International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Prediction of Electric Vehicle Charging Using Machine Learning

***[1]Konathala Divya,[2]Vankala Roshankumar, [3]Jaddu Nagamani, [4]Mamidi Yendunaidu, [5]Savara Philip***

[1,2,3,4,5] UG Student, GMR Institute of Technology, Rajam, Andhrapradesh, India

**ABSTRACT**

The most important aspect of transportation is electric vehicles. Electric vehicles are the foundation of a smart city transportation application. The lack of charging infrastructure is one of the most significant barriers to electric vehicle adoption. To address the problem of EV charging duration and consumption. To address the issue, we are employing Ml algorithms to predict charging analysis, which is beneficial to drivers. It shows how long it takes for the battery to charge fully and provides detailed information about an electric vehicle's energy consumption. The reason for using ML algorithms for predictive analytics is that they can train on even larger data sets and perform deeper analysis on multiple variables with minor deployment changes. We concentrated on the charging dataset in conjunction with energy consumption and session duration in this paper. In this case, we are combining two machine learning algorithms for prediction analysis. Ridge Regression. Logistic Regression The algorithms with the best predictive performances in terms of RMSE, MAE, R2, MSE for session duration and energy consumption. We emphasized the significance of evaluation matrices and algorithms in charging behavior.

**Keywords:** Electric vehicle, Machine learning, Charging prediction, Session duration, Energy consumption.

## 1. INTRODUCTION

The rapid expansion of electric vehicle production and demand. We are using a prediction model that it developed for managing and handling the charging needs of electric vehicles. Many studies have been conducted in recent years on the design of electric vehicle power engines, but we must now focus on the charging process and prediction analysis [1-4]. So here is a machine learning algorithm to predict the charging analysis of an electric vehicle. There are several algorithms available for predicting the outcome of a dataset analysis. To ensure the accuracy of electric vehicle charging predictions, we employ machine learning algorithms. In this project, we chose a dataset that contains complete information about the charging process with multiple instances that are collected and saved in a dataset [5-7]. Here we will introduce a charging prediction for multiple electric vehicles with multiple instances charging plugin and plug out hours, i.e. the time it takes the vehicle to fully charge and the main target is energy consumption in kwh. Based on this data, we predict a charging analysis.

### *1.1. Methodology*

We determine the process used for predicting electric vehicle charging in methodology. The problem is defined, the dataset is described, and the preprocessing steps are highlighted.

Formulate the problem: Assuming the $\boldsymbol{t_{con}}$ represents the connection time of a car which the cable plugin, $\boldsymbol{t_{discon}}$ represents the disconnect time of a car when it is plugout and $\boldsymbol{B_{session}}$ consider the session charging time. e represents the energy delivered to the car during session.v bvc

$$\boldsymbol{B_{session} = (t_{con}, t_{discon}, e)}$$

$$\boldsymbol{s_{dur} = t_{discon} - t_{con}}$$

We predict the session duration and session energy consumption of individual charging records in this dataset.

### *1.2 Bock diagram of prediction analysis*

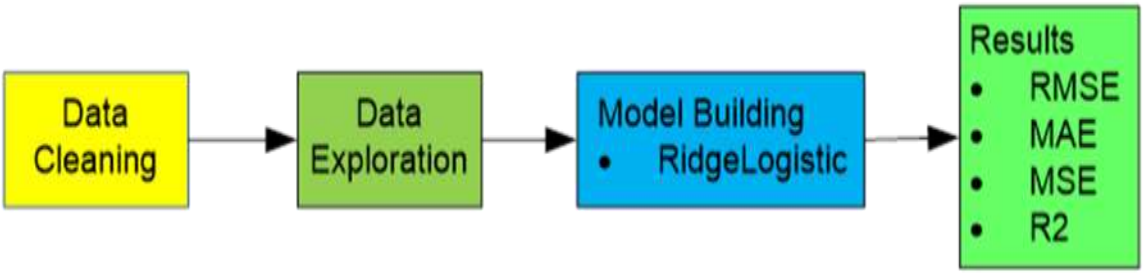


Fig. 1: Block Diagram

Above, Block diagram shows that the process of an prediction analysis of an chosen dataset which it includes four steps of processing.

- Data cleaning: Discard the unwanted data and check the null values and remove it convert the all values in int or float for better prediction analysis [8].
- EDA: Exploratory data analysis-here analyze the dataset in proper graphs for better understanding
- Model building: In this section we have to build the algorithm model for a dataset by using google Colab and Jupitor notebook
- Result: We achieved an evaluation matrix values RMSE, MAE, MSE, $R^2$

### *1.3 Dataset description:*

We will describe the dataset briefly and highlight its key features [9]. The electric vehicle charging prediction dataset describes how long it takes the vehicle to charge and how much energy is delivered to the vehicle in a given period of time. In this dataset, we primarily discuss session duration and energy consumption. The energy delivered to the car is the target variable.

**Table 1- Attributes description.**

| Attribute | Attribute description |
|---|---|
| Start plugin | The cable connects the charging station to the plug. |
| Start plugin_hour | In which hour the cable inserted to the charging station. |
| End plugout | The charging station's cable plug is removed. |
| End plugout_hour | In which hour the cable removed from the charging station |
| El_Kwh | Energy delivered to vehicle in kwh |

## 2. EXPLORATORY DATA ANALYSIS

Explore the dataset by correlating the dataset's features. The advantage of an EDA is that it provides complete dataset information in visualization format [10-12]. EDA for predicting the charging of electric vehicles by correlating the start plugin and the energy delivered to the vehicle.

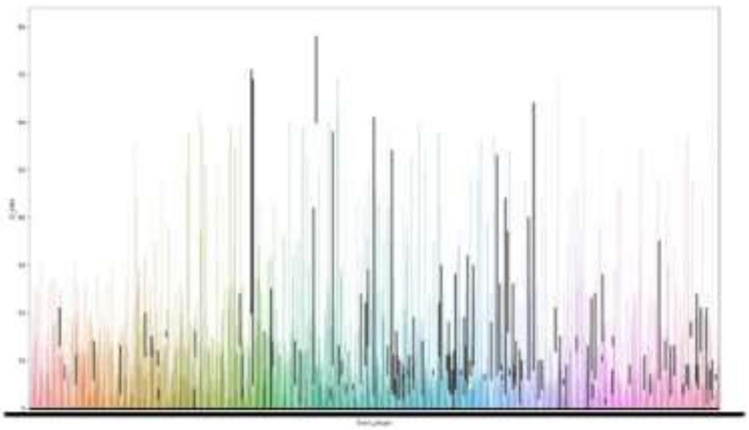

**Fig. 2: Energy Usage**

Figure -1: This figure shows that when the cable inserted into a charging station that means start plugin and how much energy delivered to a vehicle**.**

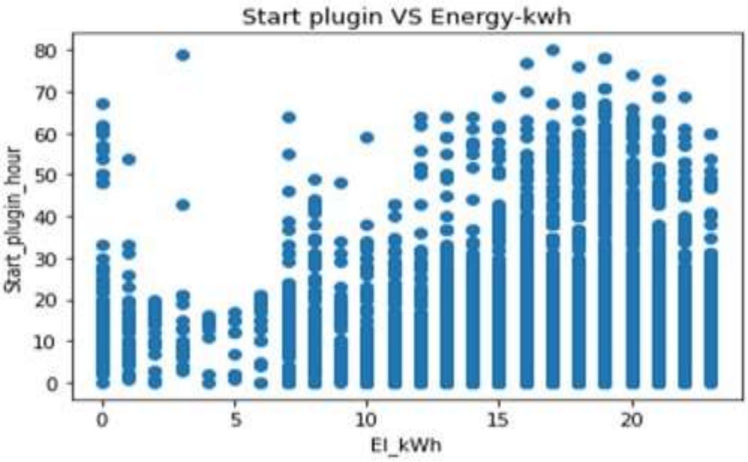


Fig. 3: Relation between the start plugin hour and session energy consumption

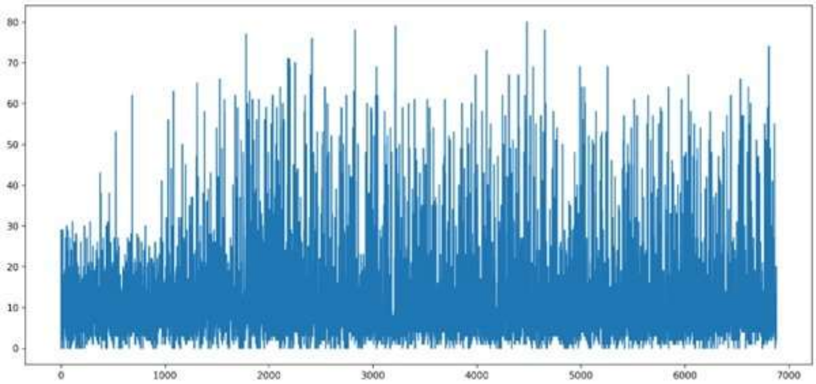


Fig. 4: Correlation of an electric vehicle and energy consumption.

## 3. MACHINE LEARNING ALGORITHM:

Machine learning can speed up data processing and analysis, making it a useful technology for predictive analytics program. Predictive analytics algorithms can train on larger data sets using machine learning [13-15].

In this paper, we use two machine learning algorithms to predict electric vehicle charging. Ridge regression

- Logistic regression

The both algorithms are regression model

***Ridge regression:***

- Ridge regression is used to create a simple model when the number of independent variables exceeds the number of samples in the dataset or when there is a correlation between the independent variables.
- Ridge regression is a technique used to estimate the coefficients of multiple regression models when the independent variables are highly correlated.
- Ridge regression is used to compare models when the dataset has multicollinearity.
- Hoerl and Kenner first proposed this theory in 1970.
- The advantage of ridge regression is that it avoids overfitting.
- It performs best on the training and testing datasets.
- Overfitting occurs when trained models perform better on training data but perform poorly on testing data.
- Linear regression and ridge regression are nearly identical, but linear regression establishes a relationship between the dependent variable and

one or more independent variables, whereas ridge regression is a technique used when the data is multicollinear.

- Ridge regression is one of the most fundamental regularisation techniques, but it is not widely used due to the complexity involved.
- Ridge regression employs a ridge estimator, which is a type of shrinkage estimator.
- It shrunk closer to the true value, reducing multicollinearity.

There are two basic regularizations: L1 and L2. L1 regularisation limits the size of the coefficient by adding an L1 penalty equal to the absolute value of the magnitude of the coefficients. Ridge regression belongs to the class of regression tools that use L2 regularisation, which equals the square of the magnitude of coefficients and does not eliminate any coefficient values.

***Regularization Techniques:***

- Penalize the magnitude of the coefficients of the features.
- Minimizing the errors between the actual and predicted observation values.
- Ridge regression model works by the performing the L2 regularization.
- Tunning factor is lamda controls the strength of the penalty value or terms.

***Logistic regression:***

Logistic regression was used in the biological sciences in the early twentieth century. It is also used in a variety of social science applications. Logistic regression is used when the dependent variable is categorical. Logistic regression is a supervised learning technique that is used several times in machine learning. There will be categorical or discrete values in the logistic regression output. For example, YES or NO, 0 or 1, true or false, and so on; those outputs are between 0 and 1.

***Types of logistic regression:***

The target variable has only two possible outcomes in binary logistic regression. For example, classifying as 0 or 1, passing or failing, and so on.

Multinominal Logistic regression: The target variable has three or more unordered categories. Predicting which food is preferred, for example.

Ordinal Logistic Regression: The target variable is divided into three or more categories that are ordered. For example, a movie can be rated from 1 to 5.

Logistic Regression equation is obtained from the Linear Regression equation.

Straight line equation:

**y**$=b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$

In logistic regression y can be lie between 0 and 1, we can divide the above equation by (1-y):

$$\frac{y}{1-y}; 0\ for\ y = 0, and\ infinity\ for\ y = 1$$

We need range between -∞ to +∞:

$$\log[\frac{y}{1-y}] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

## 4. Evaluation Matrices:

There are several measurements for the performance of regression model predictions, but here we are using dour measurements of a given dataset.

RMSE (root mean square error): The standard deviation of the prediction errors is referred to as the RMSE, and it indicates how concentrated the data is around the line of best fit.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}{n}}$$

The mean absolute error (MAE) of a model with respect to a test set is the mean of the absolute values of the individual prediction errors on all instances of the test set.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \bar{y}_i|$$

Coefficient of determination or $R^2$: The coefficient of determination is a number between 0 and 1 that represents how well a statistical model predicts an outcome.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}{\sum_{i=1}^{n}(y_i - \mu)^2}$$

## 5.Results and conclusion:

### *5.1. Results:*

Table-2:

| Evaluation | Ridge regression | Logistic regression |
|---|---|---|
| RMSE | 13.55209050208654 | 11.920842217211556 |
| MSE | 183.65915697674419 | 142.1064791676533 |
| $R^2$ | -185.78162821127086 | -483.273509979672 |
| MAE | 8.29142441860465 | 8.231631298801046 |

Graphical representation of evaluation matrices:

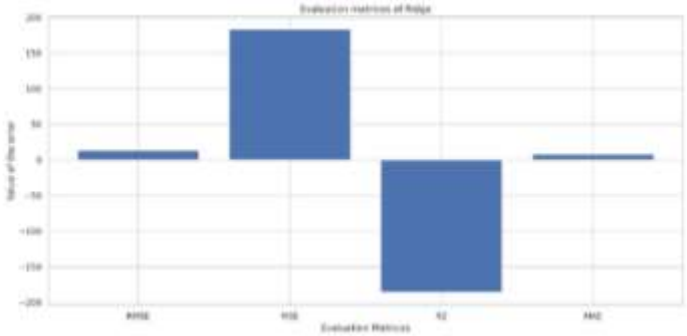

Fig. 5: Evaluation matrices of ridge regression

### *5.2. Conclusion*

In this paper, we presented a prediction model for an electric vehicle in terms of scheduling, specifically electric vehicle session duration and energy consumption (energy delivered to a vehicle while charging). We trained five popular machine learning models, as well as two ensemble learning algorithms, to predict charging behaviour. We use two regression models to predict here. We get results like RMSE, MAE, MSE, $R^2$ based on performance we conclude that Logistic regression gives best results than Ridge regression.

**References**

[1]. S. Shahriar, A. R. Al-Ali, A. H. Osman, S. Dhou, and M. Nijim, “Prediction of EV charging behavior using machine learning,” IEEE Access, vol. 9, pp. 111576–111586, 2021, doi: 10.1109/ACCESS.2021.3103119.

[2]. Y. Lu et al., “The application of improved random forest algorithm on the prediction of electric vehicle charging load,” Energies (Basel), vol. 11, no. 11, Nov. 2018, doi: 10.3390/en11113207.

[3]. Karthick, K, Premkumar, M, Manikandan, R & Cristin, R 2018, ‘Survey of Image Processing Based Applications in AMR’, Review of Computer Engineering Research, Conscientia Beam, Vol. 5, No. 1, pp. 12-19.

[4]. Y. Naitmalek, M. Najib, M. Bakhouya, and M. Essaaidi, “On the Use of Machine Learning for State-of- Charge Forecasting in Electric Vehicles,” 2019.

[5]. M. Dabbaghjamanesh, A. Moeini, and A. Kavousi-Fard, “Reinforcement Learning-Based Load Forecasting of Electric Vehicle Charging Station Using Q-Learning Technique,” IEEE Trans Industr Inform, vol. 17, no. 6, pp. 4229–4237, Jun. 2021, doi: 10.1109/TII.2020.2990397.

[6]. Li et al., “A prediction method of charging station capacity based on deep learning,” in Proceedings - 2017 4th International Conference on Information Science and Control Engineering, ICISCE 2017, Nov. 2017, pp. 82–84. doi: 10.1109/ICISCE.2017.27.

[7]. Loheswaran, K, Daniya T & Karthick K 2019, ‘Hybrid Cuckoo Search Algorithm with Iterative Local Search for Optimized Task scheduling on cloud computing environment’ Journal of Computational and Theoretical Nanoscience, American Scientific Publishers, Vol. 16, No. 5/6, pp.2065–2071

[8]. S. Khan, B. Brandherm, and A. Swamy, “Electric vehicle user behavior prediction using learning-based approaches,” Nov. 2020. doi: 10.1109/EPEC48502.2020.9320065.

[9]. X. Huang, D. Wu, and B. Boulet, “Ensemble learning for charging load forecasting of electric vehicle charging stations,” Nov. 2020. doi: 10.1109/EPEC48502.2020.9319916.

[10]. S. Deb, A. K. Goswami, R. L. Chetri, and R. Roy, “Prediction of plug-in electric vehicle’s state-of-charge using gradient boosting method and random forest method,” Dec. 2020. doi: 10.1109/PEDES49360.2020.9379906.

[11]. S. Ai, A. Chakravorty, and C. Rong, “Household EV charging demand prediction using machine and ensemble learning,” in Proceedings - 2nd IEEE International Conference on Energy Internet, ICEI 2018, Jul. 2018, pp. 163–168. doi: 10.1109/ICEI.2018.00037.

[12]. X. Gao, G. Yuan, and M. Zhang, “Fault Detection of Electric Vehicle Charging Piles Based on Extreme Learning Machine Algorithm,” in Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020, Mar. 2020, pp. 849–852. doi: 10.1109/ICCMC48092.2020.ICCMC-000157.

[13]. Karthick, K & Chitra, S 2017, ‘Novel Method for Energy Consumption Billing Using Optical Character Recognition’, Energy Engineering: Journal of Association of Energy Engineers, Taylor & Francis, vol. 114, no. 3, pp. 64-76, ISSN:1998595 https://doi.org/10.1080/01998595.2017.11863765

[14]. Kavaskar S, Sendilkumar S, Karthick, K, "Power Quality Disturbance Detection using Machine Learning Algorithm," 2020 IEEE International Conference on Advances and Developments in Electrical and Electronics Engineering (ICADEE), Coimbatore, India, 2020, pp. 1-5, doi: 10.1109/ICADEE51157.2020.9368939.

[15]. Karthick K, Aruna S K, Ravi Samikannu, Ramya Kuppusamy, Yuvaraja Teekaraman & Amruth Ramesh Thelkar, 2022, “Implementation of a Heart Disease Risk Prediction Model using Machine Learning”, Computational and Mathematical Methods in Medicine, Hindawi, vol. 2022, Article ID 6517716, 14 pages, 2022. https://doi.org/10.1155/2022/6517716