MACHINE LEARNING-BASED FORECASTING OF EV CHARGING BEHAVIOR

Rajender Kumar[1] and Himani Kumari[2]

[1]Assistant Professor and [2]B. Tech Student, Electronics and Communication Engineering Department, NIT Kurukshetra. India

ABSTRACT

The advancement of technology in transportation has made Electric Vehicles an integral part in Smart transportation within the realm of smart cities. Due to their environmentally friendly nature and contribution to reducing greenhouse emissions, electric vehicles (EVs) are increasingly being embraced as a smart mobility solution in the context of smart cities. The widespread adoption of electric vehicles (EVs) necessitates the establishment of a comprehensive network of EV charging stations. But, the expanded deployment of these stations poses challenges to existing power grid infrastructure. The strain on power grid stations is a direct consequence of the additional energy demand required for charging numerous EVs simultaneously. The proposed method in this research not only utilizes historical charging data but also weather conditions and information about vehicle in predicting EV session length and energy consumption by applying supervised, unsupervised machine learning techniques, as well as Artificial Neural Networks. The utilization of these advanced technologies holds immense potential in optimizing charging infrastructure and enhancing the overall efficiency of EV charging systems.

Supervised Machine Learning algorithms used are Random Forest, SVM and XGBoost. K-means Clustering and KNN are the unsupervised machine learning models that are applied. The model that performed best for session duration was through a voting ensemble of rf, svm, xgboost, with R2 score of 0.73 and SMAPE scores of 26.6% and for energy consumption R2 score is 0.61 and SMAPE score of 21.3% by stacking ensemble of rf and xgboost for session length and energy consumptions, respectively. This research also shows how weather and information about vehicles can influence EV charging behavior.

Keywords: *EV Charging, SVM, Machine Learning, Deep Learning, R2 score, SMAPE, MAE, XGBoost.*

I. INTRODUCTION

Transportation is an essential aspect of our daily lives, yet the use of traditional fuel-powered vehicles has detrimental effects on the environment. These vehicles contribute to pollution and consume significant amounts of natural resources, leading to their depletion. This depletion poses a threat to valuable resources that are vital for sustaining ecosystems and future generations. As a result, there is a growing need for sustainable alternatives to reduce environmental pollution and conserve valuable resources. Electric vehicles (EVs) are a promising solution for sustainability in today's world, where climate change is a major problem.

Research findings demonstrate that electric vehicles (EVs) possess the capacity to decrease carbon emissions by up to 45%. Over the years, the battery range and reliability of EVs have improved, contributing to their increasing popularity and higher satisfaction among EV owners. Government initiatives and the widespread deployment of charging stations have also enhanced driver flexibility and further encouraged EV adoption, positioning EVs as a cleaner alternative for transportation.

Notwithstanding the promising potential, a few challenges still need to be addressed, specifically the time taken by cars to charge and the charging infrastructure. Over the years, time has substantially reduced but still, it is considerably higher than time taken by Internal Combustion Engine vehicles for refueling. Although technologies like extreme fast charging and wireless charging hold promise, they are still grappling with various challenges and require additional time for widespread adoption. As a result, the limited capacity of charging infrastructure necessitates heavy reliance on public charging stations, as a result of the high power requirements of electric vehicles (EVs), power distribution grids experience strain, potentially leading to degradation and failures. To prevent such issues, it is crucial to avoid uncoordinated charging behavior. The optimal solution lies in the efficient management of charging station scheduling. Extensive research has been dedicated to smart scheduling using data-driven approaches, encompassing optimization strategies and metaheuristic methods. Additionally, factors influencing charging behavior from a psychological standpoint have been explored. For this, interviews had also been conducted with EV drivers. A thorough review, specifically centered around machine learning and data-driven approaches, has concluded that these methodologies are more effective for scheduling purposes.

In summary, as the urgency to combat climate change intensifies, the prominence of EVs and the effective management of EV charging behavior assume critical significance. The integration of machine learning and data-driven approaches presents valuable opportunities for optimizing EV charging infrastructure and facilitating the transition towards cleaner and more sustainable transportation systems.

A. LITERATURE REVIEW

Research in the field of EV charging behavior has been going on for a very long time. In this research, however, the charging behavior that has been focused on are EV charging session duration and energy consumption.

Previous studies have demonstrated that machine learning algorithms can generate accurate predictions for time series data [1], and hence can be used for predicting charging behavior.

In a different study [3], the arrival and departure times of electric vehicle (EV) commuters within a university campus was predicted by applying support vector machines. The mean absolute percentage error (MAPE) for arrival time prediction was 2.9%, and the MAPE for departure time prediction was 3.7%. In another study [4], an ensemble machine learning approach was employed to predict session length as well as energy consumption. The ensemble model consisted of three machine learning algorithms: support vector machines, random forests, and diffusion-based kernel density estimators (DKDEs). Two distinct datasets of charging information were used for training, one represented public charging and other residential. The ensemble learning model outperformed the individual models in both session length and energy consumption predictions. The session duration prediction had a reported symmetric mean absolute percentage error (SMAPE) of 10.4%, while the energy consumption prediction had a SMAPE of 7.5%. In the study presented in [6], a regression model was employed to forecast the energy demands based on data obtained from public charging stations in Nebraska, USA. Among the models evaluated, the XGBoost model exhibited superior performance compared to linear regression models: random forest, and support vector machine. In study [7], researchers used the k-nearest neighbor (k-NN) algorithm to forecast energy consumption on a university campus. The objective was to forecast energy usage for the next 24 hours by analyzing past energy consumption patterns. The best performance among the evaluated models was achieved with a k value of 1 (1-NN). The highest SMAPE recorded was 15.3%. In [8], the authors focused on predicting the energy requirements of a charging station for the following day by leveraging historical energy consumption data. Different algorithms were employed, including Support Vector Machines and Random Forest. Additionally, they explored a pattern sequence-based forecasting (PSF) approach [9], which involved clustering the days based on patterns and making predictions specific to each cluster. The method which utilized PSF gave the best results, achieving SMAPE value of 14.1%.

B. OBJECTIVES AND APPLICATION

While the aforementioned studies have demonstrated the successful application of machine learning in predicting session duration and energy consumption, the main emphasis has been on leveraging past charging data for analysis.

The suggested approach incorporates supplementary input factors such as vehicle details and weather conditions, in addition to past charging data, to enhance the analysis. Then the dataset is trained by applying multiple machine Learning algorithms(supervised as well as unsupervised) which are k-means clustering algorithm, Random Forest(RF), Support Vector Machine(SVM), XGBoost, ANN and KNN. Based on the results, we then selected the models with the best score and ensemble them in one model to give improved results. The model also shows that additional data of weather and vehicle can impact the EV charging behavior predictions.

It is also necessary to know why EV charging behavior prediction by applying ML is important. Using machine learning in this area has several practical applications. Here are a few examples:

1. **Infrastructure Planning:** Accurate predictions of energy consumption and session duration can help in planning and optimizing charging infrastructure. By analyzing the charging patterns and demand, stakeholders can determine the optimal number and placement of charging stations to ensure efficient usage and reduce the chances of congestion.
2. **Grid Management:** Predicting EV charging behavior can assist in managing the electrical grid load. By forecasting the energy demand from charging stations, utility companies can better plan and allocate resources to balance the load, avoid peak demand situations, and prevent grid instability.
3. **User Experience and Convenience:** Machine learning predictions can be used to enhance the user experience by providing accurate estimations of charging time and energy consumption. This information can be integrated into EV navigation systems or charging apps, allowing users to plan their trips and charging stops more effectively.
4. **Energy Management and Pricing:** Predicting EV charging behavior can help energy providers optimize their energy management strategies. By understanding charging patterns, they can offer dynamic pricing plans that incentivize off-peak charging, load balancing, and efficient utilization of renewable energy sources.
5. **Policy and Regulation:** Governments and regulatory bodies can leverage machine learning predictions to develop policies and regulations that support the growth of electric vehicles. Accurate insights into charging behavior can assist in determining incentives, subsidies, and infrastructure development plans.

II. BACKGROUND

In this section we shall learn about the various machine learning algorithms used to EV charging behavior. The different methods used are defined below.also the various metrics for evaluation are also described.

MACHINE LEARNING

It is a subset of artificial intelligence, which allows computers to learn without being explicitly programmed. Its applications span across various domains today. Mathematical modeling is ideal for solving problems that can be precisely modeled, offering error-free and exact solutions. However, in situations where mathematical modeling is challenging or not feasible, machine learning proves invaluable. By crafting algorithms and allowing them to uncover approximate models, machine learning facilitates the understanding of complex connections between input and the output.

I. SUPERVISED LEARNING

It is a category of machine learning where machines are trained using pre-labeled training data. In this approach, the machines learn to make predictions by observing the input data along with their corresponding correct output labels. The algorithm analyzes this labeled data to identify underlying patterns and trends, enabling it to establish the input-output relationship and select the most suitable model. Once trained, the model can be employed for making predictions. Unlabeled data is provided to the trained model, which then generates predicted labels..

In this research, the input data comprises a charging dataset, weather information, and information about vehicles, collectively referred to as features. The output parameters, also known as target variables, consist of session length and energy consumption. The target variables - session duration and energy consumption are labeled, hence supervised machine learning algorithms will be used. Also the target variable are continuous, hence regression models like Random Forest (RF), Support Vector Machine (SVM), XGBoost and KNN are being applied.

1. Random Forest (RF)

It is an ensemble technique that leverages the power of multiple decision trees to produce a single outcome. By combining the outputs of these individual trees, the algorithm enhances its overall performance, making it a powerful tool in machine learning. It utilizes the bagging method, where each tree is constructed from a different bootstrap sample (sampling with replacement). In regression tasks, as shown in fig 1 the predictions of all the trees are aggregated by taking their average value, while in classification tasks, the majority vote across the trees is considered. In this case, each decision tree in the ensemble provides a prediction, and the final result is determined by selecting the prediction with the highest accuracy.

$$Pf = \frac{(P1 + P2 + \ldots\ldots + Pn)}{T}$$

where

Pf represents the final prediction by random forest model, P1 is prediction from first decision tree, P2 is prediction from second decision tree and Pn is prediction from nth tree and T is total no. of trees.

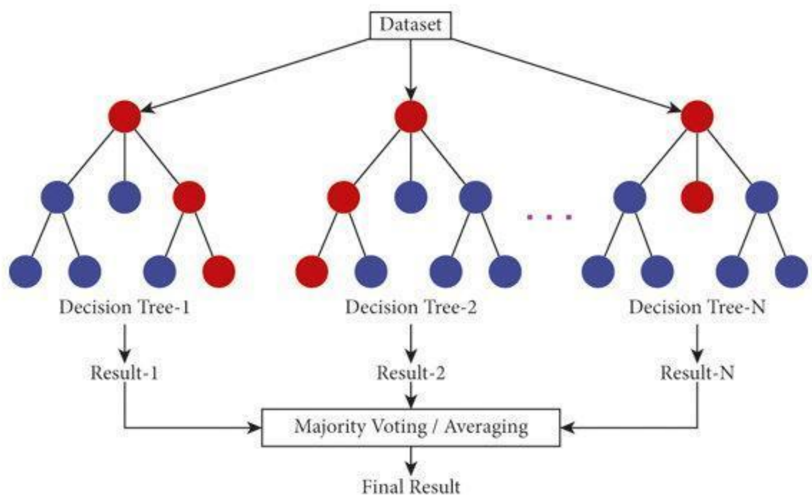


Fig 1. Illustration of Random Forest

2. Support Vector Machine

SVM is a popular classification as well as a regressor. SVR finds a hyperplane that best fits the data in a continuous plane. It finds optimal hyperplanes to maximize margins in classification and minimize errors in regression. Kernel functions like linear, polynomial, or RBF map data to higher dimensions for better separability. SVM handles complex boundaries but can be slow for large datasets. Still, SVM and SVR generalize well and are less prone to overfitting with proper regularization. Consider dataset size and computational demands when choosing SVM or SVR.

3. XGboost

XGBoost is a powerful machine learning algorithm known for its accuracy and scalability. It uses gradient boosting to train an ensemble of decision trees sequentially. Regularization methods like column subsampling and shrinkage prevent overfitting. XGBoost efficiently utilizes parallel computing for faster training. It handles missing data and provides insights into feature importance. Unlike Random Forest, XGBoost builds trees sequentially, reducing overfitting.

4. K-Nearest Neighbour

It is commonly utilized for classification and regression tasks. It is a technique that makes use of the similarity between data instances to make accurate predictions. It forecasts the value of novel data points by taking into account the values of its closest neighbors in the training dataset. Distance between a new point and all other points is calculated and k nearest neighbors are selected. The predictions for new data points are determined by finding the k nearest neighbors to the data point and using their class labels (for classification) or average/weighted values (for regression). KNN is known for its simplicity and ease of implementation, as it does not require a separate training phase. However, it can be sensitive to the choice of distance metric and the value of k. Despite its simplicity, KNN can achieve competitive performance, especially in scenarios with non-linear decision boundaries.

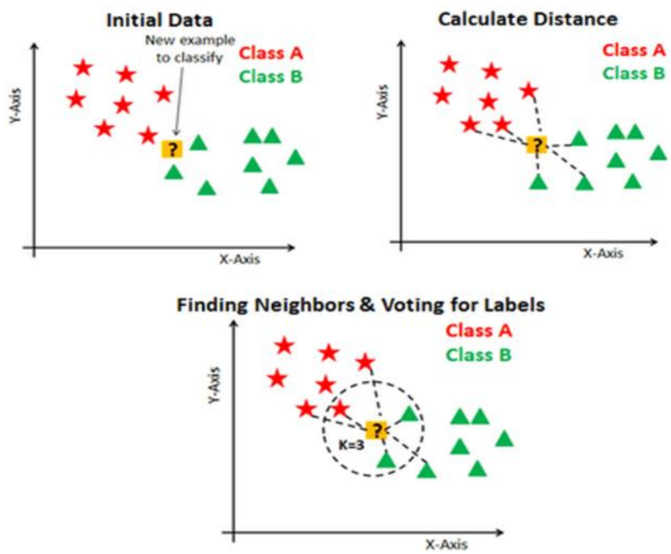


Fig 2. Illustration of KNN

II. UNSUPERVISED LEARNING

Unsupervised learning identifies patterns in datasets without labeled output variables. Cluster analysis is a technique used to group items based on common characteristics, often used within unsupervised learning to uncover underlying structures or patterns in the data without the need for labeled data.

Clustering techniques can be utilized to identify clusters of electric vehicle (EV) charging behaviors that exhibit similar patterns. For example, using features such as arriving time, avg. energy consumption, leaving time, avg session length etc, clustering algorithms can group EV charging instances into distinct clusters based on similar charging behavior. Unsupervised learning offers valuable insights into data patterns and can help identify hidden relationships and groupings. By discovering these clusters, it becomes possible to gain a deeper understanding of EV charging behaviors and develop targeted strategies or personalized recommendations for EV owners.

1. K-Means Algorithm

It is a commonly employed algorithm that strives to divide individual data points into k clusters. In the beginning, each data point is randomly assigned to one of the k center points. The algorithm then iteratively updates the cluster assignments by reassigning data points to the closest center based on new center calculations.One key consideration in K-Means clustering is determining the appropriate number of clusters, which is typically known beforehand or estimated using techniques such as the elbow method.

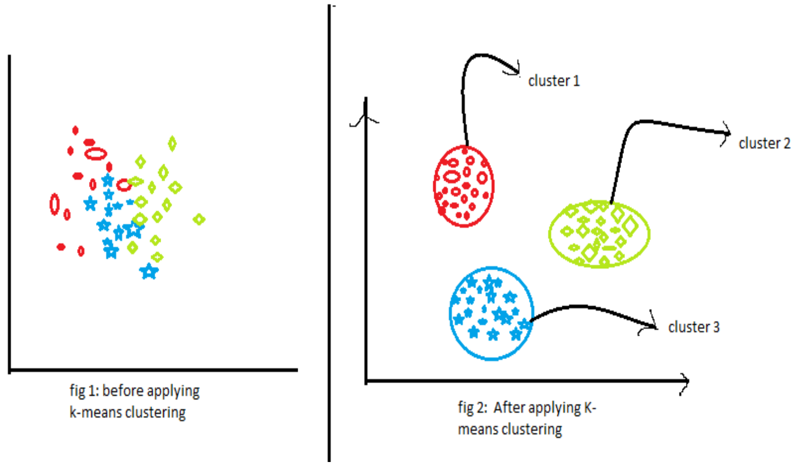


Fig 3. Illustration of K-Means Clustering

Deep Learning

DL is a distinct area within ML that utilizes artificial neural networks (ANNs) to facilitate advanced modeling capabilities. DL models differ from traditional ML models in that they are composed of multiple layers of learned functions. By utilizing a layered hierarchy of concepts, DL models can represent complex concepts in terms of simpler ones. Deep learning (DL) has made remarkable contributions to several domains, particularly in natural language processing (NLP) and computer vision. An artificial neural network (ANN) known as the multilayer perceptron (MLP) is a widely used technique in deep learning. MLPs can approximate non-linear relationships using input features and can be utilized for tasks involving classification and regression. The MLP consists of distinct layers that contribute to the overall model, including an input layer that receives input features, hidden layers that learn and extract representations, and an output layer that produces the final prediction.

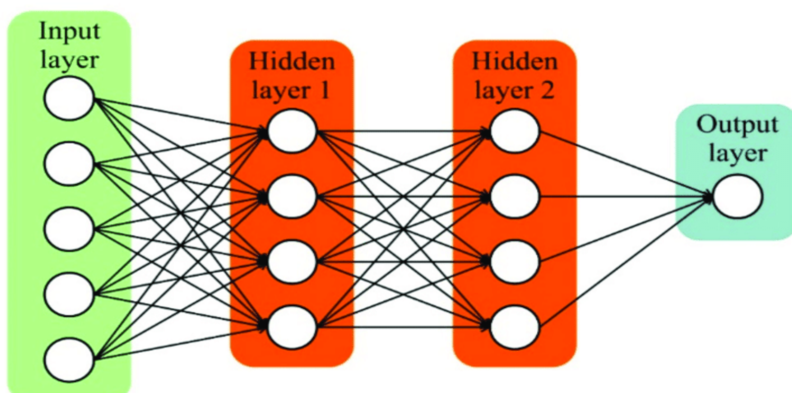


Fig 4. Illustration of ANN

Ensemble Learning

It is an approach in machine learning that combines various individual models, known as base models or weak learners, to construct a more robust and more accurate model. It leverages the principle that the combined functionality of multiple models can outperform any single model. The basic idea behind ensemble learning is to train a diverse set of base models on the same dataset and then aggregate their predictions to make a final prediction. The diversity among the base models can be achieved through various methods such as using different algorithms, varying the model parameters, or training on different subsets of the data. By combining the strengths of individual models, ensemble learning has become a key tool in data-driven decision making and has significantly advanced the field of machine learning.

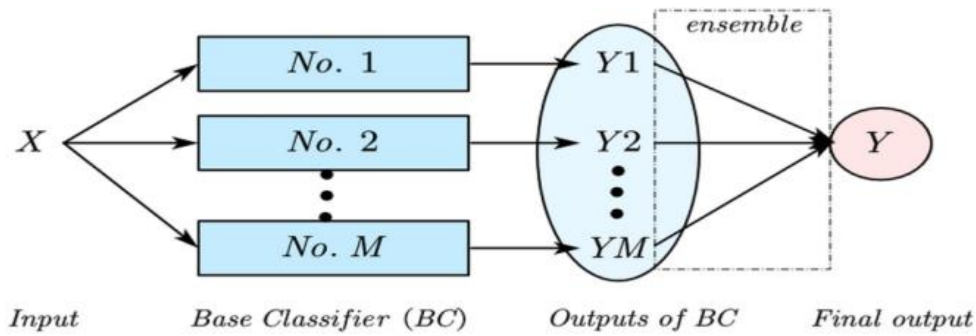


Fig 5. Illustration of Ensemble learning

III. EVALUATION METRICS

To assess the effectiveness of these models, evaluation metrics entail comparing the actual values (y) with the predicted values (ȳ) for multiple subsets of values in the dataset (n).

The metrics are described below:

1. **Mean Absolute Error (MAE):** It measures the average absolute difference between the original and predicted values across all data points. It provides a straightforward measure of the model's prediction accuracy without considering the direction of errors.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Actual(i) - Predicted(i)|$$

2. **Root Mean Squared Error (RMSE):** It is a measure of the difference between predicted and actual values. It is calculated by taking the square root of the average of the squared differences between the predicted and actual values. RMSE penalizes larger errors more heavily than mean absolute error (MAE), because it squares the differences before taking the average.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(Predicted(i) - Actual(i))^2}{n}}$$

3. **Coefficient of Determination (R^2):** R-squared (R^2) is a statistical measure that quantifies how well a regression model fits the data. It is a proportion of the variance in the dependent variable that is explained by the independent variable(s) in the model. R-squared values range from 0 to 1, with 1 indicating a perfect fit between the model and the data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Actual(i) - Predicted(i))^2}{\sum_{i=1}^{n} (Actual(i) - ActualAverage)^2}$$

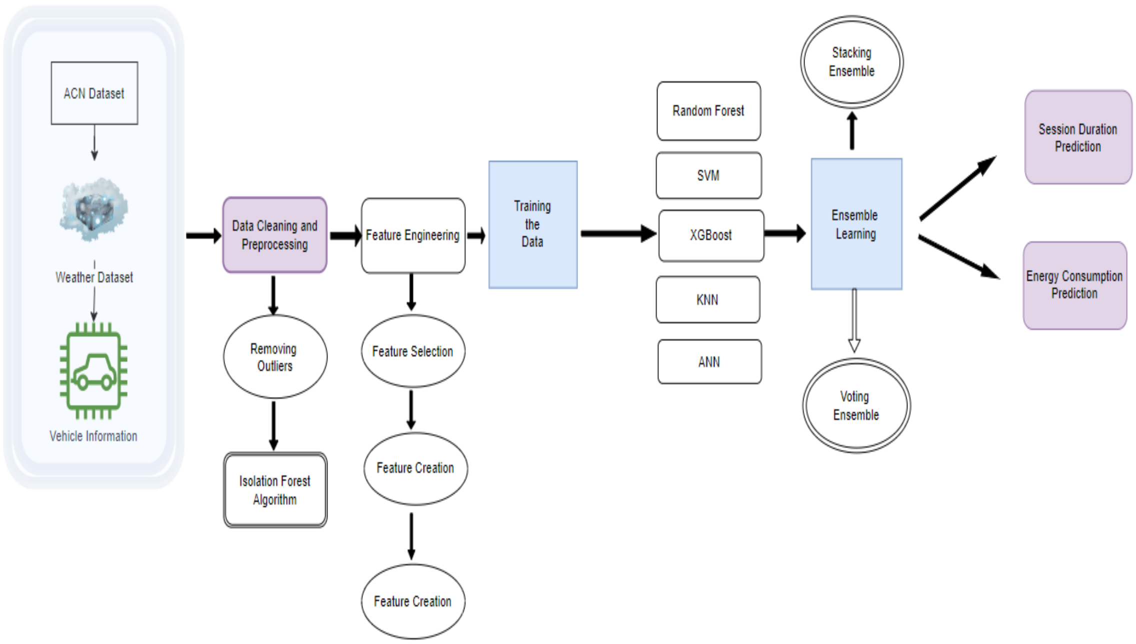
4. **The Symmetric Mean Absolute Percentage Error (SMAPE):** It is a metric used to assess the accuracy of predictions in forecasting and regression tasks. It measures the percentage difference between the predicted and actual values, taking into account the magnitude of both values. SMAPE provides a symmetric and scale-independent measure of the relative error.

$$SMAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|Actual(i) - Predicted(i)|}{(|Actual(i)| + |Predicted(i)|)/2} \times 100$$

If RMSE, MAE, and SMAPE have lower scores, it indicates more accurate predictions, as they reflect smaller differences between the predicted value (ȳ) and the actual value (y). These metrics are preferred when the goal is to minimize the prediction error. In contrast, the R^2 value measures the goodness of fit for regression models and ranges from 0 to 1, with 1 indicating a perfect fit. Higher R^2 values are desirable as they indicate better model performance in capturing the variation in the dependent variable.By utilizing these evaluation metrics, we can make informed decisions and select the most suitable model for accurate EV charging behavior prediction.

IV. METHODOLOGY

Here, the approach that is employed for predicting charging behavior has been described. We begin by formulating the problem and defining the specific objectives of our prediction task. Next, we provide details about the dataset used in our study, including its source and relevant characteristics. We then discuss the preprocessing steps applied to the dataset, which involve data cleaning, normalization, and feature selection, to ensure the quality and suitability of the data for training the learning models. Finally, we elaborate on the methods utilized for training the predictive models, outlining the specific algorithms or techniques employed and any relevant considerations in the model selection process.



A. EV Charging Dataset

Adaptive Charging Network (ACN) dataset[10] has been used int this study, which captures real-time records from charging stations in California. The ACN dataset is a valuable resource for researchers studying electric vehicle (EV) charging behavior, as it provides detailed information about the time, duration, and power of EV charging sessions. The charging events used are recorded from April 2018, enabling the analysis of charging patterns across various time periods. Weather statistics are taken from the local weather station at Caltech, USA. Satellite data can also be used but the data collected from local weather stations is more accurate than data from satellites[11]. Due to its higher accuracy, the weather data obtained from the local station is primarily given priority.

Nevertheless, the data obtained from local weather stations may contain missing values and may not include crucial variables like snowfall and rainfall. To address this, the missing values and additional variables are incorporated into the local station data [12] by integrating relevant information from the MERRA-2 weather data. By combining the local station data with the supplemental information from MERRA-2, a more comprehensive and complete weather dataset is created for use in this study. This approach ensures that the weather data used in local station and the broader coverage provided by satellite data, thus enhancing the reliability and applicability of the weather analysis in relation to EV charging behavior. The study collects additional data on electric vehicles, including battery pack capacity and fast charge speed, through a mobile application that users can access by scanning a QR code. In cases where clients don't utilize the portable application, default values are produced for these areas without connecting client identifiers. By analyzing this vast dataset, the study aims to gain insights into EV charging behavior and how external factors affect charging patterns.

B. Description of EV charging behaviour

EV charging behavior refers to the patterns, preferences, and habits exhibited by electric vehicle (EV) users in relation to charging their vehicles. In this study, two key parameters of EV user behavior that are taken into consideration are the session duration and the departure time.

$EV_{ub} \triangleq (t_{con}, t_{discon}, e_{con})$ - (1)

where, EV_{ub} : EV charging behavior

t_{con}: connection time

t_{discon}: disconnection time

e_{con}: energy consumed

(i) Session duration: session duration refers to the duration of time that an electric vehicle is connected to a charging station and actively receiving a charge.

$S_{dur} \triangleq t_{discon} - t_{con}$ - (2)

where, S_{dur} : session duration

(ii)Energy consumption: It refers to the amount of electrical energy consumed by an EV during its operation, particularly while driving and charging.

C. Data Cleaning and Pre- processing

These are crucial steps in ensuring the generalization and performance of machine learning models. One key aspect is handling missing values, as many algorithms cannot handle datasets with missing entries. In this work, missing entries are addressed by simply dropping them from the dataset.

Another important consideration is dealing with outliers, which can have a negative impact on model performance. Outliers are data points which deviate from the rest of the observations in a sample. To identify and remove outliers, various techniques can be employed. One commonly used approach is visualizing the variables using box plots, which can reveal extreme values.

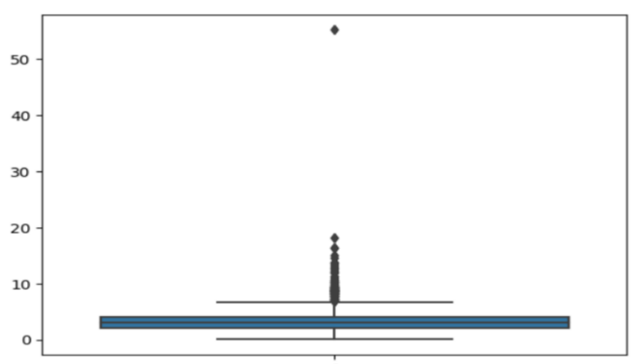


Fig 7. Boxplot for session duration

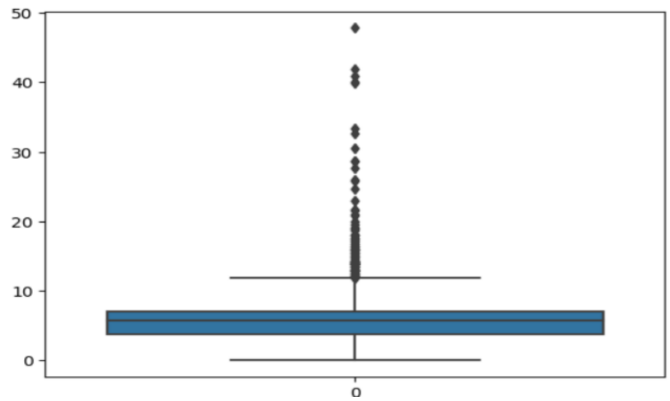


Fig 8. Boxplot for energy consumption

An effective method for outlier detection is the isolation forest algorithm[13]. Its primary purpose is to identify anomalous data points in large datasets with high-dimensional features. The algorithm operates by isolating outliers based on the intuition that they will have shorter average path lengths within a random forest.

The isolation forest algorithm follows a series of steps to accomplish outlier detection. First, a random feature is selected, and a split value is chosen randomly within the range of that feature's minimum and maximum values. The data is then recursively partitioned based on this split, with points falling below the split value sent to the left child node and points above sent to the right child node. This partitioning process continues until all data points are isolated. Multiple isolation trees are constructed by repeating the random selection and partitioning steps. The number of trees can be predefined or determined through techniques like cross-validation. Once the trees are built, the average path length from the root node to each data point is calculated. Shorter path lengths indicate that a point is easier to isolate and likely to be an outlier. The algorithm assigns an outlier score to each data point based on its average path length. Higher scores indicate a higher likelihood of being an outlier. A threshold can then be set to classify data points as outliers or inliers based on their scores. Points with scores above the threshold are considered outliers, while those below are considered inliers.

To prepare the charging dataset for analysis, the datetime fields are adjusted to be in a consistent format. The datetime fields in the weather data obtained from the Caltech local station are originally taken in UTC (GMT) time. However, they are time zone aware, meaning that they include information about the local time zone. To ensure that the data is in a standardized format that can be easily analyzed, the datetime fields are first converted to naive datetime objects. However, to ensure consistency and compatibility with the charging dataset, the datetime values are converted to the local time zone of "America/Los Angeles" (GMT-7). This conversion aligns the timestamps in the weather data with the corresponding charging events and allows for accurate analysis and comparison between the two datasets. By converting the datetime fields to the same time zone, any potential discrepancies or inconsistencies arising from the use of different time zones are mitigated, enabling meaningful analysis and interpretation of the weather-related variables in relation to the charging behavior.

D. Feature Engineering

It is a crucial step in the machine learning pipeline, where raw data is transformed and manipulated to create meaningful and informative features that can improve the performance of predictive models. It involves selecting, creating, and transforming variables or attributes from the raw dataset to enhance the representation of the underlying patterns and relationships.

Following techniques have been implemented on the dataset:

1. **Feature Selection:** In this, a subset of relevant features is chosen from the original dataset. Unnecessary or redundant features can be removed, reducing dimensionality and simplifying the model. Features like siteid, stationid, spaceid, timezone, loud cover which has no relevance has been dropped.

2. **Feature Creation:** New features can be derived from existing ones to capture additional information. This can involve mathematical operations, such as calculating ratios or differences between variables, or creating interaction terms to capture synergistic effects. Average Session Length and average energy consumption has been derived from session length and energy consumption respectively for each record.

3. **Feature Scaling:** Numerical variables may need to be scaled or normalized to ensure their values are on a similar scale. Widely used scaling techniques include standardization (z-score normalization) or min-max scaling.

In our work, we utilized standardization, a technique that transforms data which has a mean of 0 and a standard deviation of 1. Also, the mean of each feature is subtracted from its data points and then divided by the standard deviation.

4. **Machine learning** models cannot directly use datetime objects. Therefore, it is necessary to convert the connection time and disconnection time of the records into numerical features that can be understood by the models. To convert time fields into a numeric format suitable for the models, a common approach is to represent time as a decimal value. To accomplish this, you can divide the number of minutes by 60 and then add the quotient to the hour. By doing so, the time is transformed into a continuous numerical feature that captures the fractional representation of the hour.

Table 1: Features included

Feature	Explanation
session_length	charging Time duration, Target variable
kWh	energy consumed by vehicle Target variable
Num.Rep	numerical representation of arrival time
Avg. Sess.	average session length
Avg.Cons.	average energy consumption
Temperature	temp. on that particular day
Avg.Humidity	average humidity
Avg.wind speed	wind speed that particular day
Visibility	visibility at that time
Pressure	Pressure on that particular day
Battery_pack kWh	Battery pack capacity
FastCharge_KmH	Fast charging speed of battery

E. Model Selection and Evaluation

Based on the ACN dataset from 2018, K-means clustering is applied to specific features to identify sinilar charging behavior. To select the best number of clusters to train g the dataset, an elbow plot is used. The elbow plot displays the cluster Sum of Squared Error (SSE) as the

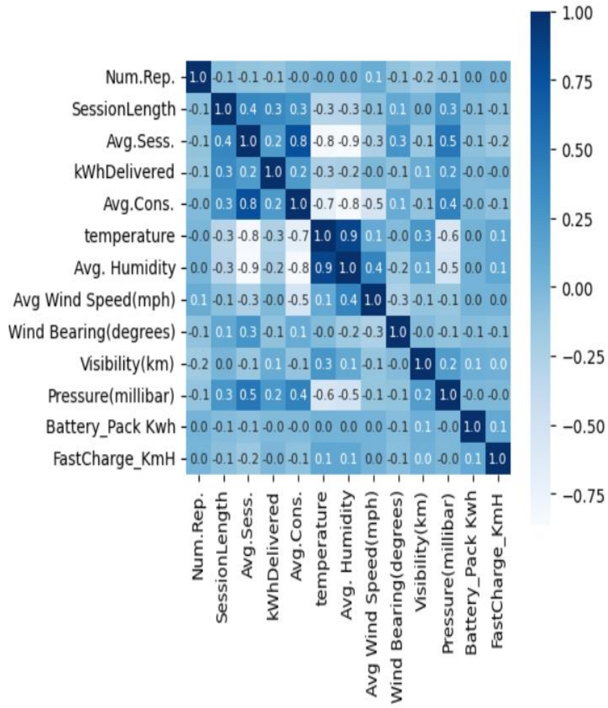


Fig. Heatmap showing correlation between features

the number of clusters varies. In Figure 9, the elbow plot shows the SSE for a range of clusters from 1 to 10.By examining the plot, the optimal number of clusters can be identified. The "elbow" in the plot represents a point where the decrease in SSE starts to level off. This point indicates a trade-off between minimizing the SSE and not overfitting the data with too many clusters.

The elbow plot suggests that the required number of clusters is two for the clustering algorithm. This means that dividing the data into two distinct clusters yields a good balance between capturing the patterns and structure in the data while avoiding excessive complexity. The elbow plot is a tool to get the required number of clusters, which can aid in selecting a clustering method that offers valuable insights and enables further analysis or decision-making based on the resulting clusters.

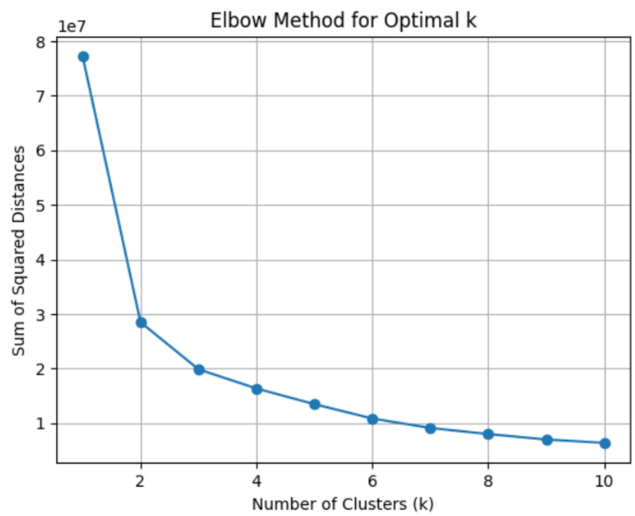


Fig 9. Elbow graph to find optimal clusters

Before training the dataset it is splitted into a ratio of 80:20. 80% data is used for training and 20% is for testing. To find the optimal hyperparameters, grid-search or random-search combined was employed with K-fold cross-validation. The information is isolated into K folds, and the model is trained on K-1 folds at the same time assessing its execution on the remaining overlay. This process ensures robustness and unbiased evaluation of the model's effectiveness. This process is repeated continuously for all folds, and the model's performance is assessed using various metrics. Random Forest, XGBoost, Support Vector Machines, KNN, ANN, as well as ensemble learning techniques[16] such as Voting Regressor and Stacking Regressor, are among the popular and successful algorithms employed. These models are trained and tested using the Scikit-learn library in Python, and evaluation metrics such as MAE, RMSE, R2 score, and SMAPE are used. The best performing models are then ensembled using stack and voting regressor.

V. RESULTS AND COMPARISON

The initial training begins with the Random Forest algorithm, which is chosen due to its ability to provide relative feature importance. By using the Random Forest algorithm, we can gain insights into the significance of different features in the dataset. After this, SVM, XGBoost algorithm, KNN and Artificial Neural Network (ANN) are applied for predicting the session duration. These models are chosen due to their effectiveness in capturing complex patterns and relationships in the data. After predicting the duration of the session, it becomes a new feature that is incorporated into the model's training process for predicting energy consumption. This approach leverages the predicted session duration as an additional input, allowing the model to incorporate this valuable information and potentially improve the accuracy of energy consumption predictions.

A. Session Duration Prediction

In order to optimize the performance of our models, a grid search approach was employed for the selection of the hyperparameters. For the deep Artificial Neural Network (ANN) in particular, we performed extensive experiments to identify the most effective architecture. After conducting extensive experiments, we determined that the optimal architecture for the deep artificial neural network (ANN) consisted of two hidden layers. One layer has 32 nodes and the other has 16 nodes. The Rectified Linear Unit (ReLU) [14] activation function was utilized for all hidden layers, while the output layer employed a linear activation function. This was because the output was expected to be a numerical value. We utilized the Adam optimization algorithm [15] and selected a training batch size of 1000 epochs. These choices were made after extensive experimentation and consideration of the characteristics of the dataset.To leverage the strengths of multiple algorithms, three best-performing models were selected: RF, SVM and XGBoost. These models were chosen based on their individual performance and ability to capture different aspects of the data.We employed two ensemble learning techniques, namely voting ensemble and stacking ensemble, to combine the predictions of these models. In the voting ensemble, the predictions from each individual model were averaged to produce the final prediction. This approach benefits from the collective wisdom of the models and can improve overall prediction accuracy. In the stacking ensemble, the predictions that were obtained from the individual models were used as inputs to a final estimator ,after this cross validation was applied. This final estimator learned to make predictions based on the patterns and relationships identified by the base models, leading to potentially enhanced performance. The test scores are shown in Table I for session duration.RF, SVM, XGboost and KNN have similar performance while ANN is slightly worse.

Table 2. Test Scores for Session Duration

Model	MAE (hr)	RMSE (hr)	R2 Score	SMAPE (%)
1. RF	0.75	1.28	0.67	26.4
2.SVM	0.91	1.31	0.65	31.0
3.XGBoost	0.70	1.18	0.69	25.7
4.KNN	1.03	1.44	0.57	32.5
5.ANN	0.93	1.52	0.54	40.4
6.Stacking Ensemble	0.74	1.25	0.71	29.0
7.Voting Ensemble	0.75	1.14	0.73	26.6

B. Energy Consumption Prediction

We followed a similar methodology as session duration prediction to predict energy consumption, but made slight changes to the deep Artificial Neural Network (ANN) architecture. For this task, the deep ANN had two hidden layers. One layer had 64 nodes and the other layer had 16 nodes. The decision to use this architecture was based on experimental evaluation, which showed that it effectively captured the underlying patterns in the data.

In the deep learning model, the Rectified Linear Units (ReLU) activation function was used for all the hidden layers, while a linear activation function was employed in the output layer to make predictions of numeric values. During training, we used a batch size of 64, indicating the number of samples processed in each iteration, and the number of epochs was set to 1000, indicating the number of times the entire training dataset was iterated. These hyperparameter values were determined through experimentation to achieve optimal performance for the energy consumption prediction task.

Among the four models evaluated, Random Forest (RF) consistently demonstrated the best cross-validation scores. Therefore, we selected RF, along with Support Vector

Machine (SVM) and XGBoost, as the top three models to construct two ensemble models: a voting ensemble and a stacking ensemble.

Interestingly, the stacking ensemble model emerged as the best-performing model, as indicated by the highlighted results.

C. Comparison

When comparing the results of session duration and energy consumption, it is observed that predicting energy consumption is generally more challenging, which aligns with previous studies on the ACN data [21].However, in some cases, such as in [24], it was found that predicting energy consumption was easier than session duration.

Table 3. Test Scores for Energy Consumption

Model	MAE (hr)	RMSE (hr)	R2 Score	SMAPE (%)
1.RF	0.99	2.12	0.59	22.3
2.SVM	1.49	2.66	0.40	32.1
3.XGBoost	0.95	2.09	0.60	24.1
4.KNN	1.85	2.74	0.32	37.8
5.ANN	1.55	2.93	0.25	50.0
6.Stacking Ensemble	0.92	2.06	0.61	21.3
7.Voting Ensemble	0.93	2.08	0.60	21.4

Additionally, it was observed that deep artificial neural network (ANN) models performed less accurately in both scenarios. While deep learning models excel in tasks such as image and audio data analysis as they have the ability to learn and extract relevant features from the data without the need for explicit feature extraction, traditional machine learning generally outperforms deep learning models where feature extraction is involved. In both scenarios it was observed that ensemble learning models performed better than individual machine learning models. However, when it came to predicting session duration, the impact of ensemble learning was more significant. This could be attributed to the three models having comparable training performance, and their predictions being combined to achieve further improvement.

When analyzing the same dataset, it is clear that the inclusion of additional weather and vehicle data has led to improved predictions of electric vehicle charging behavior. This underscores the importance of utilizing supplementary information to increase the precision and dependability of the models.

VI. FUTURE WORK

According to recent research in the field, the use of machine learning (ML) for predicting EV charging behavior poses several challenges. One of the most significant challenges is the scarcity of publicly available datasets that are required for training ML models. Currently, there are limited EV charging datasets that are accessible for research purposes, whereas many others are privately owned by business entities. Additionally, the datasets that are available represent the charging behavior of only certain geographical areas, which limits the scope of the predictions that can be made. In order to tackle this issue, it is imperative for researchers to promote data sharing by making the datasets utilized in their studies openly accessible online.

Additionally, a significant challenge lies in the insufficient incorporation of predictive models into smart scheduling practices. Efficient scheduling is crucial for managing electric vehicle (EV) charging, and this involves taking into account short-term (predictions for the next few minutes or hours) and long-term (predictions for the next day or week) machine learning (ML) forecasts. Previous research has focused on using historical charging data with limited variables, such as arrival time, departure time, and energy consumption to train the models. Therefore, it is essential to develop accurate predicting models that consider a wider range of variables for effective load management scheduling.

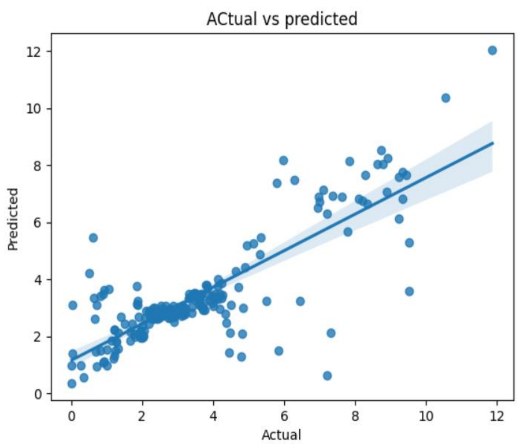


Fig 10(a). Random Forest

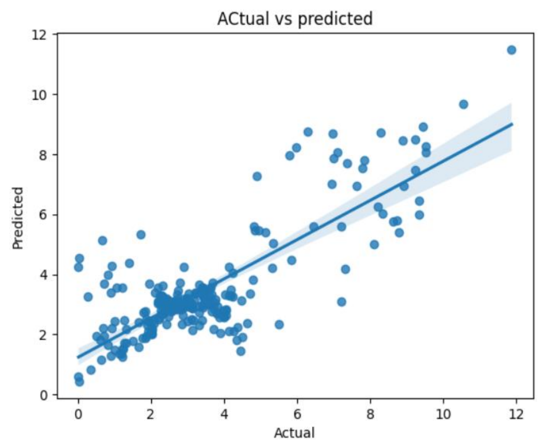


Fig 10(b). SVM

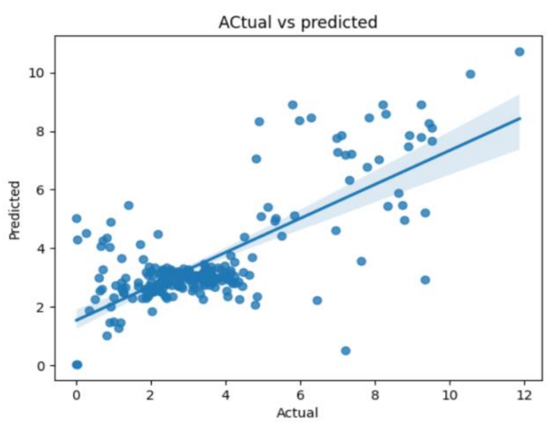


Fig 10(c) XGboost

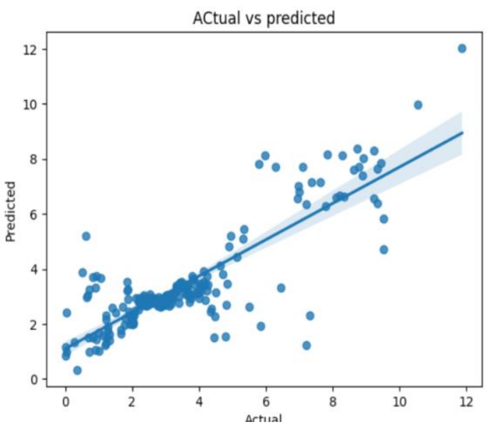


Fig 10(d) Stacking Ensemble

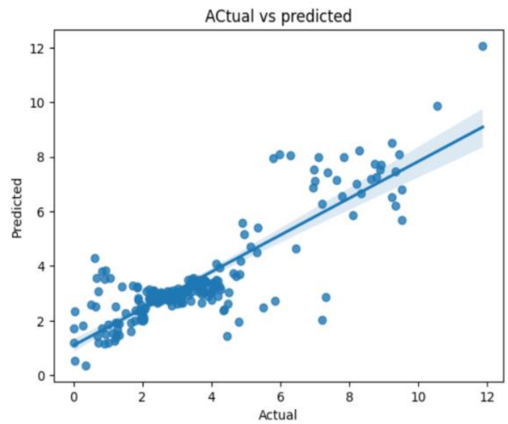


Fig 10(e) Voting Ensemble

Fig 10.Scatterplot of Models that give the best results for session duration-Voting Ensemble (RF+SVM+XGB) Stacking Ensemble (RF+SVM+XGB)

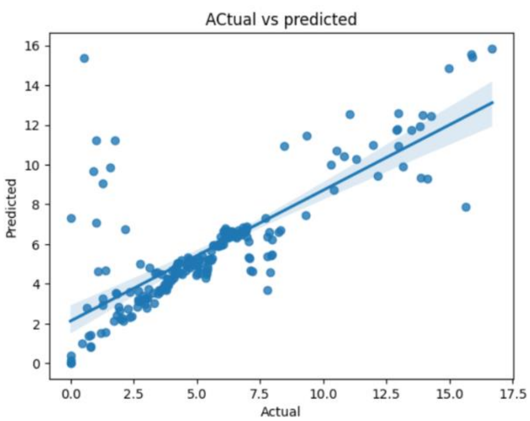


Fig 11(a). Random Forest

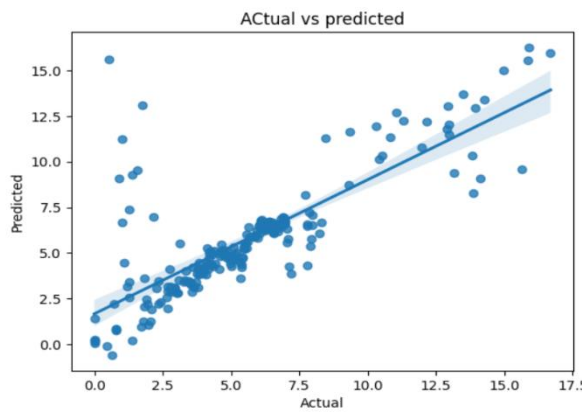


Fig 11(b). XGBoost

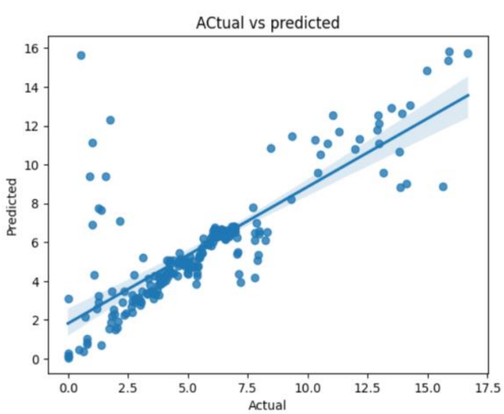


Fig 11(c). Stacking Ensemble

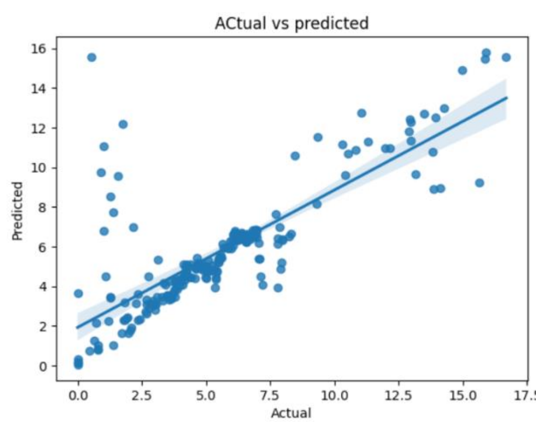


Fig 11(d). Voting Ensemble

Fig 11. Scatter Plot of Models that gave the best results- Stacking ensemble (RF + XGBoost), Voting Ensemble (RF + XGBoost)

Reinforcement learning is a subset of ML that works on learning through trial and error using reward and punishment mechanisms. Recent studies have utilized this technique for EV charging scheduling, indicating potential for further research in this field. Future research can focus on exploring various areas such as coordinating EV charging and voltage control using deep reinforcement learning, constructing new action spaces, and incorporating supplementary information such as weather and vehicle data to improve the accuracy of models for efficient scheduling of EV charging in the short and long-term.

VII. CONCLUSION

In this study, we introduced a comprehensive framework for predicting two crucial aspects of EV charging behavior: session length and energy consumption. A distinguishing feature of our work is the integration of weather and vehicle data alongside historical charging data. By training four well-established ML algorithms and two ensemble learning algorithms, we achieved better prediction performance in comparison to previously done research efforts.

According to our research, we have demonstrated the usefulness of the ACN dataset by improving the prediction of charging behavior. Our study shows that incorporating weather and vehicle information can significantly enhance the accuracy of charging behavior prediction. By considering these additional factors, we gain valuable insights and a more comprehensive understanding of the dynamics of EV charging.

REFERENCES

1. N. K. Ahmed, A. F. Atiya, N. E. Gayar, and H. El-Shishiny, ‘‘An empirical comparison of machine learning models for time series forecasting,’’ *Econ. Rev.*, vol. 29, nos. 5–6, pp. 594–621, Aug. 2010.
2. Y.-W. Chung, B. Khaki, T. Li, C. Chu, and R. Gadh,”Ensemble machine learning-based algorithm for electric vehicle user behavior prediction,” Appl. Energy, vol. 254, Nov. 2019 DOI:10.1016/j.apenergy.2019.113732
3. Z. Xu, ‘‘Forecasting electric vehicle arrival & departure time on UCSD campus using support vector machines,’’ Ph.D. dissertation, Dept. Eng. Sci., Appl. Ocean Sci., UC San Diego, San Diego, CA, USA, 2017.
4. Y.-W. Chung, B. Khaki, T. Li, C. Chu, and R. Gadh, ‘‘Ensemble machine learning-based algorithm for electric vehicle user behavior prediction,’’ Appl. Energy, vol. 254, Nov. 2019, DOI:10.1016/j.apenergy.2019.113732
5. Y. Xiong, C.-C. Chu, R. Gadh, and B. Wang, ‘‘Distributed optimal vehicle grid integration strategy with user behavior prediction,’’ in Proc. IEEE Power Energy Soc. Gen. Meeting, Jul. 2017, pp. 1–5.
6. A. Almaghrebi, F. Aljuheshi, M. Rafaie, K. James, and M. Alahmad, ‘‘Data-driven charging demand prediction at public charging stations using supervised machine learning regression methods,’’ Energies, vol. 13, no. 16, p. 4231, Aug. 2020. DOI: 10.3390/en13164231
7. Murtagh, Fionn. ”A survey of recent advances in hierarchical clustering algorithms.” The computer journal 26, no. 4 (1983): 354-359. DOI:10.1093/comjnl/26.4.354
8. M. Majidpour, C. Qiu, P. Chu, R. Gadh, and H. R. Pota, ‘‘A novel forecasting algorithm for electric vehicle charging stations,’’ in Proc. Int. Conf. Connected Vehicles Expo (ICCVE), Nov. 2014, pp. 1035–1040. DOI:10.1109/ICCVE.2014.7297504
9. N. Bokde, M. W. Beck, F. Martínez Álvarez, and K. Kulat, ‘‘A novel imputation methodology for time series based on pattern sequence forecasting,’’ Pattern Recognit. Lett., vol. 116, pp. 88–96, Dec. 2018 DOI:10.1016/j.patrec.2018.09.020
10. ACN-Data—A Public EV Charging Dataset. Accessed: Mar. 9, 2023. [Online]. Available: https://ev.caltech.edu/dataset

11. R. Gelaro, W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, and K. Wargan, ''The modern-era retrospective analysis for research and applications, version 2 (MERRA-2),'' J. Climate, vol. 30, pp. 5419–5454, Jul. 2017, DOI: https://doi.org/10.1175/JCLI-D-16-0758.1

12. R. Mendelsohn, P. Kurukulasuriya, A. Basist, F. Kogan, and C. Williams, ''Climate analysis with satellite versus weather station data,'' Climatic Change, vol. 81, no. 1, pp. 71–83, Mar. 2007, doi: 10.1007/s10584-006- 9139-x.

13. Fei Tony Liu,Kai Ming Ting : Zhi-Hua Zhou,Isolation Forest ,10472172, Feb 10,2009 DOI: 10.1109/ICDM.2008.17

14. V. Nair and G. E. Hinton, ''Rectified linear units improve restricted Boltzmann machines,'' in Proc. 27th Int. Conf. Int. Conf. Mach. Learn., 2010, pp. 807–814.

15. D. P. Kingma and J. Ba, ''Adam: A method for stochastic optimization,'' 2014, arXiv:1412.6980. [Online]. Available: http://arxiv.org/abs/1412.6980

16. D. Opitz and R. Maclin, ''Popular ensemble methods: An empirical study,'' J. Artif. Intell. Res., vol. 11, pp. 169–198, Aug. 1999.

17. Z. J. Lee, T. Li, and S. H. Low, ''ACN-data: Analysis and applications of an open EV charging dataset,'' in Proc. 10th ACM Int. Conf. Future Energy Syst., New York, NY, USA, 2019, pp. 139–149, DOI:10.1145/3307772.3328313

18. Y.-W. Chung, B. Khaki, T. Li, C. Chu, and R. Gadh, ''Ensemble machine learning-based algorithm for electric vehicle user behavior prediction,'' Appl. Energy, vol. 254, Nov. 2019, Art. no. 113732, DOI:10.1016/j.apenergy.2019.113732

19. Sebastian Raschka, About Feature Scaling and Normalization and the effect of standardization for machine learning algorithms, Jul 2014,DOI:10.13140/2.1.4245.1849