SCIENTIFIC OASIS

Decision Making: Applications in
Management and Engineering

ISSN: 2560-6018, eISSN: 2620-0104

DECISION MAKING:
APPLICATIONS IN
MANAGEMENT AND
ENGINEERING

PREDICTION OF EV CHARGING PATTERNS USING HYBIRD MACHINE LEARNING ALGORITHMS

Sathishkumar. S[1]
Department of Computer Science and Engineering,
Bharath Institute of Higher Education and Research Chennai, India.
skumarscse@gmail.com

Yogesh Rajkumar.R[2]
Department of Information Technology
Bharath Institute of Higher Education and Research Chennai, India
yogeshrajkumar.it@bharathuniv.ac.in.

Abstract**—**Electric cars are becoming increasingly popular in the transportation sector. Despite their growing acceptance, several factors still limit their widespread use. These factors include the relatively limited driving range of electric vehicles and the high costs related to battery production and maintenance. Understanding the energy usage of EVs becomes more important nowadays due to the rapid usage and emergence of EVs into the market. To address these challenges, researchers have turned to machine learning models to predict electric vehicle charging patterns accurately. Among these models, an ensemble learning approach has shown to outperform previous models significantly. This is evidenced by superior Symmetric Mean Absolute Percentage Error (SMAPE) scores, indicating more accurate and reliable forecasting of charging behaviours**.**

Keywords**—***electric vehicle, charging behavior, transportation, SMAPE ensemble learning.*

Introduction

Electric vehicles (EVs) are becoming more and more important for accomplishing sustainable transportation goals as cities around the world strive to lower carbon emissions. EVs are positioned as a major answer to the climate crisis since they can reduce carbon emissions by up to 45% when compared to traditional internal combustion engine vehicles. However, there are drawbacks to the widespread use of EVs, including increased charging times and high-power demands on the electrical system. These problems are particularly pertinent in cities where population growth is anticipated, leading to higher energy consumption and a strain on the infrastructure already in place [1]. Improving the user experience and reducing the strain on power systems require effective management of EV charging behaviour. Accurately forecasting EV charging habits, including session length and energy usage, is a major challenge. These forecasts would assist utilities in reducing peak demand, improving charging schedules, and promoting a more robust grid.

Forecasting is made more difficult by the fact that EV charging patterns depend on a number of variables, such as user behaviour, vehicle type, and time of day [2]. In order to predict EV charging trends, this study presents a machine learning-based method that makes use of techniques including Random Forest (RF), XGBoost, Support Vector Machine (SVM), and Artificial Neural Networks (ANN). The model attempts to increase the accuracy of charging predictions by utilizing historical data and sophisticated ensemble techniques, taking into account both energy use and session duration [3]. In the end, these insights are meant to help sustainable urban transportation options by improving the efficiency and coordination of EV charging infrastructure [4].

Literature Survey

A unique information regarding charging of EV, possessing around thirty thousand sessions, was presented by Lee [1]. They used GMM to predict the duration of the session and the energy needed, considering the variation of the estimated arrival times. The energy consumption and session time were found to have SMAPEs of 15.9% and 14.4%, respectively [5].

Çolak uses a machine learning technique to investigate how coolant flow and road gradient affect the electrical components of electric vehicles that run on batteries [6]. Although it acknowledges the computing resources needed for training larger datasets, the paper highlights the critical role that data quantity plays in improving prediction accuracy for Artificial Neural Networks (ANN) [7], arguing that sufficient value is essential.

Yu [8] used mean estimate to forecast the start time and length of the session. Next, using linear regression and the length of the session, energy consumption estimations were obtained [9]. In order to stabilise the system and average the loading status in such a way that it is minimized to a greater extent. Nonetheless, a quantitative assessment of the predicted performances was not conducted.[10]

Khan [11] also used a variety of algorithms, such as SVM and RF, to forecast the requirements in terms of energy for a station that is used for charging for the following day based on the energy usage of the day before, in which forecasts are generated for each day after the days are classified using clustering. The most accurate findings were obtained using the Pattern Sequence-based Forecasting (PSF)-based technique, which had an SMAPE score of 14.1% on average [12].

The application of Vehicle-to-Grid (V2G) technologies to reduce the danger of charging electric vehicles on the distribution grid was examined by Yilmaz and Krein [13] and Habib [14]. It is indicated that V2G technologies can aid in enhancing the grid's effectiveness, stability, and dependability. Additionally, load balancing, current harmonic filtering, and power management are some of the benefits of V2G technology, according to Yilmaz and Krein [12]. Nevertheless, deep discharging of EVs can be brought on by V2G technology. The deterioration of EV batteries results in a reduction in battery lifetime and a subsequent drop in consumer satisfaction.

A.Almaghrebi [13] employed many models to forecast energy consumption from data obtained from the charging stations which are publicly available in the states of US. In addition to past billing information, input elements included season, weekday, kind of place, and charging costs. The XGBoost model fared better on the test set than SVM, RF, and linear regression.

methodology

The methodology of this study closely follows standard machine learning practices. It begins with the collection of extensive data on electric vehicle (EV) charging patterns and battery longevity from various sources. This dataset includes features which are critical for building an accurate predictive model. The data undergoes thorough pre-processing to ensure quality and consistency, including cleaning to remove inaccuracies or missing values and standardization to facilitate better model performance [14]. The training set of the dataset is used to acquaint the model with patterns in the data, while the testing set assesses the model's performance using newly collected data. Feature selection guides the model by identifying significant factors that impact battery life. Feature selection identifies the most important variables impacting EV charging behavior and battery life, enabling the model to focus on these crucial features and avoid unnecessary complexity [15].

EV charging behavior is influenced by various complex factors, including user patterns, charging session lengths, energy requirements, and availability of charging infrastructure. A single algorithm may not capture all these patterns adequately. Ensemble learning, by blending models like SVM, kernel density estimators, and random forests, allows the system to benefit from each model’s unique strengths, leading to a more comprehensive and accurate prediction. The research leverages data from both public and residential charging datasets, which may have different characteristics, such as varied session durations or charging frequencies. Different algorithms, like decision trees in random forests and kernel methods in SVMs, may handle this data diversity better when combined, enabling the ensemble to generalize across different data types more effectively than a single algorithm [16]. Ensemble methods are known to reduce model bias and variance, providing more stable predictions. In this study, XGBoost achieved a strong individual performance, but the ensemble approach, combining multiple models, further enhanced the accuracy, achieving SMAPE scores of 10.4% for session duration and 7.5% for energy consumption. These lower error metrics demonstrate that the ensemble model is better at minimizing errors due to its balanced approach [17].

PROPOSED METHOD

The System Architecture consists of following stages;

- Data Collection
- Preprocessing

- Outlier Removal
- Feature Engineering
- Model Selection
- Analysis and Evaluation

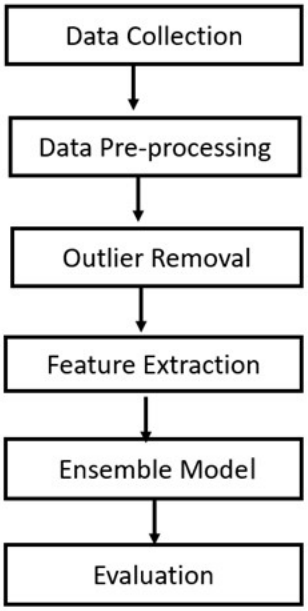

Fig. 1. System Architecture

Implementation

A.DATA COLLECTION

Data Collection phase is the initial phase of this study. Usually, data is collected from a publicly available source. This effort will make use of the ACN dataset, one of few publicly available datasets [18]. The charge records from JPL and Caltech, two university campus stations, are included in the collection. The Adaptive Charging Network (ACN) dataset is a comprehensive dataset focused on electric vehicle (EV) charging sessions, typically used in research to study and model EV charging behavior, energy consumption patterns, and charging demand forecasting. Developed by the California Institute of Technology and other collaborators, this dataset records detailed charging sessions from the ACN, a network of EV charging stations. The ACN dataset is used to train predictive models, often with machine learning techniques, to accurately estimate energy consumption and optimize charging infrastructure. The dataset is particularly valuable in EV research because it captures real-world, timestamped EV charging events across varied conditions, offering insights that support infrastructure planning, energy management, and the development of adaptive, efficient charging strategies for EV networks. Since the JPL station is solely accessible by workers, it will not be taken into consideration in this work, in contrast to other stations thar are publicly available. [19]

Despite the existence of a tiny weather centre on the Caltech campus, we did not use it because the breeze’s interval records were erratic and had missing data. Furthermore, factors like precipitation and rainfall were not recorded by this site. Consequently, we employed the meteorological information from NASA's Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2).[20]

B.PREPROCESSING

Following data gathering, the pre-processing stage starts. In this case, the dataset has undergone many processes to guarantee its accuracy and stability. In order to look for duplicate and missing values, we went over the data. The dataset was confirmed to be devoid of duplicate or missing occurrences after preprocessing. Duplicate values are often detected by comparing key attributes that uniquely identify a session, such as the session ID, vehicle ID, start time, and location. In some cases, partial duplicates may exist where entries have slight discrepancies (e.g., slight variations in timestamps). We applied additional logic, such as rounding timestamps to the nearest minute or averaging values, to ensure that only one record per charging event is kept. Pre-processing and cleaning of the data is done to guarantee the prediction model’s effectiveness and accuracy.[21][22]

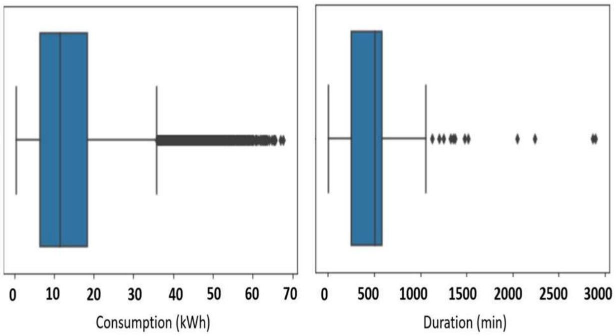

Fig. 2. Consumption of energy (left-side diagram), session duration (right-side diagram) boxplots

Outliers’ identification is a crucial phase in the approach that comes after data collection and pre-processing. By protecting the data integrity and improving the accuracy of machine learning models when predicting the battery life of electric vehicles, this procedure increases the dependability of sustainable transportation initiatives. Therefore, we decided to perform the following,

- Use the isolation forest approach.
- To conduct multivariate outlier identification.

The examples with short average path lengths on the iTrees are the outliers. The observations are "isolated" by choosing a variable at random, variable's maximum and lowest. Until every observation has been isolated, partitioning is done recursively. Following partitioning, the observations with shorter path lengths for certain sites are probably the outliers.[23][24] Figure 3 shows the procedure for identifying the target variables' outlier.

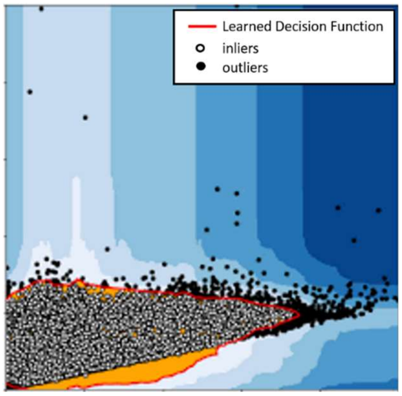

Fig. 3. Outlier detection using isolation forest

Next, the test-train splitting technique is applied to divide the Pre-Processed dataset. The test data and the train data are two distinct sets that comprise the total dataset. Test data makes up 20% of the total dataset and is mainly assessed for consideration of functionality, accuracy, and other metrics. Eighty percent of the data consists of data required for training. The model is trained using the recommended algorithmic strategies on this train set of data. A pattern found in the train data is used by the algorithm to learn.[25]

C.FEATURE EXTRACTION

The process of utilising human expertise to turn data into a meaningful representation is known as feature engineering. Despite being labor-intensive, this technique is crucial because it addresses a flaw in the learning algorithms. We then perform the following actions,

- Time = (Minute/60) + hour
- Use time to generate numeric data.
- Calculate average session duration time
- Calculate average departure time
- Calculate average energy usage

This is accomplished by obtaining the charging record's user ID and compiling all of his prior records. On the other hand, temporal data and certain properties.[26] The trigonometric translation is carried out as follows in order to depict the closeness of these values:

$$f_x = sin(2\pi f/max(f)) \qquad (1)$$

$$f_y = cos(2\pi f/max(f)) \qquad (2)$$

Where,

f_x---- Cyclic feature’s 1st component.

f_y ----Cyclic feature’s 2nd component.

f_y ----Feature that has to be modified

One-hot encoding, which converts a lonely variable having n points, k unique classes to k binary variables having n points each, was utilised to change other categorical variables.[27]

TABLE I. EXTRACTED FEATURES AND THEIR DESCRIPTIONS

Feature	**Description**
session_length	Length of charging duration, target variable
kWh_delivered	Session energy consumption, target variable
time_con	Numerical representation of the connection time (arrival time)
time_con	Day of the week, one-hot encoded
is_weekend	Binary variable indicating whether the session took place in a weekend
holiday	Binary variable indicating whether the session took place on a US federal holiday
hr_x, hr_y	Sine and Cosine components of the hour
day_x, day_y	Sine and Cosine components of the day
mnth_x, mnth_y	Sine component of the month
mean_d_time	Historical average departure time
mean_con	Historical average consumption
mean_dur	Historical average session length

max_traffic_aft_arvl	maximum traffic level after arrival
avg_temp_nxt	average temperature of next 10 hours
avg_rain_nxt	average rainfall of next 10 hours

D.MODEL EVALUATION

The most important part of the model selection process is figuring out which machine learning algorithm is most appropriate for a certain task. To make a choice, a number of models must be tested and assessed. Model training is an essential step in the production process.[28]

Below is the list of models that have been performed and analysed in the study.

- Random Forest
- XG Boost
- Support Vector Machine (SVM)
- Deep ANN

Charging sessions from the 2019 calendar year in the ACN dataset is selected to incorporate seasonal factors into the training process. The data is split into test and train data. We employed Kfold cross-validation at the time of training, repeating the training process K times while leaving out 1/K of the data for testing each time. K value of 10 is used usually. For optimization, we used the grid search method, that tests specified factors to find best set. For efficiency, we performed the grid search with 5-fold cross-validation.[29]

Drawing inspiration from the previous studies, we experimented with ensemble learning techniques. In the voting regressor, multiple bases are trained using training data, and the average is used as final output. The stacking regressor, on the other hand, applied the stacked generalization technique. [30]

In our pursuit to enhance the prediction accuracy of Electric Vehicle (EV) charging patterns, we propose an advanced ensemble learning technique. This method leverages the collective strengths of several machine learning algorithms to construct a robust and reliable forecasting model. By integrating Support Vector Machines (SVM), XGBoost, Deep Artificial Neural Networks (ANN), and Random Forest (RF), we aim to achieve superior predictive performance and address the complex, non-linear relationships inherent in EV charging data.

Random Forest, an ensemble technique, creates multiple decision trees using random subsets of features and samples. Each tree votes on the predicted class, and the majority vote determines the final prediction. This method is particularly effective for classification and regression tasks due to its ability to improve accuracy, reduce overfitting, and offer insights into feature importance. XGBoost, renowned for its high performance and scalability, handles complex data patterns efficiently and includes built-in feature importance. By iteratively correcting errors from previous models, XGBoost enhances the precision of predictions, making it ideal for EV charge forecasting. Bagging, or Bootstrap Aggregating, is an ensemble technique that improves the stability and accuracy of machine learning algorithms. It works by training multiple versions of a model on different subsets of the training data and then combining their predictions. This approach reduces variance and helps prevent overfitting, making it particularly effective for models like decision trees.

Support Vector Machines are chosen for their adeptness at managing both linear and non-linear interactions, and their resilience against overfitting. SVMs perform exceptionally well in high-dimensional spaces, crucial for capturing intricate EV charging patterns influenced by variables such as time, weather, and user behavior. Deep Artificial Neural Networks are incorporated for their ability to model complex, non-linear relationships in extensive datasets. ANNs learn from large volumes of data, identifying nuanced patterns that traditional algorithms might miss, thus ensuring highly accurate and generalizable predictions across diverse charging scenarios. By employing this ensemble learning framework, we can harness the individual strengths of these algorithms, resulting in a comprehensive and effective model for predicting EV usage patterns. This approach not only improves prediction accuracy but also enhances the robustness and adaptability of the forecasting system, ultimately contributing to more efficient and reliable EV charging infrastructure planning and management.

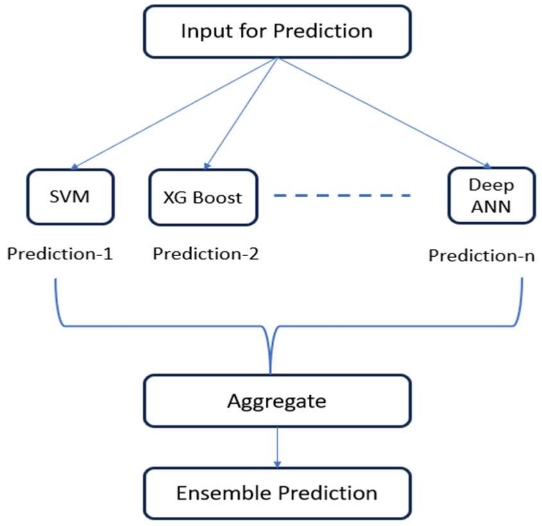

Fig. 4. Illustration of Ensemble technique

G.EVALUATION AND DISCUSSION

The assessment criteria used to determine how well the regression model works are the R^2, MAE, MSE, and RMSE.

Metric calculatory equations are provided below where

y_i --- original value

y_p --- expected value

n --- total occurrences

R^2 value, quantifies the predictability of the dependent variable's variance from independent factors.[32]

The *MAE* is given as

$$\text{MAE} = \frac{|(y_i - y_p)|}{n} \qquad (3)$$

The R*MSE* is given as

$$\text{RMSE} = \sqrt{\frac{\Sigma(y_i - y_p)^2}{n}} \qquad (4)$$

The R-Squared is given as

$$R^2 = 1 - \frac{\Sigma(y_i - y_p)^2}{\Sigma(y_i - y_i)^2} \quad - \qquad (5)$$

The RF method, which may be used to visualise the variable significance, is where we start the experiment [30]. This feature selection technique eliminates several variables that are seldom useful and frequently impair performance. Ten-fold cross-validation is a technique that may be used to obtain an accurate assessment of an ML model's capacity for generalisation as well as to select the optimal collection of hyperparameters regarding a given dataset. The effectiveness of these methods on the characteristics of the input and aim output dataset was evaluated. As a result, 10 loops are used in the training process, and the precision of the process was calculated by averaging the results from each loop. We chose to include the least significant variables in the model training since, in this instance, their inclusion resulted in a negligible performance boost. Variables can also be arranged according to their respective importance. The contribution of each characteristic in identifying the best splits determines this.[33] The top ten crucial factors for session length and energy use are displayed in Figures 5 and 6, respectively.

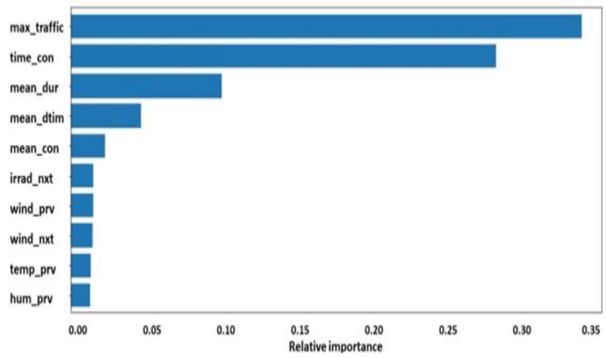

Fig. 5. Top ten features for session duration

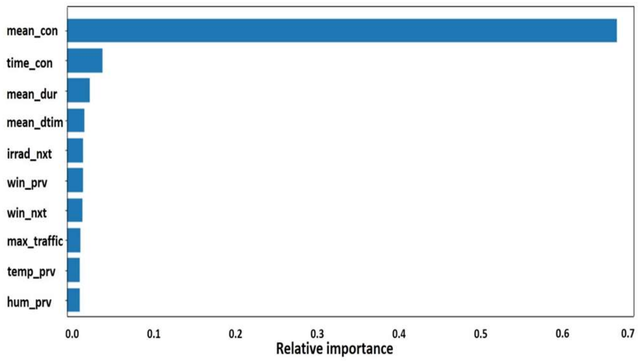

Fig. 6. Top ten features for energy consumption

SESSION DURATION PREDICTIONS

Search in grid technique was utilized to find models' best parameters. We empirically found that the best design for the deep ANN training possess 3 layers consisting of a series of nodes in order of 64, 32, 16. Since we anticipate the prediction to be a numerical value, the output layer's activation was linear, and all hidden layers' activation functions were Rectified Linear Units (Relu) [59]. There were 32 people in the training batch, and there were 15 epochs in total of iterations.[34] The training curve of loss is displayed in Figure 5, and the tenfold cross validation scores are compiled in Table 2.

TABLE II. TRAINING SCORES FOR SESSION DURATION

Model	**RMSE (mins)**	**MAE (mins)**	**R^2**	**SMAPE (%)**
RF	101	64.7	0.74	10.3
SVM	103	68.0	0.73	10.4
XG Boost	101	69.1	0.74	10.5
Deep ANN	100.5	74.3	0.73	10.8

Voting Ensemble	99.9	68.5	0.74	10.1
Stacking Ensemble	99.9	69.3	0.74	10.2

While deep ANN performs somewhat worse, the training results for RF, SVM, and XGBoost are relatively comparable. As a consequence, we combined the two ensemble models that performed the best among the three models we used in the training phase, improving the cross-validation scores. We then display the test set results. Table 3 provides a summary of the test set outcomes.

TABLE III. TEST SCORES FOR SESSION DURATION

Model	**RMSE (mins)**	**MAE (mins)**	**R²**	**SMAPE (%)**
RF	98	64.7	0.64	10.1
SVM	102	68.0	0.63	10.1
XG Boost	101	69.1	0.64	10.1
Deep ANN	100.5	74.3	0.53	10.8
Voting Ensemble	97.9	68.5	0.74	9.92
Stacking Ensemble	97.9	67.3	0.74	9.95
User predictions	430	394	-4.20	69.9

As said, the ensemble learning strategy yields the greatest outcomes.

ENERGY CONSUMPTION PREDICTIONS

This method was also used to the session length prediction. The deep ANN design, was the lone exception. There were twenty epochs. All of them having size of 64. The train set's 10-fold cross validation scores are summarised in Table 4.

The main standard metrics used in the following table are:

- RMSE
- MAE
- R^2
- SMAPE

Here RMSE and MAE are entered in terms of kWh.

TABLE IV. TRAINING SCORES FOR ENERGY CONSUMPTION

Model	RMSE (kWh)	MAE (kWh)	R^2	SMAPE (%)
RF	5.49	3.40	0.69	11.9
SVM	5.65	3.53	0.67	12.6
XG Boost	5.56	3.49	0.68	12.4
Deep ANN	5.61	3.60	0.67	12.9
Voting Ensemble	5.50	3.42	0.69	12.0
Stacking Ensemble	5.48	3.40	0.69	11.9

While the scores of remaining 3 techniques are comparable, RF has the greatest ratings. The top-most three models such as

- SVM
- XG Boost
- RF

are selected for the creation of 2 ensemble techniques. The results of the train which were produced by the ensemble techniques, were comparable to the top-performing RF model rather than outperforming it. In Table 5, the test set results are displayed.

TABLE V. TEST SCORES FOR ENERGY CONSUMPTION

Model	RMSE (mins)	MAE (mins)	R^2	SMAPE (%)
RF	5.50	3.39	0.54	11.7
SVM	5.69	3.54	0.51	12.4
XG Boost	5.61	3.48	0.51	12.1
Deep ANN	5.65	3.55	0.55	12.5
Voting Ensemble	5.54	3.41	0.69	11.8
Stacking Ensemble	5.50	3.38	0.70	11.6
User predictions	20.6	11.8	0.04	55.0

ANALYSIS AND DISCUSSION

Upon examining the SMAPE and total R2 of both forecasts, it seems that the energy consumption prediction may be more challenging. This aligns with the previous works using ACN data. On the other hand, the reverse was seen in another instance [24], i.e., it was simpler to anticipate energy usage.

Furthermore, in the two cases the anticipation of the performer about their action differed significantly from their original action, underscoring necessity of analysis. Better R2 and SMAPE values show that users' forecasts regarding their energy usage are somewhat more accurate than their predictions regarding the length of the session. Moreover, in both instances, ensemble learning predictions beat those of individual ML models, with the impact being more pronounced for session time prediction. This is due to the fact that in first scenario, the training performances of the top 3 performing models were comparable, and merging their predictions produced an improvement. Figure 7 and 8 shows the validation loss curve.[35][36].

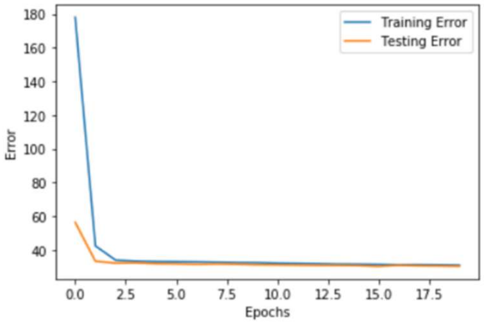

Fig. 7. Session duration's curve of validation loss

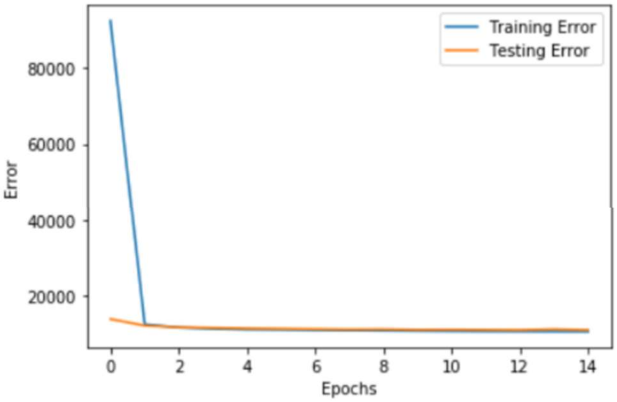

Fig. 8. Energy consumption's curve of validation loss

TABLE VI. COMPARING PERFORMANCE TO EARLIER WORK

Session	**Energy**	**Dataset**

Duration	**Consumption**	
SMAPE: 14.4%	SMAPE: 14.9%	ACN (historical charging)
MAE: 80 minute	Not considered	German charging data (historical charging, vehicle & location info)
Not considered	R2: 0.56	Nebraska public charging (historical charging, temporal & location)
SMAPE: 9.4%	SMAPE: 8.5%	UCLA campus (historical charging) and Residential charging data from UK

Conclusion

In this study, we proposed an advanced system for the scheduling-aware prediction of two critical EV charging behaviors: duration for the EV session, energy usage during these sessions. Unlike previous research efforts that primarily rely on historical charge data alone, this approach integrates additional contextual information such as weather conditions, traffic patterns, and local events. This comprehensive dataset enables a more accurate and holistic prediction of charging behaviors. To achieve this, we trained two sophisticated ensemble learning algorithms along with four well-known ML models: SVM, XGBoost, Deep ANN, and Random Forest. These results indicate that the prediction performance of this models significantly outperforms previous studies. Moreover, the machine learning methodology was applied to analyse the vast amount of test-related data, enabling the forecasting of energy use and identification of the primary variables influencing it. The inclusion of weather and traffic data has proven particularly beneficial, providing valuable insights that enhance prediction accuracy. By applying these enhanced models to the ACN dataset, we demonstrated a substantial improvement in identifying both length of EV charging sessions and associated energy consumption. This work not only advances the state of the EV charging behavior prediction, also but underscores the importance of incorporating diverse data sources to achieve more reliable and robust outcomes. In order to evaluate generalizability and scalability and enable the development of globally adaptive EV charging infrastructure, future research could also concentrate on applying these models across various geographic regions or car types.

References

[1] Çolak, B. A new study on the prediction of the effects of road gradient and coolant flow on electric vehicle battery power electronics components using machine learning approach. J. Energy Storage 2023

[2] Guo, X.; Wang, K.; Yao, S.; Fu, G.; Ning, Y. RUL prediction of lithium-ion battery based on CEEMDAN-CNN BiLSTM model. Energy Rep. 2023

[3] Li, D.; Liu, P.; Zhang, Z.; Zhang, L.; Deng, J.; Wang, Z.; Dorrell, D.G.; Li, W.; Sauer, D.U. Battery Thermal Runaway Fault Prognosis in Electric Vehicles Based on Abnormal Heat Generation and Deep Learning Algorithms. IEEE Trans. Power Electron. 2022

[4] Dataset Link. Available online: https://www.kaggle.com/datasets/ignaciovinuales/battery-remaining-useful-life-rul (accessed on 30 November 2023).

[5] S. Dias Vasconcelos *et al*., "Assessment of Electric Vehicles Charging Grid Impact via Predictive Indicator," in *IEEE Access*, vol. 12, pp. 163307-163323, 2024, doi: 10.1109/ACCESS.2024.3482095.

[6] Ali, A.; Emran, N.A.; Asmai, S.A. Missing values compensation in duplicates detection using hot deck method. J. Big Data 2021

[7] Montesinos López, O.A.; Montesinos López, A.; Crossa, J. Overfitting, Model Tuning, and Evaluation of Prediction Performance. In Multivariate Statistical Machine Learning Methods for Genomic Prediction; Springer: Cham, Switzerland, 2022.

[8] Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy 2021

[9] Yu, Y.; Zhu, Y.; Li, S.; Wan, D. Time Series Outlier Detection Based on Sliding Window Prediction. Math. Probl. Eng. 2014

[10] Krishnan, S.; Aruna, S.K.; Kanagarathinam, K.; Venugopal, E. Identification of Dry Bean Varieties Based on Multiple Attributes Using CatBoost Machine Learning Algorithm. Sci. Program. 2023

[11] Khan, F.N.U.; Rasul, M.G.; Sayem, A.S.M.; Mandal, N. Maximizing energy density of lithium-ion batteries for electric vehicles: A critical review. Energy Rep. 2023

[12] Alanazi, F. Electric Vehicles: Benefits, Challenges, and Potential Solutions for Widespread Adaptation. Appl. Sci. 2023

[13] Kumar, M.; Panda, K.P.; Naayagi, R.T.; Thakur, R.; Panda, G. Comprehensive Review of Electric Vehicle Technology and Its Impacts: Detailed Investigation of Charging Infrastructure, Power Management, and Control Techniques. Appl. Sci. 2023

[14] S. S. A. Naqvi *et al*., "Evolving Electric Mobility Energy Efficiency: In-Depth Analysis of Integrated Electronic Control Unit Development in Electric Vehicles," in *IEEE Access*, vol. 12, pp. 15957-15983, 2024, doi: 10.1109/ACCESS.2024.3356598.

[15] Uzair, M.; Abbas, G.; Hosain, S. Characteristics of Battery Management Systems of Electric Vehicles with Consideration of the Active and Passive Cell Balancing Process. World Electr. Veh. J. 2021

[16] S. Zou *et al*., "Design and Analysis of a Novel Multimode Powertrain for a PHEV Using Two Electric Machines," in *IEEE Access*, vol. 12, pp. 76442-76457, 2024, doi: 10.1109/ACCESS.2024.3406541

[17] Li, X.; Yu, D.; Byg, V.S.; Ioan, S.D. The development of machine learning-based remaining useful life prediction for lithium-ion batteries. J. Energy Chem. 2023

[18] Chou, J.-H.; Wang, F.-K.; Lo, S.-C. Predicting future capacity of lithium-ion batteries using transfer learning method. J. Energy Storage 2023

[19] Zhao, J.; Ling, H.; Liu, J.; Wang, J.; Burke, A.F.; Lian, Y. Machine learning for predicting battery capacity for electric vehicles. eTransportation 2023

[20] Zoerr, C.; Sturm, J.J.; Solchenbach, S.; Erhard, S.V.; Latz, A. Electrochemical polarization-based fast charging of lithium-ion batteries in embedded systems. J. Energy Storage 2023

[21] Najera-Flores, D.A.; Hu, Z.; Chadha, M.; Todd, M.D. A Physics-Constrained Bayesian neural network for battery remaining useful life prediction. Appl. Math. Model. 2023

[22] Dominguez, D.Z.; Mondal, B.; Gaberscek, M.; Morcrette, M.; Franco, A.A. Impact of the manufacturing process on graphite blend electrodes with silicon nanoparticles for lithium-ion batteries. J. Power Source 2023

[23] M. Abdelsattar, M. A. Ismeil, M. M. Aly and S. Saber Abu-Elwfa, "Analysis of Renewable Energy Sources and Electrical Vehicles Integration Into Microgrid," in *IEEE Access*, vol. 12, pp. 66822-66832, 2024, doi: 10.1109/ACCESS.2024.3399124.

[24] Uzair, M.; Abbas, G.; Hosain, S. Characteristics of Battery Management Systems of Electric Vehicles with Consideration of the Active and Passive Cell Balancing Process. World Electr. Veh. J. 2021

[25] Wang, Y.; Zhao, Y.; Addepalli, S. Remaining Useful Life Prediction using Deep Learning Approaches: A Review. Procedia Manuf. 2020

[26] H. Maghfiroh, O. Wahyunggoro and A. I. Cahyadi, "Energy Management in Hybrid Electric and Hybrid Energy Storage System Vehicles: A Fuzzy Logic Controller Review," in *IEEE Access*, vol. 12, pp. 56097-56109, 2024, doi: 10.1109/ACCESS.2024.3390436.

[27] Khan, F.N.U.; Rasul, M.G.; Sayem, A.S.M.; Mandal, N. Maximizing energy density of lithium-ion batteries for electric vehicles: A critical review. Energy Rep. 2023

[28] Alanazi, F. Electric Vehicles: Benefits, Challenges, and Potential Solutions for Widespread Adaptation. Appl. Sci. 2023

[29] Kumar, M.; Panda, K.P.; Naayagi, R.T.; Thakur, R.; Panda, G. Comprehensive Review of Electric Vehicle Technology and Its Impacts: Detailed Investigation of Charging Infrastructure, Power Management, and Control Techniques. Appl. Sci. 2023

[30] F. Noor, K. Zeb, S. Ullah, Z. Ullah, M. Khalid and A. Al-Durra, "Design and Validation of Adaptive Barrier Function Sliding Mode Controller for a Novel Multisource Hybrid Energy Storage System Based Electric Vehicle," in *IEEE Access*, vol. 12, pp. 145270-145285, 2024, doi: 10.1109/ACCESS.2024.3471893.

[31] Montesinos López, O.A.; Montesinos López, A.; Crossa, J. Overfitting, Model Tuning, and Evaluation of Prediction Performance. In Multivariate Statistical Machine Learning Methods for Genomic Prediction; Springer: Cham, Switzerland, 2022.

[32] Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy 2021

[33] Yu, Y.; Zhu, Y.; Li, S.; Wan, D. Time Series Outlier Detection Based on Sliding Window Prediction. Math. Probl. Eng. 2014.

[34] X. Zhang, F. Gao, X. Gong, Z. Wang, and Y. Liu, "Comparison of Climate Change Impact Between Power System of Electric Vehicles and Internal Combustion Engine Vehicles," in Advances in Energy and Environmental Materials, Singapore, 2018, pp. 739–7

[35] Jiang, Y.; Song, W. Predicting the Cycle Life of Lithium-Ion Batteries Using Data-Driven Machine Learning Based on Discharge Voltage Curves. Batteries 2023

[36] Gandhi, S.M.; Sarkar, B.C. Chapter 11—Conventional and Statistical Resource/Reserve Estimation. In Essentials of Mineral Exploration and Evaluation; Gandhi, S.M., Sarkar, B.C., Eds.; Elsevier: Amsterdam, The Netherlands, 2016

[37] Montesinos López, O.A.; Montesinos López, A.; Crossa, J. Overfitting, Model Tuning, and Evaluation of Prediction Performance. In Multivariate Statistical Machine Learning Methods for Genomic Prediction

[38] Wang, Y.; Zhao, Y.; Addepalli, S. Remaining Useful Life Prediction using Deep Learning Approaches: A Review. Procedia Manuf. 2020

[39] Wu, J.; Kong, L.; Cheng, Z.; Yang, Y.; Zuo, H. RUL Prediction for Lithium Batteries Using a Novel Ensemble Learning Method. Energy Rep. 2022