INTERNATIONAL JOURNAL OF ADVANCED RESEARCH
International Peer-Reviewed Research Journal
ISSN (Online): 3106-7980 | ISSN (Print): 3106-7972
Volume 3 • Issue 7 • November 2026

PREDICTING ELECTRIC VEHICLE BATTERY STATE OF CHARGE WITH ENSEMBLE LEARNING FOR ENERGY EFFICIENCY OPTIMISATION

Shahan Yamin Siddiqui[*1], Muhammad Farrukh Khan[2], Shahzada Atif Naveed[3], Abdullah Dar[4], Unaiza Rehman[5], Muhammad Waleed[6]

[*1,6] *Department of Computer Science, NASTP Institute of Information Technology, Lahore 54810, Pakistan*
[2] *Department of Artificial Intelligence, NASTP Institute of Information Technology, Lahore 54810, Pakistan*
[3] *Department of Computer Science, Rashid Latif Khan University, Lahore, Pakistan*
[4] *Global Reporting and Analysis Department, Ibex Global Solution, Lahore, Pakistan*
[5] *Department of Software Engineering, NASTP Institute of Information Technology, Lahore 54810, Pakistan*

[*1]drshahan@niit.edu.pk, [2]dr.farrukhkhan@niit.edu.pk, [3]shahzada.atif@rlku.edu.pk, [4]abdullah.dar@ibex.co, [5]unaizarehman@niit.edu.pk, [6]muhammad.waleed@niit.edu.pk

DOI: https://doi.org/10.5281/zenodo.24131572

Keywords
Electric Vehicles, Energy Efficiency, Energy Management System, State of Charge, Machine Learning, XGBoost, Random Forest.

Article History
Received: 05 August 2026
Accepted: 19 October 2026
Published: 11 November 2026

Copyright @Author
Corresponding Author: *
Shahan Yamin Siddiqui

Abstract

The transition from internal combustion engine vehicles to electric vehicles (EVs) is central to decarbonizing road transport, but limited driving range, battery degradation and the variability of real-world energy use continue to restrict adoption. Accurate, data-driven estimation of the energy state of the vehicle is a prerequisite for energy management strategies that can respond to changing driving and environmental conditions. This study models an intelligent energy efficiency system for EVs that organizes vehicle telemetry into a three-stage pipeline of data collection, data preprocessing and learning and evaluation, and uses the resulting predictions to drive energy optimization actions such as speed adjustment, battery optimization and power distribution. At the core of the system, three machine learning regressors, a Random Forest, an Extreme Gradient Boosting (XGBoost) model and a multilayer perceptron (MLP) neural network, were trained to predict the state of charge (SOC) of the high-voltage battery from cleaned driving, battery and environmental features. The models were evaluated on an 80:20 train–test split with fixed hyperparameters and further assessed by five-fold cross-validation. XGBoost achieved the best overall performance, with a root mean squared error (RMSE) of 0.7809, a coefficient of determination (R^2) of 0.9724 and the lowest cross-validation RMSE of 0.8002, while the Random Forest attained the lowest mean absolute error (MAE) of 0.0834 and an R^2 of 0.9673. The MLP performed markedly worse, with an RMSE of 1.6011, an R^2 of 0.8838 and a cross-validation RMSE of 1.8326, indicating weaker generalization on the available data. The results show that tree-based ensemble learners provide accurate and stable SOC predictions that can serve as the predictive layer of software-based EV energy management, improving energy awareness without hardware modification.

INTRODUCTION

Electric vehicles (EVs) are among the key solutions utilized to curb greenhouse gas emissions and fossil fuel consumption in road transport. Since they operate using electric engines powered by rechargeable batteries, EVs do not produce exhaust emissions, convert energy into motion with higher efficiency than conventional engines, and utilize electricity from renewable sources. Nonetheless, their massive diffusion is limited by many factors, including the insufficient driving range, the lengthy charging time, the price and degradation of lithium-ion batteries, and the unequal distribution of EV charging stations [1, 2].

The main core of these factors is efficiency. The driving range of an EV, so important for its owners' confidence, is determined through the effectiveness of the energy catched in the battery to provide driving in real-life conditions. Energy consumption depends on factors that interact with each other, such as driving style, speed and acceleration patterns, type of terrain and road, traffic density, temperature outside, external load, and battery condition [3, 2]. Range anxiety arises mostly because the listed factors constantly oscillate in time, and thus prediction based on static variables fails.

Conventional energy management strategies address this problem only partially. Rule-based and optimization-based controllers are typically designed around fixed operating patterns or individual components, such as the battery or the regenerative braking system, and they do not adapt well when driving conditions, traffic and driver behavior change during operation [4], [5]. Physics-based consumption models require detailed vehicle parameters that are rarely available for every vehicle and are sensitive to unmodelled effects [3]. These limitations motivate data-driven approaches that learn the relationship between operating conditions and energy use directly from vehicle data.

Machine learning has therefore become an active field in EV energy research. Supervised regressors and neural networks have been used to forecast energy consumption from trip and environmental features [6], [7], ensemble methods such as stacked generalization have been shown to improve the robustness of these forecasts [8], and classifier fusion has been applied to energy management in fuel-cell hybrid vehicles [9]. Reinforcement learning has been used to make energy-efficient driving decisions in dynamic traffic [10], and machine learning has also been applied to charge management and load forecasting at the grid interface [11]. Reviews of the field conclude that data-driven methods can capture the non-linear dependencies of EV energy use more accurately than static models, but that their application to fully electric vehicles, as opposed to hybrid powertrains, remains comparatively limited [1], [4].

Several gaps nevertheless remain. Much of the machine learning work on energy management concentrates on hybrid electric vehicles, where the central problem is the power split between engine and motor, rather than on battery-only powertrains whose energy management depends entirely on the state of the battery [4], [5]. Predictive studies frequently stop at estimating consumption without integrating the prediction into a system that acts on it [6], while advanced controllers often neglect battery state variables in their decisions [10]. Finally, comparisons between families of learners, such as bagging ensembles, boosting ensembles and neural networks, are seldom carried out under a common protocol with cross-validated error estimates, which makes it difficult to decide which model should form the predictive core of an energy management system.

This work addresses these gaps by modelling an intelligent energy efficiency system for EVs in which the battery state of charge (SOC) is predicted by machine learning and used to guide energy optimization. The objectives are to assess and compare the effectiveness of different machine learning algorithms for EV energy prediction, to identify the driving, battery and environmental factors that such models should consider, and to compare the accuracy of the learned models with that of conventional energy prediction approaches. The proposed system organizes vehicle telemetry into a three-stage

pipeline of data collection, preprocessing and learning and evaluation. Within this pipeline, a Random Forest, an XGBoost model and a multilayer perceptron (MLP) are trained to predict the SOC of the high-voltage battery [12], [13]. XGBoost achieved the highest R^2 (0.9724) and the lowest test and cross-validation RMSE, and the Random Forest achieved the lowest MAE, whereas the MLP generalized markedly worse.

The remainder of this paper is organized as follows. The Related Work section reviews machine learning approaches to EV energy prediction and management. The Proposed Methodology section describes the system architecture, data preprocessing, feature selection, learning algorithms and the optimization and decision-fusion layers. The Experimental Setup section presents the dataset, the training and validation protocol and the evaluation metrics. The Results and Discussion section reports and analyses the performance of the three models and compares them with previously published work, and the final section concludes the paper and outlines future work.

The major contributions of this work are as follows:

- A modular intelligent energy efficiency system for EVs that links telemetry acquisition, data preprocessing, machine learning prediction, energy optimization and decision fusion in a single pipeline.
- A preprocessing and feature extraction procedure that removes sensor errors, missing values and outliers and derives battery-level, driving-pattern and energy-usage features for prediction of the high-voltage battery state of charge.
- A comparative evaluation of Random Forest, XGBoost and MLP repressors under a common 80:20 protocol, reported with MAE, RMSE and R^2 and supported by five-fold cross-validation.
- Evidence that tree-based ensembles predict SOC accurately and stably, with XGBoost achieving an R^2 of 0.9724 and the lowest cross-validation RMSE (0.8002), while the MLP generalizes markedly worse on the same data.
- A comparison with previously published machine learning approaches to EV energy prediction and management, together with a discussion of how the predictive layer supports real-time optimization actions.

Related Work

Several reviews have mapped the use of machine learning for EV energy. Adnane et al. [1] surveyed machine learning for managing EV energy consumption and concluded that, despite strong results in other fields, such methods are still applied only modestly to EV energy management, which they identified as a key enabler of vehicles that make fast, autonomous decisions to reduce power consumption. Zhang et al. [2] reviewed machine learning approaches for EV energy consumption modelling in urban transport, contrasting traditional estimation models with interpretable machine learning and neural networks, and argued that accurate consumption modelling is central to overcoming limited range and charging infrastructure. Chen et al. [3] classified consumption estimation models into physics-based, statistical and data-driven families and highlighted the trade-off between the parameter requirements of physical models and the data requirements of learned ones.

Predictive models of EV energy use form the largest group of studies. Vatin and Manasa [6] applied predictive machine learning models to EVs with battery capacities of 55–75 kWh and reported MAEs of 0.13–0.18 kWh/km and R^2 values of 0.80–0.88 for energy consumption, but their models estimate consumption without adapting the vehicle's operation. Zhu et al. [7] predicted consumption from field data of a real-world EV fleet using physics-informed features and a quantile regression neural network, reducing the average error from 6.30% to 5.04% through vehicle-specific online adaptation and providing calibrated prediction intervals. Ullah et al. [8] combined decision tree, random forest and k-nearest neighbor repressors through stacked generalization to predict energy efficiency from GPS tracking data, trip characteristics, auxiliary loads and road gradient, and found the ensemble

more robust than its constituents. These studies show that learned models capture the non-linear dependence of energy use on driving and environmental conditions, but they differ widely in their targets, data and evaluation protocols.

A second group integrates learning into energy management strategies. Jui et al. [4] surveyed machine learning-based energy management for hybrid electric vehicles and showed how learning methods adapt to changing operating conditions and capture non-linear powertrain behavior, while Recalde et al. [5] reviewed machine learning, model predictive control and real-time optimization for plug-in hybrid vehicles and noted the open challenges of real-time deployment. Chatterjee et al. [9] fused support vector machine, k-nearest neighbor and naive Bayes classifiers in the energy management of a fuel-cell hybrid vehicle and raised the classification accuracy from 92–96% for the individual classifiers to 98% for the fused decision. Manivannan [14] proposed an IoT-based, machine learning-driven energy management system for hybrid vehicles. Much of this work, however, concerns hybrid powertrains, in which the main decision is how to split power between engine and motor rather than how to manage a battery-only system.

Reinforcement learning and grid-level studies extend the scope of energy management beyond the vehicle. Wu et al. [10] developed a deep reinforcement learning strategy for autonomous EVs that jointly considers car-following and lane-change decisions in dynamic traffic and markedly improves energy economy without sacrificing safety or traffic efficiency; the strategy, however, does not account for battery state variables such as capacity, temperature and ageing. Mazhar et al. [11] compared deep neural networks, k-nearest neighbors, long short-term memory networks, random forests, support vector machines and decision trees for EV charge management in the smart grid and found that an LSTM-based approach improved peak voltage, power losses and voltage stability. These methods are mighty, but because they are data-and computation-intensive, they are conceptualized for decisions given at traffic or grid level totally separate from the energy state of the vehicle.

Thus, a clear gap is left in the literature. Predictive studies often give good estimates of absolute consumption but would seldom be tucked in systems that act on the prediction [6], [7]; management studies focus on hybrid powertrains [4], [5], and [9]; and reinforcement learning strategies for fully electric vehicles often disregard the battery state in their decisions [10]. Furthermore,28 cross-validated error estimates of bagging ensembles, boosting ensembles and neural networks in the same protocol rarely exist. What is needed is an energy efficiency system for battery-only EVs in which an accurately predicted battery state forms the basis for optimization actions, and in which the predictive core is selected through a systematic comparison of learner families. The system proposed in this paper is designed to meet these requirements.

Proposed Methodology

The proposed intelligent energy efficiency system is built on a modular architecture that connects data collection, preprocessing, feature extraction, machine learning and energy optimization, as shown in Figure 1. The architecture is organized into three stages. In the data collection stage, telemetry describing traffic flow, vehicle speed, battery usage and driving behavior is gathered from EVs in operation and transmitted to a central database. In the data preprocessing stage, the raw records are cleaned, transformed into informative features and partitioned into training and testing sets. In the learning and evaluation stage, machine learning models are trained and assessed, and the selected model drives an optimization layer that adjusts power use and battery operation and issues energy efficiency actions such as speed adjustment, battery optimization and power distribution.

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH
International Peer-Reviewed Research Journal
ISSN (Online): 3106-7980 | ISSN (Print): 3106-7972
Volume 3 • Issue 7 • November 2026

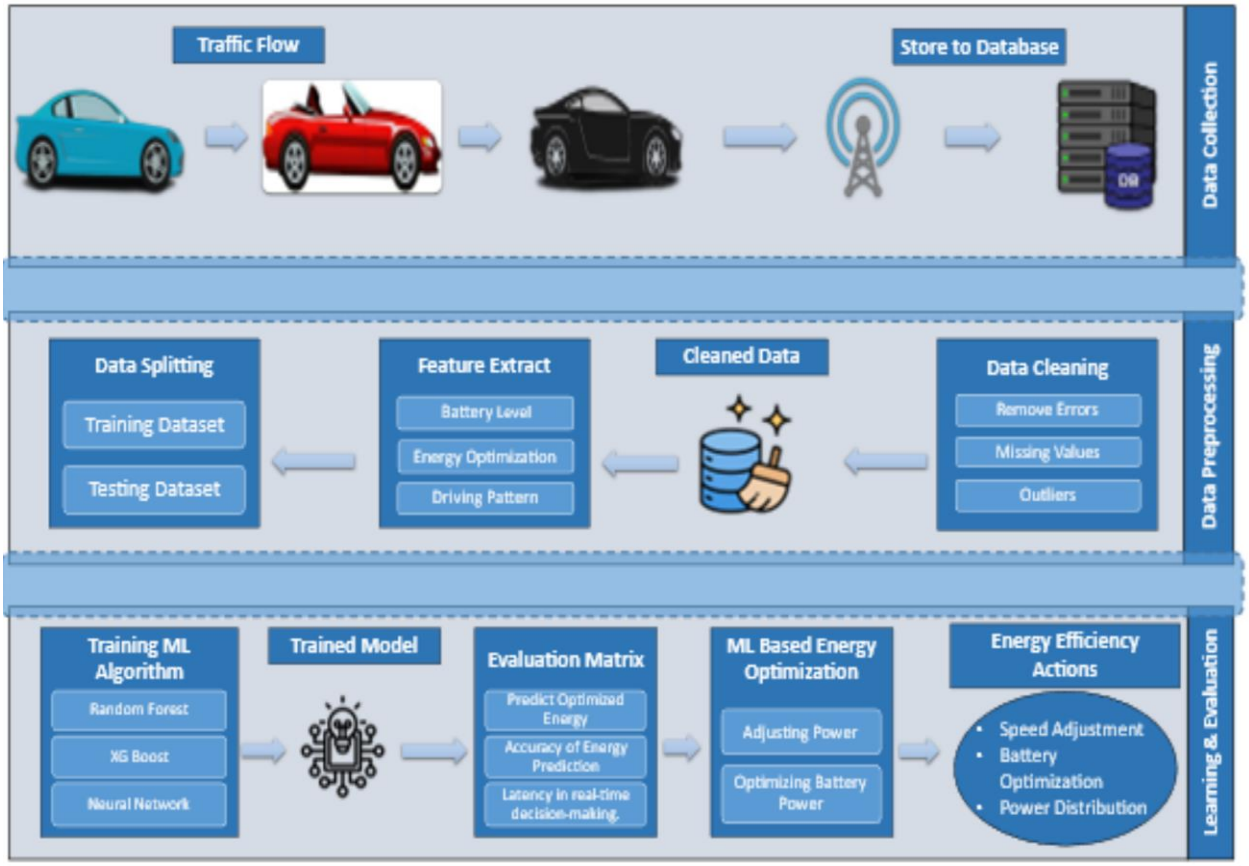


Figure 1: Proposed Intelligent Energy Efficiency System for Electric Vehicles

A. Data Collection

The process of data acquisition starts with telemetry from electric vehicles in the new system. The parameters of the data include speed, acceleration/deceleration, battery state, power consumption, road state, and traffic density. The information is then sent with wireless channels to a central database, where it is kept and processed before creating models. Since the data is received constantly, it would help to increase the size of the database.

B. Data Preprocessing

The raw telemetry is affected by sensor malfunctions, gaps in transmission and impossible data. The first stage of the preprocessing is the removal of the erroneous records, followed by filling the missing values with the help of imputation, and finally locating and eliminating the outliers that alter the training of the model negatively. The gathered non-disturbed information goes through the procedure of feature extraction and is divided into training and testing sets for supervised learning. The described procedure guarantees that the training is conducted with the proper credible data required for reliable prediction.

C. Feature Selection

Feature selection refers to keeping only the informative variables affecting energy use and battery behavior while discarding repetitions and uninformative. There are three feature groups considered: battery-level features that measure the energy available in the vehicle, driving-pattern features such as acceleration, braking rate, speed, and cruising; as well as energy-expenditure features that include average energy consumption per km

and power tendency. The final models are built based on domain knowledge of the features and their relation to the target variable, which minimizes computational expense, prevents overfitting, and makes the models useful for practical applications.

D. Random Forest

Random Forest is an ensemble of decision trees, each trained on a bootstrap sample of the data and a random subset of features, whose predictions are averaged for regression [12]. For an input x, the ensemble prediction is

$\hat{y}(x) = (1 / B) \Sigma_b T_b(x), \quad b = 1, ..., B$

where T_b is the b-th tree and B is the number of trees. Averaging decorrelated trees reduces variance and makes the model robust to the noisy, heterogeneous signals typical of vehicle telemetry, and the impurity-based feature importances of the forest indicate which variables most affect energy use.

E. XGBoost

XGBoost builds an additive ensemble of regression trees in which each new tree is fitted to the gradient of the loss with respect to the current prediction [13]. The model minimizes the regularized objective

$Obj = \Sigma_i\, l(y_i, \hat{y}_i) + \Sigma_k\, \Omega(f_k), \quad \Omega(f) = \gamma T + \tfrac{1}{2} \lambda\, \Sigma_j\, w_j^2$

where l is the squared-error loss between the observed and predicted SOC, f_k is the k-th tree, T is the number of leaves and w_j are the leaf weights. The regularization term, shrinkage and column subsampling limit overfitting, and the efficient implementation allows the model to be retrained quickly as new driving data arrive, which suits real-time energy prediction.

F. Multilayer Perceptron

The neural network component is a multilayer perceptron (MLP) that maps the input features through fully connected hidden layers with non-linear activations to a continuous SOC estimate. For a hidden layer l, the activations are

$h^{(l)} = \varphi(W^{(l)} h^{(l-1)} + b^{(l)})$

where $W^{(l)}$ and $b^{(l)}$ are the weights and biases of the layer and φ is the activation function. The MLP can represent complex, interacting dependencies between driving conditions and battery state, but it typically requires more data and more careful tuning than tree ensembles.

G. Energy Optimization and Decision Fusion

The trained predictive model supplies the optimization layer of the system. In real-time energy optimization, the predicted battery state is combined with current inputs such as speed, traffic information and battery status to adjust the acceleration profile, the use of regenerative braking and the distribution of power. Dynamic load balancing uses the same predictions to allocate power among the drivetrain, the heating and air-conditioning system and auxiliary consumers, preventing overload and reducing energy losses. A decision-fusion layer combines the outputs of the prediction modules, such as the driving state, battery state and power requirement, into a single action through rule-based or probabilistic weighting according to context and confidence; for example, a power-saving indication from the battery module can trigger the reduction of high-consumption loads, whereas a demand for immediate power switches the system to a high-power mode. The present study evaluates the predictive layer on which these actions depend.

Experimental Setup

A. Dataset

Using an EV operation dataset that spans several driving scenarios namely; urban driving, highway driving and stop-and-go driving, various models were developed. The dataset captures the battery's condition and other features like vehicle speed, energy used, vehicle acceleration and braking patterns, and traffic situation to create the target variable which is the high-voltage battery state of charge (SOC). The values were cleaned from missing observations, outliers as well as incorrect information.

B. Training and Validation Protocol

The processed dataset is divided into the training dataset that consists of approximately 80% volume of the records and the test dataset taking the rest

20% of the records. The Random Forest, MLP repressor and XGBoost were trained on the training dataset and tested on the test dataset, leaving the tuning parameters the same in all trials so that the comparison highlights the learning algorithms and excludes the differences in tuning effort. To assess generalization, each model was additionally evaluated by five-fold cross-validation, in which the data are partitioned into five folds and each fold serves once as the validation set while the remaining four are used for training; the mean RMSE across the folds is reported. The experiments were implemented in Python.

C. Evaluation Metrics

Prediction quality was assessed with three standard regression metrics. For n test samples with observed values y_i, predictions $\hat{y}_i$ and mean observed value $\bar{y}$:

$$MAE = (1 / n) \Sigma_i |y_i - \hat{y}_i|$$
$$RMSE = \sqrt{[(1 / n) \Sigma_i (y_i - \hat{y}_i)^2]}$$
$$R^2 = 1 - \Sigma_i (y_i - \hat{y}_i)^2 / \Sigma_i (y_i - \bar{y})^2$$

MAE measures the typical magnitude of the error, RMSE penalizes large errors more heavily, and R^2 expresses the proportion of the variance in SOC explained by the model. Reporting MAE and RMSE together also reveals how errors are distributed: when RMSE is much larger than MAE, most predictions are very close to the observed values but a small number of cases carry large errors.

Results and Discussion

Table 1 reports the performance of the three models on the test set, and Table 2 reports the mean RMSE obtained by five-fold cross-validation. Both are visualized in Figure 2.

International Journal of

Table 1: Performance of the Models on the Test Set

Model	MAE	RMSE	R^2
Random Forest	0.0834	0.8499	0.9673
XGBoost	0.0856	0.7809	0.9724
MLP Regressor	0.3947	1.6011	0.8838

Table 2: Five-Fold Cross-Validation RMSE

Model	Cross-Validation RMSE	Test RMSE	Difference
Random Forest	0.8759	0.8499	+0.0260
XGBoost	0.8002	0.7809	+0.0193
MLP Regressor	1.8326	1.6011	+0.2315

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH
International Peer-Reviewed Research Journal
ISSN (Online): 3106-7980 | ISSN (Print): 3106-7972
Volume 3 • Issue 7 • November 2026

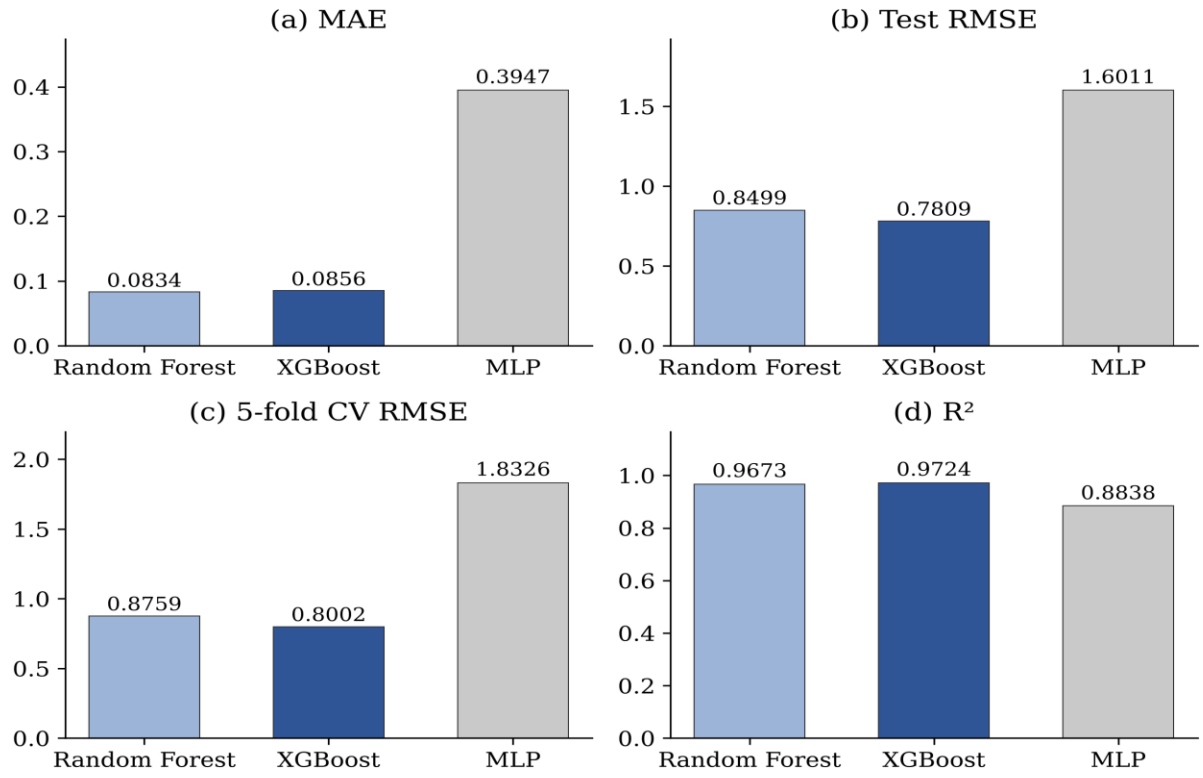


Figure 2: MAE, Test RMSE, Cross-Validation RMSE and R^2 of the Three Models

XGBoost provided the best overall fit. It achieved the highest R^2 (0.9724), the lowest test RMSE (0.7809) and the lowest cross-validation RMSE (0.8002), and its cross-validation RMSE exceeded its test RMSE by only 0.0193, indicating stable behavior across different partitions of the data. Its advantage stems from the sequential correction of residual errors, combined with regularization that limits the complexity of individual trees [13], which allows it to capture the non-linear interactions between driving, battery and environmental features without overfitting.

The Random Forest performed almost as well, with an R^2 of 0.9673 and the lowest MAE of all models (0.0834). Its lower MAE, but higher RMSE when compared to XGBoost shows that the forest can make accurate predictions for usual cases, but it produces some major errors in unusual cases, as RMSE is much more sensitive to squared errors. Both tree ensembles showed that RMSE was about ten times larger than MAE. Therefore, error distributions in these models are very compact, but with a long tail: majority of SOC are predicted correctly, while there is a minority of cases, mostly related to unexpected transitions, that is responsible for the largest squared error in predictions.

The neural network has shown worse results in comparison to both models, as the MAE (0.3947) was almost five times more than that of other two models, the test RMSE (1.6011) was almost twice as large, and the R^2 (0.8838) was significantly lower. Additionally, the difference between cross-validation and test RMSE (0.2315) was significantly bigger than the one for the tree models suggesting that results of this model depend on random partition of the data much more than the results from the other two models. Neural networks generally require larger datasets and more extensive hyperparameter tuning than tree ensembles, and with fixed hyperparameters on structured telemetry data the MLP could not

match the inductive bias of the tree-based models. Overall, the comparison shows that tree ensembles are the most suitable predictive core for the proposed system, with XGBoost preferred when large errors must be minimized and the Random Forest a close, easily interpretable alternative.

Table 3 places these results in the context of previously published machine learning approaches to EV energy prediction and management. The comparison is qualitative, because the studies differ in their prediction targets, vehicles, datasets and evaluation protocols.

Table 3: Comparison with Previously Published Studies

Study	Method	Task / Reported Result	Limitation
Vatin and Manasa [6]	Predictive ML models	Energy consumption: MAE 0.13–0.18 kWh/km, R^2 0.80–0.88	Static prediction; no adaptive optimization
Zhu et al. [7]	Physics-informed QRNN with online adaptation	Energy consumption: average error 5.04%	Requires fleet data and vehicle-specific adaptation
Ullah et al. [8]	Stacked DT, RF and KNN	Energy efficiency (kWh/km): ensemble more robust than base models	Prediction only; no battery state
Chatterjee et al. [9]	SVM, KNN and NB with classifier fusion	FCHEV energy management: 98% accuracy	Hybrid powertrain; high computation
Wu et al. [10]	Deep reinforcement learning	Energy-efficient driving decisions in dynamic traffic	Battery state variables not modelled
This study	Random Forest, XGBoost, MLP	Battery SOC: R^2 0.9724, RMSE 0.7809 (XGBoost)	Single dataset; predictive layer evaluated

Compared with the predictive studies, the proposed approach focuses on the battery state rather than on trip-level consumption, and its tree ensembles explain more than 96% of the variance in SOC. Unlike the energy management studies on hybrid and fuel-cell vehicles [9], [4], the system is designed for battery-only powertrains, in which the battery state is the decisive quantity, and unlike reinforcement learning strategies that omit battery variables [10], it places a predicted battery state at the center of its optimization layer. Because the system relies on software intelligence rather than hardware modification, it can be applied to existing vehicles and updated as new data become available.

Several limitations should be noted. The models were trained and tested on a single dataset, and validation on publicly available real-world EV driving data is required; their generalization to other vehicle models, battery chemistries, climates and terrains remains to be demonstrated. The hyperparameters were held fixed, and systematic tuning could improve all three models, particularly the MLP. The present evaluation concerns the predictive layer; the energy savings delivered by the optimization, load-balancing and decision-fusion layers in closed-loop operation, together with their decision latency, remain to be quantified in simulation or on-vehicle tests. Finally, SOC prediction errors are concentrated in a small number of cases, and analyzing these cases,

for example by driving event or SOC range, would guide further improvement.

Conclusion and Future Work

This study modelled an intelligent energy efficiency system for electric vehicles in which machine learning predicts the state of charge of the high-voltage battery and the prediction guides energy optimization. The system organizes vehicle telemetry into a pipeline of data collection, preprocessing and feature extraction, and learning and evaluation, and it couples the selected predictive model with real-time optimization, dynamic load balancing and decision fusion to support actions such as speed adjustment, battery optimization and power distribution.

Three learners were compared under a common 80:20 protocol with five-fold cross-validation. XGBoost achieved the best overall performance, with an R^2 of 0.9724, a test RMSE of 0.7809 and the lowest cross-validation RMSE of 0.8002; the Random Forest achieved the lowest MAE (0.0834) and an R^2 of 0.9673; and the MLP was markedly less accurate and less stable, with an R^2 of 0.8838 and a cross-validation RMSE of 1.8326. These results show that tree-based ensembles provide an accurate and stable predictive core for software-based EV energy management that improves energy awareness without hardware changes.

Future work will extend the dataset to additional vehicle models, battery chemistries, weather conditions and terrain types to improve generalization, and will incorporate driver behavior profiles and real-time weather information. Sequence models such as long short-term memory networks and Transformer architectures will be investigated to capture temporal dependencies in driving and energy use, and hyper parameter optimization will be applied to all models. The optimization and decision-fusion layers will be evaluated in closed-loop simulation and on-vehicle tests to quantify energy savings and decision latency. Finally, a cloud–edge deployment combining on-board inference with cloud-based retraining, together with vehicle-to-infrastructure and vehicle-to-grid communication, would enable cooperative energy optimization across the wider transport network.

REFERENCES

[1] M. Adnane, A. Khoumsi, and J. P. F. Trovão, "Efficient management of energy consumption of electric vehicles using machine learning—A systematic and comprehensive survey," Energies, vol. 16, no. 13, art. 4897, 2023.

[2] X. Zhang, Z. Zhang, Y. Liu, Z. Xu, and X. Qu, "A review of machine learning approaches for electric vehicle energy consumption modelling in urban transportation," Renewable Energy, vol. 234, art. 121243, 2024.

[3] Y. Chen, G. Wu, R. Sun, A. Dubey, A. Laszka, and P. Pugliese, "A review and outlook on energy consumption estimation models for electric vehicles," SAE International Journal of Sustainable Transportation, Energy, Environment, & Policy, vol. 2, no. 1, 2021.

[4] J. J. Jui, M. A. Ahmad, M. M. I. Molla, and M. I. M. Rashid, "Optimal energy management strategies for hybrid electric vehicles: A recent survey of machine learning approaches," Journal of Engineering Research, 2024, doi: 10.1016/j.jer.2024.01.016.

[5] A. Recalde, R. Cajo, W. Velasquez, and M. S. Alvarez-Alvarado, "Machine learning and optimization in energy management systems for plug-in hybrid electric vehicles: A comprehensive review," Energies, vol. 17, no. 13, art. 3059, 2024.

[6] N. I. Vatin and V. Manasa, "Predictive machine learning for optimal energy management in sustainable transportation systems," MATEC Web of Conferences, vol. 392, art. 01169, 2024.

[7] Q. Zhu, Y. Huang, C. F. Lee, P. Liu, J. Zhang, and T. Wik, "Predicting electric vehicle energy consumption from field data using machine learning," IEEE Transactions on Transportation Electrification, vol. 11, no. 1, pp. 2120–2132, 2025.

[8] I. Ullah, K. Liu, T. Yamamoto, M. Zahid, and A. Jamal, "Electric vehicle energy consumption prediction using stacked generalization: An ensemble learning approach," International Journal of Green Energy, 2021, doi: 10.1080/15435075.2021.1881902.

[9] D. Chatterjee, P. K. Biswas, C. Sain, A. Roy, and F. Ahmad, "Efficient energy management strategy for fuel cell hybrid electric vehicles using classifier fusion technique," IEEE Access, vol. 11, 2023.

[10] J. Wu, Z. Song, and C. Lv, "Deep reinforcement learning-based energy-efficient decision-making for autonomous electric vehicle in dynamic traffic environments," IEEE Transactions on Transportation Electrification, vol. 10, no. 1, pp. 875–887, 2024.

[11] T. Mazhar, R. N. Asif, M. A. Malik, M. A. Nadeem, I. Haq, M. Iqbal, M. Kamran, and S. Ashraf, "Electric vehicle charging system in the smart grid using different machine learning methods," Sustainability, vol. 15, no. 3, art. 2603, 2023.

[12] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[13] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[14] R. Manivannan, "Research on IoT-based hybrid electrical vehicles energy management systems using machine learning-based algorithm," Sustainable Computing: Informatics and Systems, vol. 41, art. 100943, 2024.

International Journal of Advanced Research