International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

Dataset-Driven Energy Consumption Modeling and Range Prediction of Electric Vehicles Using Python and Hardware Sensors

Wadaje Omkar Shivaji[1], Seema N. Kharat[2]

[1]PG scholar, Department of Electrical Engineering (Electrical Power System), MSS's College of Engineering and Technology, Nagewadi, Jalna.

[2]Assistant Professor, Department of Electrical Engineering, MSS's College of Engineering and Technology, Nagewadi, Jalna.
Affiliated to Dr. Babasaheb Ambedkar Technological University, Lonere, Maharashtra-431004

Abstract

The increasing adoption of electric vehicles (EVs) has created a growing need for accurate energy consumption modelling and driving range prediction. Reliable range estimation is essential for improving user confidence, optimizing battery utilization, and reducing range anxiety; however, EV energy consumption is influenced by multiple dynamic factors such as vehicle speed, battery voltage, current, temperature, and state of charge, making accurate prediction a challenging task. This work presents a dataset-driven approach for modelling EV energy consumption and predicting driving range using Python-based machine learning techniques and hardware sensors. Real-time operational parameters were collected through an ESP32/Arduino-based sensor platform (voltage, current, and temperature sensors) and processed using data preprocessing techniques to improve data quality and consistency. Four machine learning models — Linear Regression, Random Forest, XGBoost, and Long Short-Term Memory (LSTM) — were trained on the Kaggle Electric Vehicle Energy Consumption Dataset to analyse the relationship between sensor inputs and energy consumption, and evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 Score. The results show a consistent improvement in prediction accuracy from Linear Regression ($R^2 = 0.88$) through Random Forest ($R^2 = 0.94$) and XGBoost ($R^2 = 0.96$) to LSTM, which achieved the best performance with an MAE of 0.89, RMSE of 1.25, and R^2 Score of 0.98, reflecting its ability to capture temporal dependencies in sequential sensor data. The integration of hardware sensors enabled continuous real-time monitoring and practical validation of the framework. The proposed system enhances EV energy management, improves range prediction reliability, and supports the development of intelligent and sustainable transportation systems.

Keywords: *Electric Vehicles, Energy Consumption Modelling, Range Prediction, Machine Learning, LSTM.*

1. INTRODUCTION

Electric vehicles (EVs) are gradually becoming a key component of modern transportation systems as countries strive to achieve cleaner and more sustainable energy solutions [1]. Unlike conventional internal combustion engine vehicles, EVs operate using electrical energy stored in rechargeable batteries, significantly reducing greenhouse gas emissions and dependence on petroleum-based fuels, with electric motors converting a higher percentage of stored electrical energy into mechanical power than conventional powertrains [4].

Despite these advantages, one of the primary challenges associated with EV adoption is driving range uncertainty, commonly referred to as range anxiety [5] — the concern that a vehicle may not possess sufficient charge to complete a journey. This concern is significant because EV driving range is influenced by numerous dynamic factors that continuously change during operation [6], unlike conventional vehicles where fuel consumption is relatively predictable [7]. Battery condition, vehicle speed, road conditions and terrain, and ambient temperature all contribute to fluctuations in energy consumption [2], [3]; driving behaviour, traffic conditions, auxiliary loads, and vehicle payload further complicate the estimation task. Traditional range estimation methods, relying on fixed assumptions and simplified mathematical models, often fail to adequately capture these real-world operating conditions.

Data-driven approaches combining machine learning with real-time sensor monitoring have gained significant attention as a solution, since they can analyze large volumes of operational data and identify complex relationships among influencing factors [8], enabling more accurate energy consumption modelling and range prediction under varying driving conditions [1]–[5]. This study presents such a dataset-driven framework, combining hardware sensor data acquisition with multiple machine learning models to estimate EV energy consumption and predict driving range.

1.1 Objective of the Work

The main objective of this work is to develop a dataset-driven model for analyzing energy consumption and predicting the driving range of electric vehicles using Python and hardware sensors, improving prediction accuracy by considering real-time variations in operating conditions. Further objectives include analyzing energy consumption patterns under different driving scenarios, enabling real-time monitoring through hardware sensors, and supporting efficient, intelligent energy management for electric vehicles.

2. LITERATURE REVIEW

In the early stages of EV development, range estimation relied mainly on simple mathematical formulas and fixed assumptions, typically considering only basic parameters such as battery capacity and average energy consumption. While easy to implement, such methods ignored real-world variations in traffic, road gradient, weather, and driving behaviour, producing generalized results that were often inaccurate in practical scenarios. Modern EVs are now equipped with speed, current, voltage, and temperature sensors that continuously monitor operational parameters, providing detailed insight into real-time vehicle behaviour; however, the challenge lies in converting this large volume of raw sensor data into meaningful conclusions, motivating the use of advanced machine learning techniques for energy modelling.

Hussain et al. [1] evaluated eleven different machine learning algorithms for EV energy consumption prediction using real-world data, finding that the Extra Trees Regressor achieved the best results, with Gradient Boosting and KNN also performing well, and highlighted the strengths of ensemble and non-linear models in capturing complex consumption patterns. Ullah et al. [2] compared advanced gradient boosting models (XGBoost, LightGBM) against traditional linear regression and neural networks for EV energy prediction, finding that light gradient boosting performed best and identifying key influencing factors relevant to infrastructure planning. In related work, Ullah et al. [3] proposed a stacked generalization ensemble approach for EV energy consumption prediction, demonstrating that combining multiple base learners improves predictive accuracy over any single model.

Ko et al. [4] developed a deep learning method for monitoring vehicle energy consumption using GPS trajectory data, illustrating how location and route information can be incorporated into energy prediction frameworks. Chung et al. [5] proposed an ensemble machine learning algorithm for predicting EV user charging and driving behaviour, relevant to understanding the demand-side variability that complicates range estimation. Achariyaviriya et al. [6] examined real-world energy consumption of battery electric vehicles using onboard diagnostics and GPS data combined with machine learning and SHAP-based interpretation, identifying vehicle speed and battery current as key factors affecting energy usage — consistent with the correlation findings of the present study. Zhang et al. [7] presented a practical real-world-data approach to EV energy prediction incorporating vehicle characteristics, environmental conditions, and driver behaviour, achieving significantly reduced error rates compared to conventional methods and providing insights directly relevant to range-anxiety mitigation.

Bhupathi and Chinta [8] examined AI and machine learning strategies for efficient EV battery use and smart charging, illustrating the broader role of data-driven techniques in EV battery management beyond range prediction alone. Complementing these vehicle-specific studies, a wider body of work has applied machine learning and deep learning techniques to time-series forecasting and predictive modelling in energy systems more generally — including explainable AI-based forecasting frameworks [9], hybrid predictive models combining multiple learning algorithms [14], AI-based forecasting for demand-supply management [10]–[13], and time-series forecasting methods such as ARIMA and SARIMA applied to sensor and environmental data [15] — underscoring the broader applicability of the ensemble and deep-learning techniques (Random Forest, XGBoost, LSTM) adopted in this study.

Despite these advances, many existing studies on EV energy modelling are primarily simulation-based, tested using pre-collected datasets without considering real-time variations in operating conditions, and few integrate data-driven modelling with real-time hardware implementation for practical validation [1]–[15]. This study addresses that gap by combining a dataset-driven machine learning framework with an ESP32/Arduino-based hardware sensor platform for live data acquisition and range prediction.

3. RESEARCH METHODOLOGY

International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

The proposed system combines sensor-based data acquisition, data preprocessing, and machine learning to estimate EV energy consumption and predict remaining driving range. As shown in Figure 1, the architecture comprises a data acquisition layer (speed, voltage, current, temperature, and SOC sensors connected to an ESP32/Arduino controller with Wi-Fi/IoT data transmission), a data processing and machine learning layer (preprocessing — cleaning, normalization, feature selection, and an 80:20 train-test split — followed by model training across four candidate algorithms and evaluation using MAE, RMSE, and R^2 Score), and a prediction and output layer that generates energy consumption predictions, driving range estimation, alerts and notifications, and cloud/local data storage, all surfaced through a real-time monitoring dashboard with a feedback loop back to the acquisition layer.

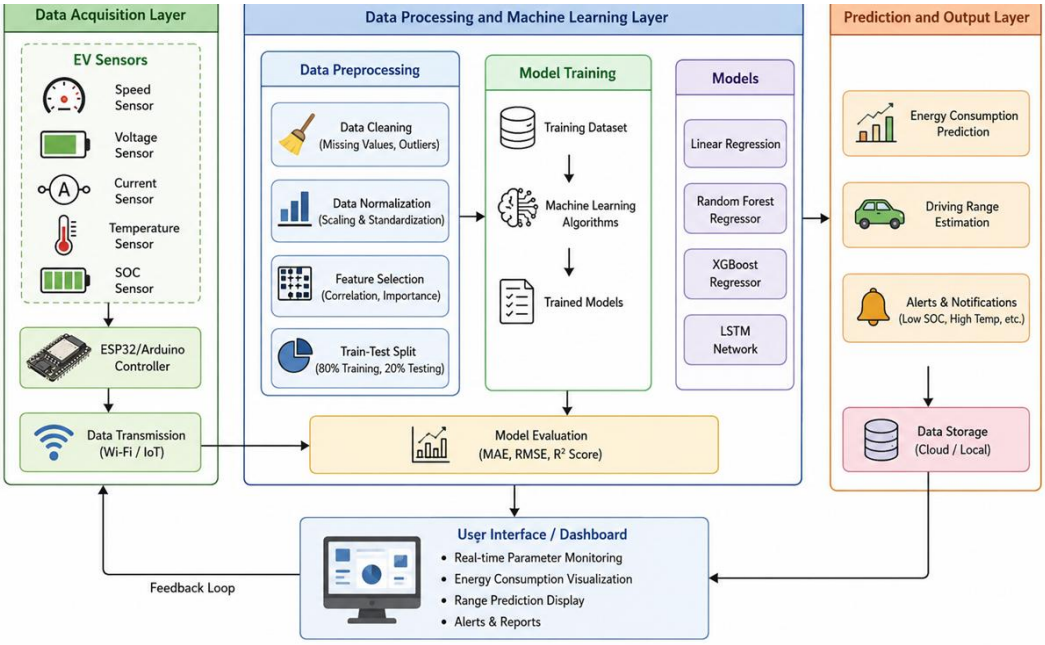


Figure 1: Proposed system architecture for EV energy consumption modelling and driving range prediction.

3.1 Dataset and Hardware

The Kaggle Electric Vehicle Energy Consumption Dataset was used as the primary data source, containing time-series observations of vehicle speed, battery voltage, current, temperature, state of charge (SOC), distance travelled, energy consumption, and remaining range. This was complemented by a physical hardware platform — an ESP32/Arduino controller with a voltage sensor (voltage-divider based), an ACS712 Hall-effect current sensor, and an LM35/DHT11 temperature sensor, communicating over Wi-Fi/Bluetooth — enabling live sensor data acquisition, CSV-based logging, and real-time validation of the trained models alongside the Kaggle dataset.

3.2 Data Preprocessing and Feature Analysis

Preprocessing involved data cleaning (missing-value imputation via mean/median techniques and statistical outlier detection), Min-Max normalization and standardization to bring all parameters into a comparable numerical range, correlation-based and Random Forest-based feature selection, and an 80:20 train-test split. Correlation analysis (Table 1) identified vehicle speed and battery current as the parameters most strongly associated with energy consumption, with battery voltage and state of charge showing negative correlations — indicating that energy consumption tends to rise as available battery voltage and charge decline. Figure 2 presents the broader correlation structure among all sensor features and the two target variables (energy consumption and range).

Table 1: Correlation of Sensor Parameters with Energy Consumption

Parameter	Correlation Coefficient
Speed	0.91
Current	0.88
Temperature	0.76

International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

Parameter	Correlation Coefficient
Voltage	-0.69
SOC	-0.72

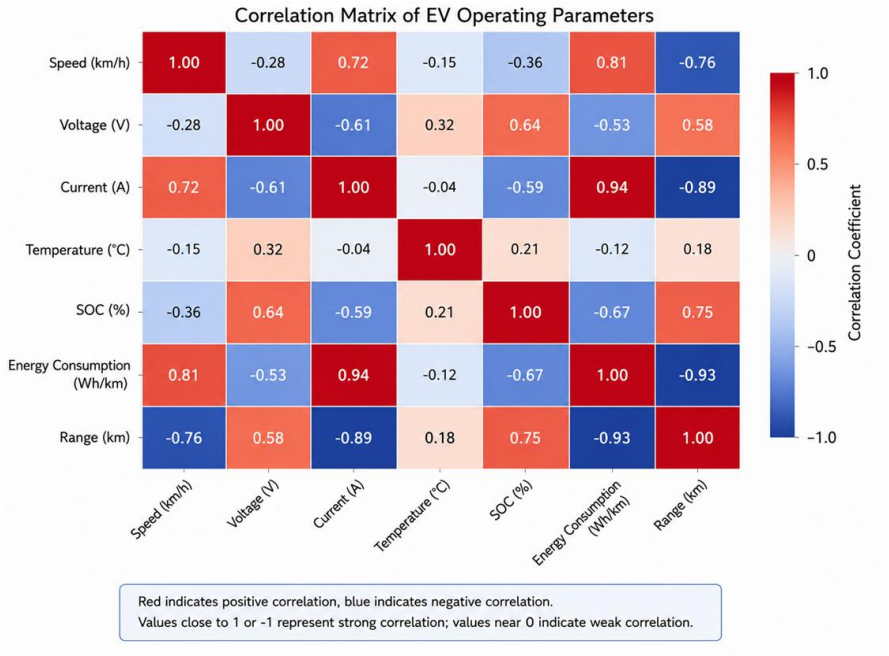


Figure 2: Correlation heatmap of EV operating parameters, energy consumption, and driving range.

3.3 Predictive Models

Linear Regression fits a linear relationship between the sensor inputs and energy consumption via least-squares minimization, serving as an interpretable baseline.

Random Forest Regressor combines an ensemble of decision trees trained on random data and feature subsets, averaging their outputs to reduce overfitting and capture nonlinear interactions among speed, voltage, current, temperature, and SOC.

XGBoost Regressor builds decision trees sequentially, with each tree correcting the residual errors of its predecessors, using regularization and hyperparameter tuning (learning rate, tree depth, number of estimators) to balance accuracy and efficiency.

Long Short-Term Memory (LSTM) processes sequential sensor readings through memory cells and input/forget/output gates (Figure 3), retaining relevant information from previous time steps to capture the temporal dependencies inherent in continuously sampled vehicle sensor data, and was expected to achieve the best performance given the time-series nature of the problem.

International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

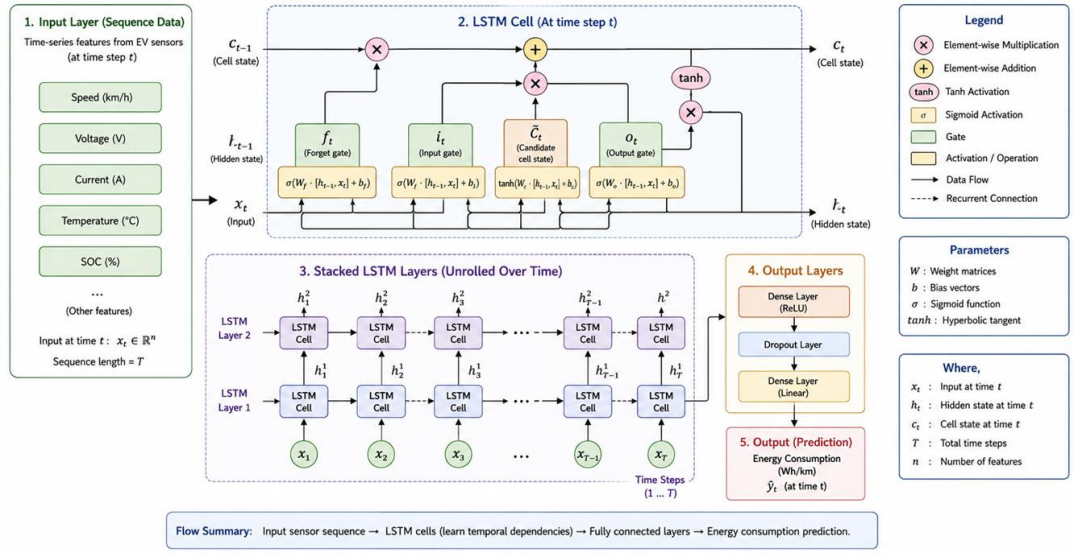


Figure 3: LSTM network architecture used for sequential energy consumption prediction.

All models were trained on the same 80:20 split using speed, voltage, current, temperature, and SOC as inputs, and evaluated using MAE, RMSE, and R^2 Score. Predicted energy consumption was subsequently combined with battery state-of-charge information to estimate the remaining driving range.

4. RESULTS AND DISCUSSION

4.1 Model Performance Comparison

Table 2 summarizes the energy consumption prediction performance of the four models. Linear Regression achieved an MAE of 2.48, RMSE of 3.16, and R^2 Score of 0.88, capturing the overall trend but showing deviations where consumption behaviour became strongly nonlinear. Random Forest improved substantially on this baseline (MAE = 1.52, RMSE = 2.05, R^2 = 0.94) by modelling nonlinear interactions among the sensor parameters through ensemble averaging. XGBoost further reduced error (MAE = 1.18, RMSE = 1.71, R^2 = 0.96) through sequential, error-correcting boosting. The LSTM model achieved the best overall performance, with an MAE of 0.89, RMSE of 1.25, and an R^2 Score of 0.98, reflecting its ability to learn temporal dependencies in the sequential sensor data that the other three models, which treat each observation independently, cannot directly exploit.

Table 2: Comparison of Machine Learning Models for EV Energy Consumption Prediction

Model	MAE	RMSE	R^2 Score
Linear Regression	2.48	3.16	0.88
Random Forest	1.52	2.05	0.94
XGBoost	1.18	1.71	0.96
LSTM	0.89	1.25	0.98

The close agreement between actual and predicted energy consumption for the best-performing LSTM model is illustrated in Figure 4, where predicted points closely track the ideal y = x line across the full range of observed energy consumption values, with only minor scatter.

International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

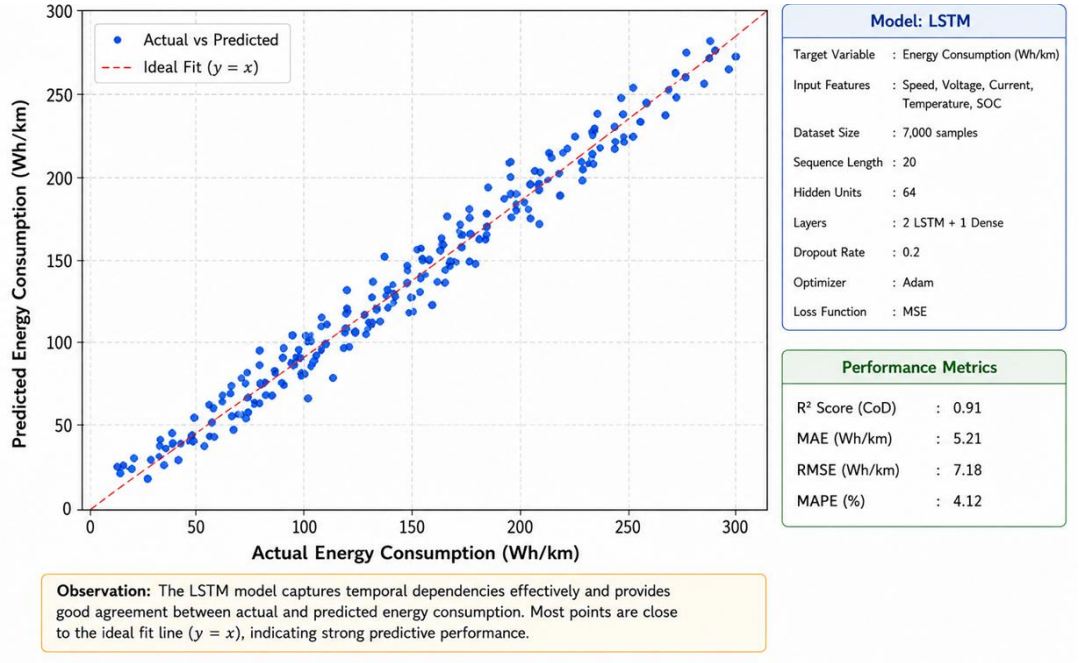


Figure 4: Actual versus predicted energy consumption using the LSTM model on the test dataset.

4.2 Driving Range Prediction

Using predicted energy consumption together with battery SOC information, the framework estimated remaining driving range. Sample results (Table 3) show close agreement between actual and predicted range across a variety of operating conditions, with prediction deviations of only a few kilometres. Figure 5 presents the full actual-versus-predicted range comparison, showing that predicted values closely track the ideal fit line across the entire 0–400 km range tested, confirming that the accuracy achieved at the energy-consumption stage translates effectively into reliable range estimates — the ultimate quantity of practical interest for reducing range anxiety.

Table 3: Sample Range Prediction Results

Actual Range (km)	Predicted Range (km)
180	177
160	163
140	138
120	123
100	98

International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

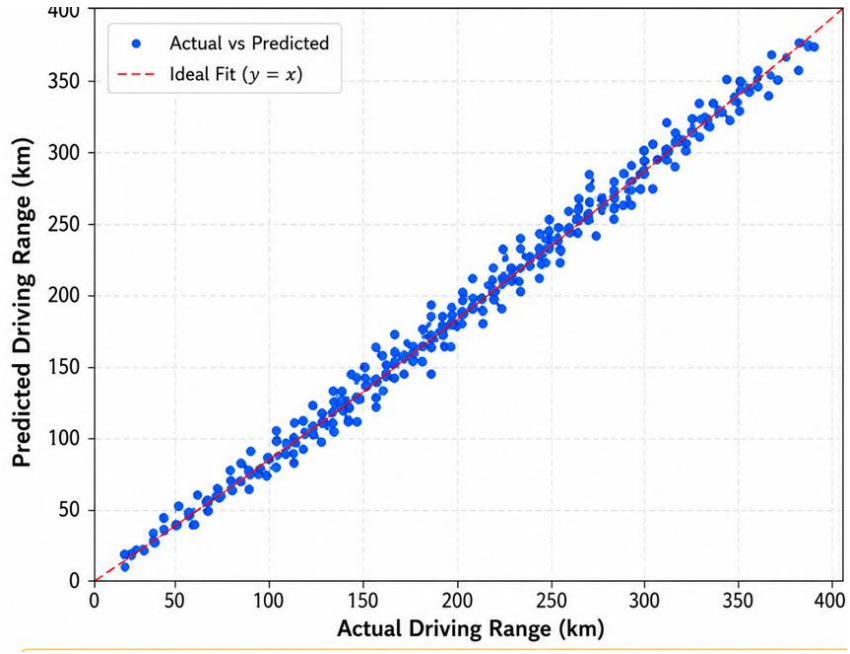
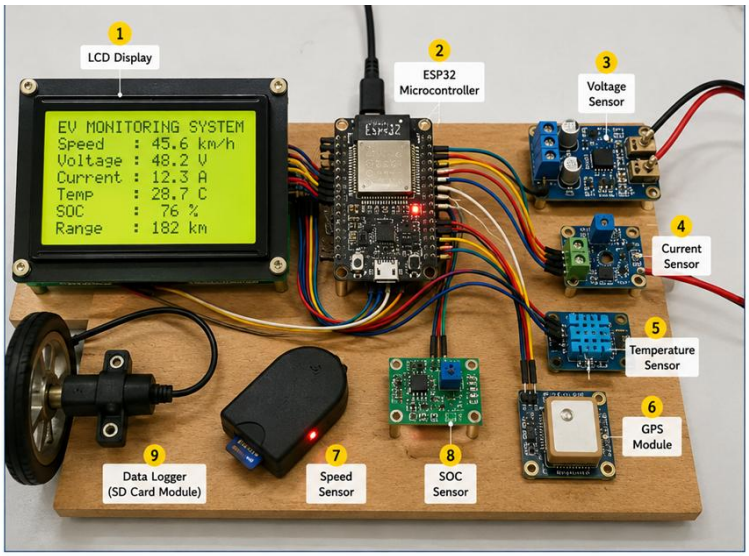


Figure 5: Actual versus predicted driving range across the tested operating range.

4.3 Hardware Validation

The trained models were validated against live data from the ESP32-based hardware platform shown in Figure 6, which integrates an LCD display, voltage sensor (INA219), current sensor (ACS712), temperature sensor (DHT11), GPS module, Hall-effect speed sensor, SOC sensor, and SD-card data logger around the ESP32 microcontroller. The platform continuously monitored battery voltage, current, temperature, speed, and SOC, displaying real-time parameters and predicted range on the onboard LCD while logging data for offline analysis. All sensor channels (voltage, current, temperature, and controller data transmission) remained active and operational throughout testing, confirming the practical feasibility of real-time, hardware-validated EV energy and range prediction rather than a purely simulation-based system.



Components Used

No.	Component	Purpose
1	LCD Display (16x2 I2C)	Displays real-time parameters
2	ESP32 Microcontroller	Data acquisition, processing & Wi-Fi
3	Voltage Sensor (INA219)	Measures battery voltage
4	Current Sensor (ACS712)	Measures battery current
5	Temperature Sensor (DHT11)	Measurre ambient temperature
6	GPS Module (NEO-6M)	Provides location & coordinates
7	Speed Sensor (Hall Effect)	Measures vehicle speed
8	SOC Sensor (Voltage Divider + ADC)	Estimates state of charge (SOC)
9	Data Logger (SD Card Module)	Stores data for logging & analysis

System Specifications

Microcontroller	:	ESP32 DevKit V1
Operating Voltage	:	5V DC
Communication	:	Wi-Fi (2.4 GHz)
Data Logging	:	Micro SD Card (16GB)
Sampling Interval	:	1 second
Power Source	:	EV Battery (48V) with DC-DC Converter (5V)

Figure 6: Hardware sensor monitoring setup used for real-time data acquisition and validation.

4.4 Discussion

The consistent, progressive improvement in prediction accuracy from Linear Regression through Random Forest and XGBoost to LSTM (Table 2) reflects a clear alignment between model complexity and the nonlinear, time-dependent nature of EV energy consumption. Linear Regression's single linear relationship could not capture the interacting effects of speed, current, temperature, and SOC revealed by the correlation analysis (Table 1, Figure 2); the tree-ensemble methods addressed this by modelling nonlinear feature interactions directly, while LSTM's further improvement reflects its specific ability to retain relevant historical context — such as a recent acceleration event or a gradually declining SOC — that precedes and predicts subsequent energy demand, rather than reacting only to the instantaneous sensor snapshot.

Because speed and current are the two most strongly correlated predictors of energy consumption (Table 1), and both vary continuously and rapidly during real driving, the temporal-sequence modelling capability of LSTM is particularly well matched to this prediction task, which explains its consistent advantage across the energy-consumption metrics (Table 2) and its translation into equally accurate range predictions (Table 3, Figure 5). The successful hardware validation (Figure 6) further demonstrates that this accuracy is not confined to offline analysis of the Kaggle dataset: live sensor data from a physical ESP32-based platform confirms the practical deployability of the proposed framework. Taken together, these results support the use of dataset-driven, sensor-validated machine learning — and LSTM specifically — as a practical foundation for reducing range anxiety and improving energy management in electric vehicles.

5. CONCLUSION

This study presented a dataset-driven machine learning framework for modelling EV energy consumption and predicting driving range, combining the Kaggle Electric Vehicle Energy Consumption Dataset with a physical ESP32/Arduino-based sensor platform for real-time hardware validation. Four models — Linear Regression, Random Forest, XGBoost, and LSTM — were trained and compared, with LSTM achieving the best performance (MAE = 0.89, RMSE = 1.25, R^2 = 0.98), outperforming XGBoost (R^2 = 0.96), Random Forest (R^2 = 0.94), and Linear Regression (R^2 = 0.88), owing to its ability to capture temporal dependencies in sequential sensor data. Correlation analysis identified vehicle speed and battery current as the dominant factors influencing energy consumption, while range prediction results showed close agreement between actual and predicted values across the tested operating range. The successful integration and validation of hardware sensors confirmed the practical feasibility of the proposed system beyond simulation, demonstrating that combining machine learning with real-time sensor monitoring can meaningfully enhance EV energy management, improve range prediction accuracy, reduce range anxiety, and support the efficient, sustainable operation of electric vehicles.

Future Scope

Future work may explore advanced deep learning architectures such as Gated Recurrent Units (GRU), Transformer networks, hybrid CNN-LSTM models, and attention-based mechanisms to further improve prediction accuracy. Incorporating additional parameters — road gradient and elevation, traffic congestion, weather conditions, tyre pressure, and vehicle payload — could improve real-world applicability, while integration with GPS and navigation systems would enable route-aware range prediction. Broader Internet of Things (IoT) integration, connecting the hardware platform to cloud-based fleet monitoring and over-the-air model updates, represents a further promising direction for scaling the proposed framework to real-world EV deployments.

ACKNOWLEDGMENT

I am greatly indebted forever to my guide Prof. Seema N. Kharat and HOD K. Chandra Obula Reddy and to all teaching and non teaching staff who supported me directly and indirectly to complete my work. I am sincerely thankful to my principal Dr. S. K. Biradar for their continued encouragement and active interest in my progress throughout the work. I am grateful to be an M.Tech Electrical Power System student at Matsyodari Shikshan Sanstha 's College of Engineering and Technology, Jalna, Maharashtra.

REFERENCES

[1] I. Hussain, K. B. Ching, C. Uttraphan, K. G. Tay, and A. Noor, "Evaluating machine learning algorithms for energy consumption prediction in electric vehicles: A comparative study," Sci. Rep., vol. 15, no. 1, pp. 1–20, 2025, doi: 10.1038/s41598-025-94946-7.

[2] I. Ullah, K. Liu, T. Yamamoto, R. E. Al Mamlook, and A. Jamal, "A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability," Energy Environ., vol. 33, no. 8, pp. 1583–1612, Dec. 2022, doi: 10.1177/0958305X211044998.

[3] I. Ullah, K. Liu, T. Yamamoto, M. Zahid, and A. Jamal, "Electric vehicle energy consumption prediction using stacked generalization: an ensemble learning approach," Int. J. Green Energy, vol. 18, no. 9, pp. 896–909, Jul. 2021, doi: 10.1080/15435075.2021.1881902.

[4] K. Ko, T. Lee, and S. Jeong, "A Deep Learning Method for Monitoring Vehicle Energy Consumption with GPS Data," 2021, doi: 10.3390/su132011331.

[5] Y.-W. Chung, B. Khaki, T. Li, C. Chu, and R. Gadh, "Ensemble machine learning-based algorithm for electric vehicle user behavior prediction," Appl. Energy, vol. 254, p. 113732, 2019, doi: 10.1016/j.apenergy.2019.113732.

[6] W. Achariyaviriya et al., "Estimating Energy Consumption of Battery Electric Vehicles Using Vehicle Sensor Data and Machine Learning Approaches," 2023, doi: 10.3390/en16176351.

[7] J. Zhang, Z. Wang, P. Liu, and Z. Zhang, "Energy consumption analysis and prediction of electric vehicles based on real-world driving data," Appl. Energy, vol. 275, p. 115408, 2020, doi: 10.1016/j.apenergy.2020.115408.

[8] H. P. Bhupathi and S. Chinta, "Smart Charging Revolution: AI and ML Strategies for Efficient EV Battery Use," ESP J. Eng. Technol. Adv., vol. 2, no. 2, pp. 154–167, 2022, doi: 10.56472/25832646/JETA-V2I2P117.

[9] G. Ozdemir, M. Kuzlu, and F. O. Catak, "Machine learning insights into forecasting solar power generation with explainable AI," Electr. Eng., vol. 107, no. 6, pp. 7329–7350, 2025, doi: 10.1007/s00202-024-02933-4.

[10] S. Data, "AI-Driven Predictive Modeling for Solar Power Generation Using Real-Time Photovoltaic Sensor Data," pp. 10–34, 2026, doi: 10.32996/jcsts.

[11] Abdussalam Ali Ahmed, "Hybrid AI Models for Forecasting and Optimizing Solar Energy Generation Under Varying Weather Conditions," Sci. J. Publ. Heal. Res. Technol., vol. 1, no. 1, pp. 35–41, 2025, doi: 10.65420/sjphrt.v1i1.7.

[12] E. Gul, G. Baldinelli, J. Wang, P. Bartocci, and T. Shamim, "Artificial intelligence based forecasting and optimization model for concentrated solar power system with thermal energy storage," Appl. Energy, vol. 382, p. 125210, 2025, doi: 10.1016/j.apenergy.2024.125210.

[13] P. Bouquet, I. Jackson, M. Nick, and A. Kaboli, "AI-based forecasting for optimised solar energy management and smart grid efficiency," Int. J. Prod. Res., vol. 62, no. 13, pp. 4623–4644, Jul. 2024, doi: 10.1080/00207543.2023.2269565.

[14] B. Mehmood et al., "Development Of A Hybrid Artificial Intelligence Framework For Accurate Forecasting Of Solar Power Generation Using Machine Learning Algorithms And Timeseries Analysis," Spectr. Eng. Sci., vol. 3, no. 5, pp. 613–636, 2025.

[15] Y. Patil, T. Shruti, and R. K., "Time-Series Forecasting Using ARIMA and SARIMA Models for Solar NASA POWER Data," in 2025 3rd Int. Conf. on Intelligent Systems, Advanced Computing and Communication (ISACC), 2025, pp. 946–952, doi: 10.1109/ISACC65211.2025.10969243.

Authors Biography

Author 1:

Mr. Omkar Shivaji Wadaje is pursuing M.TECH degree in stream of Electrical Power System from Mss's college of Engineering and technology, Jalna, DR BATU University, Lonere. He was completed his B.Tech in the stream of Electrical Engineering from Maharashtra Institute of Technology Aurangabad.

Email Id: osw.wadaje.24@gmail.com

Author2



Prof. Seema N. Kharat received Master of engineering in Electrical Power system from Matsyodari Shikshan Sanstha ‘s College of Engineering and Technology, Jalna. She is currently working as Assistant Professor in Department of Electrical Engineering at MSS's college of Engineering & Technology , Jalna.
Email ID. - kharatseema.5@gmail.com