.7274 | 314.2621 | 0.8451 |
| Random Forest | 13.6498 | 186.3196 | 0.9082 |
| Deep MLP | 11.8738 | 140.9893 | 0.9305 |

1. PERFORMANCE EVALUATION OF PROPOSED MODELS



1. LINEAR REGRESSION’s ACTUAL VS PREDICTED PLOT



1. RANDOM FOREST’S ACTUAL VS PREDICTED PLOT
2. ACTUAL VS PREDICTED PLOT BY DEEP MLP

# Conclusion

Over the years, the adoption of battery electric vehicles (BEVs) has been growing, but a major hindrance to their promotion and usage is the issue of inaccurate display of residual power. This problem contributes to "range anxiety" among drivers, caused by uncertainties in battery performance and other factors. The goal of this study is to tackle this problem by creating a model that can precisely predict the driving range of BEVs.This study introduces advanced ML techniques for accurately estimating the mileage of electric vehicles (EVs) by considering both internal and external factors. These factors include the use of heating, average speed, air conditioning, energy consumption, and the route type. By integrating these factors, the model is able to predict the most efficient values for different conditions, which have not been thoroughly examined in previous studies.

We carried out a comparative analysis on the use of machine learning algorithms for predicting electric vehicles (EVs) driving range. In order to achieve this, we examined a real-world dataset that included various factors affecting the EV range. To enhance the quality of our data and facilitate model training, we incorporated exploratory data analysis techniques during the data pre-processing phase. These methods allowed us to successfully prepare the data and develop a thorough grasp of it. We then implemented and assessed the performance of several regression models, which included linear regression, multilayer perceptron (MLP) and random forest (RF). Finding the best machine learning strategy for precisely forecasting the range of electric cars (EVs) was the main goal of this work. With the help of this study, we were able to determine the strategy that provides the best predictive performance for estimating EV driving range.

Our study yielded insightful results regarding the use of advanced models to forecast the Mileage of electric vehicles (EVs). We evaluated the performance of different regression models, including Linear Regression, Random Forest, and, Deep MLP on a real-world dataset consisting of various factors that affect EV range. Our findings indicated that the Deep MLP and Random Forest models outperformed the traditional Linear Regression algorithm, with higher R2 scores and lower MAE and RMSE values.

EV technology is continually advancing, and future advancements may lead to the development of vehicles with longer ranges. Future research could focus on incorporating additional variables, such as battery health and charging infrastructure, traffic patterns, road slope, and driver behaviour to further enhance the accuracy of EV range prediction models.

Furthermore, XGBOOST and LightGBM methods provide distinct opportunities for researchers and practitioners to develop precise, efficient, and trustworthy data-driven approaches for EVs energy consumption studies.[8].

# ACKNOWLEDGEMENT

We wish to express our sincere appreciation to Professor Dr. Debani Prasad Mishra, “Head of the Department of Electrical and Electronics Engineering”, for his support and invaluable guidance throughout the duration of this study. Additionally, we would also extend our gratitude to our esteemed institution, the International Institute of Information Technology, Bhubaneswar, for providing us with the opportunity to conduct this research. Their unwavering commitment to academic excellence and innovation has greatly contributed to the success of this project.

##### References

1. M. Eisel, I. Nastjuk, and L. Kolbe, “Understanding the inﬂuence of in-vehicle information systems on range stress – Insights from an electric vehicle ﬁeld experiment,” Transportation Research Part F:Trafﬁc Psychology and Behaviour, vol. 43, pp. 199–211, 2016.
2. T. Franke, I. Neumann, F. Bühler, P. Cocron, and J. Krems, “Experiencing Range in an Electric Vehicle - Understanding Psychological Barriers,” Applied Psychology: An International Review, vol. 61, no. 3, pp. 368–391, 2012
3. Z. Li, S. Jiang, J. Dong, S. Wang, Z. Ming, and L. Li, “Battery capacity design for electric vehicles considering the diversity of daily vehicles miles traveled,” Transportation Research Part C: Emerging Technologies, vol. 72, pp. 272–282, 2016.
4. J. G. Hayes, R. P. R. de Oliveira, S. Vaughan, and M. G. Egan, “Simplified electric vehicle power train models and range estimation,” in 2011 IEEE Vehicle Power and Propulsion Conference, pp. 1–5, 2011.
5. J. Dong and Z. Lin, “Stochastic modeling of battery electric vehicle driver behavior,” Transportation Research Record: Journal of the Transportation Research Board, vol. 2454, no. 1, pp. 61–67, 2014.
6. J. Brady and M. O’Mahony, “Modelling charging profiles of electric vehicles based on real-world electric vehicle charging data,” Sustainable Cities and Society, vol. 26, pp. 203–216, 2016.
7. Y. Huang, H. Wang, A. Khajepour, H. He, and J. Ji, “Model predictive control power management strategies for HEVs: A review,” J. Power Sources, vol. 341, pp. 91–106, Feb. 2017.
8. Zhao, Liang & Yao, Wei & Wang, Yu & Hu, Jie. (2020). Machine Learning-Based Method for Remaining Range Prediction of Electric Vehicles. IEEE Access. 8. 212423-212441. 10.1109/ACCESS.2020.3039815.
9. M. Bohlsen, “EV Company News for The Month of January 2019,” Seeking Alpha, 2019.
10. Bolovinou A, Bakas I, Amditis A, et al. Online prediction of an electric vehicle remaining range based on regression analysis. In: 2014 IEEE International Electric Vehicle Conference (IEVC). IEEE, 2014, pp. 1–8.
11. B. Zheng, P. He, L. Zhao, and H. Li, “A Hybrid Machine Learning Model for Range Estimation of Electric Vehicles,” in 2016 IEEE Global Communications Conference (GLOBECOM), pp. 1–6, 2016.
12. J. Hong, S. Park, and N. Chang, “Accurate remaining range estimation for Electric vehicles,” in 2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC), pp. 781–786, 2016.
13. R. O. Duda, P. E. Hart, and D. G. Stork, Pattern classiﬁcation. Hoboken, New Jersey, USA: John Wiley & Sons, 2012.
14. S. Grubwinkler, T. Brunner, and M. Lienkamp, “Range prediction for evs via crowd-sourcing,” in 2014 IEEE Vehicle Power and Propulsion Conference (VPPC), 2014, pp. 1–6.
15. C.-M. Tseng and C.-K. Chau, “Personalized prediction of vehicle energy consumption based on participatory sensing,” IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 11, pp. 3103–3113, Nov. 2017. [Online]. Available: <https://doi.org/10.1109/tits.2017.2672880>
16. T. Straub, M. Nagy, M. Sidorov, L. Tonetto, M. Frey, and F. Gauterin, “Energetic map data imputation: A machine learning approach,” Energies, vol. 13, no. 4, p. 982, Feb 2020. [Online]. Available: <http://dx.doi.org/10.3390/en13040982>
17. Finn, T. and McGee, R. (2012). Driving pattern identification for EV range estimation. Electric Vehicle Conference (IEVC), 2012 IEEE International. Greenville. USA, pp, 1-7.
18. Nilsson, M. (2011). “Electric Vehicle: The Phenomenon of Range Anxiety”. ELVIRE. Sweden.
19. <https://spritmonitor.de/>
20. He H, Cao J, Cui X. Energy optimization of electric vehicle’s acceleration process based on reinforcement learning. J Clean Prod 2020; 248: 119302.
21. Braun A, Rid W. Assessing driving pattern factors for the specific energy use of electric vehicles: A factor analysis approach from case study data of the Mitsubishi i–MiEV minicar. Transp Res Part Transp Environ 2018; 58: 225–238.
22. Felipe J, Amarillo JC, Naranjo JE, et al. Energy consumption estimation in electric vehicles considering driving style. In: 2015 IEEE 18th International Conference on Intelligent Transportation Systems. IEEE, 2015, pp. 101–106.
23. Wang B, Wang Y, Qin K, et al. Detecting transportation modes based on LightGBM classifier from GPS trajectory data. In: 2018 26th International Conference on Geoinformatics. IEEE, 2018, pp. 1– 7.
24. Yavasoglu HA, Tetik YE, Gokce K. Implementation of machine learning based real time range estimation method without destination knowledge for BEVs. Energy 2019; 172: 1179–1186.
25. S. Sautermeister, M. Falk, B. Baker, F. Gauterin, and M. Vaillant, “Influence of Measurement and Prediction Uncertainties on Range Estimation for Electric Vehicles,” IEEE Trans. Intell. Transp. Syst., vol. 19, no. 8, pp. 2615–2626, Aug. 2018.
26. De Cauwer C, Van Mierlo J, Coosemans T. Energy consumption prediction for electric vehicles based on real-world data. Energies 2015; 8: 8573–8593.
27. Lee C-H, Wu C-H. A novel big data modeling method for improving driving range estimation of EVs. IEEE Access 2015; 3: 1980–1993.
28. S. A. Birrell, A. McGordon, and P. A. Jennings, “Defining the accuracy of real-world range estimations of an electric vehicle,” in 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 2590–2595, 2014.
29. J. Hong, S. Park, and N. Chang, “Accurate remaining range estimation for Electric vehicles,” in 2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC), pp. 781–786, 2016.
30. Fiori C, Ahn K, Rakha HA. Power-based electric vehicle energy consumption model: Model development and validation. Appl Energy 2016; 168: 257–268.