Smart EV Charging: Forecasting and Allocation

Harshit Pathak
Department of Computer Science and Engineering
B.M. Institute of Engineering and Technology
Haryana, India
harshitpathak.pro@gmail.com

Tarun Attri
Department of Computer Science and Engineering
B.M. Institute of Engineering and Technology
Haryana, India
attriitarun@gmail.com

Saiyam Jain
Department of Computer Science and Engineering
B.M. Institute of Engineering and Technology Haryana, India
Veersaiyam5@gmail.com

Rajdeep Choudhury
Department of Computer Science and Engineering
B.M. Institute of Engineering and Technology
Haryana, India
choudhuryrajdeep975@gmail.com

Paramjeet Ruhal
Assistant Professor
Department of Computer Science and Engineering
B.M. Institute of Engineering and Technology
Haryana, India
paramjeetruhal@gmail.com

Babita Antil
Associate Professor
Department of Computer Science and Engineering
B.M. Institute of Engineering and Technology
Haryana, India
babita.antil@gmail.com

Abstract — The growing adoption of electric vehicles (EVs) demands smarter charging infrastructure planning. This research presents an AI-powered system for forecasting EV charging demand and recommending optimal station allocation. Using large-scale datasets, time-series models (LSTM, Prophet), and clustering techniques, the system predicts future charging needs and identifies high-demand regions. An interactive dashboard provides stakeholders with real-time insights for infrastructure expansion and offers EV drivers personalized station recommendations based on location, availability, and wait time. The proposed system is data-driven, scalable, and suitable for real-world deployment to accelerate sustainable EV adoption.

Keywords — Electric Vehicles, Charging Demand Forecasting, Smart Station Allocation, Machine Learning, Sustainable Transportation

I. INTRODUCTION

Rapid adoption of electric vehicles (EVs) is reshaping global transportation due to the need to reduce carbon emissions and dependence on fossil fuels. However, large-scale EV adoption introduces challenges in maintaining a reliable, well-distributed, and intelligent charging infrastructure. Key issues include the lack of predictive tools to estimate station availability, range anxiety, long waiting times, fragmented charging network data, and rising demand. Traditional dashboards and monitoring systems mainly rely on historical data and therefore provide limited support for future decision-making.

To address these challenges, this research proposes allocating AI-driven EV charging demand forecasting and smart station, a system that takes advantage of machine learning, predictive analytics and clustering techniques that adapt to demanding demand and deployment of infrastructure. Open Charge Map and U.S. Using large-scale datasets from sources such as DOE AFDC, the system applies advanced models (LSTM, Prophet, XGBoost) to the time-range forecasts and clustering methods (K-Means, DBSCAN) to identify high-demand areas.

Beyond the prognosis, the platform provides real-time, working insight through an interactive AI-operated dashboard, enabling data-operated infrastructure to make decisions and city planners to make decisions, the optimal charging stations for drivers based on location, availability and waiting time. Scalable, intelligent and user-focused, proposed system bridges the gap between current boundaries and future needs, accelerating the transition towards a clever, more durable EV ecosystem.

II. LITERATURE REVIEW

2. Literature Review and Existing Solutions

Research on customized electric vehicles (EV) charging has progressed in several domains, which have developed from basic statistical trend analysis, developed for demand forecast and grid management from AI-operated solutions. The literature review is organized into four key categories relevant to the proposed system: EV charging data analysis, demand forecasting models, geospatial and clustering techniques, and recommendation or adaptation-based systems.

2.1 EV Charging Data Analysis and Infrastructure Assessment

Preliminary studies were mainly focused on descriptive figures of charging sessions to understand the original use pattern. Research often analyzes variables like
To mark the duration of the session, energy consumption (KWH), and connector type charge behavior.
Early infrastructure evaluation depends on identification of major metrics
Station usage rate and peak hours. However, traditional analysis of fragmented data in many networks offers challenges, making integrated trend analysis and hard forecast. Lack of standardization in data formats also complicates comparative analysis in various fields.

2.2 Time-Series Demand Forecasting Models

Accurate prediction of future charging demand is essential for proactive planning and maintaining grid stability. Early approaches relied on basic statistical models such as ARIMA for time-series forecasting. More advanced studies have adopted specialized models capable of capturing complex, non-linear patterns and seasonal variations.

- **Prophet:** developed by Facebook, this model is effective for predicting time-series data with strong seasonal components, which is highly relevant to daily and weekly charging cycles.
- **Deep Learning Model (LSTM):** Long short-term memory (LSTM) network, a type of recurrent nervous network (RNN), rapid energy load and long-term temporary dependence in the use pattern are used to catch long-term temporary dependence, which offers better performances for complex forecasting tasks such as station-level demand.
- **XGBoost:** A gradient-boosting framework such as XGBoost is also applied to predict the demand of the station by integrating various characteristics such as days, weather and external phenomena.

2.3 Geospatial Analysis and Clustering for Optimal Allocation

It is as important to identify where to keep new charging stations as they need. Geospatial techniques are employed to map the existing infrastructure and highlight areas of high demand. Like equipment GeoPandas and Folium support geospatial visualization of charging hotspots. To advance beyond basic mapping, the system employs unsupervised learning algorithms such as:

- **Clustering (K-Means, DBSCAN):** These techniques effectively identify demand clusters and underserved areas based on usage intensity and similarity in charging behavior (Bui et al., 2023). This clustering provides a data-driven foundation for infrastructure expansion and optimal station allocation.

2.4 Optimization and Recommendation Systems

To enhance the user experience and maximize efficiency, research has shifted toward prescriptive systems. While basic navigation systems recommend the nearest station, advanced frameworks incorporate multiple criteria for a true optimization model [8]:

- **Multi-Criteria Recommendation:** Systems are being developed to suggest optimal stations by integrating factors beyond distance, including predicted occupancy/wait time, cost, connector type/speed, and real-time availability.
- **Load Balancing:** Optimization models are essential for introducing dynamic pricing and routing drivers to less-congested stations, thereby balancing demand and reducing strain on the power grid.
- **Smart Grid and Renewable Integration:** An emerging trend is the integration of renewable energy inputs (solar, wind) with charging demand forecasts to promote sustainable charging and grid load management.

2.5 Research Gap and Positioning of Proposed System

From literature, the existing solutions are often limited: they either make basic descriptive analysis, focus on forecast completely without optimizing the allocation, or offer simple recommendations, which lacks the future waiting-time. The primary difference is the lack of an integrated, AI-manual decision-support platform that integrates future modeling with real-time user-inserts. The current ecosystem is prone to fragmented data, dependence on non-transparent pricing and the absence of active planning devices.

The proposed project, the AI-managed EV charging demand forecasting and allocation of smart stations, pulls these intervals directly:

1. Integrated data platform: integrating diverse global data sets and standardization format for comprehensive analysis.
2. Advanced future modeling: Demand for forecasting and geopolitical hotspot identity for both LSTM, Prophet, and Clustering.
3. Real-time, multi-stakeholder intelligence: providing actionable insight through policy makers (expansion reports), operators (load management), and drivers (smart station recommendations).
4. Stability integration: Including renewable energy and grid data for environmental impact evaluation.

This hybrid approach represents a step towards intelligent, connected and scalable EV infrastructure management, which addresses both a high-level plan and daily user feature.

2.6 Recent Studies (2021-2025)

Recent studies (2021–2025) have advanced EV charging demand forecasting and station management using machine learning. For example, Liu et al. (2021) employed an LSTM-based model for short-term charging load prediction at the station level. Chen and Zhang (2022) integrated queueing theory with deep learning for wait-time estimation. Park et al. (2023) proposed a price- based load balancing system to reduce charging congestion. Wang et al. (2024) utilized hybrid Prophet–XGBoost models for spatial-temporal EV demand forecasting.

Our work differs by combining station-level time-series forecasting, clustering, and a prescriptive recommendation engine within a unified pipeline. Additionally, we emphasize operational evaluation through simulation and MLOps deployment, offering a comprehensive and reproducible framework.

III. PROPOSED METHODOLOGY

The proposed system employs a robust, multi-stage data science and engineering methodology, structured to ensure accuracy, scalability, and real-time responsiveness for all stakeholders.

3.1 Data Acquisition and Preprocessing Pipeline

The foundation of the system is the consolidation of diverse, fragmented data sources into a unified platform.

- **Data Collection:** Charging session data is collected from global and regional APIs and public portals, including Open Charge Map, U.S. DOE AFDC, and European providers (ElaadNL, EVA). Auxiliary data, such as geospatial coordinates, EV model specifications, and real-time station availability, are also ingested.
- **Data Cleaning and Feature Engineering:** Raw data is processed to handle missing values, correct inconsistencies, and normalize formats across different providers. Key features are then extracted, including session frequency, energy consumed (kWh), station utilization rates, and time-of-day/day-of-week indicators for seasonality.
- **Geospatial Integration:** Latitude and longitude coordinates are integrated with usage metrics, enabling spatial querying and the identification of high-demand zones using tools like GeoPandas and Folium.

3.2 Advanced Analytical and Predictive Modelling

The system utilizes a hybrid approach combining time-series forecasting and unsupervised learning to generate predictive insights.

i. Time-Series Demand Forecasting

Forecasting models are applied to predict station-level energy load and usage patterns to anticipate demand surges and grid requirements:

- **LSTM Networks:** Long Short-Term Memory (LSTM) models, a deep learning technique, are used to capture complex, non-linear dependencies and long-term temporal patterns in the charging time-series data [10]. This is particularly effective for predicting volatile, station-level demand.
- **Prophet:** The Prophet model is implemented in parallel to robustly model and forecast trends, seasonality (daily, weekly, yearly), and the impact of holiday effects, leveraging its specialized structure for structured time-series components [1].
- **XGBoost:** Extreme Gradient Boosting (XGBoost) is used as a powerful ensemble method to integrate various features (e.g., weather data, energy prices) for refined forecast accuracy [4].
- **Model Validation:** Forecast accuracy is continuously monitored against baseline models (like ARIMA) and real-world usage to ensure reliability.

ii. Clustering and Geospatial Hotspot Detection

Unsupervised learning is applied to categorize stations and users based on their charging behaviour, providing structural insights for optimal allocation:

- **K-Means and DBSCAN Clustering:** These algorithms group stations based on metrics like average session duration, energy consumption, and peak usage timing [3]. The resulting clusters define distinct demand profiles (e.g., residential vs. commercial hubs), which informs targeted infrastructure expansion.
- **Hotspot Mapping:** Geospatial analysis is used to visualize high-density usage areas in real-time and predict future hotspots based on demographic and urban planning data.

3.3 Prescriptive and Recommendation Systems

Moving from prediction to action, the system implements multi-criteria optimization algorithms to provide actionable intelligence for both operators and end-users.

i. Smart Station Recommendation (Driver View)

For EV drivers, the system generates real-time station recommendations based on a multi-objective function that minimizes factors such as [7]:

- Distance and Route Proximity
- Predicted Wait Time (Occupancy): Calculated using real-time availability and forecast demand.
- Charging Cost and Connector Type
- Energy Source: Exploring the availability of renewable inputs.

ii. Load Balancing and Optimization (Operator View)

For charging operators and energy providers, the system provides capacity alerts and suggests load-balancing strategies, including the potential for dynamic pricing using Reinforcement Learning (RL) models (as proposed in future scope) to smooth peak demand curves and reduce grid strain.

3.4 System Integration and Deployment (MLOps)

The complete system is integrated through a tiered architecture and deployed using MLOps principles:

- **Architecture:** A microservices architecture separates the data ingestion, forecasting engine, and recommendation engine.
- **Dashboard Development:** An interactive AI-driven dashboard (using frameworks like Plotly Dash or Streamlit) serves as the primary interface, providing role-based views for drivers, operators, and policymakers.
- **Deployment and Monitoring:** Models are served via REST APIs for near-real-time performance. Monitoring pipelines track data drift, forecast accuracy, and recommendation usage to ensure continuous adaptation and scheduled retraining, guaranteeing the system's long-term operational feasibility and reliability.

IV. DATASET AND EXPERIMENTAL SETUP

- We used publicly available electric vehicle (EV) charging session data collected from the Open Charge Map API and the U.S. Department of Energy Alternative Fuels Data Center (AFDC). The dataset contains information about charging sessions recorded between January 2019 and December 2024 across major cities across California.
- Each record includes timestamp, station ID, geographic coordinates (latitude and longitude), connector type, session duration, and energy consumed (kWh). Additional contextual data such as weather, holidays, and population density were merged to improve demand prediction accuracy.
- Preprocessing: Duplicate and incomplete records were removed. Missing values for energy and duration were filled using median imputation. The session data was aggregated into daily station-level time series, capturing the number of charging sessions, average kWh consumed, and average session duration per station.
- Train/Test Split: The dataset was split chronologically — 70% for training, 15% for validation, and 15% for testing. The forecasting models were evaluated using walk-forward validation to account for temporal dependency.
- Evaluation Metrics: To evaluate forecasting performance, we used Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and R^2-score.
- Baseline Models: Persistence model (previous-day value), ARIMA, and Prophet were used as baselines. Advanced models included LSTM and XGBoost. Hyperparameters were tuned using grid search on the validation set.

V. WAIT-TIME MODELLING AND STATION RECOMMENDATION

- To assess operational efficiency, each charging station was modeled as an M/M/c queue [5], where 'c' represents the number of chargers available at that station. The arrival rate (λ) was estimated from historical session frequency, and the service rate (μ) from the average session duration.

- The expected waiting time was then derived using standard queueing theory equations.
- The system recommends the optimal charging station for drivers using a multi-objective score function:

$$S = w^1 T_{ww} + w_2\, D + w_3\, C$$

$$w_1 + w_2 + w_3 = 1, \qquad w_i \geq 0$$

$$S = w^{\mathrm{T}}\, x$$

$$x = [\; T_{ww}\; D\; C \;]^{\mathrm{T}}$$

$$w = [\; w_1\; w_2\; w_3 \;]^{\mathrm{T}}$$

The system recommends the station for which S is minimized, where w_1, w_2, $w_3$are empirically chosen weights.

The recommendation aims to **minimize driver wait time and detour distance** while balancing station utilization.
A **discrete-event simulator** was implemented to replay actual arrival timestamps and evaluate average waiting times, service rates, and utilization levels under varying demand conditions.

VI. SYSTEM DESIGN AND IMPLEMENTATION

The proposed EV charging system is designed as a scalable, multi-layered framework that integrates large-scale data processing, advanced machine learning models, and real-time visualization tools to create a comprehensive decision-support and user-optimization platform. The system's design is conceptually organized into Frontend, Backend, Database, and AI/ML layers.

4.1 Hardware and Infrastructure Requirements

The system's feasibility is supported by readily available hardware and scalable cloud infrastructure.

- **Computer/Server:** The system requires a multi-core processor (CPU≥2.5 GHz) and a minimum of 8GB RAM, with 16GB recommended for ML training. A GPU is recommended for deep learning models like LSTM. Sufficient storage (SSD recommended, 100GB+) is needed depending on the dataset size.
- **Network and Cloud:** A stable internet connection is necessary for API data collection and dashboard hosting. Cloud server or database support (AWS, Google Cloud, or Azure) is critical for scalability. Optional GPU cloud services can be utilized for faster ML model training.
- **Databases:** The system uses a Relational Database Management System (RDBMS) like PostgreSQL or MySQL for structured data. A Geospatial extension (PostGIS) is required to handle spatial queries and hotspot detection. An optional NoSQL Database (MongoDB) can be used for semi-structured data like API responses and logs.

4.2 Software Modules

The system's intelligence and functionality are organized into four primary software layers:

- **Frontend (User Interface Layer):** This layer provides role-based dashboards (for policymakers, operators, and drivers). It features interactive data visualizations, geospatial heatmaps, real-time trends, and predictive insights. Technologies include Tableau, Power BI, Plotly Dash, and Kepler.gl/Leaflet for mapping.
- **Backend (Application & Logic Layer):** This layer implements a Microservices architecture to separate engines for forecasting, recommendation, and pricing. It uses
- REST APIs to serve models and a potential GraphQL API for flexible client queries. Real-time stream processing is also supported for live station status updates.
- **Database (Storage Layer):** This layer supports the storage and retrieval of various data types.
- **Charging Session Data:** Stores session duration, energy consumption, connector type, and station location.
- **Forecasting Data:** Stores time-series data for future demand predictions.
- **Geospatial Data:** Stores latitude/longitude and region identifiers for mapping.
- **User and Vehicle Data:** Includes EV model, user profile, and charging preferences for personalization.
- **AI/ML Layer (Prediction & Recommendation):** This layer executes the core intelligence.
- It runs time-series models (LSTM, Prophet) for demand forecasting.
- It hosts the Recommendation System for smart station allocation.
- It contains Anomaly Detection AI and Explainable AI (XAI) modules.

4.3 End-to-End System Workflow

The integrated system operates through a structured workflow across its components:

1. **Data Ingestion & Cleaning:** Data is collected from external sources (e.g., Open Charge Map) and cleaned, normalized, and integrated.
2. **Exploratory Data Analysis (EDA):** Patterns, peak loads, and regional hotspots are identified.
3. **Advanced Analytics:** Clustering algorithms group stations by behaviour, and time-series models (LSTM, Prophet) forecast future demand and energy load.
4. **Real-Time Monitoring:** Live station status and current availability are continuously tracked.
5. **Recommendation & Optimization:** The Recommendation Engine receives driver inputs (location, vehicle) and current availability/forecasts to suggest the optimal station based on predicted wait time, cost, and distance.
6. **Dashboard Display:** The AI-driven dashboard updates in real-time, showing forecasts, heatmaps, and station recommendations tailored to the user's role (Driver/Operator/Policymaker).
7. **MLOps and Monitoring:** Models are served via APIs, and continuous monitoring ensures forecast accuracy and scheduled retraining to handle data drift.

4.4 Advantages of System Integration

The fully integrated system provides significant advantages over fragmented approaches:

- **Non-Intrusive and Data-Driven:** It uses publicly available and aggregated charging data, minimizing privacy concerns.
- **Predictive over Descriptive:** It provides forecasting and prescriptive recommendations, enabling proactive planning instead of just reporting historical trends.
- **Multi-Stakeholder Value:** It simultaneously supports infrastructure planning for policymakers, load management for grid operators, and convenience for EV drivers.
- **Scalable and Reproducible:** Cloud deployment and MLOps practices (Docker, GitHub) ensure the system can handle increasing EV adoption and data volume while maintaining version control.

4.5 Implementation and MLOps Details

The proposed system was developed using Python, integrating TensorFlow/Keras (for LSTM), Prophet, and XGBoost for demand forecasting. scikit-learn and GeoPandas were used for clustering and spatial analysis.

The entire workflow was containerized with Docker and managed using MLflow for experiment tracking and model versioning. The models were served via FastAPI REST APIs with endpoints /predict demand and /recommend station.

Monitoring modules track data drift, model accuracy, and inference latency using Prometheus and Grafana dashboards. Automatic retraining is triggered monthly or when drift exceeds predefined thresholds.

VII. EXPECTED OUTCOMES AND DISCUSSIONS

The AI-in-operated EV charging demand is a multi-level system designed to adapt to the extended Electric Vehicle (EV) charging infrastructure through predictive and prescriptive analytics.

Fundamental EV charging data (from sources like Open Charge Map and U.S. DOE AFDC) and the functioning centre of the system on upgrading and implementing advanced Machine learning models including LSTM and Prophet, future charging demand and exact forecast of energy load.

Concurrent Clustering algorithms (K-Means, DBSCAN) and geospatial analysis are used to identify high- demand areas and optimal locations for infrastructure expansion [6].

The core is a wider deliverable AI-driven interactive dashboard that provides real-time, actionable intelligence to many stakeholders. Policy makers receive insight for strategic planning and station allocation, while charging operators receive forecasts and capacity alerts for efficient load balance. Most importantly, EV drivers receive smart station recommendations based on a multi-criteria optimization approach that reduces waiting time while factoring in availability, distance, and estimated charging cost [2].

This solution enhances operating efficiency, reduces stress on the power grid, and promotes stability by searching [9].

VIII. CONCLUSION

The AI-driven EV charging demand system is a multi-level framework developed to strengthen the existing electric vehicle (EV) charging infrastructure by leveraging both predictive and prescriptive analytics. The system avails a large scale from global sources to create an integrated platform for system analysis and plan, avails diverse charging session data. The main capacity lies in the use of advanced time- series models such as LSTM and Prophet, combined with clustering techniques (KMeans, DBSCAN) to identify high-demand geographical hotspots and clusters of clusters. It transfers the system beyond static analysis to provide prognosis and prescriptive intelligence. The central component of the system is an AI-powered interactive dashboard that transforms complex data into actionable insights for multiple stakeholders. Policymakers receive information for strategic planning and station allocation, energy providers gain forecasts of grid load variations for efficient load balancing, and EV drivers receive smart station recommendations based on multicriteria optimization. The system also promotes sustainability by integrating renewable energy sources and estimating potential carbon savings. Future development includes incorporating reinforcement learning for dynamic pricing and enabling deeper integration with smart grid technologies, ensuring a robust, efficient, and user-centered platform that accelerates long-term EV adoption.

IX. ACKNOWLEDGEMENT

We, the project team, would like to express our sincere gratitude to the Department of Computer Science and Engineering, B.M. Institute of Engineering and Technology, for providing us with the opportunity to conceptualize and develop this proposed research project. We are especially thankful to our project supervisor, Ms. Paramjeet, for her continuous guidance, valuable suggestions, and encouragement throughout the planning and design of this study. We also extend our appreciation to our peers and colleagues for their feedback and support during the proposal phase, which helped refine our ideas. Finally, we gratefully acknowledge the support and motivation from our families and friends, whose encouragement has been instrumental in completing the preparation of this proposed project.

X. REFRENCES

[1] Cheng, N., Zheng, P., Ruan, X., & Zhu, Z. (2023). "Electric vehicle charging load prediction based on variational mode decomposition and Prophet-LSTM." Frontiers in Energy Research, 11.

[2] Jamil, U., Alva, R. J., Ahmed, S., & Jin, Y. F. (2025). "Artificial Intelligence-Driven Optimal Charging Strategy for Electric Vehicles and Impacts on Electric Power Grid." Electronics, 14(7), 1471.

[3] Bui, T. P., Le, T. V., & Nguyen, D. C. (2023). "Clustering-Based Optimal Operation of Charging Stations Under High Penetration of Electric Vehicles." Sustainable Energy Grids and Networks, 36, 101178.

[4] Wang, S., Chen, A., Shao, C., Wang, P., Yang, X., & Wang, S. (2023). "Short-term electric vehicle charging demand prediction: A deep learning approach." Applied Energy, 340, 121032

[5] Mohamed, A. M. H., & El-Dardiry, S. E. Z. I. (2025). "Smart Charging Solutions for Electric Vehicles: An AI-Driven Approach to Load Balancing and Grid Integration." Journal of Information Systems Engineering and Management, 10(29s).

[6] Bhupathi, H. P., & Chinta, S. (2021). "Integrating AI with Renewable Energy for EV Charging: Developing Systems That Optimize the Use of Solar or Wind Energy for EV Charging." ESP Journal of Engineering & Technology Advancements, 1(2), 260–271.

[7] Saba, S., Mosalli, H., Yang, Y., Yeh, H. G., & Aghdam, A. G. (2025). "Demand Forecasting for Electric Vehicle Charging Stations using Multivariate Time-Series Analysis." arXiv preprint.

[8] Tian, Z., Jung, T., Wang, Y., Zhang, F., Tu, L., Xu, C., & Li, X. Y. (2016). "Real-time charging station recommendation system for electric-vehicle taxis." IEEE Transactions on Intelligent Transportation Systems, 17(11), 3098–3109.

[9] Borah, S. (2024). "Optimization of Electric Vehicle Charging Infrastructure with AI." Research Journal of Computer Systems and Engineering, 5(1), 59–70.

[10] Yi, Z., Liu, S., Zhang, D., & Zhao, C. (2022). "Electric vehicle charging demand forecasting using deep learning model." Journal of Intelligent Transportation Systems, 26(6), 690–703