



# Electric Vehicle Customer Churn Prediction

Vigneshwari V<sup>1</sup>, Gopikrishna S V<sup>2</sup>, Ms. M. Preethi<sup>3</sup>

Student, Department of Computer Science with Data Analytics, Dr N.G.P. Arts and Science Collage, Coimbatore,  
Tamil Nadu, India<sup>1</sup>

Student, Department of Computer Science with Data Analytics, Dr N.G.P. Arts and Science Collage, Coimbatore,  
Tamil Nadu, India<sup>2</sup>

Assistant Professor, Department of Computer Science with Data Analytics, Dr N.G.P. Arts and Science College,  
Coimbatore, Tamil Nadu, India<sup>3</sup>

**ABSTRACT:** The rapid expansion of electric vehicle (EV) adoption has increased competition among service providers, making customer retention an essential operational priority. Unlike traditional automotive markets, EV ecosystems generate continuous interaction data through charging behavior, subscription billing, maintenance activity, and digital platform usage. These behavioral signals provide valuable insight into customer engagement and potential service discontinuation. This study proposes a machine learning-based framework for predicting EV customer churn and supporting proactive retention strategies. Operational indicators such as charging frequency, service interactions, payment reliability, and vehicle activity recency are transformed into meaningful behavioral features through structured feature engineering. Logistic Regression and Random Forest models are developed and comparatively evaluated, with Random Forest demonstrating stronger predictive performance across accuracy, recall, and ROC-AUC metrics.

To improve real-world usability, the trained model is integrated into an interactive Streamlit dashboard that enables risk segmentation and intervention simulation. Organizations can analyze churn probability and evaluate retention strategies before implementation. The proposed framework demonstrates how predictive analytics can be combined with decision support to enhance customer lifecycle management within evolving EV service environments.

**KEYWORDS:** Electric Vehicles, Customer Churn Prediction, Machine Learning, Risk Segmentation, Predictive Analytics, Customer Retention

## I. INTRODUCTION

The global shift toward electric mobility has accelerated in recent years due to environmental concerns, technological innovation, and supportive governmental policies. As adoption increases, EV companies are transitioning from a growth phase focused on customer acquisition toward sustaining long-term engagement and service continuity. EV ownership differs significantly from conventional vehicle usage because customers interact regularly with digital platforms, charging infrastructure, and subscription-based services.

Customer churn presents a major financial challenge, as acquiring new users often requires significantly higher investment than retaining existing customers. Early identification of disengagement enables organizations to intervene before customers discontinue services. Traditional monitoring approaches rely mainly on descriptive statistics and manual analysis, which are limited in their ability to capture evolving behavioral patterns.

Machine learning techniques provide an effective alternative by learning relationships between operational usage patterns and churn outcomes. By analyzing historical interaction data, predictive models can detect subtle changes in customer behavior that indicate declining engagement.

This research introduces a churn prediction framework specifically designed for EV operational datasets. The study combines behavioral feature engineering, supervised learning models, and interactive visualization to support proactive decision making.

The main contributions of this work are:

1. Development of a behavioral churn scoring mechanism using EV operational metrics,
2. Integration of temporal and efficiency-based feature engineering,
3. Comparative evaluation of predictive machine learning models, and
4. Deployment of an interactive dashboard enabling risk interpretation and intervention planning.

## II. RELATED WORK

Customer churn prediction has been widely explored in domains such as telecommunications, banking, and subscription-based services. Early approaches relied on rule-based systems and statistical thresholds derived from historical cancellation data. Although these methods provided initial insights, they lacked adaptability to changing customer behavior.

Recent studies increasingly apply machine learning algorithms capable of modeling nonlinear relationships between customer attributes and churn outcomes. Logistic Regression remains popular because of its interpretability and computational efficiency, while ensemble approaches such as Random Forest improve prediction accuracy through aggregation of multiple decision trees. Prior research highlights the effectiveness of ensemble learning in handling complex and heterogeneous datasets.

Behavior-driven analysis has gained importance with the availability of large operational datasets. Variables such as service usage frequency, payment history, and engagement intensity are commonly used as predictive indicators. However, relatively limited research focuses on churn prediction in EV ecosystems, where operational characteristics differ substantially from traditional industries.

Existing solutions frequently emphasize prediction accuracy without addressing business usability. Modern retention systems require analytical outputs that support operational decisions. The present study extends prior work by integrating predictive modeling with risk segmentation and intervention simulation tailored to EV service environments.

## III. PROPOSED METHODOLOGY

The proposed system is designed as an end-to-end analytical pipeline that converts raw EV operational data into actionable retention insights. The framework includes feature engineering, data preprocessing, machine learning modeling, and intervention analysis. Unlike conventional churn prediction studies that focus solely on classification performance, the proposed framework integrates predictive modeling with risk segmentation and intervention simulation, enabling practical decision support for EV service management.

### A. Feature Engineering

Feature engineering transforms raw operational records into indicators that better represent customer engagement. Important attributes include average travel distance, charging frequency, billing behavior, and maintenance history. Temporal features such as customer tenure, inactivity duration, and recency of vehicle activity are derived from timestamp information to capture behavioral trends over time.

Additional calculated metrics—including cost per kilometer, charging efficiency, and service intensity—provide contextual understanding of how customers interact with EV services. These features help identify gradual reductions in engagement that may not be visible through raw data alone.

### B. Data Preprocessing

Data preprocessing ensures dataset consistency before model training. Missing values and duplicate records are handled to improve reliability, and a chronological train-test split is applied to reflect real-world prediction scenarios. Categorical attributes such as vehicle model and subscription plan are encoded numerically, while numerical variables are standardized to maintain balanced learning behavior.

A rule-based churn labeling strategy is introduced using behavioral thresholds such as inactivity periods, payment failures, and reduced usage frequency.

### C. Machine Learning Models

Two supervised learning algorithms are implemented for comparative analysis. Logistic Regression is used as a baseline model due to its interpretability and simplicity. Random Forest is selected as the primary model because it captures nonlinear relationships and feature interactions effectively.

Model performance is evaluated using accuracy, precision, recall, and ROC-AUC metrics. Experimental results show that Random Forest provides stronger predictive capability, particularly in identifying high-risk customers with fewer false negatives.

### D. Risk Segmentation and Intervention Logic

Predicted churn probabilities are categorized into Low, Medium, and High risk groups to simplify interpretation for business users. Based on dominant risk factors, intervention strategies such as financial assistance, prioritized service scheduling, and engagement incentives are mapped to each category. This approach links analytical predictions with actionable retention planning.

## IV. EXPERIMENTAL SETUP AND RESULTS

### A. Dataset

The proposed system is evaluated using a structured Electric Vehicle (EV) operational dataset containing customer behavioral and service interaction records. The dataset includes attributes such as vehicle usage patterns, charging frequency, subscription details, billing history, maintenance visits, and customer activity timelines. These variables collectively represent customer engagement within EV service ecosystems.

The dataset represents a realistic EV operational scenario generated from structured behavioral records designed to simulate real service interactions.

To simulate real-world deployment conditions, the dataset is divided chronologically into training and testing subsets. Approximately 80% of the records are used for model training, while the remaining 20% are reserved for testing. Temporal separation ensures that the model learns from past behavioral patterns and predicts churn risk for future customers, thereby avoiding information leakage and improving practical reliability.

### B. Evaluation Metrics

System performance is evaluated using multiple classification metrics, including accuracy, precision, recall, F1score, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC).

Accuracy measures the overall correctness of predictions across both churn and non-churn classes. Precision evaluates the reliability of churn predictions by measuring how many predicted churn cases are truly churn customers. Recall represents the model's ability to correctly identify customers who are actually at risk of leaving the service. The F1-score provides a balanced measure by combining precision and recall. ROC-AUC assesses the model's capability to distinguish between churn and active customers across varying probability thresholds.

These metrics collectively ensure balanced evaluation, particularly because churn datasets often exhibit class imbalance.

### C. Performance Analysis

Table I presents the comparative performance of the implemented machine learning models for EV customer churn prediction. Logistic Regression provides stable baseline performance with strong recall but relatively lower precision due to overlapping behavioral patterns. The Random Forest classifier significantly improves predictive capability by capturing nonlinear relationships among operational features.

**Table I - PERFORMANCE COMPARISON OF CHURN PREDICTION MODELS**

Model	Accuracy (%)	Precision (%)	Recall (%)	ROC-AUC
Logistic Regression	85.0	35.8	86.0	0.931
<b>Random Forest</b>	<b>94.5</b>	<b>62.2</b>	<b>98.2</b>	<b>0.999</b>

Comparative analysis indicates that ensemble-based learning improves discrimination between active and disengaged customers by learning complex feature interactions such as activity recency, payment reliability, and service frequency. Feature importance analysis further confirms that behavioral engagement indicators contribute more strongly to prediction performance than static attributes.

#### **D. Deployment and Real-Time Performance**

For practical usability, the trained Random Forest model is integrated into an interactive Streamlit dashboard. The system enables real-time churn risk prediction and customer-level analysis through dynamic visualization modules. Risk segmentation categorizes customers into Low, Medium, and High risk groups, allowing decision makers to prioritize retention strategies.

Prediction time for individual customer records is minimal, enabling near real-time analysis during dashboard interaction. Intervention simulation recalculates churn probability instantly after applying selected retention strategies, supporting rapid evaluation of business decisions. The deployment demonstrates that the proposed framework is suitable for operational environments requiring continuous monitoring and proactive customer retention planning.

#### **V. DISCUSSION**

Results indicate that EV customer churn is strongly associated with behavioral engagement rather than demographic characteristics alone. Declining vehicle usage and repeated billing issues consistently emerge as early indicators of potential disengagement. Feature engineering contributes significantly by embedding operational context into the predictive process.

The integration of predictive analytics with visualization tools enhances practical usability. Decision makers can explore intervention outcomes before implementation, reducing uncertainty in retention planning. Nevertheless, evolving customer expectations and infrastructure availability may influence long-term model performance.

#### **VI. CONCLUSION**

This study presents a machine learning-driven framework for predicting customer churn in electric vehicle service ecosystems. Operational usage data are transformed into behavioral indicators through structured feature engineering and preprocessing techniques.

Comparative evaluation shows that the Random Forest model achieves superior performance in identifying high-risk customers. Deployment through an interactive dashboard demonstrates how predictive analytics can support real-time decision making and proactive retention strategies.

By combining prediction, visualization, and intervention simulation, the proposed system moves churn analysis beyond passive reporting toward actionable business intelligence. Future work may incorporate real-time data integration, explainable AI methods, and scalable deployment architectures to support continuously evolving EV environments.

The presented framework highlights how machine learning-driven behavioral analytics can evolve from predictive modeling into an operational decision-support tool, enabling data-informed customer retention strategies in emerging electric mobility ecosystems.

#### **VII. ACKNOWLEDGMENT**

The authors would like to express their sincere gratitude to the faculty members and mentors for their continuous guidance, encouragement, and valuable suggestions throughout the development of this research work. Special appreciation is extended to the project supervisor for providing consistent technical support and constructive feedback during each stage of the study.

The authors also acknowledge the support received in terms of academic guidance and computational resources, which contributed significantly to the successful design, implementation, and evaluation of the Electric Vehicle Customer Churn Prediction system.

**REFERENCES**

- [1] T. Verbeke, D. Martens, C. Mues, and B. Baesens, “Building Comprehensible Customer Churn Prediction Models,” *Expert Systems with Applications*, vol. 38, no. 3, pp. 2354–2364, 2011.
- [2] J. Huang, T. Kechadi, and B. Buckley, “Customer Churn Prediction in Telecommunications,” *Expert Systems*, vol. 29, no. 5, pp. 447–454, 2012.
- [3] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [4] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Morgan Kaufmann, 2016.
- [5] F. Provost and T. Fawcett, *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. O’Reilly Media, 2013.
- [6] S. Burez and D. Van den Poel, “Handling Class Imbalance in Customer Churn Prediction,” *Expert Systems with Applications*, vol. 36, no. 3, pp. 4626–4636, 2009.
- [7] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, Springer, 2013.
- [8] Streamlit Inc., “Streamlit Documentation: Building Interactive Machine Learning Applications,” 2024. [Online]. Available: <https://streamlit.io>



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details