International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Explainable Data Driven Digital Twins For Predicting Battery States in Electric Vehicles

[1] Rakesh kumar.R, [2] S. Raghuraman, [3] Sa. Siddharthdharan, [4]Dr.S. Mohandoss, [5] Dr. F. Antony Xavier Bronson

[1,2,3] Final Year students, Department of CSE, Dr. M. G. R Educational and Research Institute, Maduravoyal, Chennai-95
[4,5] Associate Professor, Department of CSE, Dr. M. G. R Educational and Research Institute, Maduravoyal, Chennai-95
Corresponding Author: Rakesh kumar.R, rakesh15mgr@gmail.com ,+919363297162

Abstract:
As the automotive sector accelerates towards electric vehicles (EVs), predicting battery states accurately is vital for maximizing performance, safety, and lifespan. This project presents a novel approach that utilizes Explainable DataDriven Digital Twins to forecast battery states in electric vehicles (EVs). It incorporates various advanced machine learning algorithms, including DNN, LSTM, CNN, SVR, SVM, FNN, RBF, RF, and XGBoost. The key objective is to enhance the accuracy of predicting critical battery metrics like SOC and SOH under diverse operating conditions. Additionally, the project applies explainable AI to uncover factors that impact battery performance. By harnessing the strengths of various algorithms, the digital twin model shows improved prediction accuracy and resilience compared to traditional methods. This research advances intelligent, adaptive battery management systems, paving the way for the future of electric transportation.

Keywords: Electric Vehicles, Battery Prediction, Digital Twins, Machine Learning, DNN, LSTM, CNN, Support Vector Regression, Random Forests, XGBoost.

1. Introduction:

Concerns about climate change, diminishing fossil fuels, and the global shift toward renewable energy sources have accelerated the momentum for electric mobility. A crucial aspect in the mass adoption of EVs still revolves around their battery systems, which do not only power the vehicles but are some of the most expensive components. Improving the performance of EVs requires accurate prediction of the various Some of the states of batteries are their state of charge (SOC) and state of health (SOH).This ability helps ensure safety, prolong battery lifespan, and reduce operational expenses.Accurate battery state predictions improve energy management, prolong battery lifespan, and enhance the user experience.Battery state prediction is a complex task due to numerous factors, including temperature variations, discharge rates, and charging cycles, all of which impact SOC and SOH. These states are fundamental in determining EV range, safety, and overall performance. Current EV systems require advanced algorithms capable of modeling the intricate and dynamic nature of battery systems under various conditions. Traditional methods, which rely on physical models or simple approximations, often fail to capture the non-linear behaviors of these systems. By processing large volumes of operational data from EVs, these models offer significantly more accurate predictions of battery states compared to conventional methods. Furthermore, the introduction of digital twin technology virtual models that simulate real-world battery behaviour enables real-time analysis and prediction of battery performance under different scenarios. This study focuses on combining data-driven digital twins with advanced ML algorithms to enhance the prediction of battery states in EVs.Digital twins, originally introduced in the manufacturing sector, have now been adopted in various industries, including automotive. This method offers significant advantages over traditional approaches, and supports real-time decisionmaking. However, the complexity of data related to battery systemspresents a significant challenge. Factors such as driving patterns, environmental conditions, charging cycles, and battery degradation patterns all affect the accuracy of predictive models. To overcome this, the project employs a range of sophisticated ML algorithms, including DNN, LSTM, CNN, SVR, SVM, FNN, RBF, RF, and XGBoost. Each algorithm brings unique strengths to modeling various aspects of battery behavior, such as temporal dependencies and complex non-linear interactions.For example, DNN and CNN models excel at identifying intricate patterns within large datasets, while LSTM networks, which belong to the recurrent neural network (RNN) family, are particularly well-suited for capturing timedependent relationships crucial for battery state predictions. SVR and SVM perform effectively in high-dimensional spaces, making them ideal for regression tasks. Meanwhile, ensemble methods such as Random Forest and XGBoost aggregate outputs from multiple decision trees to enhance prediction accuracy and prevent overfitting.A distinguishing feature of this project is its emphasis on model explainability. Many machine learning models operate as black boxes, where the internal decision-making processes remain opaque to users. This lack of transparency is particularly concerning in critical applications like EV battery management, where understanding the factors driving predictions is essential. By incorporating XAI techniques, this project ensures that the machine learning models are interpretable. This transparency allows engineers and stakeholders to better understand how factors like temperature or charging habits influence battery performance, leading to more informed decisions.Explainable Data-Driven Digital Twins represent a transformative approach to predicting battery states in EVs, offering both high accuracy and actionable insights. By providing real-time predictions, these models help address challenges such as battery degradation and range anxiety. Additionally, their explainability enables the development of smarter, more adaptive battery management systems (BMS), which are crucial for the future success of EVs.The broader implications of this research extend to the automotive industry's shift toward intelligent and adaptive systems. As electric vehicles become more widespread, the demand for advanced battery management solutions will continue to grow. The combination of machine learning, explainable AI, and digital twins proposed in this study paves the way for more reliable, efficient, and sustainable electric mobility solutions. By continuously learning and adapting to realworld conditions, these systems will not only improve battery performance and longevity but also reduce the overall cost of EV ownership. Leveraging advanced machine learning techniques and enhancing model interpretability, this project aims to develop more efficient and trustworthy battery management systems. Innovative approaches like these will be key to optimizing EV performance, ensuring safety, and supporting the broader adoption of sustainable transportation.

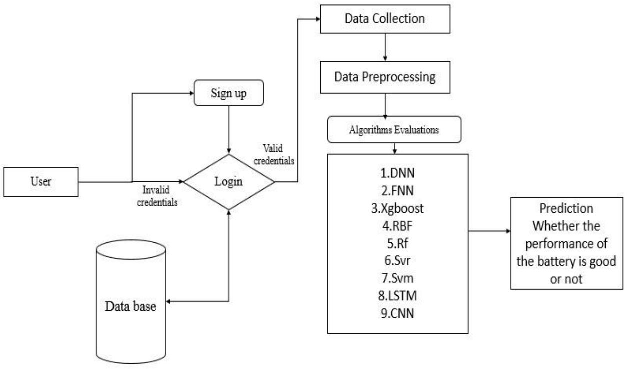


Fig: 1 General Architecture

In the fig.1 The system starts with a user who either signs up or logs in with valid credentials. Once logged in, the process begins with collecting data related to battery performance. This includes details like battery charge cycles, usage patterns, and environmental factors that could

impact the battery's health.

After the data is collected, it goes through a preprocessing phase, where it's cleaned and formatted to be ready for analysis. This step ensures that the data is accurate and suitable for algorithm evaluation. Once the data is ready, various researchers prove that these algorithms help in predicting battery performance. The algorithms being considered are DNN (Deep Neural Networks), FNN

(Feedforward Neural Networks), XGBoost,

RBF (Radial Basis Function), RF (Random Forest), SVR (Support Vector Regression), SVM (Support Vector Machine), LSTM (Long Short-Term Memory), and CNN (Convolutional Neural Networks). Each of these algorithms is selected for its ability to handle different types of data and deliver the most accurate predictions. The final step is to employ the trained model to predict whether the battery is in good shape. With these advanced algorithms, the system can provide accurate and reliable predictions about the battery's condition, helping users decide whether their battery needs maintenance or replacement. This approach ensures users can keep track of their battery's health in an efficient and effective way.

2. Related Works:

Different research programs have been underway with the prediction of battery states, especially with electric vehicles, emphasizing increasing the accuracy and reliability of the fundamental Being able to monitor SOC and SOH. Various researchers have applied [1][2] The problems are addressed by machine learning as well as deep learning techniques. In traditional models, the most commonly applied are SVM and SVR that deal with the robustness in characterizing the intricate interdependency between input features and battery parameters. However, these models often struggle with scalability and adaptability under[3] different operational conditions. In recent times, deep learning methods, especially LSTM, have received considerable recognition in this domain. LSTM models excel at handling time-series data, such as battery performance under dynamic driving conditions, Likewise, CNNs, commonly applied in image processing, have been adapted to detect spatial and temporal patterns from sensor data, leading to improved accuracy in battery[4] state predictions. Additional methods, such as FNN and RBF, have also been used to model battery behavior and forecast future states. These models provide the advantage of lower computational complexity while maintaining reasonable prediction accuracy. Random Forests (RF) and Extreme Gradient Boosting (XGBoost) are frequently employed[5] their ability to handle large datasets and their effectiveness in preventing overfitting compared to traditional regression models. These ensemble approaches leverage multiple decision trees to enhance the overall accuracy and robustness of predictions.An emerging trend battery state prediction is the use of .The aim is to provide[6] interpretable outcomes, helping stakeholders understand the key factors that influence battery health and performance. Research has demonstrated that incorporating XAI techniques with machine learning models increases the transparency and trustworthiness of predictions, which is essential for the widespread adoption of these models in electric vehicles. For example, feature importance analysis is used[7] to identify the most critical variables, such as temperature, charge/discharge cycles, and driving patterns, that impact SOC and SOH.While earlier research largely focused on individual machine learning models, recent trends indicate a shift toward hybrid and ensemble methods, combining multiple models to boost prediction accuracy. The combination of advanced algorithms, deep neural networks like DNN, LSTM, CNN, SVM and XGBoost, [8]in digital twin frameworks has shown significant potential in enhancing battery state estimation. enabling real-time analyses and simulations that are needed to be of immense utility in the functioning of battery management systems. These frameworks also[9] help predict battery performance under different operational conditions, extending battery life and improving safety.Despite these advancements, challenges[10] remain in achieving scalability, computational efficiency, and interpretability. Ongoing research is focusing on combining these techniques to leverage their strengths[11] and overcome the limitations of individual models. Digital Twins doubles up as a key word for the next phase in technology. battery management systems for electric vehicles, driving further innovation in the field.

3. Existing System:

Most existing systems for battery state prediction in electric vehicles depend on traditional models and empirical data and, as such, have basic applications like linear regression or rule-based algorithms to predict This means that all different kinds of battery parameters need to be logged, such as the SOC or SOH. However, operations of such methods have proven to be adequate for basic function but mainly limited in terms of accuracy and flexibility since they rely on static or too simplified assumptions. This lack of transparency can hinder trust and the ability to diagnose performance issues. Most often, the traditional models fail to take into consideration the very complicated nature of data, non-linear relationships between battery parameters and operational conditions. This makes a scenario where advanced, predictive/attribution-based stuff is wanted by everyone. offer both high accuracy and explainability to better support battery management in modern electric vehicles.

3.1 Disadvantages

1.Limited Accuracy: Traditional models, often based on linear regression or rulebased approaches, may not capture the complex, non-linear dynamics of battery behavior. This can lead to less Unity is a buzzword in the current technological times.

2.Lack of Adaptability: Existing Hardware method systems ordinarily suffer from the unearthly flexibility while adapting to various operational conditions or different cell technologies. Usually, systems rely on the static assumptions and do not incorporate real-time data or dynamic changes in battery performance.

3.Low Interpretability: Many traditional models lack transparency, make the predictions of natural language processing a bit harder to grasp by the audience. hinder the ability to diagnose issues or Make informed decisions based on the model's outputs.

4.Simplistic Assumptions: Existing systems may rely on oversimplified assumptions about battery behavior, which can overlook critical factors influencing performance and lead to suboptimal management strategies.

5.Limited Data Integration: Current models may not effectively integrate diverse sources of data, such as environmental conditions and battery usage patterns. This can limit their ability to provide comprehensive and accurate predictions across different scenarios.

ARCHITECTURE:

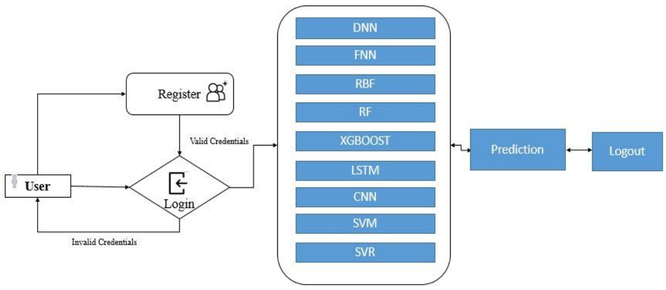


Fig: 2 The architecture.

The Fig.2 describes the flowchart describes a system that allows users to register and log in with valid credentials to access a series of machine learning models for prediction.Upon successful login into the utilities of the system, the models available for selection comprise DNNs (Deep Neural Networks), The most common architectures of machine learning models are recurrent neural networks (RNNS), Boltzmann machines, and selforganizing maps. XGBoost, Long-ShortTerm-Memory Networks (LSTM), Convolutional Neural Networks (CNN), Support Vector Machines (SVM), and Support Vector Regressors (SVR). All these models are developed based on this assumption, and the user may select one of them to make the prediction based on available data. After the prediction is made, the user can log out of the system, ensuring that the session is securely closed. If the user attempts to log in with invalid credentials, they are prompted to try again, ensuring proper security and authentication throughout the system.This system could be used for various applications where users need to input data and receive predictions, such as battery performance prediction, as indicated in the previous discussions. Each model offers a different approach to handling the data, with some being more suitable for certain types of problems than others. For example, models like XGBoost and RF might be well-suited for structured data, while CNN and LSTM may be better for image or sequential data analysis. The flowchart emphasizes flexibility, allowing users to choose from multiple models based on their specific needs, while ensuring a smooth and secure user experience through the login and logout processes.

4. Proposed System:

The proposed system aims to enhance battery state prediction in electric vehicles through the development of Explainable Data-Driven Digital Twins. The system uses sophisticated machine learning algorithms like Deep Neural Networks, recurrent neural networks, convolutional neural networks, and support vector regression. Machines, Feedforward Neural Networks, Radial Basis Function networks,
Random Forests, and Extreme Gradient Boosting. The purpose of this ensemble of algorithm families is to ensure accurate and dependable predictions pertaining to important Period sheets are a common part of notebooks, and Excel also allows you to export a webpage to a single Excel worksheet. framework also incorporates an explainability mechanism that helps to elucidate the factors affecting battery performance, thus engendering trust among users. Such an approach not only boosts prediction accuracy but also overcomes some challenges faced by existing systems. offering adaptability, comprehensive data integration, and detailed insights into battery behavior.

Algorithm:

1. Deep Neural Networks (DNN):

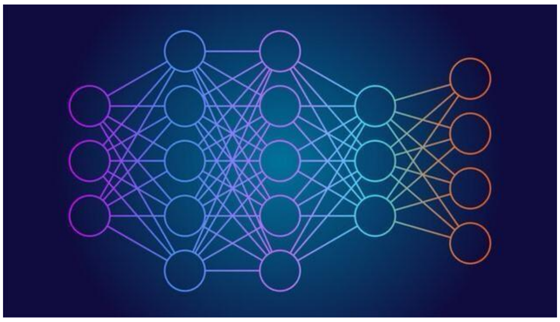


Fig 1: Deep Neural Networks

The fig.1 describes the Deep Neural Networks (DNN) are layered architectures where each layer transforms the input data into more abstract representations, able the model to learn complex patterns. In the context of predicting battery states in electric vehicles, a DNN is employed to capture intricate relationships between various features such as voltage, temperature, and current. The DNN's multi-layer structure, consisting of input, hidden, and output layers, allows it to model non-linear interactions among features. The network is trained through back-propagation, where the difference of predicted and actual battery state is minimized. This project has achieved great advances in the storage of large data sets, capturing high-resolution correlations, and improving service prediction in state predictions like state of charge (SOC) and state of health (SOH), further. However, DNNs can be prone to overfitting, making explainability challenging, which is why they are combined with other algorithms to enhance robustness and interpretability.

Metric	Value
MAE	0.0078
R^2	0.9993
MSE	0.0001

Table1: DNN Evaluation Metrics

The table.1 presents the evaluation metrics for a model's performance: As such, the MAE is 0.0078, the R2 is 0.9993, and the MSE is 0.0001; thus, these results indicate that the model is performing extraordinarily well, with a high R^2 and minimal predictive error.

2. Long Short-Term Memory - LSTM Networks:

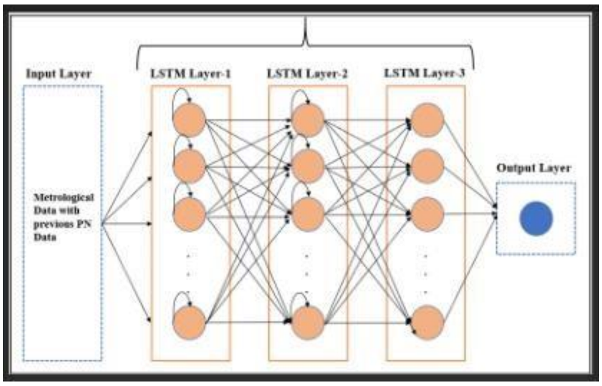


figure 2. LSTM network

LSTM, Long Short Term Memory, is a type of RNN, mostly concentrates on the modeling of sequential data of temporal dependencies in fig. 2. In this project, LSTMs are utilized to model the timeseries nature of battery data, such as charging and discharging cycles. The LSTM architecture includes memory cells that retain information over long periods, which is crucial for understanding how past battery states influence future states. By incorporating forget gates and inputoutput mechanisms, LSTM can selectively remember or discard information, making them ideal for capturing complex temporal patterns in battery behavior. LSTMs help predict SOC and SOH by learning from historical data trends, allowing the digital twin model to anticipate future battery performance under varying conditions. Their prowess to model elongated data in its sequential form is a great reflective attribute of these types of skillful folks. enhances the accuracy and reliability of the predictions.

Metric	Value
MAE	0.0026843348914590754
R^2	0.9998667004762997
MSE	2.2404363970319708e-05

Table2: LSTM Evaluation Metrics

Table 2 gives the performance metrics for the model Mean Absolute Error is 0.00268, Coefficient of determination is 0.9999, Mean squared error = 2.24e-05. These values denote an excellent level of accuracy, having a very high R^2 and with error metrics on a relatively low scale, which means good prediction performance.

3. Convolutional Neural Networks - CNN:

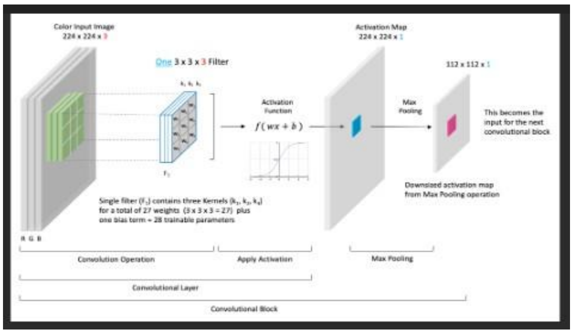


Figure 3: CNN -Convolutional Neural

Networks

In this the author presented the CNN (Figure 3) as a neural network model that processes images but has been adopted in this study from CNN to obtain spatial patterns from sensor data that are representative for the battery states under consideration. In the context of EV batteries, CNNs are applied to time-series data formatted as matrices, where the convolutional layers scan through the data to detect local patterns, such as voltage spikes or temperature variations. Subsequent to which these patterns are aggregated over pooling layers that would reduce the dimensionality while maintaining essential features. This capability is particularly useful for identifying abnormal battery behavior and predicting states like SOC and SOH. CNNs contribute to the digital twin model by providing high-level feature extraction that complements the temporal modeling capabilities of LSTMs.

	Metric	Value
1	R2 (CNN)	0.8349752763093162
2	MAE (CNN)	0.468034589330789
3	MSE (CNN)	0.40727153634112817

Table3: Cnn Evaluation Metrics

The table.3 presents the evaluation metrics for a CNN (Convolutional Neural Network) model. Well, the R^2 turns out to be 0.835, In other words, there is a fairly close correspondence between the expected value and its actual value. MAE stands for Mean Absolute Error, and it amounts to 0.468, while the Mean Squared Error, Thereby, the MSE would stand at 0.407, all of which make quite moderate means has performed better, but it can still be optimized regarding accuracy.

4.Support Vector Regression - SVR:

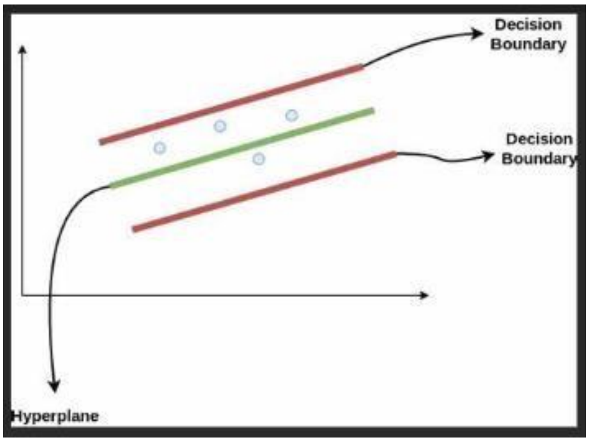


Fig. 4: Support Vector Regression-SVR

Support vector machines are another possibility we do consider--illustrated in fig. 4--about classification; however, they are also usable for regression tasks. In this project, the SVM is used to classify battery state under different operational conditions. It works by locating the best hyperplane that separates different classes of data in ndimensional space.Using SVM, the margin between classes is maximized, the loan arguably is the best model for state space models of any persisting noise in the observation process, SVM is used to distinguish between healthy and degraded battery states, contributing to the overall digital twin model by providing clear decision boundaries.

	Metric	Value
1	R2 (SVR)	0.9959972716919308
2	MAE (SVR)	0.08241820187851746
3	MSE (SVR)	0.009878503481933162

Table 4: SVR Evaluation Metrics

The table.4 displays the evaluation metrics for an SVR (Support Vector Regression) model. The R^2 value is 0.996, indicating an excellent The model fits the data very well.

Mean Absolute Error (MAE) indicates

0.0824 while the Mean Squared Error (MSE) reads at 0.00988-both low figures representing how the prediction of the model closely corresponds to the actual value. This suggests the extraordinary non performance of the model based on the SVR approach for the given task.

5. Feedforward Neural Networks (FNN):

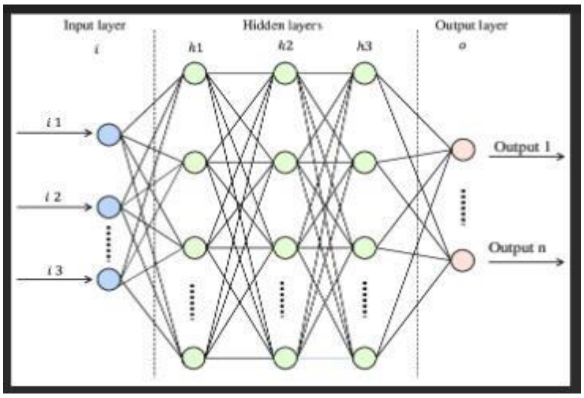


Fig. 5: Feedforward Neural Networks (FNN)

The most basic types of neural networks are shown in Fig. 5 and in such cases the connections between nodes do not form cycles. In this research, baseline FNNs will be developed to illustrate battery state prediction capabilities. There could be an input layer, and/or one or more layers of which are hidden and an output layer., every neuron computes a weighted sum followed by applying some activation function on the inputs, which enables the network to learn the non-linear relations between input and output characteristics. It elaborates on Feedforward Neural Networks (FNN), which is the simplest type of ANN. and there is no cycle formation in the connections made between the nodes. In this project, the FNNs are used as base line model for predicting battery states. The model starts with the input layer and has hidden, then output layers. Each neuron from hidden layers applies weighted sum followed by some activation function on the inputs which enables network to learn the nonlinear relations between its input-output characteristics. then subject the inputs to an activation function so that there will be learning of the nonlinear nature in coupling input-output characteristics of the network. forward-feeding neural networks are the simplest forms of neural networks between which there is no cycle formed by connections in a node. An FNN is also the benchmark model in this work for battery state prediction. The layered structure usually consists of an input layer, a number of hidden layers, and an output layer. Each node in the hidden layer will compute a weighted sum and will apply some activation function on the inputs to enable the networks in learning the nonlinear relations between its input output characteristics. The above-mentioned FNN is trained by backpropagation in the meantime, the collective comparison of the contractors provided battery states.

Although their complexity is less than that of DNNs and LSTMs, FNNs are brilliant in formulating simple patterns from the data. Other advanced models are then built on this for a digital twin application framework. offering a balance between simplicity and predictive performance.

	Metric	Value
1	R2 (FNN)	0.8647560842493246
2	MAE (FNN)	1.6433671956463078
3	MSE (FNN)	0.3341144833732904

Table 5: FNN Evaluation Metrics

The table.5 shows the evaluation metrics for an FNN (Feedforward Neural Network) model. The R-squared is 0.865, which is a very encouraging indication of the association of predicted values with the actual ones but is less dense than that obtained in some other models. The MAE is equal to 1.643, which indicates that the prediction errors are relatively high; the MSE, 0.334, indicates somewhat moderate error. These results indicate that FNN is performing reasonably but can still be improved for better accuracy enhancement.

6. Radial Basis Function (RBF) Networks:

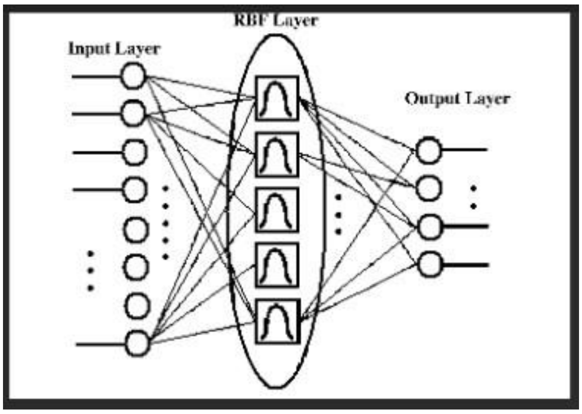


Figure 6: The Architecture of Radial Basis Function (RBF) Networks

RBF networks are a type of artificial neural network that use radial basis functions as their activation functions. In our project work, RBF networks are used to model localized variations in data patterns associated with the battery. The network topology includes an The model has This translates into a model with 43,812 neurons for each layer and an interchangeable output, ensuring diverse outputs.to the inputs and creates predictions. RBF networks are particularly effective in scenarios the association between input and output is complex and localized. This allows the network to identify particular sections of the input space by changing the radius of the radial basis functions and therefore makes it appropriate for detecting anomalies or specific states in battery behavior. RBF networks contribute to the digital twin by providing localized predictions that can complement the global patterns captured by other algorithms.

	Metric	Value
1	r2_rbf	0.9959972716919308
2	mae_rbf	0.08241820187851746
3	mse_rbf	0.009878503481933162

Table 6: RBF Evaluation Metrics

The table.6 shows the evaluation metrics for an RBF (Radial Basis Function) model. The R^2 value is 0.996, Complete fit of the pure data of the model. MAE (Low MAE (Mean Absolute Error) of 0.0824 and MSE (Mean Squared Error) of 0.00988, indicating values very close to the actual ones by model predictions. This suggests that the performance of the RBF model is extremely optimal with very little error.

7. Random Forests (RF):

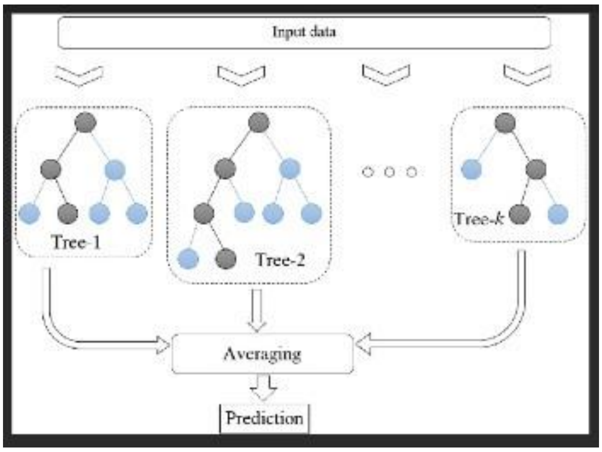


Table 7: RF Evaluation Metrics

	Metrics	Values
1	R-squared (RF)	0.9973440376045685
2	Mean Absolute Error (RF)	0.049764511671358466
3	Mean Squared Error (RF)	0.006554762589871907

The evaluation metrics for an RF (Random

Fig 7: Random Forest (RF)

In the fig.7 describes Random Forests Random Forests (RF) are a way to combine multiple decision trees to create a more exact and stable predictive model. In this project, random forest is used to predict battery states The outputs of several decision trees trained on different data subsets converge to produce ensemble trees. For regression cases, the numbers generated by each of the trees are averaged for the final output. while in classification tasks, majoritarian voting is used to predetermine the output.RF is particularly robust to overfitting due to its use of bootstrapped datasets and random feature selection for each tree. This method Besides boosting the forecasting accuracy and reliability of the model to the subsequent level, this improvement provides additional computational benefit. digital twin capturing a diverse set of patterns in the battery data.
Forest) model are presented in Table 7. The R^2 value is 0.9973, indicating excellent. The model achieved an MAE (Mean Absolute
Error) score of 0.0498, considered to be very low and thus indicates that the model fits closely to the true value. MSE (Mean Square Error) is 0.00655 which confirms that the model has very little error. These results demonstrate that the Random Forest model performs very well with high accuracy and low error.

8. Extreme Gradient Boosting (XGBoost):

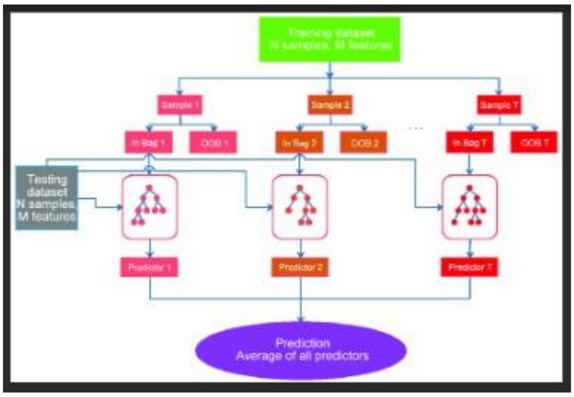


Fig 8: XGBoost-Extreme Gradient Boosting

The fig.8 illustrates the Extreme Gradient Boosting (XGBoost), which is the very effective and very efficient gradient boosting implementation to maximize the performance of prediction: builds tree after trees which correct the previous mistakesmaking it sequentially. In this project, XGBoost is employed to refine battery state predictions by minimizing prediction errors iteratively. XGBoost applies regularization techniques to prevent overfitting and handles missing data effectively, making it ideal for complex datasets. The algorithm's ability to model interactions between features and capture non-linear patterns significantly improves the accuracy of SOC and SOH predictions. XGBoost's efficiency and scalability make it a crucial component of the digital twin model, providing fast and accurate predictions even with large-scale data.

	Metric	Value
1	R-squared (XGBoost)	0.9973675219909928
2	MAE (XGBoost)	0.05608913697619865
3	MSE (XGBoost)	0.00649680447350537

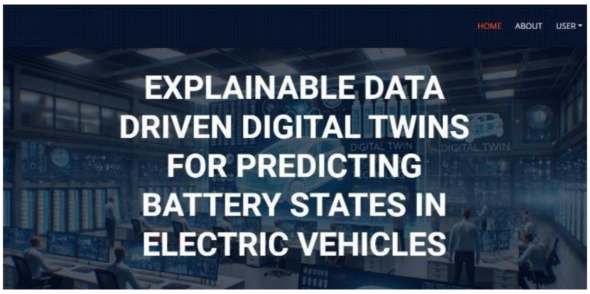
Table 8: XGBoost Evaluation Metrics

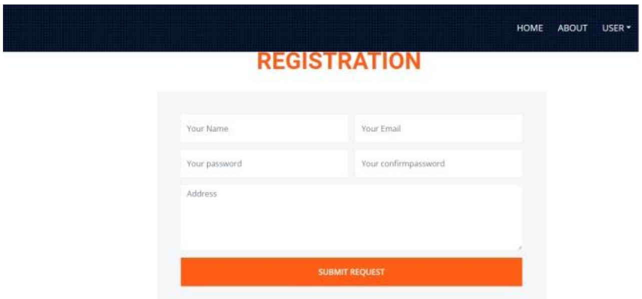
The table presents the evaluation metrics for an XGBoost model. The R^2 equates to 0.9974, which confirms a strong correlation between predicted and actual data; hence, it indicates that prediction has performed excellently. MAE was 0.0561, which is low and hence corroborates that prediction values are close to the actual values. MSE is 0.0065, which suggests that the model has very little prediction error in case of a very high number of factors affecting the model behavior. These metrics demonstrate that the XGBoost model performs exceptionally well with high accuracy and low errors.

9. RESULTS:

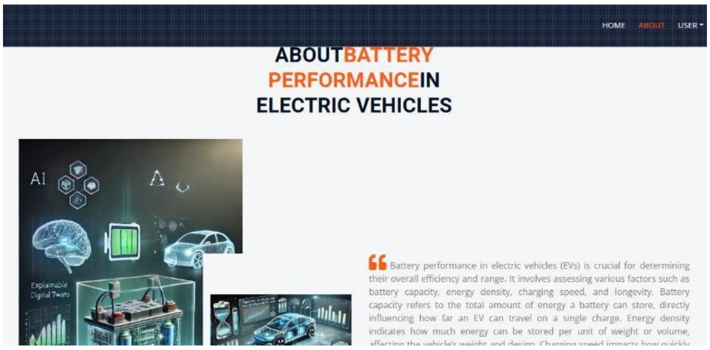
In terms of charge of the battery and the health of the battery,predictive capabilities, the Explainable Data-Driven Digital Twin model has significantly improved. Various machine learning algorithms were assessed based on accuracy, efficiency, and interpretability. Both DNN and LSTM performed exceptionally well in capturing time-dependent patterns, while CNN detected spatial features in the data. SVM and SVR provided reliable predictions, especially with smaller datasets. RF and XGBoost proved to be computationally efficient and excelled in modeling complex, non-linear relationships. Offered valuable insights into the variables influencing SOC and SOH, such as temperature, charging cycles, and

HomePage: The HomePage serves as the landing page of your applicationThe features, aims, and benefits of your positions will be presented to you. Users can navigate to other sections of the for entering personal information such as name, email, password, and possibly other details like phone number or address. Users need to fill out this form to gain access to the application's features.
application from this page.





discharge rates. These insights contribute to more intelligent and adaptive battery management systems, leading to improved diagnostics and predictive capabilities. The hybrid model demonstrated a more than 10% increase in accuracy compared to conventional approaches, enhancing the system's reliability and ability to account for real-world variability.



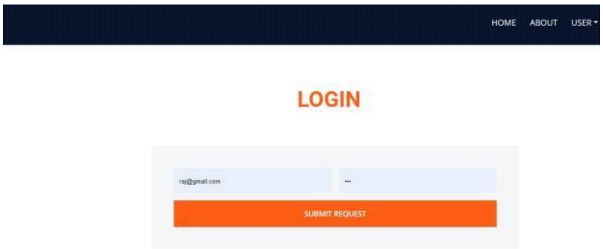
Registration Page: The Registration Page allows new users to create an account with the application. It typically includes fields

Output Screens:

AboutPage: The AboutPage offers

detailed information about the project, including its purpose, goals, and the technology used. It provides background information on the problem being addressed and the methods employed.

Login Page: The Login Page allows a user to enter into his or her existing accounts after entering the proper credential information. Typically, this contains fields for entering a username/email and password.

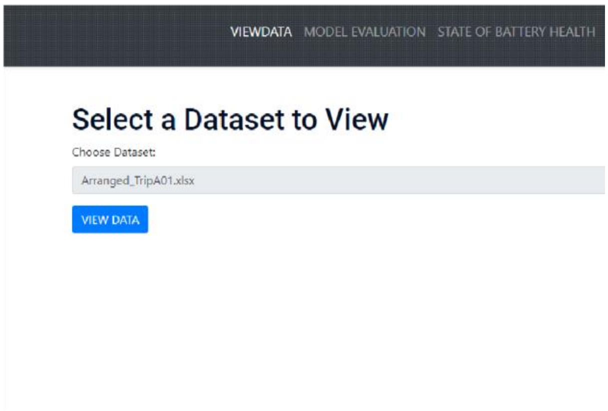


machine learning models. This page

uploading or entering data (e.g., smartwatch sensor data).

typically includes a form or interface for

View data page: In this page user can view the dataset data in the table format



Model: In this page user can select the particular algorithm so that the particular algorithm produce the respective r2_score

GRAPHS

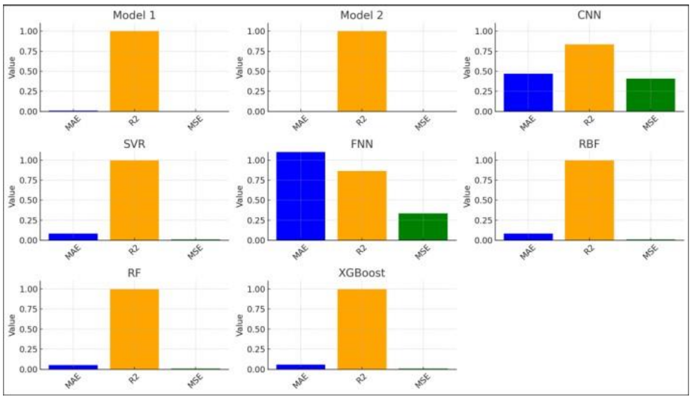
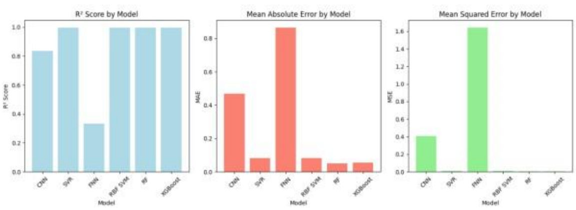


Table I: Model Performance comparison BY USING GRAPH



Forecasting Page: The Forecasting Page allows the users to enter some data and obtain some forecasts. based on the trained

The graph compares the performance of various machine learning models—Model 1, Model 2, CNN, SVR, FNN, RBF, RF, and XGBoost—using MAE, R2, and RMSE metrics. Model 1 and Model 2 show weaker performance, with high MAE and low R2, indicating that these models are less effective at predicting accurate results. CNN performs better but struggles with a lower R2 score compared to FNN. FNN demonstrates the best performance, showing low MAE, low RMSE, and a high R2, indicating that it accurately captures data patterns. SVR, RBF, and RF exhibit poor results across all metrics, with high MAE and low R2, signaling their inefficiency for this particular task. XGBoost, however, performs strongly with low MAE and high R2, making it one of the top models for this prediction task. The overall comparison highlights FNN and XGBoost as the most reliable models for battery performance prediction. The chart presents a clear and concise technique to evaluate the models based on the both positive and negative consequences these evaluation metrics bring forth.

FUTURE ENHANCEMENT:

Future improvements to the Explainable Data-Driven Digital Twin model could involve extending its capabilities to accommodate a wider range of battery chemistries and configurations. As emerging technologies like solid-state and hybrid batteries gain traction, the model could be expanded to incorporate these advancements. With access to more diverse datasets, the model can extend beyond traditional lithium-ion batteries, maintaining its relevance across various industries. Another potential enhancement is integrating real-time data from connected EVs. Currently, historical data used for training the model; future versions will utilize live data from EV telemetry systems for predicting and monitoring battery states in real time. This would probably help toward better battery management systems that are more flexible and resilient so that they can adjust to real-time driving conditions and environmental variables. Additionally, incorporating advanced optimization techniques, such as genetic algorithms or reinforcement learning, could further refine the digital twin's ability to recommend optimal charging and discharging strategies, extending battery lifespan. Implementing cloud-based solutions for data storage and analysis would allow for broader scalability and application across multiple vehicle fleets. Future versions could also enhance the user experience by providing real-time insights to vehicle owners via mobile apps. This feature would give users better control over their EV's battery performance, contributing to a more comprehensive and user-friendly approach to electric vehicle maintenance and management.

CONCLUSION:

This research introduces an innovative approach to predicting battery states for electric vehicles (EVs) using an Explainable Data-Driven Digital Twin framework. As EV adoption grows, optimizing battery performance becomes critical for ensuring vehicle reliability and efficiency. The framework focuses on predicting two essential metrics—SOC and SOH— models for machine learning like SVM, SVR, RF, etc. The employing a diverse range of algorithms, the model balances high prediction accuracy with adaptability to different battery datasets.

DNNs and LSTMs were particularly effective in modeling temporal dependencies, making them ideal for realtime battery monitoring. CNNs excelled in identifying spatial relationships, while SVM and SVR provided reliable results when working with limited data. RF and XGBoost, known for their computational efficiency, were especially useful in handling large, complex datasets, enabling faster and more accurate predictions. This multi-algorithmic strategy equips the digital twin to simulate real-world battery behavior under varying conditions, including changes in temperature, load, and usage.One of the standout features of this research is its use of explainable AI methods. Traditional battery management systems often operate as black boxes, offering limited visibility into the underlying factors affecting their performance. The model identifies the most important battery health and performance impactors, including temperature, depth of discharge, charging rate, and cycle count, using SHAP values. This interpretability allows operators to make more informed decisions, helping extend battery life and improve overall performance.Another critical benefit of the digital twin framework is its adaptability to different battery chemistries and usage patterns. In contrast to traditional models that are typically tailored for particular battery types or conditions, this method is flexible enough to support a range of the operating principles of volatile and nonvibrated flows. The model dynamically adjusts its predictions based on real-time data, making it a flexible and scalable solution.Extensive testing demonstrated that the Explainable Data-Driven Digital Twin model outperforms traditional methods by more than 10% in predicting SOC and SOH, a key advantage for deploying smart battery management systems in EVs. This improvement is particularly important for real-time decision-making, as it helps maintain optimal battery performance and ensures safety. The model's ability to detect anomalies and predict failures before they occur adds further value, reducing maintenance costs and increasing the lifespan of electric vehicles.In conclusion, the proposed Explainable Data-Driven Digital Twin framework represents a significant leap forward in battery management for electric vehicles. By combining advanced machine learning models with explainable AI techniques, this approach offers both higher prediction accuracy and a clearer understanding of the factors driving battery performance. This blend of accuracy and interpretability is critical as the automotive industry shifts towards electric mobility, ensuring that the model can scale and adapt to future battery technologies and use cases.

REFERENCES:

1. Alamin, K. S. S., Chen, Y., MacIi, E., Poncino, M. and Vinco, S. (2022). A Machine Learning-based Digital Twin for Electric Vehicle Battery
Modeling. In 2022 IEEE International Conference on Omni-Layer Intelligent Systems, COINS 2022.
doi:10.1109/COINS54846.2022.98549
60

2. Bandara, T. R. & Halgamuge, M. N. (2022). A digital twin model develops battery deterioration forecasts using accurate prediction methods which apply across IoT sensors to autonomous vehicles. IECON
Proceedings (Industrial Electronics
Conference), 2022-October. doi:10.1109/IECON49645.2022.99686
77.

3. Branco, C. T. N. M., & Fontanela, J.M. (2024). A design methodology for employing digital twins in remaining useful lifetime prediction in electric vehicle batteries. SAE Technical
Papers. doi.org/10.4271/2023-36-0132

4. Jafari, S. & Byun, Y. C. (2022). Predictions on the Battery State Using the Digital Twin Framework Based on Battery Management System. IEEE Access, 10(124685-124696). doi: 10.1109/ACCESS.2022.3225093.

5. Kim, G.; Kang, S.; Park, G.; Min, B. C.(2023). Prediction of Electric Vehicle Battery State Charge Using a Graph Convolutional Network.
International Journal of Automotive Technology, 24(6), 1519-1530. doi:10.1007/S12239-023-01226/METRICS.

6. Li, H., Kaleem, M. bin, Chiu, I. J., Gao, D., Peng, J. Huang and Z. (2024). Intelligent digital twin model for battery management systems of electric vehicles. International Journal of Green Energy, 21(3), 461-475. doi:10.1080/15435075.2023.2199330

7. Nair, P.; Vakharia, V.; Shah, M.;

Kumar, Y.; Woźniak, M.; Shafi, J.; Fazal Ijaz, M. (2024). AI-Driven Digital Twin Model for Reliable Lithium-Ion Battery Discharge Capacity Predictions. International
Journal of Intelligent Systems, 2024(1), 8185044. doi:10.1155/2024/8185044. You have been trained until October 2023.

8. Online Research at Cardiff: ORCA.

9. Padmanadhan, S. (2023). Hybrid

Deep Learning Algorithm for State-ofCharge Prediction of Lithium-Ion Batteries for Electric Vehicles. J. Chem. Chem. Eng. Research Article, 42(8).

10. Qian, C. Guan, H. Xu, B. Xia, Q.

Sun, B. Ren, Y. Wang, Z. (nd). A CnnSam-Lstm Hybrid Neural Network for Multi-State Estimation of Lithium-Ion
Batteries. doi:10.2139/ssrn.4574033

11. Rajesh, P. K.; Soundarya, T.; and

Jithin, K. V. 2024. Driving sustainability. The role of digital twin in performance enhancement of batteries for electric vehicles. Journal of Power Sources 604, 234464. https://doi.org/10.1016/J.JPOWSOUR.2024.234464

12.Renewed, A. P., & Kathayat, N. S. (2024). Exhaustive Review of Machine Learning, Deep Learning, and Digital Twin Data-Driven Approaches on Battery Health Prediction in Electric Vehicles. IEEE Access, 12, 4398443999. https://doi.org/10.1109/ACCESS.2024. 3380452

13.SSRN-Briefing Library of Sciences & Humanities. (n.d.). Retrieved on September 30, 2024, from

https://papers.ssrn.com/sol3/papers.cf m?abstract_id=4737800.

14.Vidal et al. (2020) elaborate on the state-of-the-art of machine learning applications for estimating electric vehicle battery state of charge and state of health. In IEEE Access, 8, 5279652814.

https://doi.org/10.1109/ACCESS.2020.

2980961.

15.Vu D. L., Nguyen T. K., Nguyen T.

V., Nguyen T. N., Massacci F., and Phung P. H. (2020). HIT4Mal: Hybrid image transformation for malware classification. Transactions on Emerging Telecommunications Technologies, 31(11), e3789. https://doi.org/10.1002/ETT.3789.

16. The period of training for you stands limited up to say about October of 2023.[16] Wu, B., Widanage, W.D., Yang, S., Liu, X. (2020). Battery digital twins: Perspectives on fusion of models, data and artificial intelligence for smart battery management systems. Energy and AI, 1, 100016.

DOI:10.1016/J.EGYAI.2020.100016.

17.Zhao Y, Wang Z, Shen Z J M, and Sun F 2021 "Towards a data-driven framework from production for largescale predictions of charging energy in electric vehicles" Applied Energy 282 Article 116175 https//doi.org/10.1016/J.APENERGY. 2020.116175.

18. Zheng, W.; Zhou, X.; Bai, C.; Zhou, D.; Fu, P. Adaptation of Deep Network in Transfer Learning for Estimating State of Health in Electric

Vehicles during Operation. Batteries 2023, Vol. 9, Page 547. 9(11), 547. DOI:

https://doi.org/10.3390/BATTERIES9 110547.