International Journal of
Computational Learning & Intelligence
Vol. 04 | Issue 04 | 2025

RESEARCH ARTICLE OPEN ACCESS

Predicting EV Battery Lifespan Using Machine Learning

N Vasavi . A Akshith Reddy . K Poorna Chandra . K S S Ramakrishna . P Prasanthi

Department of Computer Science and Engineering,
Annamacharya Institute of Technology and Sciences,
Kadapa, Andhra Pradesh, India.

DOI: **10.5281/zenodo.15250347**
Received: 27 January 2025 / Revised: 21 February 2025 / Accepted: 27 March 2025

Abstract – The continuous advancement of electric vehicle (EV) technology has heightened the emphasis on sustainable energy storage, making lithium-ion batteries a crucial component. Ensuring battery reliability and longevity is essential for optimizing EV performance and reducing maintenance costs. This study explores the prediction of Remaining Useful Life (RUL) for lithium-ion batteries using advanced Machine Learning (ML) models, specifically Random Forest (RF) and Support Vector Machine (SVM). Accurate RUL estimation enhances battery management, prevents failures, and improves safety.A comprehensive dataset from the NASA Ames Prognostics Center of Excellence is preprocessed, with the One-way ANOVA method applied for optimal feature selection. Data normalization techniques are employed to enhance model consistency, while hyperparameter tuning (HPT) optimizes predictive performance. Real-time factors such as temperature fluctuations and usage cycles are incorporated to analyze their impact on battery degradation. The proposed system provides deeper insights into battery aging trends, enabling proactive maintenance strategies.Model performance is evaluated using R2 score and Mean Squared Error (MSE), where the RF model achieves an R2 score of 0.83 and an MSE of 1.67, demonstrating high reliability. The results contribute to improving battery efficiency and safety through predictive modeling, facilitating better battery management in EVs. By leveraging ML-driven predictive analytics, this research supports the advancement of sustainable and cost-effective energy solutions, promoting wider EV adoption and a greener future.

Index Terms – Electric Vehicle Batteries, Machine Learning (ML), Remaining Useful Life (RUL), Random Forest (RF), Support Vector Machine (SVM).

I. INTRODUCTION

Electric vehicles (EVs) are emerging as a solution to reduce carbon emissions and dependence on fossil fuels, addressing global environmental challenges. India, as part of its sustainable energy goals, has introduced initiatives such as the FAME scheme to boost EV adoption. Statistics indicate that India aims to electrify 30% of its vehicle fleet by 2030, with EV sales growing at a compound annual growth rate (CAGR) of 49% over the past five years. However, the key challenge lies in ensuring battery reliability, as unexpected battery failure not only increases costs but also impacts consumer trust. Predicting the Remaining Useful Life (RUL) of EV batteries through machine learning is a transformative approach to overcome these challenges, ensuring enhanced operational efficiency, better resource utilization, and sustainability.

Traditional methods for predicting EV battery life rely on rule-based systems and empirical models, which often fail to account for dynamic battery behaviors and real-time usage patterns. Machine learning models, on the other hand, offer precise predictions by analyzing complex degradation patterns, leading to improved EV fleet operations, battery recycling, and cost-effective maintenance. Applications include optimizing EV fleet logistics, managing battery warranty programs, and enabling predictive maintenance for commercial EV fleets.

This study presents a machine learning-based framework for predicting the RUL of EV batteries. The system is designed to analyze real-time battery performance indicators, such as voltage, current, temperature, and charge-discharge cycles, to provide accurate RUL estimations. Advanced machine learning techniques, including Bagging with Decision Trees, are employed to enhance prediction accuracy. Research papers, such as "Machine Learning Approaches for Battery Lifetime Prediction" and "Data-Driven Predictive Models for EV Battery Health," provide a strong foundation for implementing such a system.

The goal of this research is to:

1. Develop an AI-driven model that improves battery life prediction accuracy over traditional methods.
2. Evaluate the effectiveness of machine learning algorithms in predicting the Remaining Useful Life (RUL) of EV batteries.
3. Provide real-time battery health insights to optimize battery performance and lifecycle management.

Section II reviews existing literature on EV battery health prediction and machine learning in battery management. Section III describes the methodology used, including data collection, preprocessing, and model development. Section IV presents the results and analysis, highlighting the model's predictive accuracy. Section V discusses the findings, their implications, and future research directions. Finally, Section VI concludes the study by summarizing the contributions and potential applications of the proposed RUL prediction model in sustainable energy solutions. This research aims to bridge the gap between AI-driven predictive analytics and EV battery management, contributing to the broader goal of improving battery efficiency, reliability, and sustainability in electric transportation.

II. LITERATURE SURVEY

As the era of electrification progresses, characterized by the rapid expansion of renewable energy-based electric vehicles (EVs), the significance of energy storage systems has increased considerably. Li-ion batteries are considered superior to other energy storage technologies due to their high energy density, robust power output, and low self-discharge rates [1], [2]. The growing demand for efficient battery management technologies has necessitated advancements in battery management systems (BMS). A well-rounded BMS includes functions such as data gathering, status analysis and forecasting, charging and discharging control, safety mechanisms, thermal management, power balancing, and communication systems [3]. The efficiency of a BMS is assessed by its accuracy in state estimation, ensuring stable energy storage system operation and prolonged battery life [4]. An ideal BMS should possess multitasking capabilities and integrate a real-time operating system for continuous monitoring and rapid adjustments.

Despite their advantages, lithium-ion batteries have limitations, particularly in terms of lifespan and cost, which hinder widespread adoption [5]. Battery performance deteriorates over time due to calendar ageing and cycle ageing, leading to various degradation events [6], [7]. Ageing increases operational costs, reduces equipment lifespan, and affects safety [8], [9], [10]. A battery is generally considered to have reached its end-of-life when its capacity declines to 80% of its initial value [11]. The Remaining Useful Life (RUL) refers to the predicted operational period from the current state until the battery reaches its end-of-life [12], [13], [14]. Battery longevity is influenced by its composition and internal chemical changes occurring during charge-discharge cycles. This ageing process is intricate and non-linear, affected by factors such as temperature, charge/discharge rates, and environmental conditions [15], [16], [17], [18]. Accurately predicting RUL in complex scenarios is a significant challenge. In industrial settings, precise RUL predictions can reduce investment costs and improve profitability [19], [20]. They also enhance energy storage system safety, stability, and longevity [21], [22].

RUL prediction methods can be categorized into three main approaches: model-based, data-driven, and hybrid methods [23], [24]. The model-based approach formulates a mathematical representation of internal battery physical and electrochemical reactions to generate predictive models that assess battery status [25], [26], [27]. However, developing these models is computationally intensive and requires extensive parameterization, making practical implementation challenging. Nevertheless, once established, they offer high accuracy [28], [29]. Conversely, data-driven methods utilize historical data, making them more practical for lithium-ion battery applications due to the complexity of battery behavior [30], [31], [32].

Accurately predicting the RUL of EV batteries is crucial for optimizing both performance and safety. This study aims to enhance RUL predictions by modeling historical battery data using supervised learning techniques such as Random Forest (RF) and Support Vector Machine (SVM). The historical battery data is sourced from NASA Ames Prognostics Center of Excellence (PCoE). Model accuracy is improved using feature selection techniques, including one-way ANOVA and hyperparameter tuning. The effectiveness of the models is assessed based on R^2 and Mean Squared Error (MSE) values. A key novelty of this study is obtaining real-time data on battery capacity decline and temperature fluctuations.

Additionally, it establishes the relationship between battery consumption cycles and the percentage of capacity decline. The research was conducted using Google Colab with a 12GB NVIDIA Tesla K80 GPU for hardware acceleration. The dataset is split into training and testing groups for preprocessing and essential feature extraction, ensuring a robust framework for accurate RUL prediction.

III. METHODOLOGY

This research provides an intricate model aimed at forecasting the Remaining Useful Life (RUL) of electric vehicle (EV) batteries. The methodology, as represented in our workflow diagram, begins with painstaking data pre-processing, involving the imputation of missing values from a comprehensive dataset acquired from the NASA Ames Prognostics Centre of Excellence. Feature selection and training of machine learning models follow, with a focus on Random Forest and Support Vector Regressor, and performance-enhancing Hyperparameter adjustment serving to reinforce the training as shown in fig-1.

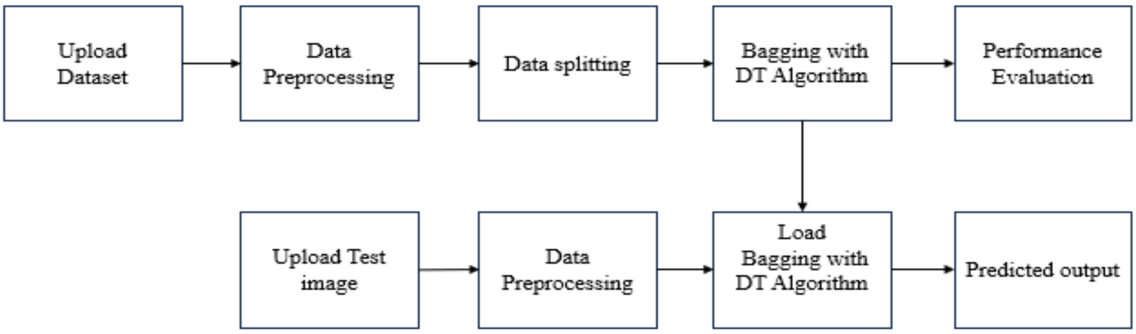

Fig. 1: System Architecture

1. Collection and structure

The dataset used in this work is acquired from a custom-built battery prognostics laboratory at the NASA Ames Prognostics Center of Excellence (PCoE). The data set is structured in cycles for 29 batteries in .mat file format, each having charge, discharge, & impedance operations. The data comprises multiple parameters, such as voltage, current, temperature, and impedance, acquired during different battery operations.

2. Data Preprocessing

The originally iterated through each .mat file, representing individual battery cycles, to extract pertinent data. This data was then methodically formatted into a CSV file, boosting ease of handling and analysis. To resolve gaps in the dataset, missing values in important features, including voltage, current, and temperature, underwent mean imputation, assuring dataset completeness. The main objective of our investigation was determining the capacity %age for each cycle, using the formula in Equation (1):

$$\%age = (C_1C_1 - C_2) * 100 \qquad (1)$$

International Journal of
Computational Learning & Intelligence
Vol. 04 | Issue 04 | 2025

Where C_1C_1 is the maximum available capacity of the current cycle, and C_2 is the maximum available capacity of the beginning cycle. This computation is essential for understanding the degradation patterns in EV battery life over multiple cycles.

Correlation Matrix Insights

The correlation matrix gives substantial insights into the relationship between various features and battery health. Temperature change and discharge indicate a considerable association with battery capacity, affirming the decision to do regression tasks on these variables, as shown in Figure 2.

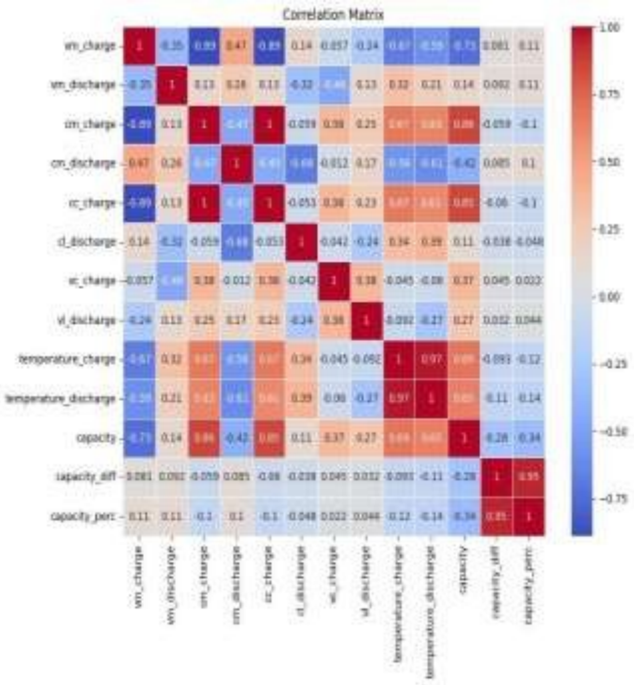

Fig2: Correlation Matrix

3. Feature Selection

Selecting the most important features is essential for constructing precise predictive models. The study applies the One-Way Analysis of Variance (ANOVA) method as a credible statistical method to find features that have a substantial impact on battery degradation. In this process, the important features of battery life deterioration are selected carefully for analysis. It minimised the complexity of the model and enhanced the overall accuracy of prediction using IEEE Transaction Access on Machine Learning,Volume:12,Issue Date:16.September.2024 the ANOVA (Analysis of variation) method. The process is based on comparing the variation within groups to the variance between groups [39] based on F-statistic calculation using Equation (2):

$$F = (\sigma\ 2Btn\ grps) / (\sigma 2Within\ grps) \qquad (2)$$

Where σ^2 represents the variance. The importance of the feature is decided by the F-value, as shown in Figure 3. The more important feature has a higher F-value. The impact of these features is more on the dependent variable.

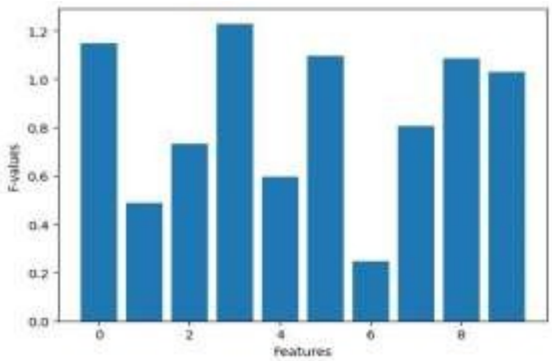

Fig 3: F-value using ANOVA

4. Train And Split

The datasets are split into two parts for analysis using machine learning. 70% of datasets are allocated for training the model, and 30% of datasets are allocated for testing the model finalised based on accuracy during the training.

5. Machine Learning Models

There are many machine learning techniques that are used for data analysis. The model is decided on the basis of the type of dataset. RF and SVM are more suitable models for battery Remaining Useful Battery robust prediction. These models can handle high-dimensional data and capture complex relationships. Hyperparameter tuning is used to improve the model's performance.

IV. RESULTS AND DISCUSSIONS

This study developed a machine learning-based framework to predict battery capacity degradation using key factors such as temperature, discharge cycles, voltage, and current. The Random Forest Regressor model achieved an R^2 score of 0.83 with a Mean Squared Error (MSE) of 1.67 for predicting capacity loss due to temperature changes. The Support Vector Regressor (SVR) demonstrated R^2 scores between 0.90 and 0.99, effectively capturing the relationship between cycle count and battery degradation. When multiple features were considered, the Random Forest model achieved an R^2 score of 0.73 and an MSE of 1.9, highlighting the combined impact of various factors on battery health. These results confirm the proposed model's effectiveness in accurately predicting battery life, enabling better battery management and optimizing EV performance. Future work aims to improve real-time adaptability, integrate additional influencing factors, and expand datasets for enhanced generalization.

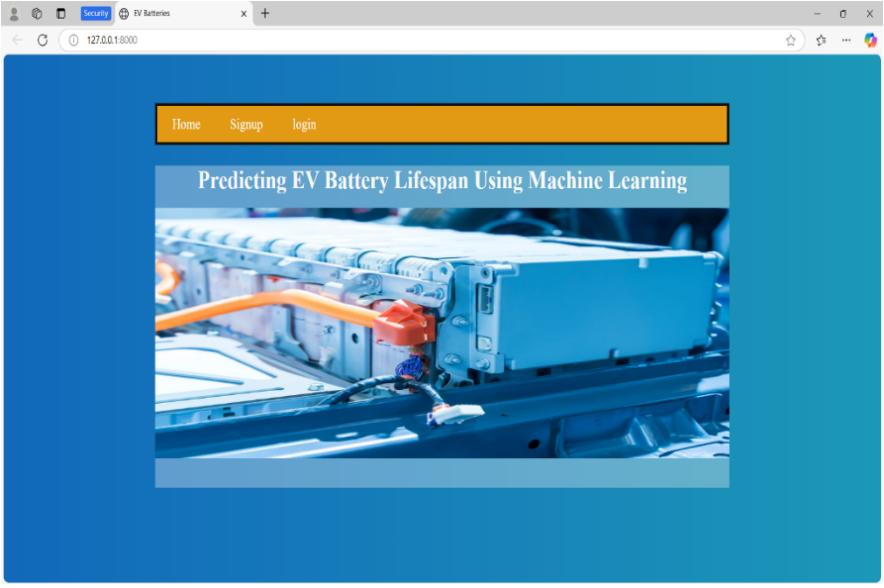

Fig 4: Home page

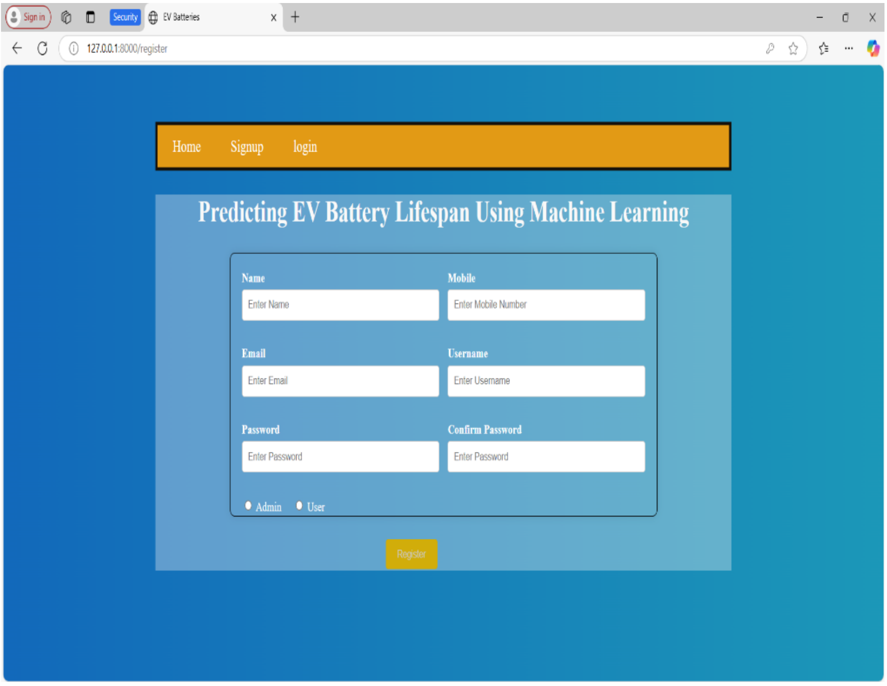

Fig 5: Registration page for admin and user

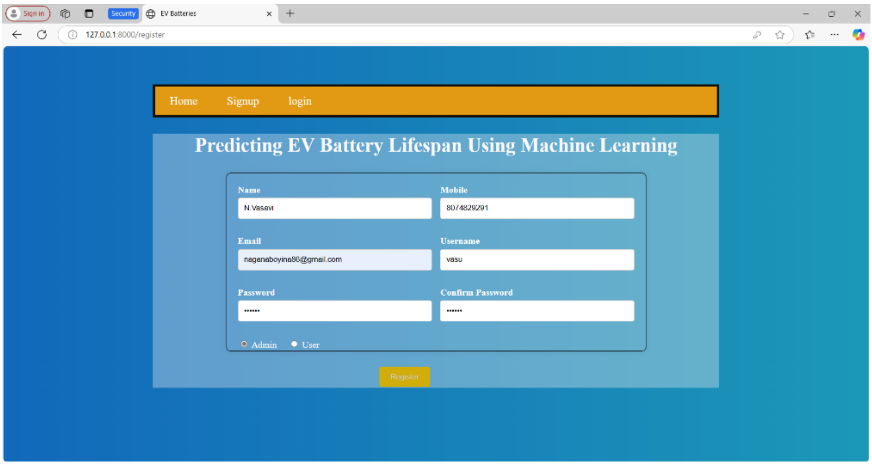

Fig 6: Admin registration details

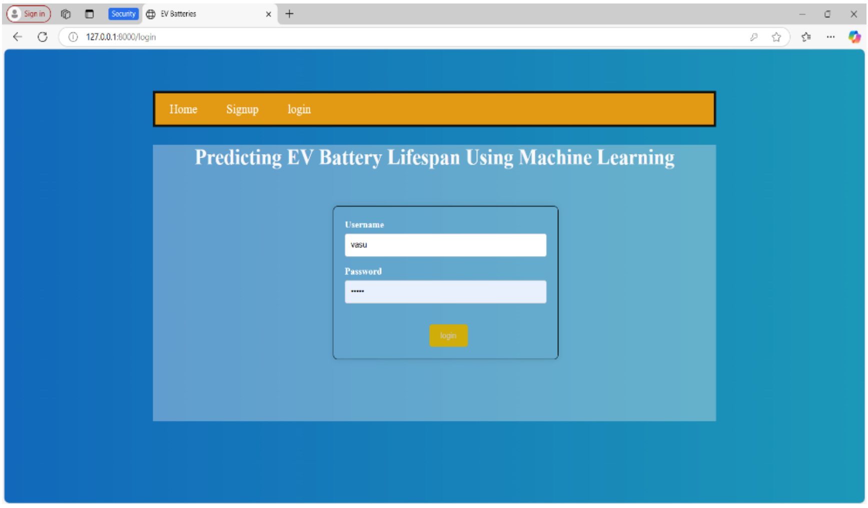

Fig 7: Login page

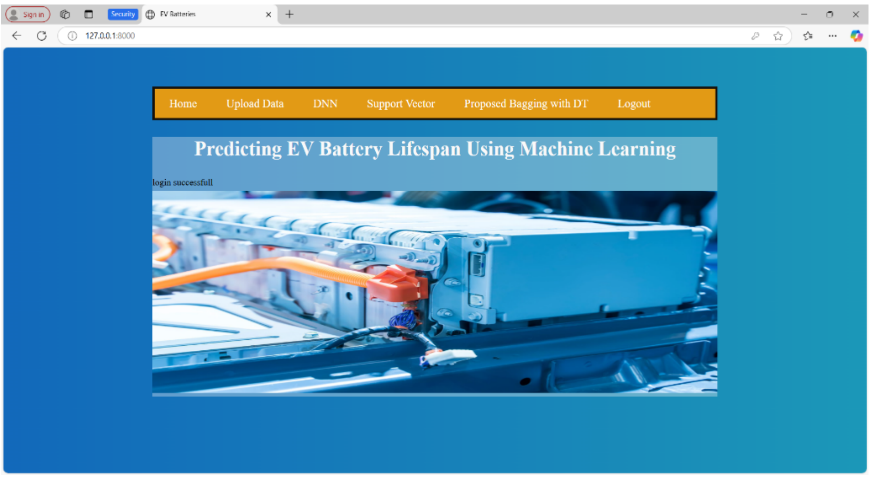

Fig 8: Login successful page

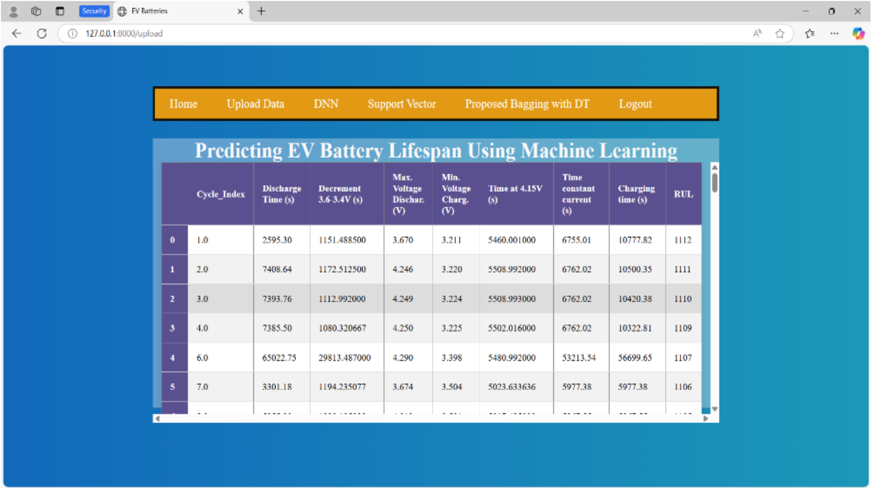

Fig 9: Uploading data

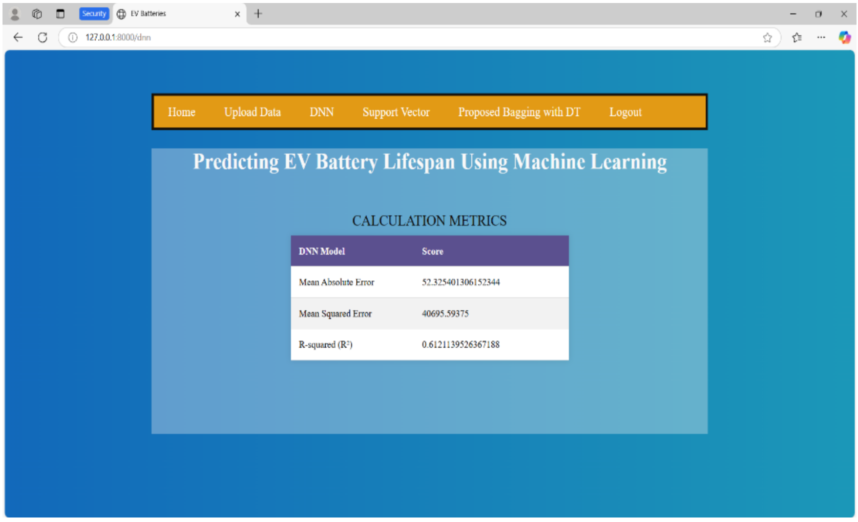

Fig 10: DNN calculation metrics

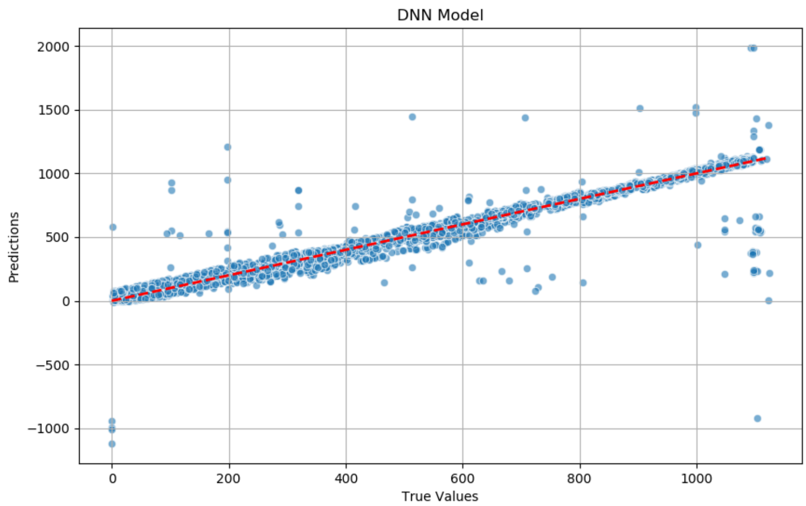

Fig 11: DNN model

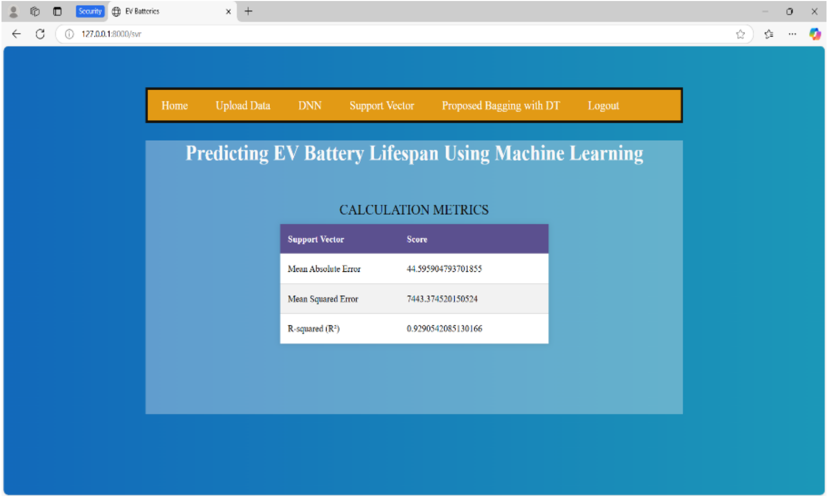

Fig 12: Support vector calculation metrics

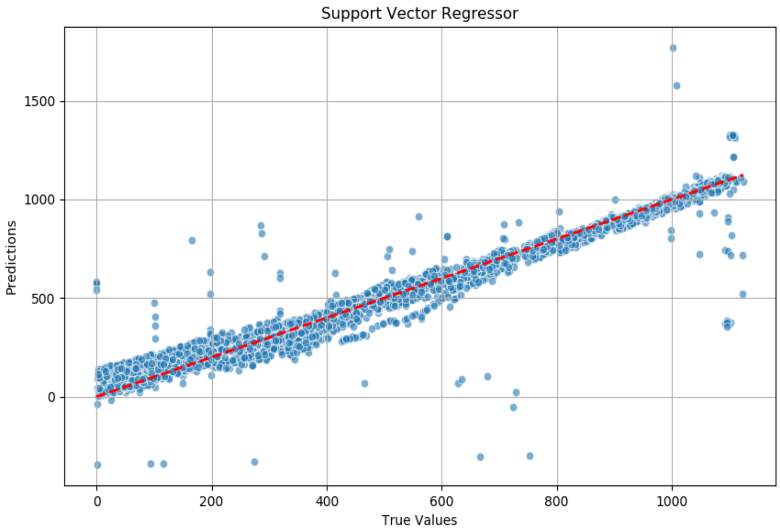

Fig 13: Support vector regressor

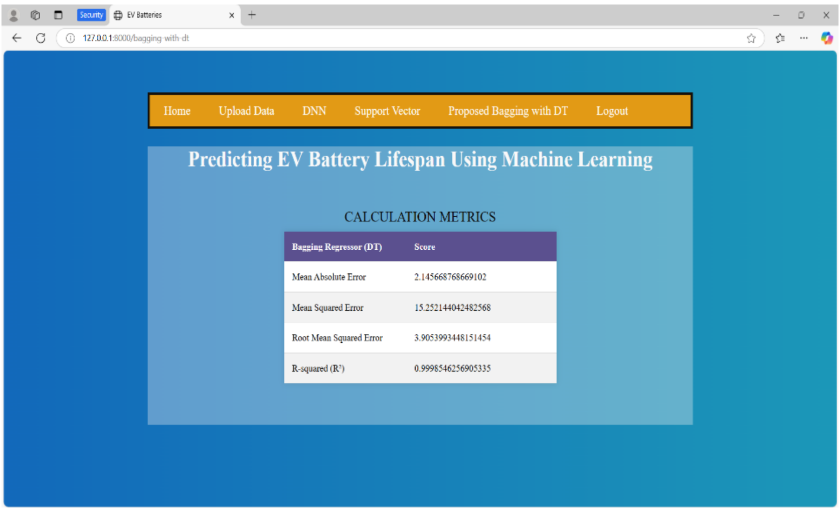

Fig 14: Bagging regressor calculation metrics

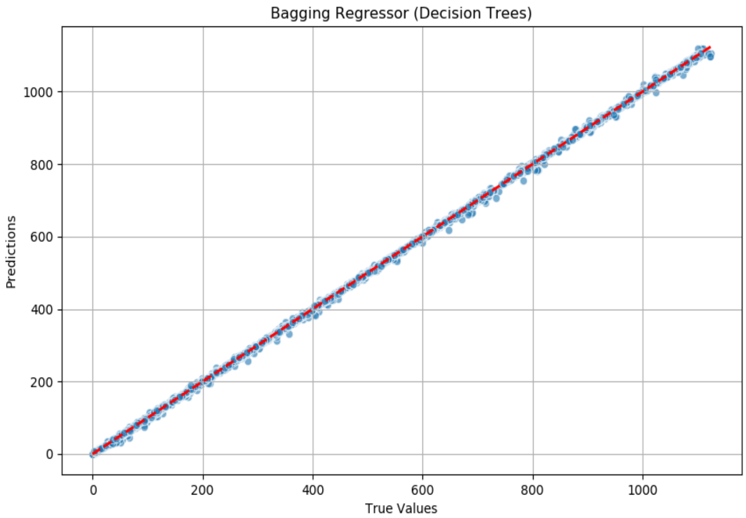

Fig 15: Bagging regressor

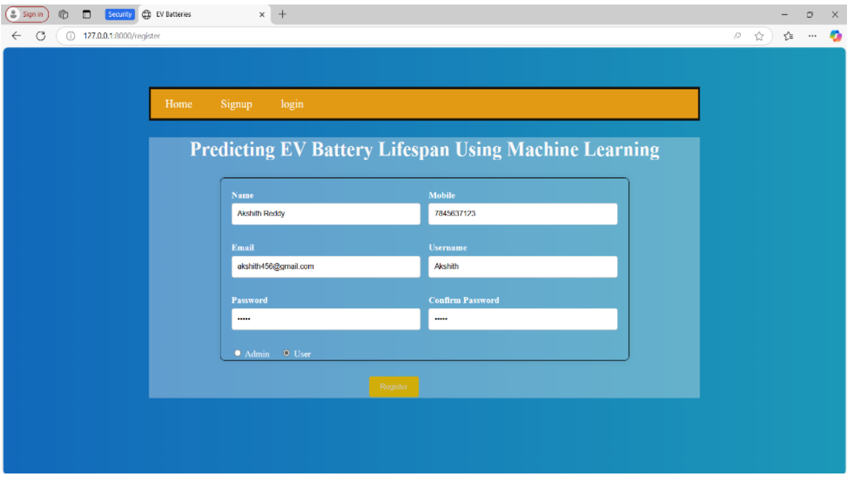

Fig 16: User registration details

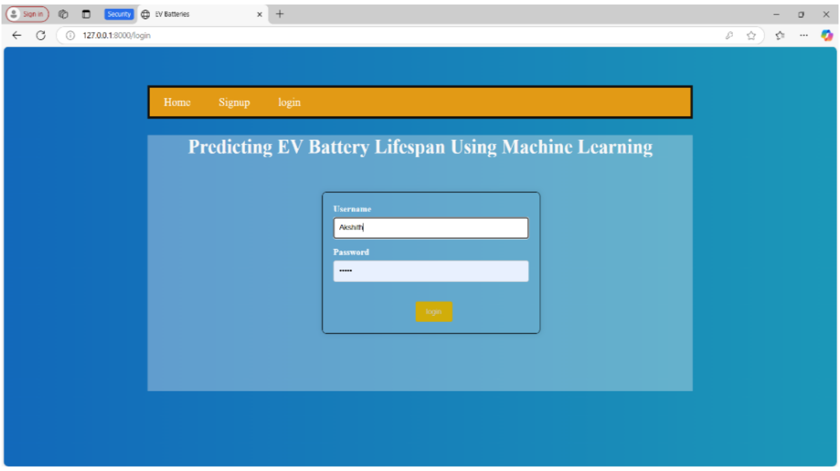

Fig 17: Login page

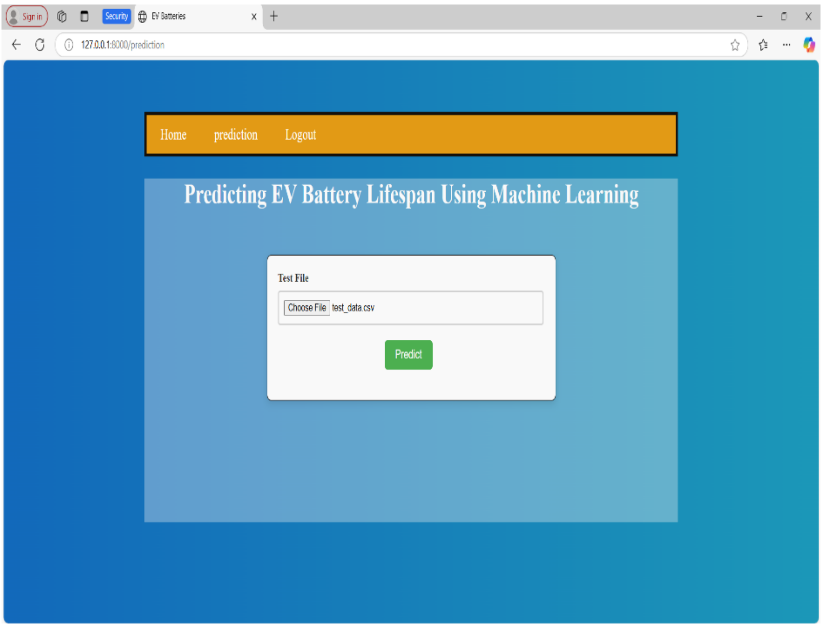

Fig 18: Prediction page

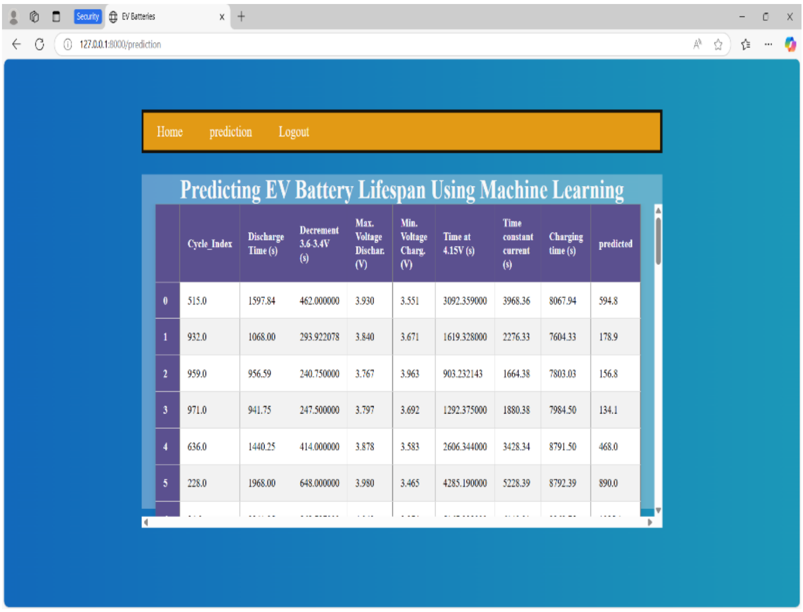

Fig 19: Prediction of RUL

V. CONCLUSION AND FUTURE WORK

This study analyzed a dataset that provides valuable insights into EV battery performance and health across different charging and discharging cycles. By examining key parameters such as discharge time, voltage levels, and charging durations, we can model and predict the RUL of EV batteries. This predictive capability enables timely maintenance and replacements, ensuring optimal performance and efficiency. The dataset is essential for researchers and engineers working on battery degradation analysis, health monitoring systems, and predictive maintenance models, contributing to improved battery longevity and reliability in applications like electric vehicles and energy storage systems. Despite its promising insights, this study has certain limitations. The dataset was collected under specific operating conditions, which may affect its generalizability to broader EV applications. Incorporating more diverse datasets, including different battery chemistries, environmental conditions, and real-world driving patterns, will enhance the robustness of predictive models.

Another key area for improvement is the integration of real-time battery health monitoring. Implementing an IoT-based dynamic learning approach that continuously updates battery models based on live sensor data could improve prediction accuracy and responsiveness. Additionally, expanding the dataset to include factors like thermal management, charging patterns, and user behavior could provide a more comprehensive understanding of battery performance.

REFERENCES

1. Gao, Y., Zhang, X., Guo, B., Zhu, C., Wiedemann, J., Wang, L., & Cao, J. (2020). Health-aware multiobjective optimal charging strategy with coupled electrochemical-thermal-aging model for lithium-ion battery. *IEEE Transactions on Industrial Informatics, 16*(5), 3417–3429. https://doi.org/10.1109/TII.2019.2949815
2. Peng, J., Zhou, Z., Wang, J., Wu, D., & Guo, Y. (2019). Residual remaining useful life prediction method for lithium-ion batteries in satellite with incomplete healthy historical data. *IEEE Access, 7*, 127788–127799. https://doi.org/10.1109/ACCESS.2019.2939480
3. Gabbar, H., Othman, A., & Abdussami, M. (2021). Review of battery management systems (BMS) development and industrial standards. *Technologies, 9*(2), 28. https://doi.org/10.3390/technologies9020028
4. Xing, Y., Ma, E. W. M., Tsui, K. L., & Pecht, M. (2011). Battery management systems in electric and hybrid vehicles. *Energies, 4*(11), 1840–1857. https://doi.org/10.3390/en4111840
5. Ahmed, S. T., Kumar, V. V., Singh, K. K., Singh, A., Muthukumaran, V., & Gupta, D. (2022). 6G enabled federated learning for secure IoMT resource recommendation and propagation analysis. *Computers and Electrical Engineering*, *102*, 108210.
6. Ahmed, S. T., Singh, D. K., Basha, S. M., Abouel Nasr, E., Kamrani, A. K., & Aboudaif, M. K. (2021). Neural network based mental depression identification and sentiments classification technique from speech signals: A COVID-19 Focused Pandemic Study. *Frontiers in public health*, *9*, 781827.
7. Ahmed, S. T., Vinoth Kumar, V., Mahesh, T. R., Narasimha Prasad, L. V., Velmurugan, A. K., Muthukumaran, V., & Niveditha, V. R. (2024). FedOPT: federated learning-based heterogeneous resource recommendation and optimization for edge computing. *Soft Computing*, 1-12.
8. Fathima, A. S., Basha, S. M., Ahmed, S. T., Khan, S. B., Asiri, F., Basheer, S., & Shukla, M. (2025). Empowering consumer healthcare through sensor-rich devices using federated learning for secure resource recommendation. *IEEE Transactions on Consumer Electronics*
9. Fathima, A. S., Basha, S. M., Ahmed, S. T., Mathivanan, S. K., Rajendran, S., Mallik, S., & Zhao, Z. (2023). Federated learning based futuristic biomedical big-data analysis and standardization. *Plos one*, *18*(10), e0291631.
10. Fathima, A. S., Reema, S., & Ahmed, S. T. (2023, December). ANN based fake profile detection and categorization using premetric paradigms on instagram. In *2023 Innovations in Power and Advanced Computing Technologies (i-PACT)* (pp. 1-6). IEEE
11. Guha, A., & Patra, A. (2018). Online estimation of the electrochemical impedance spectrum and remaining useful life of lithium-ion batteries. *IEEE Transactions on Instrumentation and Measurement, 67*(8), 1836–1849. https://doi.org/10.1109/TIM.2018.2803202
12. Hu, X., Xu, L., Lin, X., & Pecht, M. (2020). Battery lifetime prognostics. *Joule, 4*(2), 310–346. https://doi.org/10.1016/j.joule.2019.11.018
13. Jamshidi, M. B., & Rostami, S. (2017). A dynamic artificial neural network approach to estimate thermal behaviors of Li-ion batteries. In *Proceedings of the IEEE 2nd International Conference on Automatic Control and Intelligent Systems (I2CACIS)* (Vol. 7, pp. 116–121). IEEE. https://doi.org/10.1109/I2CACIS.2017.8239036

14. Jamshidi, M. B., Alibeigi, N., Lalbakhsh, A., & Roshani, S. (2019). An ANFIS approach to modeling a small satellite power source of NASA. In *Proceedings of the IEEE 16th International Conference on Networking, Sensing and Control (ICNSC)* (pp. 459–464). IEEE. https://doi.org/10.1109/ICNSC.2019.8743159
15. Jamshidi, M. B., Jamshidi, M., & Rostami, S. (2017). An intelligent approach for nonlinear system identification of a Li-ion battery. In *Proceedings of the IEEE 2nd International Conference on Automatic Control and Intelligent Systems (I2CACIS)* (pp. 98–103). IEEE. https://doi.org/10.1109/I2CACIS.2017.8239021
16. Kumar, A., Satheesha, T. Y., Salvador, B. B. L., Mithileysh, S., & Ahmed, S. T. (2023). Augmented Intelligence enabled Deep Neural Networking (AuDNN) framework for skin cancer classification and prediction using multi-dimensional datasets on industrial IoT standards. *Microprocessors and Microsystems*, *97*, 104755.
17. Lipu, M. S. H., Hannan, M. A., Hussain, A., Hoque, M. M., Ker, P. J., Saad, M. H. M., & Ayob, A. (2018). A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations. *Journal of Cleaner Production, 205*, 115–133. https://doi.org/10.1016/j.jclepro.2018.09.077
18. Liu, K., Li, Y., Hu, X., Lucu, M., & Widanage, W. D. (2020). Gaussian process regression with automatic relevance determination kernel for calendar aging prediction of lithium-ion batteries. *IEEE Transactions on Industrial Informatics, 16*(6), 3767–3777. https://doi.org/10.1109/TII.2019.2952906
19. Lui, Y. H., Li, M., Downey, A., Shen, S., Nemani, V. P., Ye, H., VanElzen, C., Jain, G., Hu, S., Laflamme, S., & Hu, C. (2021). Physics-based prognostics of implantable-grade lithium-ion battery for remaining useful life prediction. *Journal of Power Sources, 485*, Article 229327. https://doi.org/10.1016/j.jpowsour.2020.229327
20. Ma, J., Xu, S., Shang, P., Ding, Y., Qin, W., Cheng, Y., Lu, C., Su, Y., Chong, J., Jin, H., & Lin, Y. (2020). Cycle life test optimization for different Li-ion power battery formulations using a hybrid remaining-useful-life prediction method. *Applied Energy, 262*, Article 114490. https://doi.org/10.1016/j.apenergy.2020.114490
21. She, C., Wang, Z., Sun, F., Liu, P., & Zhang, L. (2020). Battery aging assessment for real-world electric buses based on incremental capacity analysis and radial basis function neural network. *IEEE Transactions on Industrial Informatics, 16*(5), 3345–3354. https://doi.org/10.1109/TII.2019.2949264
22. Shi, Y., Xu, B., Tan, Y., Kirschen, D., & Zhang, B. (2019). Optimal battery control under cycle aging mechanisms in pay for performance settings. *IEEE Transactions on Automatic Control, 64*(6), 2324–2339. https://doi.org/10.1109/TAC.2018.2867720
23. Wang, Y., Tian, J., Sun, Z., Wang, L., Xu, R., Li, M., & Chen, Z. (2020). A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renewable and Sustainable Energy Reviews, 131*, Article 110015. https://doi.org/10.1016/j.rser.2020.110015
24. Xiong, R., Li, L., & Tian, J. (2018). Towards a smarter battery management system: A critical review on battery state of health monitoring methods. *Journal of Power Sources, 405*, 18–29. https://doi.org/10.1016/j.jpowsour.2018.10.019
25. Xu, B., Zhao, J., Zheng, T., Litvinov, E., & Kirschen, D. S. (2018). Factoring the cycle aging cost of batteries participating in electricity markets. *IEEE Transactions on Power Systems, 33*(2), 2248–2259. https://doi.org/10.1109/TPWRS.2017.2749260
26. Zhang, C., Wang, Y., Gao, Y., Wang, F., Mu, B., & Zhang, W. (2019). Accelerated fading recognition for lithium-ion batteries with nickel-cobalt-manganese cathode using quantile regression method. *Applied Energy, 256*, Article 113841. https://doi.org/10.1016/j.apenergy.2019.113841