International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 04 | Apr 2026 www.irjet.net p-ISSN: 2395-0072

PREDICTIVE MAINTENANCE IN BATTERY HEALTH FOR ELECTRIC VEHICLES USING MACHINE LEARNING

Dr. L. Rasikannan[1], B. Priyanka[2], M. Murugeswari[3], S. Saranya[4], P. Yazhini[5]

[1]*Associate Professor, Department of CSE, Government College of Engineering, Srirangam, Tamilnadu, India*
[2,3,4,5]*UG student, Department of CSE, Government College of Engineering, Srirangam, Tamilnadu, India*

---***---

Abstract - *Electric Vehicles (EVs) rely on lithium-ion batteries as their primary energy source, making battery health monitoring a critical aspect of vehicle performance, safety, and longevity. Conventional battery management systems are limited to threshold-based monitoring and fail to provide predictive insights into potential failures. This paper proposes a machine learning-based predictive maintenance framework for EV battery health analysis using a Random Forest classifier. The system integrates battery dataset acquisition, preprocessing, feature selection, and classification to detect early-stage faults and categorize battery conditions into normal, warning, and critical states. Key parameters such as State of Charge (SOC), State of Health (SOH), temperature, voltage, and current are utilized for prediction. The proposed model enhances fault detection accuracy and enables proactive maintenance through automated email alerts. Furthermore, a Streamlit-based web dashboard provides real-time visualization and user interaction. Experimental results demonstrate that the proposed approach significantly improves prediction accuracy and reduces the risk of unexpected battery failures, thereby enhancing the reliability and efficiency of EV systems.*

Keywords**: Electric Vehicles, Battery Health Monitoring, Predictive Maintenance, Random Forest, Fault Detection, Streamlit, Machine Learning**

1. INTRODUCTION

The global transition towards sustainable and eco-friendly transportation has significantly accelerated the adoption of Electric Vehicles (EVs) in recent years. As governments and industries aim to reduce carbon emissions and dependency on fossil fuels, EVs have emerged as a promising alternative to conventional internal combustion engine vehicles. At the core of every EV lies the lithium-ion battery, which plays a crucial role in determining the vehicle's performance, efficiency, safety, and overall lifespan.

Despite the rapid advancements in battery technology, battery degradation remains a major challenge in EV systems. The performance of lithium-ion batteries is influenced by multiple factors such as charging cycles, temperature variations, depth of discharge, and current fluctuations. Over time, these factors lead to capacity fading, internal resistance increase, and eventual battery failure. Therefore, accurate monitoring and timely prediction of battery health have become essential to ensure reliability and prevent unexpected breakdowns.

Traditional Battery Management Systems (BMS) are primarily designed to monitor basic parameters such as voltage, current, and temperature. These systems rely on predefined threshold values to detect abnormalities. However, such rule-based approaches are reactive in nature and fail to provide early warnings about potential failures. As a result, faults are often detected only after significant degradation has occurred, leading to reduced battery lifespan, increased maintenance costs, and potential safety risks.

In recent years, the emergence of data-driven techniques and machine learning has opened new possibilities for predictive maintenance in EV systems. Machine learning models have the ability to analyze large volumes of historical and real-time data, identify hidden patterns, and make intelligent predictions about future system behavior. These capabilities make them highly suitable for battery health monitoring applications, where the relationship between parameters is complex and nonlinear.

To address these challenges, this paper proposes a predictive maintenance framework for EV battery health monitoring using a Random Forest-based machine learning model. Random Forest, an ensemble learning technique, is known for its high accuracy, robustness, and ability to handle high-dimensional data. By combining multiple decision trees, it reduces overfitting and improves generalization performance. The proposed system integrates multiple modules, including data acquisition, preprocessing, feature selection, fault prediction, and alert generation. Key battery parameters such as State of Charge (SOC), State of Health (SOH), temperature, voltage, and current are analysed to classify battery conditions into normal, warning, and critical states. In addition, an automated email alert mechanism is incorporated to notify administrators when critical conditions are detected, enabling timely intervention.

Furthermore, a Streamlit-based web dashboard is developed to provide real-time visualization and user interaction. The dashboard allows users to upload datasets, input parameters manually, and view prediction results in an intuitive manner. This enhances the usability and practical applicability of the system in real-world scenarios. The main objective of this research is to develop an intelligent and scalable system capable of predicting battery faults at an early stage, thereby improving battery lifespan, reducing maintenance costs, and ensuring operational safety. By integrating machine learning with real-time monitoring and alert mechanisms, the proposed system contributes to the advancement of smart EV battery management solutions.

The remainder of this paper is organized as follows: Section 2 presents the related work in battery health prediction. Section 3 describes the proposed system architecture and methodology. Section 4 discusses the experimental evaluation and results. Finally, Section 5 concludes the paper and outlines future research directions.

2. REALTED WORK

Researchers Battery health monitoring and predictive maintenance have gained significant attention in recent years due to the rapid growth of Electric Vehicles (EVs). Various research efforts have been made to estimate battery degradation and predict failures using both traditional and data-driven approaches.

Early studies primarily relied on electrochemical models and equivalent circuit models, which require precise domain knowledge and are computationally expensive. These models, although accurate, are not scalable for real-time applications due to their complexity.

With the advancement of machine learning, researchers have explored regression-based and classification-based techniques for battery health prediction. Linear regression and support vector machines (SVM) have been used to estimate State of Health (SOH), but these models often fail to capture nonlinear relationships present in battery data.

Deep learning approaches such as Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN) have demonstrated strong performance in modelling time-series battery data. These models can capture temporal dependencies and degradation trends. However, they require large datasets, longer training times, and high computational resources, which limit their practical deployment in real-time systems.

Recent studies have also investigated hybrid models that combine statistical and machine learning techniques to improve prediction accuracy. While these approaches provide better results, they increase system complexity and are difficult to implement in lightweight environments. Ensemble learning methods, particularly Random Forest, have shown promising results in classification tasks due to their ability to handle high-dimensional data and reduce overfitting. Random Forest models are less sensitive to noise and provide stable predictions compared to single models.

However, most existing works focus only on prediction accuracy and lack integration with real-time monitoring systems, automated alert mechanisms, and user-friendly interfaces. There is a clear gap in developing a complete system that combines prediction, visualization, and alerting. To overcome these limitations, this paper proposes a comprehensive predictive maintenance framework that integrates Random Forest-based classification with real-time alert generation and an interactive dashboard for EV battery health monitoring.

3. PROPOSED SYSTEM

3.1 System Architecture

The proposed system is designed as a modular and scalable framework that enables efficient monitoring and prediction of EV battery health. The architecture consists of multiple interconnected components that work together to process data and generate predictions. The system begins with data acquisition, where battery data is collected either through dataset upload or manual input. The collected data is then passed through a preprocessing module to remove inconsistencies and improve quality.

Next, the feature selection module identifies the most relevant parameters affecting battery health. These features are then used to train the Random Forest model, which performs classification of battery conditions.

The prediction output is further analyzed to generate recommendations and trigger alerts in case of critical conditions. Finally, all results are displayed through a Streamlit-based dashboard for user interaction. This architecture ensures efficient data flow, scalability, and real-time decision-making capability.

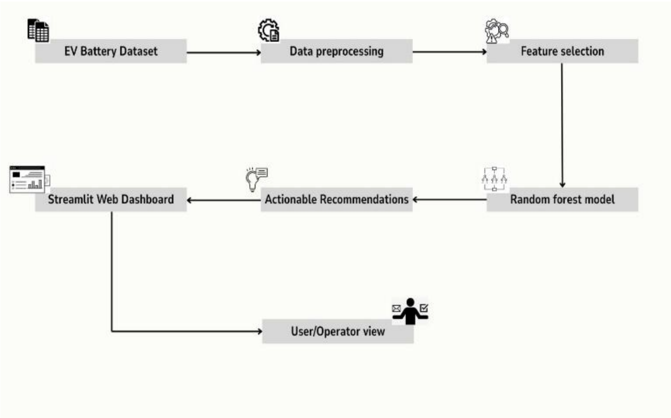


Fig -1: System Model

3.2 Methodology

3.2.1 Data Acquisition Module

The data acquisition module is responsible for collecting and managing the battery-related data used for analysis and model training. This module allows users to either upload historical battery data in CSV format or provide real-time input manually through the dashboard. The dataset consists of critical battery parameters such as voltage, current, temperature, State of Charge (SOC), and State of Health (SOH), which play a significant role in determining battery performance. The module ensures that the collected data is structured properly and made ready for further processing. By supporting both real-time and historical data input, the system provides flexibility and adaptability for different usage scenarios.

3.2.2 Data Preprocessing Module

The Data Preprocessing Module plays a crucial role in improving the quality and reliability of the data input. Raw battery data often contains missing values, inconsistencies, noise, and duplicate records, which can negatively impact the performance of machine learning models. This module applies various preprocessing techniques such as handling missing values using statistical methods, removing duplicate entries, and converting data into a consistent format. Additionally, normalization and scaling techniques are applied to ensure uniformity across all features. Outlier detection mechanisms are also implemented to eliminate abnormal values that may distort the analysis. These preprocessing steps ensure that the dataset is clean, accurate, and suitable for effective model training.

3.2.3 Feature Selection Module

The Feature Selection Module is designed to identify the most relevant parameters that significantly influence battery health. Since not all features contribute equally to prediction accuracy, this module helps in reducing dimensionality and improving model efficiency by selecting only the important attributes. The key features identified include State of Charge (SOC), State of Health (SOH), battery temperature, voltage, and current. These parameters are selected based on their direct impact on battery degradation and performance. By eliminating redundant or less significant features, this module enhances computational efficiency, reduces model complexity, and improves prediction accuracy.

3.2.4 Random Forest Model Module

The Random Forest Model Module is the core component of the proposed system, responsible for predicting battery health conditions. Random Forest is an ensemble machine learning algorithm that constructs multiple decision trees and combines their outputs to produce accurate predictions. This model is particularly effective in handling complex and nonlinear relationships among battery parameters. It is also robust against noise and overfitting, making it suitable for real-world datasets. The model is trained using the processed dataset and classifies battery conditions into three categories: Normal, Warning, and Critical. Due to its high accuracy and stability, Random Forest is chosen as the primary prediction model in this system.

3.2.5 Fault Prediction Module

The Fault Prediction Module analyzes the output generated by the machine learning model and determines the health condition of the battery. This module uses both model predictions and predefined threshold rules to identify abnormal battery behavior. For instance, conditions such as low State of Charge (SOC), high temperature, or reduced State of Health (SOH) are used as indicators of potential faults. The module categorizes battery conditions into Normal, Warning, and Critical levels, enabling early detection of issues. This proactive approach helps in preventing severe battery damage and ensures safe operation of the electric vehicle.

3.2.6 Email Alert Module

The Email Alert Module is designed to provide real-time notifications when critical battery conditions are detected. This module is implemented using Python's SMTP protocol, which enables automatic email generation and transmission. When the system identifies a critical condition, an alert message is sent to the administrator containing details such as battery status, parameter values, and warning messages. This ensures that immediate action can be taken to prevent further damage. The alert system enhances the responsiveness and reliability of the overall monitoring framework.

3.2.7 Streamlit Web Dashboard Module

The Streamlit Web Dashboard Module provides an interactive and user-friendly interface for monitoring battery health and viewing prediction results. This module allows users to upload datasets, input battery parameters manually, and visualize results in real time. The dashboard displays key information such as battery condition, parameter values, alerts, and recommendations in a clear and organized manner. Graphical representations are also used to enhance understanding and analysis. By integrating all functionalities into a single interface, this module improves usability and enables efficient decision-making for users and operators.

4. EXPERIMENTAL EVALUATION

4.1 Environment Specification

The proposed system is implemented in a Python-based development environment on a system equipped with an Intel x86 processor, 8GB of RAM, and 1TB of storage. Python 3.x is used as the primary programming language due to its flexibility and extensive library support. Data processing is performed using Pandas and NumPy, while Scikit-learn is utilized for building and training the machine learning model. The user interface is developed using Streamlit, which enables real-time interaction and visualization. Additionally, the Simple Mail Transfer Protocol (SMTP) is integrated into the system to send automated alerts during critical battery conditions. This combination of tools ensures efficient processing, scalability, and real-time monitoring capability.

4.2 Dataset

The dataset used in this study consists of electric vehicle battery parameters collected from historical data sources and simulated inputs. The dataset includes important attributes such as voltage, current, temperature, State of Charge (SOC), and State of Health (SOH). These parameters are essential for analyzing battery performance and detecting anomalies. The dataset is preprocessed to remove noise and inconsistencies before being used for model training. For evaluation purposes, the dataset is divided into training and testing sets in an 80:20 ratio, ensuring that the model is trained effectively and validated accurately.

4.3 Experimental Setup

The experimental setup defines the workflow and configuration used to train and evaluate the model. It includes data collection, preprocessing, feature selection, model training, and prediction stages.

Table -1: Experimental Setup

Process Stage	Description
Input	EV battery dataset (CSV / manual input)
Data Preprocessing	Cleaning, normalization, missing value handling
Feature Selection	SOC, SOH, Temperature, Voltage, Current
Model Used	Random Forest Classifier
Training Ratio	80% Training, 20% Testing
Output	Normal / Warning / Critical

4.4 Performance Metrics

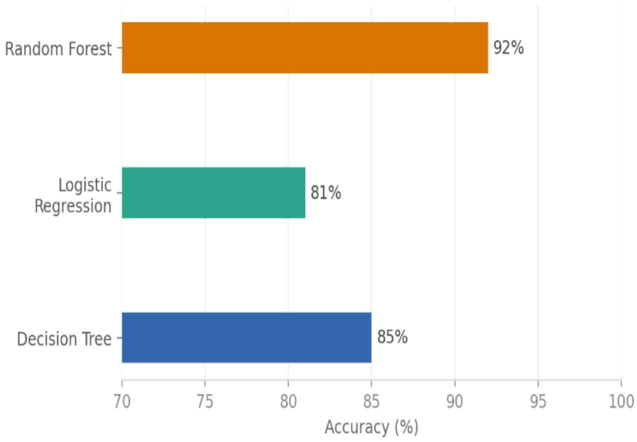
The performance of the proposed model is evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the model, while precision and recall evaluate the model's ability to correctly identify positive cases. The F1-score provides a balance between precision and recall, making it a reliable performance indicator. These metrics ensure that the model's prediction capability is thoroughly assessed.

4.5 Analysis

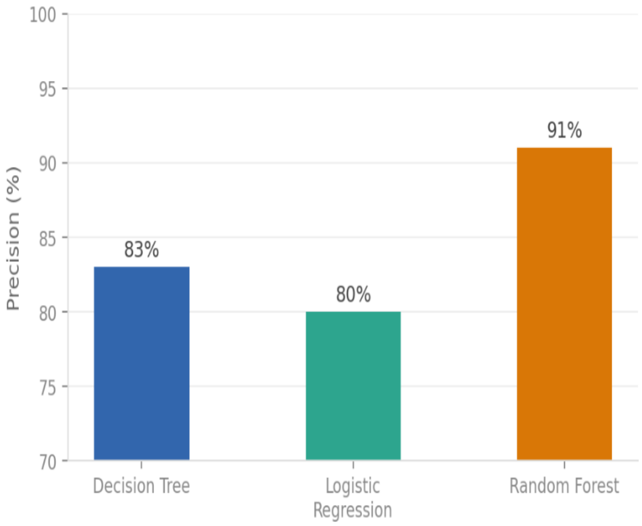
The Random Forest model is trained and evaluated using the prepared dataset. The results indicate that the model achieves high accuracy and performs better compared to traditional classification models. The model effectively captures nonlinear relationships between battery parameters and accurately classifies battery conditions.

Table -2: Model Performance Comparison

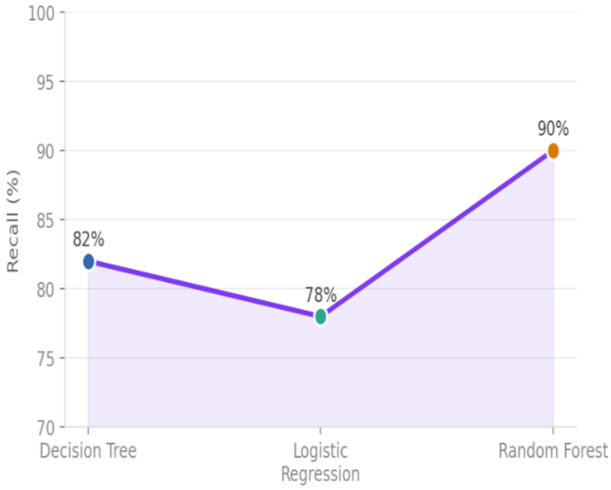
Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	85%	83%	82%	82%
Logistic Regression	81%	80%	78%	79%
Random Forest	92%	91%	90%	90%



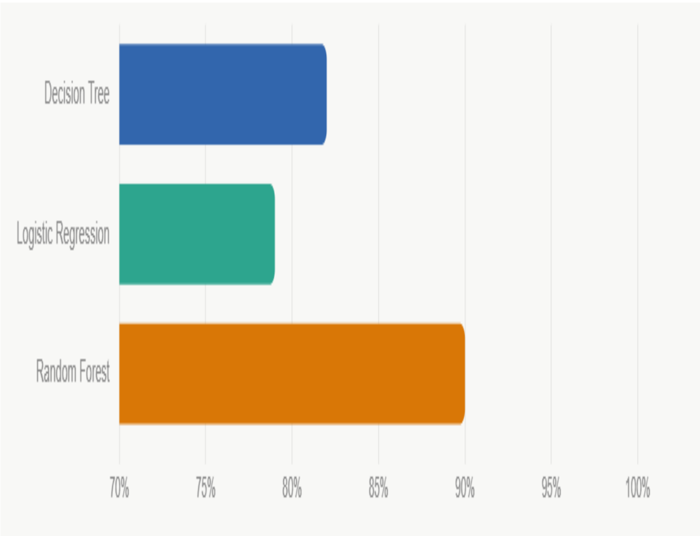
a)Accuracy Graph



b) Precision Graph



c)Recall Graph



d)F1-score Graph

6. Results and Discussion

The proposed system was implemented using a Random Forest model and a Streamlit dashboard to monitor EV battery health. The model analyzes parameters such as voltage, current, temperature, SOC, and SOH to classify battery conditions into Normal, Warning, and Critical states.The results show that the model achieves high accuracy with balanced precision and recall, indicating reliable prediction performance. It effectively identifies abnormal conditions and supports early fault detection. Compared to traditional methods, the system provides better accuracy and reduces the chances of unexpected battery failure.

The Streamlit dashboard enables real-time data input and displays prediction results instantly, improving usability. Graphical outputs further help in understanding model performance and battery behavior. Additionally, the email alert system ensures timely notification during critical conditions.Overall, the system demonstrates an efficient and practical approach for predictive maintenance, improving battery safety and lifespan.

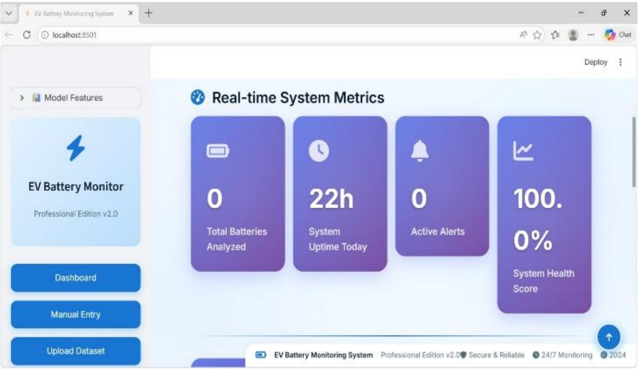


Fig - 2: Streamlit Dashboard

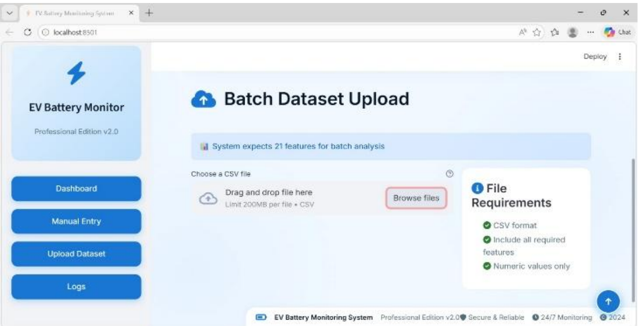


Fig - 3: Dataset Upload

Fig -4: Model Features

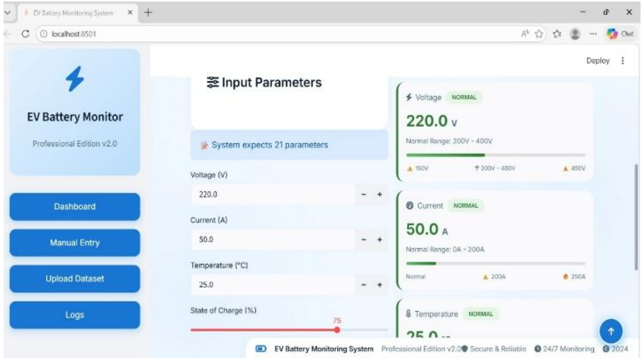


Fig -5: Manual Entry

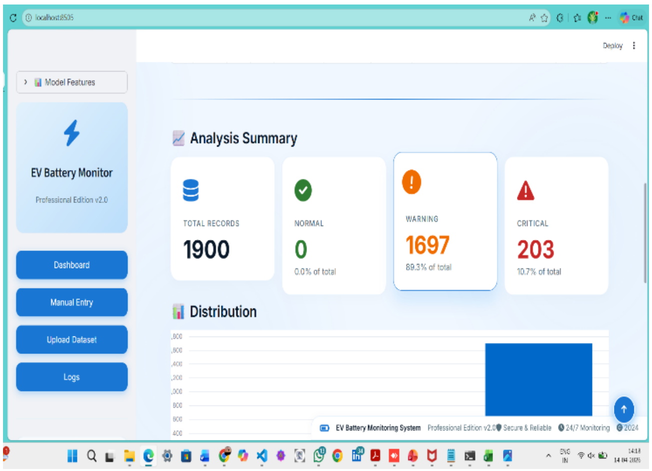


Fig -6: System Model

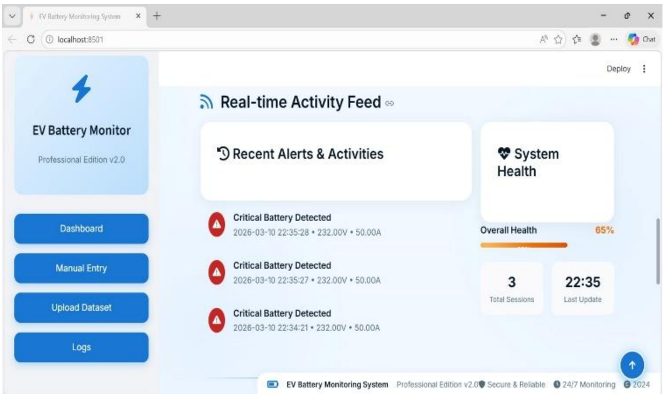


Fig - 7: Real Time Activity Feed

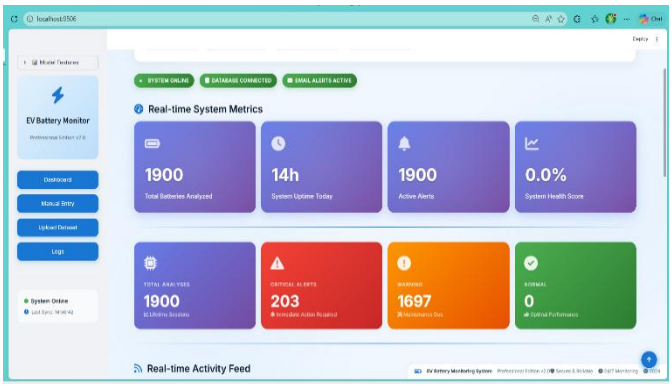


Fig – 8: System Metrics

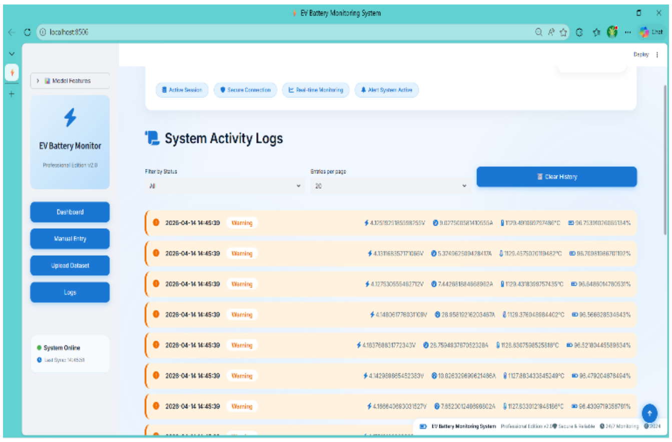


Fig - 9: Activity logs

7. CONCLUSION

This study introduces a machine learning-driven approach for monitoring and maintaining the health of electric vehicle batteries. By utilizing key operational parameters such as voltage, current, temperature, State of Charge (SOC), and State of Health (SOH), the system is capable of identifying battery conditions and categorizing them into Normal, Warning, and Critical levels. The use of data preprocessing and feature optimization techniques enhances the quality of input data, which in turn improves the overall prediction capability of the model. The Random Forest algorithm proves to be effective in handling complex relationships among battery parameters, resulting in consistent and dependable classification outcomes. The integration of a Streamlit-based dashboard further strengthens the system by enabling real-time interaction and easy visualization of results.

In addition, the implementation of an automated alert mechanism ensures that critical battery conditions are communicated promptly, allowing necessary actions to be taken at the right time. This proactive approach contributes to minimizing sudden failures and supports better maintenance planning. The outcomes of this work indicate that the proposed system offers a practical and efficient solution for battery health monitoring when compared to conventional methods. It not only improves prediction reliability but also contributes to extending battery life and enhancing operational safety. Future enhancements can include the incorporation of live sensor data through IoT technologies, as well as the application of advanced learning models to further increase system accuracy and adaptability.

REFERENCES

[1] L. Breiman, “Random forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

[2] A. Barré, B. Deguilhem, S. Grolleau, M. Gérard, F. Suard, and D. Riu, “Review of lithium-ion battery ageing mechanisms and estimation techniques for automotive applications,” J. Power Sources, vol. 241, pp. 680–689, 2013.

[3] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, and P. Van den Bossche, “State-of-health estimation methods for lithium-ion batteries: A comprehensive review for practical applications,” Renew. Sustain. Energy Rev., vol. 56, pp. 572–587, 2016.

[4] G. L. Plett, “Extended Kalman filtering techniques for battery management systems in hybrid electric vehicles,” J. Power Sources, vol. 134, no. 2, pp. 252–261, 2004.

[5] H. He, R. Xiong, J. Fan, and X. Li, “An adaptive Kalman filtering approach for estimating the state of charge in lithium-ion batteries using a modified Thevenin model,” IEEE Trans. Veh. Technol., vol. 60, no. 4, pp. 1461–1469, 2011.

[6] S. Saxena, C. Le Floch, J. MacDonald, and S. Moura, “Evaluation of electric vehicle battery end-of-life using travel pattern analysis and powertrain modelling,” J. Power Sources, vol. 282, pp. 265–276, 2015.

[7] J. Zhang and J. Lee, “A review of prognostics and health monitoring approaches for lithium-ion batteries,” J. Power Sources, vol. 196, no. 15, pp. 6007–6014, 2011.

[8] X. Hu, F. Feng, K. Liu, L. Zhang, and S. Xie, “Machine learning-based state-of-health estimation methods for lithium-ion batteries: A review,” Energy Rep., vol. 6, pp. 120–131, 2020.

[9] Y. Li, K. Liu, X. Hu, and F. Sun, "Data-driven approaches for remaining useful life prediction of lithium-ion batteries," IEEE Access, vol. 7, pp. 8281–8290, 2019.

[10] F. Yang, X. Song, F. Xu, and K. Tsui, "Machine learning and big data approaches for lithium-ion battery state-of-health estimation," Appl. Energy, vol. 256, 2019.