

RESEARCH ARTICLE

Applicability of Regression-Based Machine Learning Models in Energy Consumption Prediction for Electric Two-Wheelers Based on Micro-Trip Approach

Azhaganathan Gurusamy¹, A. Janieshbhavan², Bragadeshwaran Ashok^{3*}, S. Priyanka⁴, C. Kavitha⁵

¹Department of Automobile Engineering, B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai-600048, Tamil Nadu, India

²School of Computer Science and Engineering, Vellore Institute of Technology, Amaravati-52223, Andhra Pradesh, India

³School of Mechanical Engineering, Vellore Institute of Technology, Vellore-632014, Tamil Nadu, India

⁴School of Electronics Engineering, Vellore Institute of Technology, Vellore-632014, Tamil Nadu, India

⁵Department of Electronics and Communication Engineering, Sreenivasa Institute of Technology and Management Studies, Chittoor-517127, India

ABSTRACT – Machine learning (ML) models are crucial in foreseeing the energy consumption (EC) of alternative powertrain vehicles, which is necessary for determining the driving range (DR) and eliminating the ‘range anxiety’ of commuters. This research aims to develop an appropriate ML model for predicting the EC of electric two-wheelers (E2Ws) using a dataset from real-world driving tests. Primarily, data on vehicle, geography, and atmosphere are compiled from 42 driving trips using a data acquisition system installed on an electric two-wheeler along a predetermined route within Vellore city. The data from these driving trips are segmented into 1815 micro-trips, and the input dataset is created with 20 feature variables and one target variable to estimate and validate the EC or DR of E2Ws. In addition, the seven regression-based ML models are trained and validated with the derived dataset. Furthermore, the selected ML models are compared in terms of their prediction ability using various assessment measures, including error matrices (mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE) and coefficient of determination (R^2). The comparison assessment shows that the ExtraTree regression model outperformed other ML models in predicting the EC of E2Ws. Moreover, this article offers valuable insights for commuters, manufacturers, and other stakeholders to achieve sustainable, energy-efficient electric mobility.

ARTICLE HISTORY

Received : 12th Dec. 2024
 Revised : 23rd Sept. 2025
 Accepted : 13th Nov. 2025
 Published : 21st Jan. 2026

KEYWORDS

Electric vehicle
Electric two-wheeler
Energy consumption
Driving range
Regression models

1. INTRODUCTION

Worldwide, the movement to reduce environmental impact through public transportation has gained momentum in recent years. Additionally, there is an ongoing debate over transitioning public transportation to clean and green alternatives, including the adoption of clean fuels and alternative powertrains [1], [2]. Among the various alternatives, Electric vehicles (EVs) have recently gained more attention in public transit due to their lower carbon emissions and reduced noise pollution compared to conventional vehicles. Also, various governments around the world provide additional incentives for purchasing EVs [3], [4]. However, the rate of penetration and adoption of EVs in public transportation is hindered by several factors, including longer charging times and higher capital costs, as well as short driving ranges and inadequate charging infrastructure [5], [6]. In particular, the driving range is one of the vital parameters to consider for commuters when purchasing EVs [7], [8]. Moreover, it is limited by the usable energy capability of the power source in EVs. Also, the driving range of EVs varies non-linearly with EC level across different dynamic conditions [9], [10].

Furthermore, these issues are collectively perceived as “range anxiety,” which can be alleviated by enhancing the prediction accuracy of the driving range and upgrading the energy-efficient powertrain and charging infrastructure for EVs [11], [12]. The accurate prediction of DR for EVs remains essential for optimizing and identifying ways to improve the energy economy and powertrain efficiency [13], [14]. Additionally, the results of the driving range prediction can assist customers in selecting an appropriate EV by providing a pricing comparison. Furthermore, it is essential to provide reliable information to commuters by improving the accuracy of predictions of the total or remaining DR of EVs, thereby enhancing their practicality and reliability [15], [16]. More importantly, DR prediction relies on the accuracy of energy consumption estimates, which are influenced by factors such as load, road conditions, driving conditions, and ambient conditions [17]. Accurate forecasting of EV energy consumption enables customers to plan their trips effectively, thereby reducing ‘range anxiety’. It is vital to meticulously assess the EC patterns of EVs and the mechanisms underlying them, along with potential affecting factors [18], [19].

Overall, various mathematical strategies have been employed to enhance the accuracy and trustworthiness of EV energy consumption prediction [20], [21]. The EC prediction model of EVs is categorized into three main groups based

on the computation approaches and model inputs. They are generally denoted as analytical, statistical, and computational energy consumption models [22]. The energy consumption prediction using analytical models is based on the efficiency of powertrain components and vehicle dynamic parameters. Furthermore, this model estimates the EC over the typical or provincially developed driving cycle based on the power required at the wheel under distinctive driving, road, and atmospheric conditions. Additionally, the statistical EC model is used to elucidate the relationship between EC and its relevant influential factors [23], [24]. The polynomial relations are developed for EC based on the vehicle, environmental, driver, road and traffic-related factors of EVs [25], [26]. Ultimately, Data-driven EC models often employ data analytical techniques to analyze arithmetical relationships among input factors and output EC without requiring an investigation into the actual process of electricity generation and usage of EVs [27], [28]. An increasing number of researchers are constructing data-driven models for estimating energy use using data mining ML approaches, due to advancements in computer science and artificial intelligence [29], [30]. In several studies, ML algorithms have been developed by analyzing data on traffic, driving, roads, and environmental factors to classify and quantify the degree of relationships among the EV's kinematic factors and its energy use [31], [32], [33], [34]. Every model consolidated input factors at a diverse level, enabling estimates to be generated with a range of input rates. In addition, the use of ML algorithms in predicting EC or DR helps identify information related to associated factors, which is particularly important for original equipment manufacturers (OEMs) in efficiently redesigning their EV powertrains[35].

Overall, the issue of predicting energy consumption or driving range for EVs has primarily been investigated using conventional ML models such as “multiple linear regression (MLR), random forest (RF), decision tree (DT) and support vector machine (SVM)” [36], [37]. The persistence of developing MLR models is to assess the correlation along with the kinematic parameters of EVs and their EC. This type of assessment is based on data regarding driver behavior, road features, traffic conditions, and environmental factors. Every ML model collected input features at a distinct level, enabling estimates to be generated with a range of input features [38], [39]. An investigation of EC breakdown is undertaken using data-driven methods to examine two novel, complex compounds: positive kinetic energy (PKE) and negative kinetic energy (NKE). The EC prediction model was constructed based on the deconstruction and feature selection analysis of these two components to showcase their utility [40]. Moreover, computational models of artificial neural networks (ANNs) are established to forecast the relationship between the EC of EVs and the various factors that affect it. This predictive model could be utilized as a prime function of input components, with “the apiece factor being assigned weights based on its relative prominence due to its learning ability [41]. Artificial neural networks (ANNs) can anticipate driving behaviour by categorizing driving patterns based on actual driving data [42], [43]. Then, a neural network (NN) is used to estimate unidentified microscopic driving characteristics before departure, followed by an MLR algorithm, to predict energy consumption. The utilization of this approach yielded a mean absolute error (MAE) ranging from 12% to 14% for the average energy consumption throughout trips [44]. Further, the state-continuous reinforcement learning (DQN) algorithm is proposed as an optimization strategy for the EC of EVs to optimize the EV acceleration process [45]. Further, k-nearest neighbours (KNN) and RF models are utilized to forecast the EC of electric buses in Shenzhen, China, using a five-month dataset. Their results demonstrated that the new RF model offers better predictions compared to existing models [46]. Other than this, advanced and novel machine learning algorithms like gradient boost decision tree (GBDT), extension to “Gradient Boosting (XGBoost) and light gradient-boosting machine (LightGBM)” have proven their superior performance compared to conventional approaches in the prediction of EC or DR [47], [48], [49]. Moreover, the choice of an estimation method relies on the availability of data and the specific application being addressed.

Prior research has thoroughly investigated the forecasting of the EC using diverse data-driven methodologies, which have shown enhanced accuracy relative to conventional modeling techniques. EC is often predicted by utilizing a limited number of significant parameters and a variety of regression and neural network-based ML techniques. The precision of EC prediction is fundamentally reliant on the quantity and nature of key factors, including environmental conditions, road characteristics, and driving behaviours, which are processed through the chosen ML models. As far as we know, no studies have specifically investigated the capacity of different regression-based ML models to accurately estimate the EC of E2Ws since most of the studies focused on estimating the EC for other EV classes. Additionally, most methods employ conventional cyclic methods, which require a large amount of data, thereby increasing computational complexity. To resolve these issues in EC prediction and address the evident gap in the current literature, this research aims to assess the efficacy of regression-based ML models in anticipating the EC or DR of E2Ws using a micro-trip approach and specified input characteristics. This study considers various driving-related parameters, such as route, mode and behavioural characteristics, to evaluate the overall precision of predicting the EC or DR. An appropriate ML model for real-time EC prediction is found by evaluating the performance matrices of the chosen ML models. Overall, the principal conclusions of this work provide valuable information about E2W energy utilization and prediction, along with its prominent factors, for different stakeholders, including original equipment manufacturers, users, and lawmakers.

Consequently, this will enhance the adoption of E2Ws in road transportation through the development of an energy-efficient powertrain and the provision of adequate charging facilities. After elaborating on the brief literature and research gap in the field of EC or DR prediction for EVs in Section 1.0, Section 2.0 describes the E2W used for driving tests with a data acquisition system, driving route, data collection procedure, and preprocessing of acquired data, and explains how the selected regression models are trained and validated with the derived datasets. In section 3.0, the performance of selected regression models is assessed with training and validation datasets. Then, their suitability in predicting the EC

of E2W is evaluated using different performance indicators. Ultimately, Section 4.0 presents the key outcomes, limitations, and future research directions in the prediction of EC for E2Ws (as carried out in the research study).

2. EXPERIMENTAL METHODOLOGY

The primary objective of this work is to develop an effective ML model that enhances the accuracy of forecasting the EC in E2Ws by considering multiple influential factors. This section provides a comprehensive description of the vehicle used, route information, data collection and preprocessing methods employed, and a detailed description of the selected ML models. The methodology for EC prediction is presented as a flowchart in Figure 1.

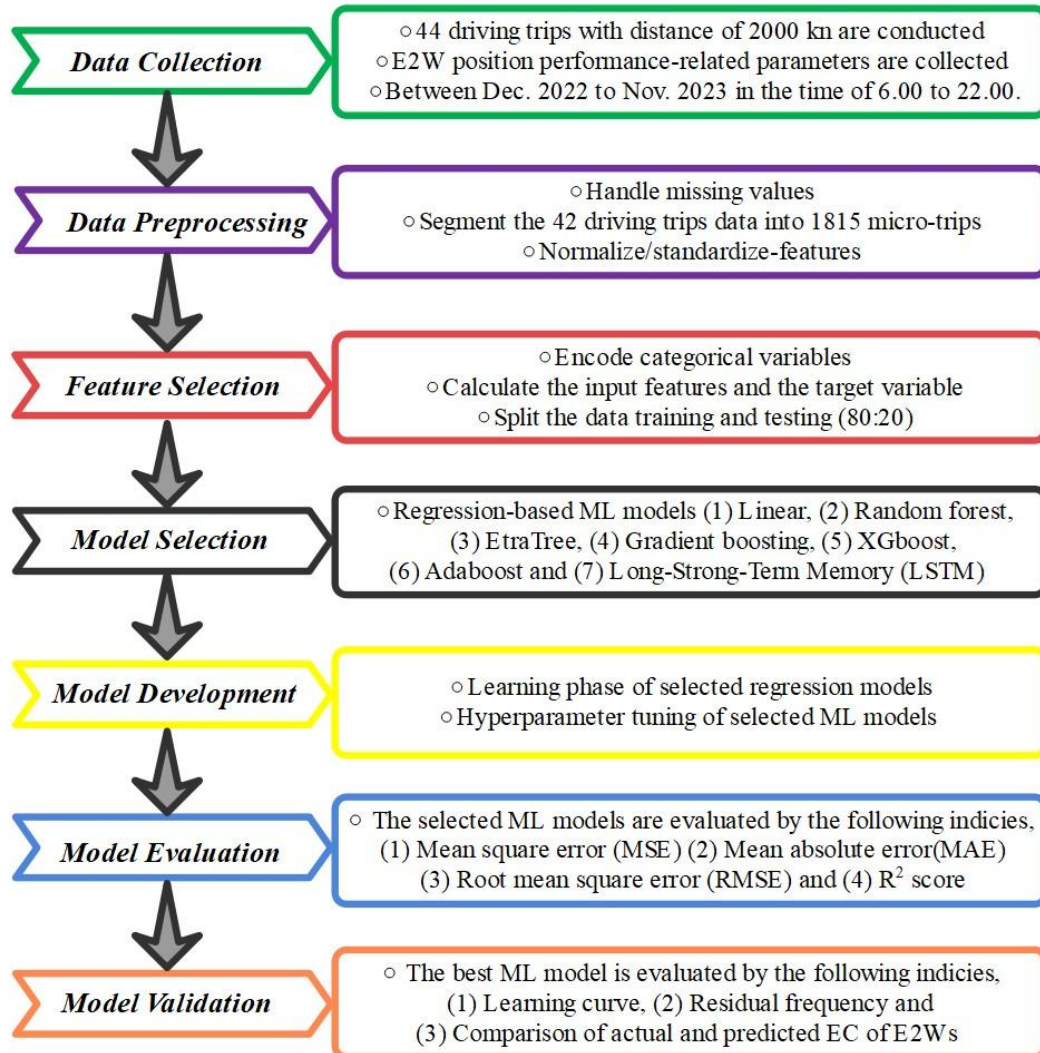


Figure 1. Flowchart for EC prediction using ML algorithms

2.1 Description of Vehicle and Data Acquisition Systems

Sensors and diagnostic tools have been used in EVs more often lately. To monitor the vehicle's performance and ensure safe and effective operation, this equipment is essential. They enable the assessment and tracking of critical vehicle components, including various controllers, sensors, and actuators that are required for their respective functions in vehicles. On-board diagnosis-II (OBD-II) is a sophisticated diagnostic platform that establishes communication with various systems in the EVs. A variety of study investigations have employed OBD-II devices to gather sensor-based data within EVs. Therefore, it is the most preferred tool in various real-time automotive research projects. This study utilizes the OBD-II for the real-time acquisition of driving data from electric scooters, which is India's most popular E2W segment among different users. Table 1 lists the technical parameters of the E2W (electric scooter) used in this investigation. To retrieve the vehicle data and atmospheric-related characteristics, the standalone data acquisition system (DAS) is interfaced with the electric scooter's 2-pin OBD adaptor. Additionally, GPS coordinates, altitude, and vehicle velocity data are collected using the MATLAB mobile application, which is installed on the smartphone to minimize energy utilization from the traction battery of electric scooters.

Table 1. Technical description of an electric scooter

Categories of attributes	Attribute type	Attributes	Values	Unit
Vehicle	-	Segment	E-scooter	-
	Physical	Dimensions (L*W*H)	189.1*73.9*111.4	cm
		Weight (Curb+load)	111.6+150	Kg
Battery	-	Chemistry	Li-ion	-
	Electrical	Total capacity	3.7	kWh
		Voltage	51.1	Volts
		'Charge capacity	72.4	Ah
Motor	-	Model	PMSM	-
	Performance	Torque (Maximum)	22	Nm
		Power - Max and rated	5.4 and 3.3	kW
Transmission	-	Model	Belt drive	-
	Mechanical	Ratio	7.8:1	-

2.2 Driving Route

To ensure the precise assessment of the EC or DR for E2Ws in different real-world driving scenarios, the driving routes should be carefully selected to replicate the diversified nature of road, environmental, and traffic conditions. For the driving test in Vellore City, India, three primary categories of road types (rural, freeway and urban) are chosen, and their topological view is shown in Figure 2. The driving route's overall length is approximately 48.65 km, comprising 14.6 km of rural (A to B), 9.7 km of freeway (B to C), and 24.35 km of urban (C to A) road segments, respectively.

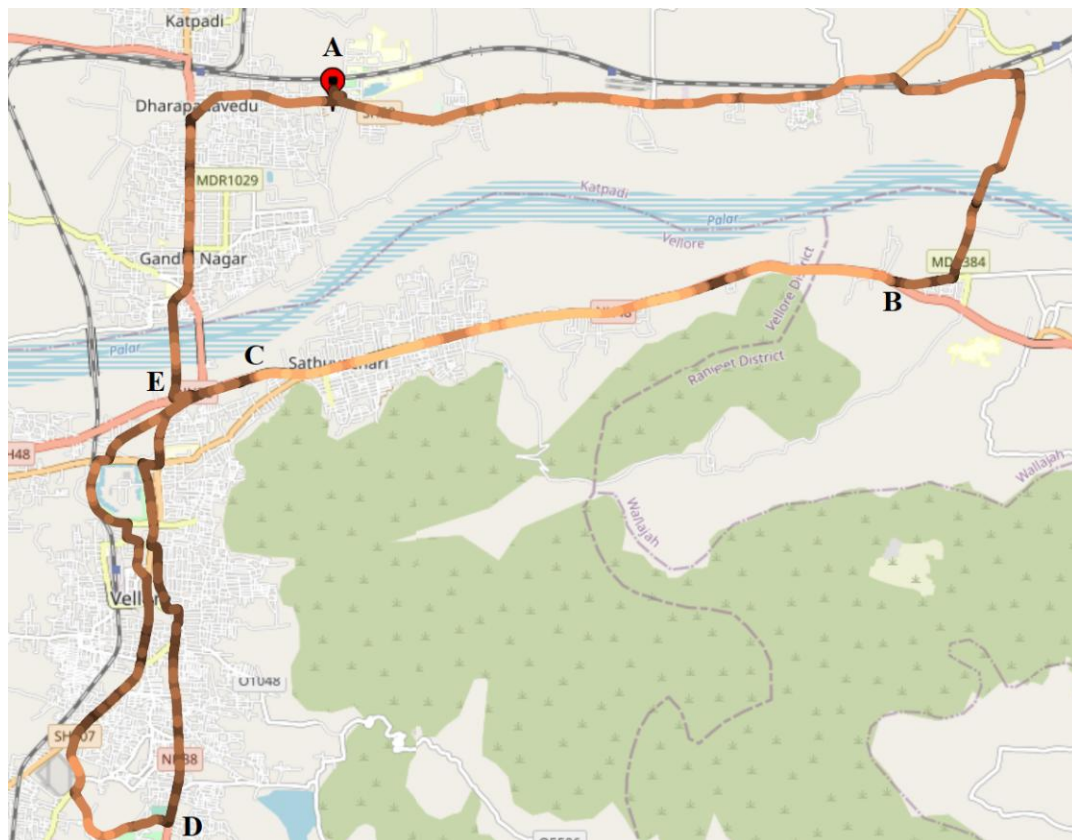


Figure 2. Topological view of the chosen driving route

The longer urban driving route is chosen over alternative road segments due to its greater traffic variations, which significantly influence the EC of E2Ws. The minimum, average and maximum elevations of the chosen driving route are 108, 209 and 259 m, with the maximum upgrade and downgrade values of 5.4 and 8.3 %. Also, the speed limit for two-wheelers is 80 km/h on highway roads and 60 km/hr on rural and urban road segments, which is regulated by the government of India. Moreover, the selection of these routes was based on the objective of obtaining a representative sample that encompasses a variety of road conditions and speed ranges, including elevation changes and curves during driving testing.

2.3 Data Collection

The vehicle position (GPS) and performance-related parameters are collected over 44 driving trips with an approximate total driving distance of 2000 km. To consider all the climatic variations in the selected driving regions, 11 driving tests are conducted in each season, like winter, summer, monsoon and autumn from December 2022 to November 2023, including weekdays and weekends. Also, the driving test is performed with a time span of 6.00 to 22.00. The MATLAB drive collects the vehicle speed, longitude, latitude, and altitude from the GPS device at a data acquisition rate of 1 Hz. Meanwhile, the atmospheric conditions and the data associated with the E2W performance parameters are collected in the data logger through DAS at a frequency of 2 Hz. Ambient temperature and humidity are collected to indicate the atmospheric conditions during the particular driving trips. Meanwhile, the E2W performance parameters associated with the battery like voltage (V_b), current (I_b), temperature (T_b), state-of-charge (SOC) and health (SOH) are collected with motor parameters like motor current (I_m), temperature (t_m), speed (N_m) and torque (T_m) along with the respective timestamp during all the driving trips. Further, the collected data are cleaned, and some of the performance, operating and road characteristics are derived at the preprocessing stage.

2.4 Training, Evaluation and Comparison of Regression-Based Predictive Models

In the pursuit of predicting battery energy consumption in E2Ws through regression-based ML models, a systematic research framework unfolds. This comprehensive journey encompasses pivotal stages, each contributing to the development and refinement of an adept model.

2.4.1 Dataset Extraction, Processing and Description

The very first part involves careful data extraction, a process fundamental to the empirical foundation of this research. Diverse datasets are amassed from an array of sources, encompassing E2W sensors, environmental variables, and driving patterns, with the overarching objective of crafting a robust dataset reflective of multifaceted scenarios influencing the EC of E2Ws. Then, each dataset entry is carefully labelled. Labels are carefully selected to coincide with real energy consumption values under certain driving situations. This accurate labelling serves as the foundation for providing the ML model with verifiable ground truth data during the training phase. Further, the raw data is systematically preprocessed. This critical stage involves noise correction, handling of missing values, and complex treatment of outliers. Such thorough preprocessing aims to deliver a clean dataset suitable for the keen eye of an ML model. In the end, the normalization step emphasizes the need for feature standardization. This step is required to avoid the dominance of specific characteristics due to built-in scale deviations.

The overall data preprocessing phase is divided into two stages: cleaning the data and preparing the dataset for the prediction of EC using selected ML algorithms. Moreover, the aggregated driving data might contain anomalies that comprise zero, outmoded, mislaid, and incorrect data attributable to various factors like failures of transmission lines and DAS, as well as loss of signal in the GPS device. To accurately forecast the EC, the gathered data must be cleaned before being analyzed by the selected ML algorithms. Initially, the data preprocessing process encompasses several processes, like examining and cleaning the noise and outliers, followed by normalization and standardization. Following the data cleansing process, the dataset is configured using in-house developed Python code to depict the target variable (EC/km) alongside the input parameters related to the road. Atmosphere, and driving styles with characteristics. For that, the driving trips are segmented as micro-trips based on the driving length, which is considered to be 1 km for each micro-trip. The entire diving trips generate 1815 micro-trips in total, which are categorized based on the different classifications of roads, including rural, motorway, and urban. For these micro-trips, the input features (equations 1-20) along with the respective target variable (equation 21) are derived based on Equations 1 to 21 inferred from our previous work [35]. The Equations (1)–(5), (6)–(9) and (10)–(13) are related to time (in seconds), speed (in m/s) and acceleration (in m/s^2)-related driving parameters. Further, Equations (14), (15)–(18), (19), and (20) are the energy, atmosphere(temperature in $^{\circ}C$ and humidity in %), road and driving mode-related parameters, and Equation (21) is a target variable (in Wh/km).

Total driving time,

$$T_{tot} = (t_2 - t_1) + \sum_{k=2}^n (t_k - t_{k-1}) \quad (1)$$

Cruising time,

$$T_{cru} = \left\{ \begin{array}{l} t_2 - t_1 \text{ } (|a_1| < 0.1 \& V_1 > 5) \\ 0 \text{ } (else) \end{array} \right\} + \sum_{k=2}^n \left\{ \begin{array}{l} t_k - t_{k-1} \text{ } (|a_1| < 0.1 \& V_1 > 5) \\ 0 \text{ } (else) \end{array} \right\} \quad (2)$$

Acceleration time,

$$T_{acc} = \left\{ \begin{array}{l} t_2 - t_1 \text{ } (a_1 > 0.1) \\ 0 \text{ } (else) \end{array} \right\} + \sum_{k=2}^n \left\{ \begin{array}{l} t_k - t_{k-1} \text{ } (a_1 > 0.1) \\ 0 \text{ } (else) \end{array} \right\} \quad (3)$$

Deceleration time,

$$T_{dec} = \left\{ \begin{array}{l} t_2 - t_1 \text{ } (a_1 < -0.1) \\ 0 \text{ } (else) \end{array} \right\} + \sum_{k=2}^n \left\{ \begin{array}{l} t_k - t_{k-1} \text{ } (a_1 < -0.1) \\ 0 \text{ } (else) \end{array} \right\} \quad (4)$$

Idling time,

$$T_{idle} = \left\{ \begin{array}{l} t_2 - t_1 \quad (V_1 = 0 \ \& \ a_1 = 0.1) \\ 0 \quad \quad \quad (else) \end{array} \right\} + \sum_{k=2}^n \left\{ \begin{array}{l} t_k - t_{k-1} \quad (V_1 = 0 \ \& \ a_1 = 0.1) \\ 0 \quad \quad \quad (else) \end{array} \right\} \quad (5)$$

Average trip speed,

$$\overline{V}_{trip} = \frac{Dist_{tot}}{T_{tot}} \quad (6)$$

Maximum trip speed,

$$V_{max} = \max\{V_1, V_n\} \quad (7)$$

Average driving speed,

$$\overline{V}_{drive} = \frac{Dist_{tot}}{T_{drive}} = \frac{Dist_{tot}}{T_{tot} - T_{idle}} \quad (8)$$

Standard deviation of speed,

$$\sigma_V = \sqrt{\frac{1}{n-1} \sum_{k=1}^n V_k^2} \quad (9)$$

Average acceleration,

$$\bar{a} = \frac{1}{T_{tot}} \sum_{k=1}^n a_k \quad (10)$$

Average '+ve' acceleration,

$$\overline{a}_{+ve} = \left(\sum_{k=1}^n \left\{ \begin{array}{l} 1 \quad (if \ a_i > 0) \\ 0 \quad (else) \end{array} \right\} \right)^{-1} * \left(\sum_{k=1}^n \left\{ \begin{array}{l} a_i \quad (if \ a_i > 0) \\ 0 \quad (else) \end{array} \right\} \right) \quad (11)$$

Average '-ve' acceleration,

$$\overline{a}_{-ve} = \left(\sum_{k=1}^n \left\{ \begin{array}{l} 1 \quad (if \ a_i < 0) \\ 0 \quad (else) \end{array} \right\} \right)^{-1} * \left(\sum_{k=1}^n \left\{ \begin{array}{l} a_i \quad (if \ a_i < 0) \\ 0 \quad (else) \end{array} \right\} \right) \quad (12)$$

Standard deviation of acceleration,

$$\sigma_a = \sqrt{\frac{1}{n-1} \sum_{k=1}^n a_k^2} \quad (13)$$

Positive kinetic energy,

$$P.K.E = \frac{1}{Dist_{tot}} * \sum_{k=2}^n \left\{ \begin{array}{l} V_i^2 - V_{i-1}^2 \quad (if \ V_i > V_{i-1}) \\ 0 \quad \quad \quad (else) \end{array} \right\} \quad (14)$$

Maximum humidity,

$$H_{max} = \max\{h_1, h_n\} \quad (15)$$

Average humidity,

$$H_{av} = \frac{1}{n} \sum_{k=1}^n h_k \quad (16)$$

Maximum temperature,

$$T_{atm_max} = \max\{T_{atm_1}, T_{atm_n}\} \quad (17)$$

Average temperature,

$$T_{atm_av} = \frac{1}{n} \sum_{k=1}^n t_k \quad (18)$$

Road types,

$$R_{type} = 0, 1 \ \& \ 2. \ for \ rural, \ highway \ and \ urban \quad (19)$$

Driving modes,

$$R_{type} = 1 \ \& \ 2 \ for \ ride \ and \ sport \ mode \quad (20)$$

Energy consumption per km,

$$EC_{km} = \frac{0.5}{3600} \sum_{k=1}^n (V_b * I_b) \quad (21)$$

where speed (V), acceleration (a), temperature (T), humidity (h), battery voltage (V_b), and current (I_b) are m/sec, m/sec², °C, %, Volts and Ampere.

2.4.2 Selection, Training and Performance Evaluation of Machine Learning Models

Choosing an ML model requires careful consideration. Whether it is the analytical elegance of Linear Regression, the delicate insights of Decision Trees, or the computational durability of Random Forests, model selection is customized to the needs of the research challenge. Additionally, the dataset has been appropriately partitioned into training (learning) and test (validation) sets. The initial strategy is employed to train the ML model, whilst the subsequent approach is utilized to systematically rate the model's predicted accuracy. Further, the model training phase occurs as the ML model delves into the subtle patterns embedded in the training data. Through the assimilation of these patterns, the model can predict energy usage depending on a set of input features. Then, the selected ML models are evaluated using different evaluation indices, such as mean squared error (MSE) and root mean squared (RMSE) errors, with R^2 used to determine how well the model's predictions fit actual energy consumption values. Followed by, in the event of unsatisfactory performance measurements, the research path leads to hyperparameter adjustment. This sophisticated process includes adjustment of the model's hyperparameters, which is aimed at producing more refined predictions. The completion of an acceptable model announces the appearance of the working model, which is balanced and effective in estimating battery energy usage using input features that are ready for deployment in real-time applications. In the end, this research concludes with a complete evaluation of the irrevocable ML model on the test set and validates its dependability and consistency, resulting in a solid assessment of its performance. The following assumptions related to data and ML models are considered for the EC prediction for E2Ws. They are,

- Acceptable and representative data - The larger and more diversified data is considered for training to include all the variations in real-time driving conditions.
- Linearity of relationships and stationarity - Input features and target variable have a simple functional relationship and are constant over time.
- Independence of observations and no multicollinearity - Data related to each micro-trip is considered as independent, and their input features are highly correlated.
- Linear Regression - Feature and target variable have a linear relationship, normally distributed errors, and constant variance throughout the predictions.
- Tree-Based Models - Data are split into regions to handle complex and non-linear relationships with comparable objectives
- Long Short-Term Memory (LSTM) – The time order is interrupted whenever the performance drops for sequential data with time-based dependencies.

The dataset is assigned into an 80:20 ratio for training and testing tenacities to conduct EC predictive analysis. The EC of E2Ws is first predicted using regression-based ML models such as Linear, random forest, EtraTree, gradient boosting, XGboost, Adaboost and Long-Short-Term Memory (LSTM). Then, their effectiveness on EC prediction in terms of accuracy is evaluated. The constructed models are assessed using numerous performance metrics by means of test data to determine their capacity to forecast the EC or DR of E2Ws.

3. RESULTS AND DISCUSSION

This work is intended to identify the most appropriate ML model for predicting the EC of E2Ws, which is a key feature in determining the remaining DR of E2Ws. In this section, the comparative analysis of different regression-based ML models is presented using different performance indices.

3.1 Performance Indices of ML Models

To evaluate the capabilities of selected ML models and find a suitable ML model for predicting the EC of E2Ws with a given dataset, the performance matrices, such as mean square (MSE) and absolute (MAE) errors and root mean square error (RMSE) with R^2 score were calculated for the selected ML models.

3.1.1 Mean Squared Error

The MSE is a vital statistic used in the field of ML analysis to assess the effectiveness of prediction models. Here, the MSE quantifies the mean squared variance between the predicted and actual EC data in the chosen dataset, as shown in Equation (22), and it assesses the accuracy of a model's estimates by quantifying the degree of alignment with the actual EC values.

$$\text{Mean square error (MSE)} = \frac{1}{n} \sum_{i=1}^n (EC_i - \widehat{EC}_i)^2 \quad (22)$$

Here, n is the number of data sets, EC_i is measured EC/km and \widehat{EC}_i is the predicted EC/km. The values of MSE for different regression-based ML models are compared in Figure 3. Ideally, an ML model with a lower MSE value is desirable for predictive analysis. 0.0191 Wh/km is the lowest MSE value, which is obtained by ExtraTree regression ML models. Also, it indicates that the ExtraTree model successfully avoids significant mispredictions. Next to the ExtraTree regression model, random forest and gradient boosting regression models perform better with lower MSE values of 0.0705 and 0.079 Wh/km. Meanwhile, the XGBoost regression model exhibits poor performance in the prediction of EC of E2Ws, with the highest MSE value of 1.1299 Wh/km

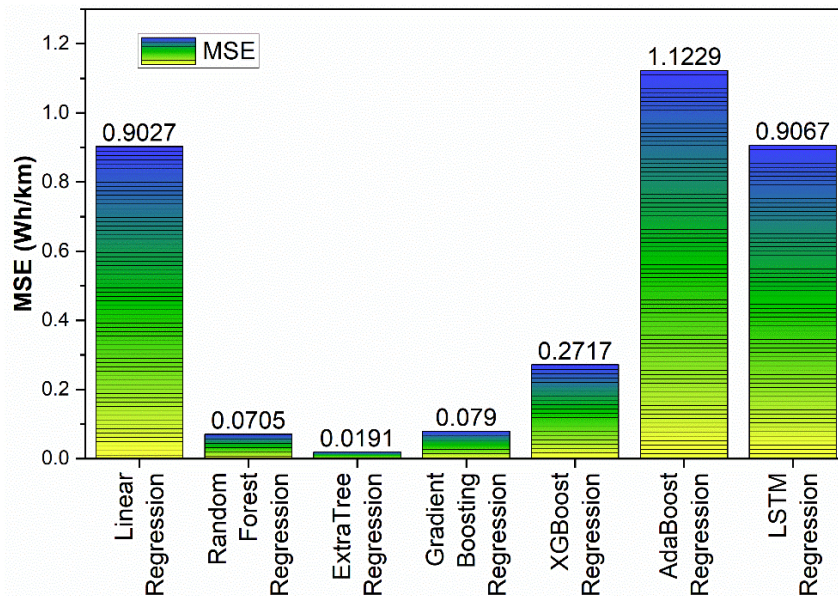


Figure 3. The MSE values of the regression models

3.1.2 Root Mean Square Error (RMSE)

The RMSE is mathematically represented as the “principal square root” of the mean of the squared differences between (\widehat{EC}_i) and (EC_i) as seen in Equation 23. The values of RMSE are primarily used in regression analysis and forecasting, with a strong emphasis on accuracy.

$$\text{Root mean square error (RMSE)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (EC_i - \widehat{EC}_i)^2} \tag{23}$$

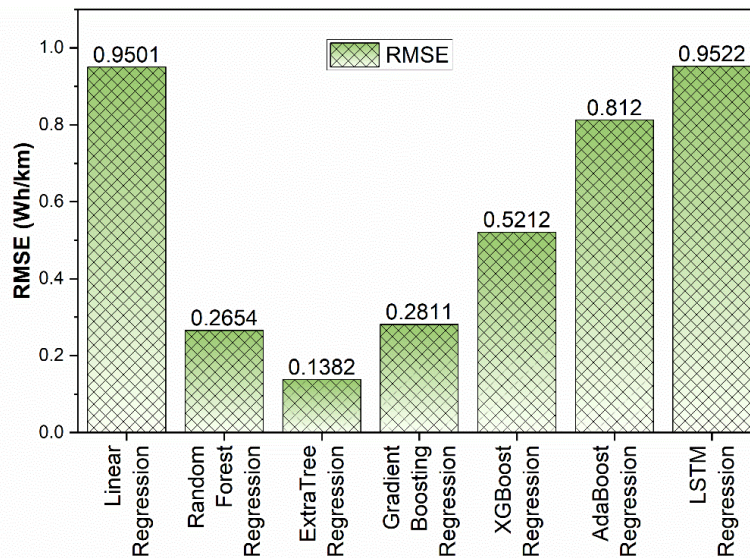


Figure 4. The RMSE values of the regression models

Figure 4 portrays the RMSE values of selected ML models after training the models with the selected dataset. For the best predictive ML model, the RMSE values should be lower than those of other ML models. It is understood that the ExtraTree regression model obtained the lowest RMSE value of 0.1382 Wh/km with a higher degree of accuracy than other regression models. Similar to MSE values, the random forest and gradient boosting regression models have the second and third lowest RMSE values of 0.2654 and 0.2811 Wh/km. Meantime, the LSTM regression model displays a poor protection performance with higher RMSE values of 0.9522 Wh/km than other regression models.

3.1.3 Mean Absolute Error

In supervised ML models. MAE defines the average of absolute errors for a group of predicted and observed EC values as a measurement of the magnitude of errors for the chosen dataset, as given in Equation 24.

$$\text{Mean absolute error (MAE)} = \frac{1}{n} \sum_{i=1}^n |EC_i - \widehat{EC}_i| \tag{24}$$

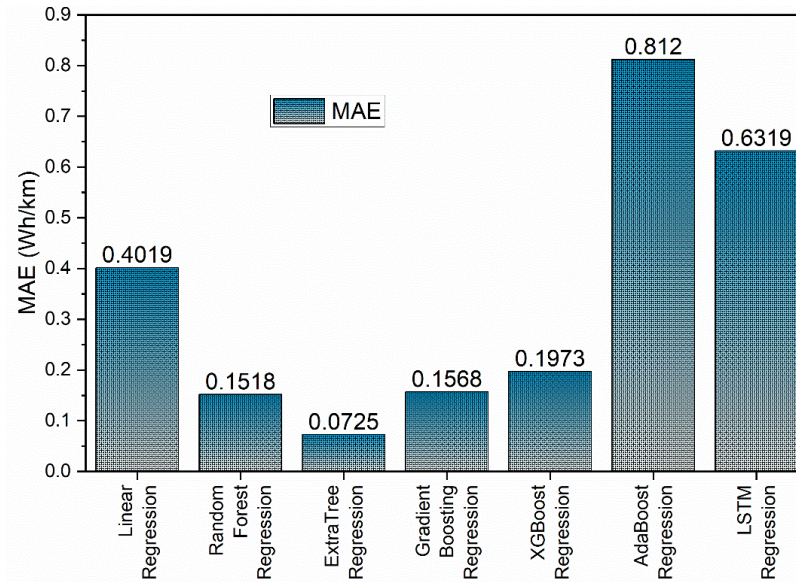


Figure 5. The MAE values of the regression models

The comparison of MAE values of different regression-based ML models is represented in Figure 5. Similar to evaluation indices like MSE and RMSE, the ML model with a lower MAE value is selected as the best predictive model with the chosen dataset. It is observed that the ExtraTree regression model has accomplished the lowest MAE value of 0.0725 Wh/km. Also, the MAE value is lower than the RMSE value of the ExtraTree model, and it is understood that the ML model has a low number of large errors. Next to that, the random forest and gradient boosting show better prediction accuracy with the MAE indices of 0.1518 and 0.1586 Wh/km, respectively. Meanwhile, the AdaBoost regression model displays a poor performance in the prediction of EC of E2Ws with a higher MAE value of 0.812 Wh/km.

3.1.4 R² Score

The R² score is another evaluation metric that defines how effectively a selected ML model fits data and anticipates appropriate EC values with a similar data set using Equation (25).

$$R^2 \text{ score} = 1 - \frac{\sum_{i=1}^n (EC_i - \widehat{EC}_i)^2}{\sum_{i=1}^n (EC_i - \overline{EC})^2} \tag{25}$$

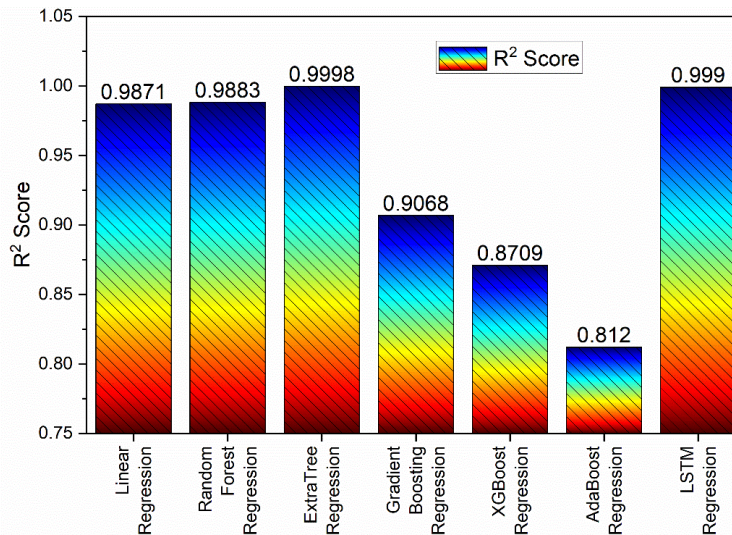


Figure 6. The R² score values of the regression models

The R² scores for ML models based on regression analysis are compared in Figure 6. The range of R² scores for the selected ML models is 0.812 to 0.9998 during the EC prediction analysis. Unlike other evaluation indicators, the ML model with the higher R² score is considered the most suitable predictive model for the particular predictive analysis. Here, the ExtraTree regression model achieves a higher R² value of 0.9998, or 99.98%, compared to other models. Also, it explains that 99.98% of the variance in the selected dataset. Additionally, LSTM and random forest regression models are the second and third best predictive models, with R² values of 0.999 (99.9%) and 0.9888 (98.88%), respectively. On the contrary, the AdaBoost regression model exhibits poor performance in predicting the EC of E2Ws, with an R² value

of 0.812, or 81.2%. Overall, an ML algorithm is deemed appropriate for the chosen EC or DR predictive analysis if it exhibits the slightest RMSE, MSE, and MAE values, along with the highest R^2 score. The ExtraTree regression model has lower RMSE, MSE, and MAE indices of 0.1382, 0.0191, and 0.0725, respectively, with a higher R^2 score of 0.9998 (or 99.98%) compared to other regression-based ML algorithms. From this, it is understood that the ExtraTree regression model outperforms other regression-based ML models in the estimation of EC of E2Ws. Next to that, the random forest and gradient boosting regression models perform well, except for the gradient boosting regression model (R^2 score is 0.9068). Meanwhile, it is determined that the AdaBoost model secured the last rank in all evaluation indices, such as RMSE, MSE, and MAE, as well as during the prediction of EC of EVs.

3.2 Validation of the Regression Models

After assessing the suitability of regression models with the selected dataset, the best regression model should be validated using several validation indices, such as the learning curve, residual distribution, and comparison of predicted EC with the actual EC of E2Ws. The entire validation process is carried out with the dataset allotted for the validation. A learning curve shows the response of a selected ML model metric to an increase in the amount of training instances. The ExtraTree regression ML model's learning curve for EC prediction is shown in Figure 7, utilizing 1460 training and validation dataset samples. It is observed that the learning curve of the selected dataset is flattened as the number of training samples considered for validation increases. Also, this model converges after 200 samples of the dataset, and it has less variation after 200 samples. Additionally, it is apparent that the breadth of both curves, as indicated by the shaded region, is reduced. The learning curve exhibits less bias and variation, achieving enhanced accuracy and convergence across both the training and testing datasets.

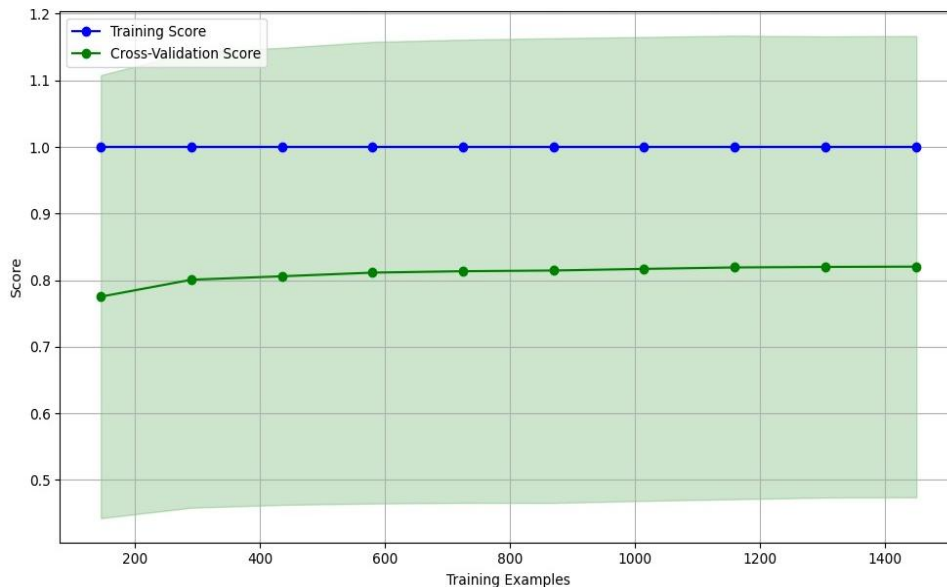


Figure 7. The ExtraTree regression ML model's learning curve

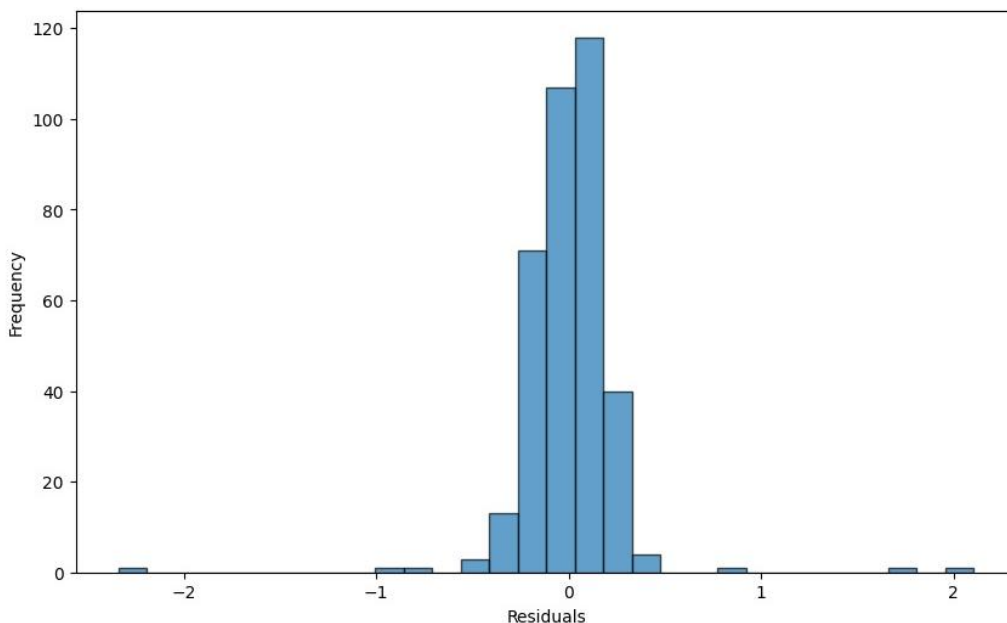


Figure 8. The results of the residual frequency for the best regression ML model

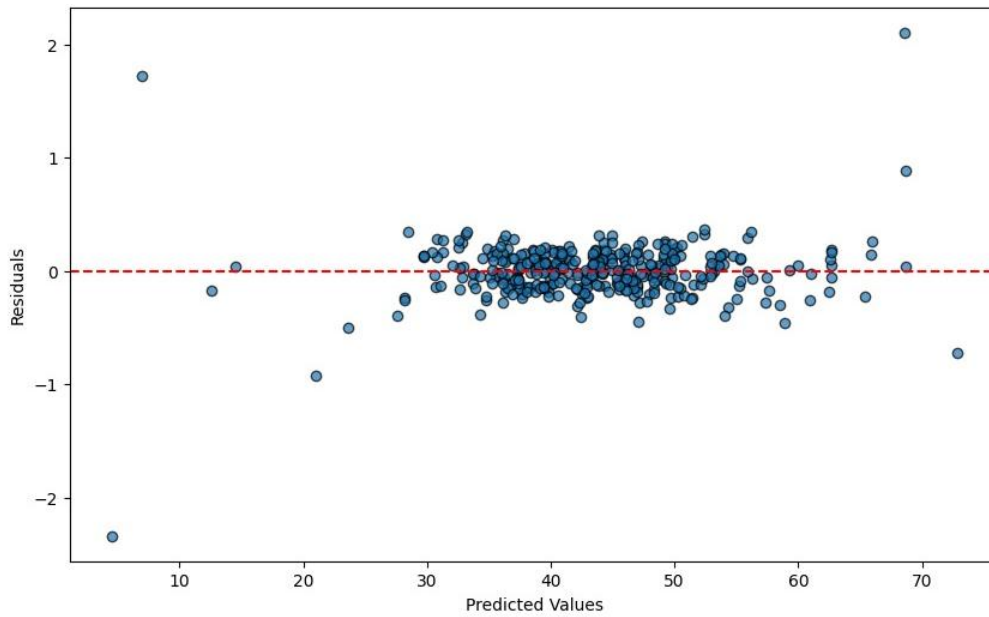


Figure 9. The ExtraTree regression model's residual plot

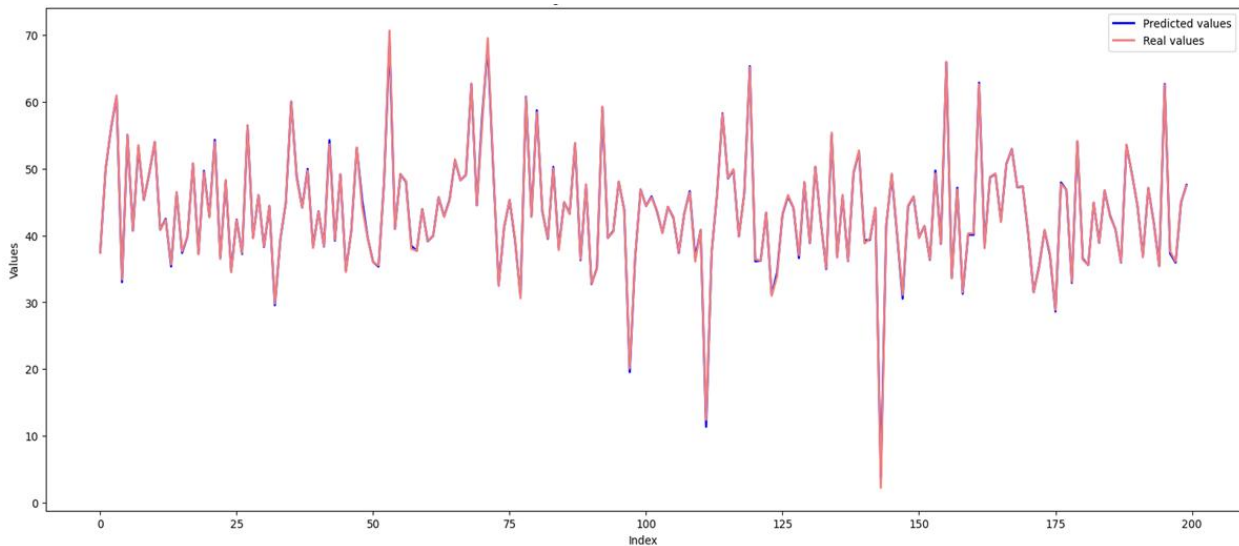


Figure 10. Comparison of actual and predicted values of EC for micro-trips

A series of statistical tests is implemented to authenticate the model's permanency and the efficacy of the model in accurately predicting the anticipated results. Figure 8 displays the outcome of a statistical test, which is determined by analyzing the inaccuracy (error) and frequency of the test data using the most accurate EC prediction models (ExtraTree regression model). The error sequence of the prediction model, with a mean of 0.024 and a standard deviation of 0.546, clearly resembles a normal distribution. The errors of the best EC predictive model are random and steadily biased. Additionally, Figure 9 illustrates the graphical representation of residual values about predicted EC values, alongside the distribution of residuals for both datasets. The ExtraTree regression model demonstrates a notable accuracy of 99.98%, surpassing other regression models in predicting the EC or DR of E2Ws. Furthermore, the residual values for both datasets trail a normal distribution pattern, with the peak of the distribution curve situated nearer to the distribution curve's highest point. Additionally, the residuals for the best EC predictive models are dispersed, which ensures model consistency in EC prediction, assuming homoscedasticity. However, the EC prediction accuracy is high for the ExtraTree ML model, and there is a minor outlier in Figure 9. This might be attributed to sensor noise, misalignment of the selected dataset, and the absence of the EC influential factors like battery temperature, road grade, and tyre pressure. Also, undefined driving events and limitations in deriving micro-trips might be influencing the outlier in EC prediction for E2Ws. Nevertheless, the predictive model exhibits higher robustness and accuracy, as these residuals are insignificant and random. To alleviate overfitting in the final model, the projected EC is juxtaposed with the observed EC of the micro-trips. Figure 10 illustrates the discrepancies between the projected and measured EC values for 200 micro-trips from the dataset used to evaluate the trained ML models. The green and orange lines denote the recorded and projected EC values, respectively. Generally,

the anticipated EC estimates nearly correspond with the actual numbers. It is noted that the ExtraTree regression model exhibits significant accuracy and demonstrates proficiency in predicting the EC or DR of E2Ws.

Table 2. Comparison of the best predictive model performance with other predictive models

Best predictive models	R ²	MAE (%)	RMSE (%)	MSE (%)	MAPE (%)	Ref.
LGBM-F	0.98	1.68	1.93	3.75	2.30	[22]
Tree structure-based XGBoost	-	-	15.50	-	12.6	[36]
Multiple linear regression	0.97	-	39.67	-	-	[47]
LightGBM	0.98	13.93	16.34	-	-	[48]
Deep neural network	-	6.6	12.20	1.80	15.8	[50]
ExtraTree regression	0.99	0.0725 (Wh/km)	0.1382 (Wh/km)	0.019 (Wh/km)	-	This study

Table 2 compares the performance of the best EC or DR predictive model in this study with that of the best predictive ML models from the existing literature. It is observed that the use of ExtraTree regression mode for EC or DR prediction for E2Ws yields better performance in terms of all fit and error indicators. Especially, the ExtraTree regression model outperforms other predictive models with a higher R² score of 0.99. Furthermore, the amalgamation of extremely low error metric values (MSE, RMSE, MAE) and diagnostic plots, with a near-perfect R² score showing no signs of overfitting or heteroscedasticity, effectively validates the model as the most ideal choice for predicting the EC of E2Ws based on the selected dataset. Overall, it is observed that the ExtraTree regression model is the most suitable predictive model with the selected feature variables. Additionally, compared to conventional cycle average methods, the micro-trip approach for EC prediction in E2Ws has provided better adaptability and flexibility, as it efficiently represents the non-linear driving behavior of E2Ws across all road segments. Also, it improves the E2Ws' EC prediction accuracy by incorporating the effects of all the driving scenarios like stop-and-go traffic, cruising and a variety of acceleration patterns and driving styles.

4. CONCLUSIONS

The feature variables, along with their respective target variable, are first obtained from the micro-trips, which consist of segmented driving data obtained during the driving test. Determining whether regression-based ML models are suitable for EC or DR prediction analysis of E2Ws is the primary goal of the research. The subsequent general conclusions emerge from the findings of the research conducted:

- i) The input comprises 20 feature variables related to road conditions, atmospheric factors, driving modes, and other relevant conditions. The identified and estimated variables, along with the target variable (EC/km), encompass 1815 micro-trips. The micro-trips originate from 42 driving trips, categorized by road segment: rural, freeway, and urban.
- ii) The performance measures, including MSE, RMSE, MAE, and R², are compared using both the test and training datasets for linear, random forest, ExtraTree, gradient boosting, XGBoost, AdaBoost, and LSTM regression-based ML models. The selected variables are used to evaluate the effectiveness of regression-based ML models in predicting the EC or remaining DR of E2Ws.
- iii) Overall, the ExtraTree regression-based ML model outperforms the other selected regression ML models, with MSE, RMSE, MAE, and R² values of 0.0191 Wh/km, 0.1382 Wh/km, 0.0725 Wh/km, and 0.998, respectively.

This paper aims to forecast the EC of E2Ws using micro-trip datasets and several regression-based ML models, as the EC directly affects the remaining or total DR of E2Ws. Nevertheless, the conducted research has some limitations, outlined below. (i) only one E2W variant is utilized, (ii) The impacts of driver-related attributes such as gender, age, educational and financial circumstances are not considered, (iii) the data collection is limited to specific regions and time periods in this study and the utilization of regression-based ML models for prediction of EC for E2Ws might be incapable to analyze the non-undeviating relationship between other EC influential parameters such as battery health, tyre pressure, HVAC energy usage road elevation or grade and battery degradation factors. Future studies should address these gaps to enhance prediction accuracy by identifying the most influential EC features in E2Ws and using hybrid ML algorithms. In summary, the study effort, presented in a structured manner, has the potential to enhance the accuracy of real-time EC prediction in on-road driving scenarios and significantly alleviate commuters' concerns about range anxiety. This is achieved by providing a solution for efficiently designing energy-saving routes in contemporary transportation.

ACKNOWLEDGEMENTS

The corresponding author, Bargadeshwaran Ashok, acknowledges the funding support for carrying out this research activity. This research work is supported and carried out as part of the funding from organizations like "Indian Council of Social Science Research (ICSSR/RPD/MN/2023-24/G/39)" and "British Council-Going Global Partnerships Industry Academia Grant 2023-24 (Contract Number: IND/CONT/G/23-24/20)".

CONFLICT OF INTEREST

The authors declare that they do not have any conflict of interest.

AUTHORS CONTRIBUTION

Azhaganathan Gurusamy (Conceptualization; Methodology; Writing - original draft; Writing - review & editing)

A. Janieshbhavan (Validation; Formal analysis; Data curation; Visualisation)

S Priyanka (Investigation; Resources; Software)

Bragadeshwaran Ashok (Conceptualization; Funding acquisition; Project administration; Supervision)

C Kavitha (Writing - original draft; Writing - review & editing)

REFERENCES

- [1] T. Pamuła and D. Pamuła, "Prediction of electric buses energy consumption from trip parameters using deep learning," *Energies*, vol. 15, no. 5, pp. 1747, 2022.
- [2] I. Hussain, K. B. Ching, C. Uttraphan, K. G. Tay, and A. Noor, "Evaluating machine learning algorithms for energy consumption prediction in electric vehicles: A comparative study," *Scientific Reports*, vol. 15, no. 1, pp. 1–20, 2025.
- [3] H. Vidhya and S. Allirani, "A literature review on electric vehicles: Architecture, electrical machines for power train, converter topologies and control techniques," in *2021 International Conference on Computational Performance Evaluation*, pp. 565–575, 2021.
- [4] J. A. Sanguesa, V. Torres-Sanz, P. Garrido, F. J. Martinez, and J. M. Marquez-Barja, "A review on electric vehicles: Technologies and challenges," *Smart Cities*, vol. 4, no. 1, pp. 372–404, 2021.
- [5] B. P. Adedeji, "Electric vehicles survey and a multifunctional artificial neural network for predicting energy consumption in all-electric vehicles," *Results in Engineering*, vol. 19, p. 101283, 2023.
- [6] X. Ma, J. Li, H. Cui, L. Cheng, Y. Ji, and J. Wang, "Electric vehicle range prediction considering real-time driving factors and battery capacity index," *Transportation Research Part D: Transport and Environment*, vol. 144, p. 104795, 2025.
- [7] W. Zhou, C. J. Cleaver, C. F. Dunant, J. M. Allwood, and J. Lin, "Cost, range anxiety and future electricity supply: A review of how today's technology trends may influence the future uptake of BEVs," *Renewable and Sustainable Energy Reviews*, vol. 173, p. 113074, 2023.
- [8] A. Gurusamy, B. Ashok, F. Alsaif, and V. Suresh, "Multifaceted multi-criteria decision making framework to prioritize the electric two-wheelers based on standard and regional driving cycles," *Energy*, vol. 305, p. 132401, 2024.
- [9] E. M. Szumska, R. S. Jurecki, "Parameters influencing on electric vehicle range," *Energies (Basel)*, vol. 14, no. 16, p. 4821, 2021.
- [10] E. Nagy and Á. Török, "Comparison of simulation- and regression-based approaches to estimating electric car power consumption," *Applied Sciences*, vol. 15, no. 2, p. 513, 2025.
- [11] X. He and Y. Hu, "Optimal mileage of electric vehicles considering range anxiety and charging times," *World Electric Vehicle Journal*, vol. 14, no. 1, p. 21, 2023.
- [12] C. Dong, Z. Xiong, C. Zhang, N. Li, Y. Li, N. Xie, et al., "A transformer-based approach for deep feature extraction and energy consumption prediction of electric buses based on driving distances," *Applied Energy*, vol. 380, p. 123941, 2025.
- [13] A. Gurusamy, B. Ashok, and B. Mason, "Prediction of electric vehicle driving range and performance characteristics: A review on analytical modeling strategies with its influential factors and improvisation techniques," *IEEE Access*, vol. 11, pp. 131521–131548, 2023.
- [14] K. Unni and S. S. Thale, "Regression algorithm-based residual range prediction and validation on EV travel data," *Australian Journal of Electrical and Electronics Engineering*, pp. 1-11, 2025.
- [15] A. Gurusamy and B. Ashok, "Establishment of electric two-wheeler driving cycle for energy economy and life cycle assessment under Indian Tier-2 city driving environments," *Sustainable Energy Technologies and Assessments*, vol. 66, p. 103826, 2024.
- [16] A. Gurusamy and B. Ashok, "Road segment and driving schedule effects in real-time operation on electric vehicle performance and operating cost analysis," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 239, no. 4, pp. 1181–1205, 2023.
- [17] Y. Jiang, J. Guo, D. Zhao, and Y. Li, "Intelligent energy consumption prediction for battery electric vehicles: A hybrid approach integrating driving behavior and environmental factors," *Energy*, vol. 308, p. 132774, 2024.
- [18] Z. Feng, J. Zhang, H. Jiang, X. Yao, Y. Qian, and H. Zhang, "Energy consumption prediction strategy for electric vehicle based on LSTM-transformer framework," *Energy*, vol. 302, p. 131780, 2024.
- [19] S. A. Hasib, M. M. Gulzar, A. Shakoore, S. Habib, and A. F. Murtaza, "Optimizing electric vehicle driving range prediction using deep learning: A deep neural network (DNN) approach," *Results in Engineering*, vol. 24, pp. 103630, 2024.
- [20] S. M. Miraftebadeh, M. Longo, and F. Foadelli, "Estimation model of total energy consumptions of electrical vehicles under different driving conditions," *Energies (Basel)*, vol. 14, no. 4, p. 854, 2021.
- [21] A. Di Martino, S. M. Miraftebadeh, and M. Longo, "Strategies for the modelisation of electric vehicle energy consumption: A review," *Energies*, vol. 15, no. 21, p. 8115, 2022.

- [22] Y. Wang, W. Chenlong, and H. Hongwen, "Estimating the energy consumption and driving range of electric vehicles with machine learning," *Journal of Physics: Conference Series*, vol. 2005, no. 1, p. 012131, 2021.
- [23] S. L. Prasad and A. Gudipalli, "An effective range estimation and state-of-charge to mitigate range anxiety in electric vehicles," *Heliyon*, vol. 11, no. 1, p. e41494, 2025.
- [24] J. E. G. Bustos, C. Baeza, B. B. Schiele, V. Rivera, B. Masserano, M. E. Orchard, et al., "A novel data-driven framework for driving range prognostics in electric vehicles," *Engineering Applications of Artificial Intelligence*, vol. 142, p. 109925, 2025.
- [25] H. Mediouni, A. Ezzouhri, Z. Charouh, K. El Harouri, S. El Hani, and M. Ghogho, "Energy consumption prediction and analysis for electric vehicles: A hybrid approach," *Energies*, vol. 15, no. 17, p. 6490, 2022.
- [26] W. Achariyaviriya, W. Wongsapai, K. Janpoom, T. Katongtung, Y. Mona, N. Tippayawong, et al., "Estimating energy consumption of battery electric vehicles using vehicle sensor data and machine learning approaches," *Energies*, vol. 16, no. 17, p. 6351, 2023.
- [27] D. Yang, H. Liu, M. Li, and H. Xu, "Data-driven analysis of battery electric vehicle energy consumption under real-world temperature conditions," *Journal of Energy Storage*, vol. 72, p. 108590, 2023.
- [28] J. Liu, Z. Wang, Y. Hou, C. Qu, J. Hong, and N. Lin, "Data-driven energy management and velocity prediction for four-wheel-independent-driving electric vehicles," *eTransportation*, vol. 9, p. 100119, 2021.
- [29] L. Zhao, W. Yao, Y. Wang, and J. Hu, "Machine learning-based method for remaining range prediction of electric vehicles," *IEEE Access*, vol. 8, pp. 212423–212441, 2020.
- [30] M. Adnane, A. Khoumsi, and J. P. F. Trovão, "Efficient management of energy consumption of electric vehicles using machine learning—A systematic and comprehensive survey," *Energies (Basel)*, vol. 16, no. 13, p. 4897, 2023.
- [31] A. Skuza, R. Jurecki, and E. Szumska, "Influence of traffic conditions on the energy consumption of an electric vehicle," *Communications - Scientific Letters of the University of Žilina*, vol. 25, no. 1, pp. 22-33, 2023.
- [32] A. Donkers, D. Yang, and M. Viktorović, "Influence of driving style, infrastructure, weather and traffic on electric vehicle performance," *Transportation Research Part D: Transport and Environment*, vol. 88, p. 102569, 2020.
- [33] F. Morlock, B. Rolle, M. Bauer, and O. Sawodny, "Forecasts of electric vehicle energy consumption based on characteristic speed profiles and real-time traffic data," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 1404–1418, 2020.
- [34] Y. Chen, G. Wu, R. Sun, A. Dubey, A. Laszka, and P. Pugliese, "A review and outlook on energy consumption estimation models for electric vehicles," *SAE International Journal of Sustainable Transportation, Energy, Environment, & Policy*, vol. 2, no. 1, pp. 79–96, 2021.
- [35] A. Gurusamy, A. Bokdia, H. Kumar, B. Ashok, and C. Gunavathi, "Appositeness of automated machine learning libraries on prediction of energy consumption for electric two-wheelers based on micro-trip approach," *Energy*, vol. 320, p. 135199, 2025.
- [36] J. Zhang, Z. Wang, P. Liu, and Z. Zhang, "Energy consumption analysis and prediction of electric vehicles based on real-world driving data," *Applied Energy*, vol. 275, p. 115408, 2020.
- [37] Y. Al-Wreikat, C. Serrano, and J. R. Sodré, "Driving behaviour and trip condition effects on the energy consumption of an electric vehicle under real-world driving," *Applied Energy*, vol. 297, p. 117096, 2021.
- [38] I. Miri, A. Fotouhi, and N. Ewin, "Electric vehicle energy consumption modelling and estimation—A case study," *International Journal of Energy Research*, vol. 45, no. 1, pp. 501–520, 2021.
- [39] T. Palit, A. B. M. M. Bari, and C. L. Karmaker, "An integrated Principal Component Analysis and Interpretive Structural Modeling approach for electric vehicle adoption decisions in sustainable transportation systems," *Decision Analytics Journal*, vol. 4, p. 100119, 2022.
- [40] Y. Xie, Y. Li, Z. Zhao, H. Dong, S. Wang, J. Liu, et al., "Microsimulation of electric vehicle energy consumption and driving range," *Applied Energy*, vol. 267, p. 115081, 2020.
- [41] I. Ullah, K. Liu, T. Yamamoto, M. Zahid, and A. Jamal, "Electric vehicle energy consumption prediction using stacked generalization: an ensemble learning approach," *International Journal of Green Energy*, vol. 18, no. 9, pp. 896–909, 2021.
- [42] S. Modi, J. Bhattacharya, and P. Basak, "Estimation of energy consumption of electric vehicles using Deep Convolutional Neural Network to reduce driver's range anxiety," *ISA Transactions*, vol. 98, pp. 454–470, 2020.
- [43] F. Marzbnai, A. H. Osman, and M. S. Hassan, "Electric vehicle energy demand prediction techniques: An in-depth and critical systematic review," *IEEE Access*, vol. 11, pp. 96242–96255, 2023.
- [44] S. K. Hora, R. Poongodan, R. P. de Prado, M. Wozniak, and P. B. Divakarachari, "Long short-term memory network-based metaheuristic for effective electric energy consumption prediction," *Applied Sciences*, vol. 11, no. 23, p. 11263, 2021.
- [45] H. He, J. Cao, and X. Cui, "Energy optimization of electric vehicle's acceleration process based on reinforcement learning," *Journal of Cleaner Production*, vol. 248, p. 119302, 2020.
- [46] R. Basso, B. Kulcsár, and I. Sanchez-Diaz, "Electric vehicle routing problem with machine learning for energy prediction," *Transportation Research Part B: Methodological*, vol. 145, pp. 24–55, 2021.
- [47] S. A. Hasib, Di. K. Saha, S. Islam, M. Tanvir, and M. S. Alam, "Driving range prediction of electric vehicles: A machine learning approach," in *2021 5th International Conference on Electrical Engineering and Information and Communication Technology*, 2021, pp. 1-6.
- [48] I. Ullah, K. Liu, T. Yamamoto, R. E. Al Mamlook, and A. Jamal, "A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability," *Energy & Environment*, vol. 33, no. 8, pp. 1583–1612, 2021.

- [49] P. Mei, H. R. Karimi, C. Huang, F. Chen, and S. Yang, "Remaining driving range prediction for electric vehicles: Key challenges and outlook," *IET Control Theory & Applications*, vol. 17, no. 14, pp. 1875–1893, 2023.
- [50] Y. Pan, W. Fang, and W. Zhang, "Development of an energy consumption prediction model for battery electric vehicles in real-world driving: A combined approach of short-trip segment division and deep learning," *Journal of Cleaner Production*, vol. 400, p. 136742, 2023.