iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** XI **Month of publication:** November 2026

DOI: https://doi.org/10.22214/ijraset.2026.85577

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue XI Nov 2026- Available at www.ijraset.com

EV Charging Station Demand Analysis

Mohanapreetha S[1], Hemanth Babu N[2], Bhuvaneshwaran M[3], Dharshan V[4]
Department of Computer Science and Engineering, Adhiyamaan College of Engineering, Hosur, India

Abstract: ***Electric vehicle (EV) charging networks are usually monitored through separate hardware logs and spreadsheets, so demand surges, queue delays and capacity needs are noticed only after they have occurred. This paper presents the EV Charging Demand Analytics Platform, which brings the session records and station metadata of a charging network into one interactive web application. The data of 8,354 charging sessions across 20 urban and highway hubs are validated, merged and enriched with time, revenue, utilization and demand-level features. Six scikit-learn pipelines are trained: a Random Forest queue waiting-time regressor, a Gradient Boosting port-capacity advisor based on a 70% target-utilization rule, a Random Forest peak-hour classifier, a Gradient Boosting session-energy forecaster, a Random Forest high-demand hotspot classifier and a revenue regressor. A Streamlit and Plotly interface gives network key performance indicators, a station leaderboard, a day-of-week by hour-block density heatmap, a six-way non-restrictive filter and what-if prediction forms. The network totals are 8,354 sessions, Rs. 3,495,296 revenue, 226,603 kWh energy, 53.2% average load and 9.5 min average waiting time. The peak-hour and demand-tier classifiers reach 99.3% and 94.0% accuracy, while the waiting-time model does not beat the mean baseline ($R^2 = -0.008$). The reasons for these results, including target leakage in several models, are discussed openly, and a plan for a stricter evaluation is given.***

Keywords: Electric Vehicle Charging, Demand Analysis, Queue Waiting Time, Capacity Planning, Random Forest, Gradient Boosting, Streamlit, Scikit-learn.

I. INTRODUCTION

The shift from internal combustion vehicles to battery electric vehicles brings new operating problems for charging point operators (CPOs) and for the power grid. A petrol refill takes three to five minutes, whereas an EV session takes from twenty minutes to several hours depending on the battery state of charge (SoC), the charger rating (AC Level 2 or DC fast) and thermal limits. This long duration makes queues, uneven arrivals and local congestion likely. Highway hubs see sharp weekend surges, while urban hubs see midday and evening commuter peaks. When many fast chargers start together, the power drawn can reach several megawatts and stress the distribution transformer, while hubs that are idle off-peak return too little on their capital cost.

In practice, operators use separate telemetry dashboards, static spreadsheets or vendor portals. These tools show past transactions but give no cross-station view, no flexible slicing of the data and no forecast of queue delay, port need, energy or revenue. Capacity decisions, such as adding fast-charging ports at a highway plaza, are therefore often taken by intuition or after customer complaints.

This paper presents the EV Charging Demand Analytics Platform, in which one dataset and one dashboard serve the whole network. Session logs and station master data are merged and feature-engineered; six machine learning pipelines give predictions for waiting time, port capacity, peak hours, energy, demand hotspots and revenue; and a Streamlit interface gives key indicators, heatmaps, filters and what-if simulators to operators and planners.

II. LITERATURE SURVEY

1) S. Deb, K. Tammi, K. Kalita and P. Mahanta, “Review of Recent Trends in Charging Infrastructure Planning for Electric Vehicles”, IEEE Access (2018). The survey covers optimization and spatial clustering methods for station placement, grid integration and sizing, considering traffic density, user convenience and distribution loss. It gives a strong basis for hub placement, but it deals with static macro-planning and presents no operational telemetry platform or machine learning prediction pipeline. [1]

2) M. B. Arias and S. Bae, “Electric Vehicle Charging Demand Forecasting Model Based on Big Data Technologies”, IEEE Trans. Smart Grid (2016). A spatial-temporal forecasting framework combines traffic flow, travel surveys and weather, and shows that urban demand correlates strongly with time of day and weather. It needs large external transport data and offers no station-level queue estimate or interactive dashboard. [2]

3) M. Majidpour, C. Qiu, P. Chu, H. R. Pota and R. Gadh, “Fast Charging Station Demand Forecasting: A Machine Learning Approach”, IEEE Trans. Transp. Electrif. (2016). Support Vector Regression and Random Forests forecast fast-charging demand on a campus testbed, and ensemble tree models outperform conventional time-series models. The study covers a single campus and has no capacity sizing or multi-dimensional filtering. [3]

©IJRASET: All Rights are Reserved

4) Y. Zhang, L. Cai and Z. Meng, "Queueing Analysis and Capacity Sizing of Fast-Charging Stations Under Heterogeneous Arrival Rates", Applied Energy (2020). M/M/s and M/G/s queueing models relate port count to peak-hour delay in a non-linear way. The theoretical assumptions fail when arrivals are not Poisson, which motivates data-driven empirical models. [4]

5) B. Al-Hanahi, I. Ahmad, D. Habibi and M. A. S. Masoum, "Charging Infrastructure for Commercial Electric Vehicles: Challenges and Future Works", IEEE Access (2021). The review identifies demand charges, uncoordinated schedules and low port utilization as problems of fleet depots and calls for real-time monitoring with predictive capacity allocation. It is a review and delivers no software. [5]

6) N. Sadeghianpourhamami, N. Refa, M. Strobbe and C. Develder, "Quantifying Flexibility in Electric Vehicle Charging With Real Data", IEEE Trans. Smart Grid (2018). Records from over 1,000 public stations in the Netherlands show that vehicles stay parked long after charging ends, which inflates occupancy. It confirms the need to treat queue length, charging duration and port utilization separately. [6]

7) E. Xydas, C. Marmaras, L. M. Cipcigan, N. Jenkins, S. Carroll and M. Barker, "A Data-Driven Approach for Characterising the Charging Demand of Electric Vehicles: A UK Case Study", Applied Energy (2016). Real telemetry from UK trials gives SoC distributions, arrival hours and energy patterns for weekdays and weekends. It does not provide multi-station drilldown tools or predictive regression. [7]

8) H. Moon, S. Park, C. Jeong and J. Lee, "Forecasting Electric Vehicle Charging Demand Using Machine Learning and Time-Series Decomposition", IEEE Trans. Veh. Technol. (2021). A hybrid of Seasonal Trend Decomposition and Random Forests predicts hourly energy draw accurately by isolating day-of-week seasonality, but it models only aggregate energy and not queue time, port sizing or revenue. [8]

9) L. Breiman, "Random Forests", Machine Learning (2001). Random Forests combine bagged trees with random feature selection, reduce overfitting, handle mixed numeric and categorical inputs and need little tuning, which suits tabular session data. [9]

10) J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine", Annals of Statistics (2001). Gradient boosting fits base learners to the pseudo-residuals of a loss function and captures non-linear feature interactions, which supports its use for energy and capacity regression. [10]

11) F. Pedregosa et al., "Scikit-learn: Machine Learning in Python", J. Mach. Learn. Res. (2011). The library offers column transformers, a uniform fit-predict interface and serializable pipelines, which keep preprocessing identical in training and inference. [11]

12) S. Wang, Z. Y. Dong, C. Chen and X. Fan, "Multi-Factor Electric Vehicle Charging Load Forecasting Under Adverse Weather Conditions", IEEE Trans. Ind. Informat. (2022). Cold and wet weather lowers battery efficiency and lengthens charging, which supports the use of weather inputs in the waiting-time and demand models. [12]

13) A. M. Morro-Mello, A. Padilha-Feltrin and J. D. Melo, "Dynamic Tariff Design and Financial Risk Assessment for Public EV Fast-Charging Hubs", Renew. Sustain. Energy Rev. (2021). Profitability models show that predictable revenue is essential for the solvency of fast-charging hubs, which supports automated revenue projection. [13]

14) T. D. Chen, K. M. Kockelman and M. Khan, "Locating Electric Vehicle Charging Stations with Optimization Modeling and Queuing Considerations", IEEE Trans. Intell. Transp. Syst. (2018). Stations that run consistently above about 70–75% load see queues grow sharply, which supports the 70% target utilization used here for port sizing. [14] 15) P. Kumar, S. Bera and A. Maiti, "Real-Time Web-Based Telemetry and Analytics Framework for Distributed EV Charging Hubs", IEEE Access (2023). A web dashboard framework for distributed charging points, built on Streamlit and WebGL tools, shows that low-latency interfaces with multi-variable filters improve operator awareness. [15]

III. RESEARCH GAP

The surveyed works each give a forecast, a queueing model, a review or a visualization framework, as summarised in Table I. The following gaps remain.

A. Planning without Operations

Placement and sizing studies [1], [4], [14] work at the planning level and do not provide a tool that operators can use daily on their own session data.

B. *Forecasts without Queue, Capacity and Revenue*

Demand forecasting studies [2], [3], [8] predict energy or load but do not predict driver queue time, the number of ports to add or the revenue of a hub.

C. *Dashboards without Predictions*

Telemetry dashboards [15] improve awareness but are not coupled to a set of trained models and what-if simulators.

D. *Rigid Views*

Operator portals give fixed, pre-aggregated summaries, and a user cannot combine station, date, location type, vehicle class, day and time slot freely in one view.

TABLE I.
CAPABILITIES OF RELATED WORK AND THE PROPOSED PLATFORM (YES = PROVIDED; PART. = PARTIALLY; – = NOT DESCRIBED)

Work	Demand / energy forecast	Queue and capacity	Revenue model	Interactive dashboard	Multi-filter drilldown	Multi-model ML suite
[1], [14]	–	Part.	–	–	–	–
[2], [3], [8]	Yes	–	–	–	–	–
[4]	–	Yes	–	–	–	–
[5], [13]	–	Part.	Part.	–	–	–
[6], [7], [12]	Part.	Part.	–	–	–	–
[15]	–	–	–	Yes	Part.	–
Proposed	Yes	Yes	Yes	Yes	Yes	Yes

IV. PROBLEM STATEMENT

Charging operators lack a unified, low-cost platform that (i) imports, cleans and merges the session and station data of a multi-hub network, (ii) forecasts queue waiting time, port needs, peak periods, session energy, demand hotspots and revenue, (iii) shows network and station indicators with flexible, multi-dimensional filtering on one dashboard, and (iv) runs on ordinary hardware without specialised infrastructure.

V. OBJECTIVES

1) To ingest, validate and harmonize session logs and station master data and to engineer temporal, operational and commercial features.
2) To forecast driver queue waiting time from queue length, station load, demand and weather.
3) To recommend additional charging ports per hub against a 70% target utilization.
4) To classify peak-hour intervals and demand tiers and to forecast session energy and revenue.
5) To benchmark the 20 hubs with KPIs, a leaderboard and a day-of-week by hour-block heatmap.
6) To provide a non-restrictive multi-dimensional filter that starts unfiltered and never returns an empty default view.

VI. METHODOLOGY

A. *System Architecture*

The platform has three layers, as shown in Fig. 1. The data ingestion and preprocessing layer reads the session file and the station master file, parses timestamps, merges the two on the station identifier and generates features. The predictive layer holds six scikit-learn pipelines. The presentation layer is a Streamlit and Plotly application with four modules: the home cockpit, overall station analysis, station-location-time analysis with global filters, and the machine learning prediction suite. On start-up, an auto-bootstrap step checks whether the merged dataset and the trained models exist and builds them if absent; results are cached with Streamlit data and resource caches so that page changes are fast.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

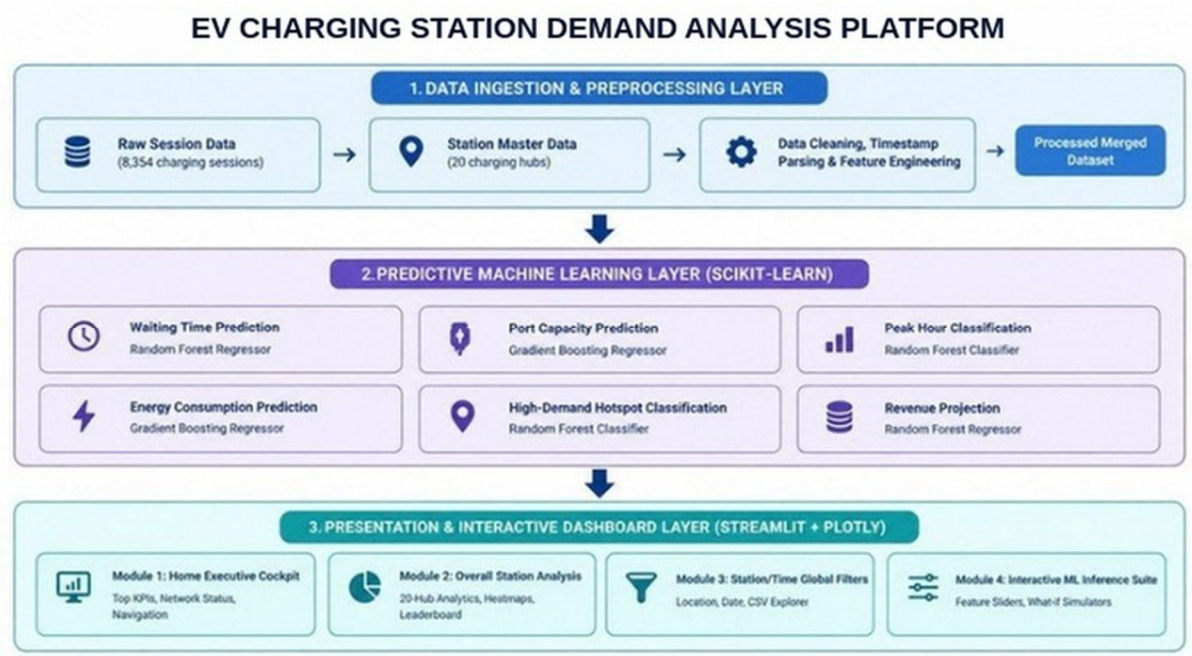


Fig. 1. Architecture of the EV Charging Demand Analysis Platform: data ingestion and preprocessing, predictive machine learning and interactive dashboard layers.

B. Dataset and Preprocessing

The session file contains 8,354 records with the identifiers, vehicle type and battery capacity (15 to 300 kWh), initial and final SoC, charging power (7.4 to 150 kW), arrival, start and end times, waiting time (0 to 45 min), charging duration (30 to 240 min), energy consumed, queue length (0 to 15), station load (10% to 95%), weather, day of week, time slot, charging demand score and renewable energy ratio. The station master file describes 20 hubs (6, 8, 10 or 12 ports; DC Fast or AC Level 2; customer tariff Rs. 14 to 18 per kWh; utility tariff Rs. 7.5 to 9.5 per kWh). Timestamps are parsed in the format DD-MM-YYYY HH:MM, and the two files are merged on the station identifier. The engineered features are listed in Table II.

TABLE II. ENGINEERED FEATURES

Feature	Formulation
arrival_hour, arrival_date	Hour (0–23) and date from the arrival timestamp
hour_block	Eight 3-hour intervals, 00–03 to 21–24
is_weekend	1 for Saturday or Sunday, else 0
revenue_inr	energy_consumed_kWh × customer_charging_rate
cost_inr, profit_inr	Electricity cost and (revenue − cost)
utilization_pct	(queue_length / total_ports), capped at 3, × 100
demand_level	Low, Medium, High from the 33rd and 66th percentiles of charging_demand
is_high_demand, is_peak	Binary flags for classification and alerts

C. Machine Learning Pipelines

Every model is a scikit-learn Pipeline with a ColumnTransformer that applies StandardScaler to numeric features and OneHotEncoder (unknown values ignored) to categorical features, followed by the estimator. Session-level models use an 80/20 train-test split with a fixed random seed of 42, and the classifiers are stratified. The six models are summarised in Table III. The trained pipelines are saved with Joblib, so that the application loads them without retraining.

©IJRASET: All Rights are Reserved

TABLE III.
THE SIX MACHINE LEARNING PIPELINES

Model	Algorithm	Target
Waiting time	Random Forest Regressor, 150 trees, depth 10	waiting_time (min)
Port sizing	Gradient Boosting Regressor, 150 trees, depth 3	recommended_ports (station level)
Peak hour	Random Forest Classifier, 150 trees	time_slot (Peak / Off-Peak)
Energy	Gradient Boosting Regressor, 150 trees	energy_consumed_kWh
High demand	Random Forest Classifier, 150 trees	demand_level (Low / Medium / High)
Revenue	Random Forest Regressor, 150 trees	revenue_inr (aggregated per station in the app)

D. Capacity Sizing Rule

The raw data has no ground-truth ideal port count, so a target is derived with a standard capacity-planning rule. With a target utilization U_t = 70% [14], the recommended number of ports of a station is

$$P_{rec} = ceil(P_{cur} \times (L_{avg} / U_t)),\ P_{add} = max(P_{rec} - P_{cur}, 0) \quad (1)$$

where P_{cur} is the installed number of ports and L_{avg} is the average station load. A Gradient Boosting regressor is then trained on station-level signals (queue length, load, demand, waiting time, sessions, duration) to reproduce this target, so that the advisor can be applied to a new station that has only these signals.

E. Filtering Engine and Dashboard

Six global filters are provided: station, date range, location type (Urban or Highway), vehicle type (2-wheeler, 3-wheeler, car, bus, truck), day of week and time slot. All filters start as empty lists, and an empty filter means no restriction, so the first view shows 100% of the data. A selection narrows the dataset immediately, and the session count, coverage, daily trend with a 7-day moving average, 24-hour demand curve, station efficiency (kWh per minute) and a raw-data explorer with CSV export are updated. The network average load factor shown on the home page is

$$Average\ load = (1 / N)\ \Sigma L_j\ ,\ j = 1 \ldots N \quad (2)$$

F. Software Environment

The platform runs on Python 3.14 with Streamlit 1.32 or later, Pandas, NumPy, Scikit-learn, Plotly and Joblib, on Windows 10/11, Ubuntu 20.04/22.04 or macOS. The minimum hardware is a 1.5 GHz dual-core processor and 4 GB RAM (8 GB recommended), and the client needs only a modern web browser.

VII. RESULTS AND DISCUSSION

A. Network Indicators

The home cockpit and the station intelligence page report the indicators of Table IV for the whole network. Fig. 2 shows the station intelligence page with the network KPIs, Fig. 3 the station, location and time cockpit with the six global filters, and Fig. 4 the machine learning prediction suite. The density heatmap shows that Friday and Saturday evenings (18:00 to 21:00) are the most congested periods of the network.

Table IV. Network-Wide Indicators

Indicator	Value
Charging sessions	8,354
Charging hubs	20
Total revenue	Rs. 3,495,296
Total energy delivered	226,603 kWh
Average revenue per hub	Rs. 174,765
Average revenue per session	Rs. 418.40
Average station load	53.2%
Average waiting time	9.5 min
Average charging duration	128 min

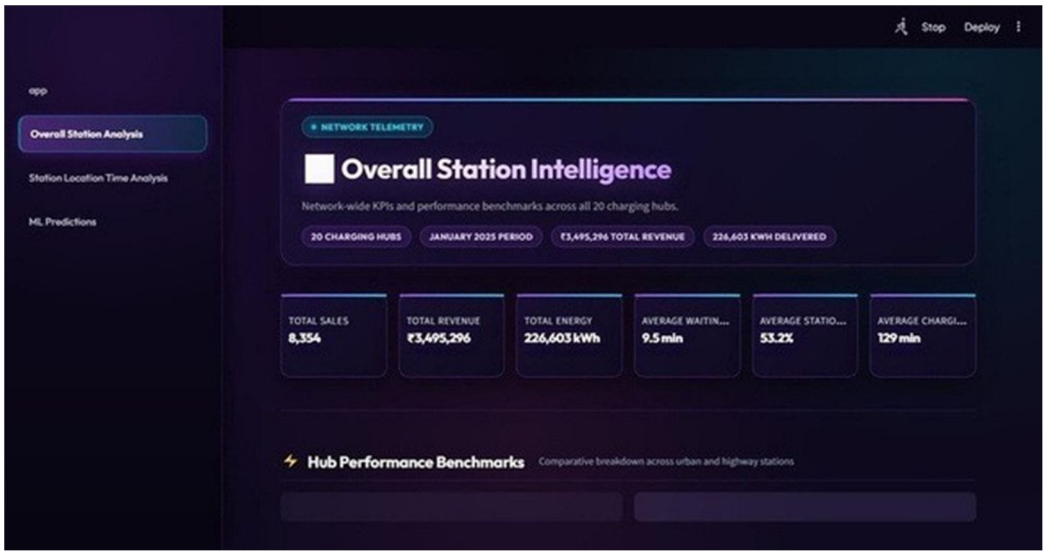


Fig. 2. Overall station intelligence: network KPIs and hub performance benchmarks.

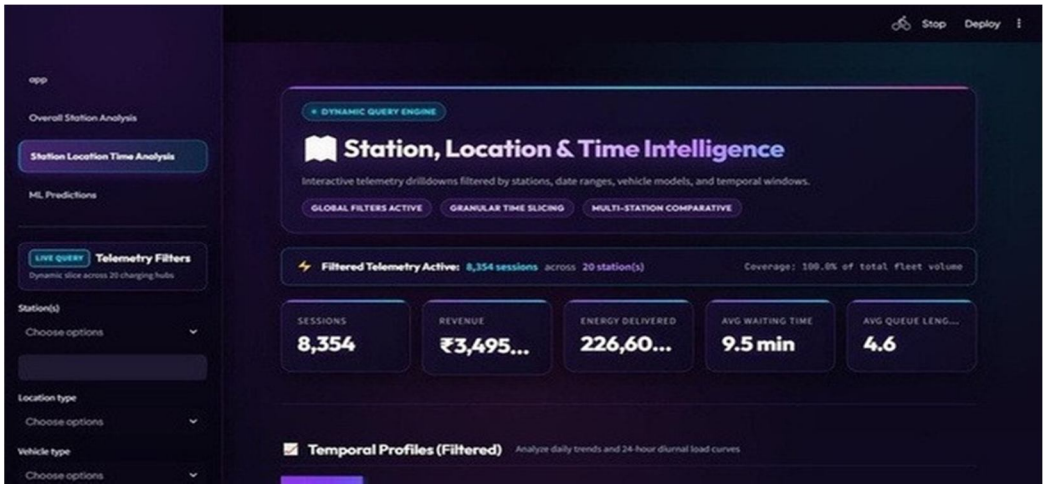


Fig. 3. Station, location and time cockpit with the six global telemetry filters.

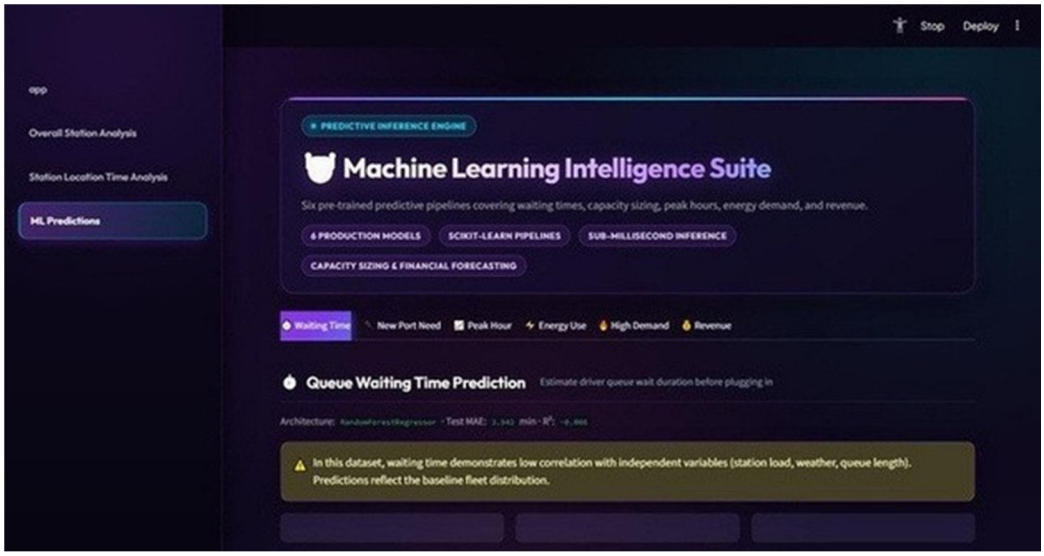


Fig. 4. Machine learning prediction suite with the six model tabs.

B. Model Performance

Table V gives the metrics saved by the training script. Regression models are scored on the held-out 20% of sessions, except the port-sizing model, which is scored in-sample because only 20 stations exist.

TABLE V. PERFORMANCE OF THE SIX MODELS

Model	Metric	Value
Waiting time	MAE / R^2	3.942 min / −0.008
Port sizing (in-sample, n = 20)	MAE / R^2	0.000 / 1.000
Peak hour	Accuracy / F1	99.3% / 0.993
Energy	MAE / R^2	0.184 kWh / 1.000
High demand	Accuracy / F1	94.0% / 0.939
Revenue	R2	approx. 1.00

C. Discussion

The platform shows that the data of a charging network can be merged, explored with free multi-dimensional filters and combined with six predictors in one low-cost web application, and that the first view never starts with an empty or biased selection.

The model results must be read with care. The waiting-time model has an R^2 of −0.008, which means it does no better than predicting the mean waiting time; the application states this on the prediction page, because in this dataset waiting time has little correlation with queue length, load or weather. The port-sizing model reproduces a formula that was itself used to create its target, and it is scored on the same 20 stations, so its R^2 of 1.000 shows that the formula can be learned and not that new stations can be predicted. The energy and revenue models have R^2 near 1.000 because energy is nearly determined by battery capacity, the SoC difference and charging power, and revenue is energy multiplied by the tariff, which is itself an input. The peak-hour and demand-tier classifiers are likely to be helped by input features that are close to their targets (the arrival hour defines the time slot, and the charging demand score defines the demand level). These high scores are therefore best regarded as a check that the pipelines work and not as evidence of forecasting skill.

Other limits are that the data come from one period, the 70% rule is a benchmark and not a measured optimum for each site, and the weather and demand scores should be validated against field data. The tool supports an operator's decision and does not replace it.

D. Evaluation Plan

To give stronger evidence, the following will be done: remove target-defining inputs from the classifiers; use a time-based split in which later days are used for testing; compare every model with a naive baseline such as the station mean; report cross-validated results with confidence intervals; and evaluate the capacity advisor by leaving one station out. Precision, recall and F1 will be reported per class for the demand tiers.

VIII. CONCLUSION AND FUTURE SCOPE

This paper presented an EV charging demand analytics platform that unifies session data of 20 hubs, offers network and station dashboards with a non-restrictive six-way filter, and provides six machine learning pipelines for waiting time, port capacity, peak hours, energy, demand hotspots and revenue, together with a 70% utilization rule for capacity planning. The network indicators are reproduced on the dashboard, and the limits of the present model results are stated openly. Future work includes the stricter evaluation described above, live integration with chargers through OCPP 1.6J and 2.0.1, time-of-use tariff optimization, multi-tenant cloud deployment with Docker and Kubernetes, solar and battery storage sizing at highway plazas, transformer thermal and grid load modelling, and a mobile application with maintenance alerts.

IX. ACKNOWLEDGMENT

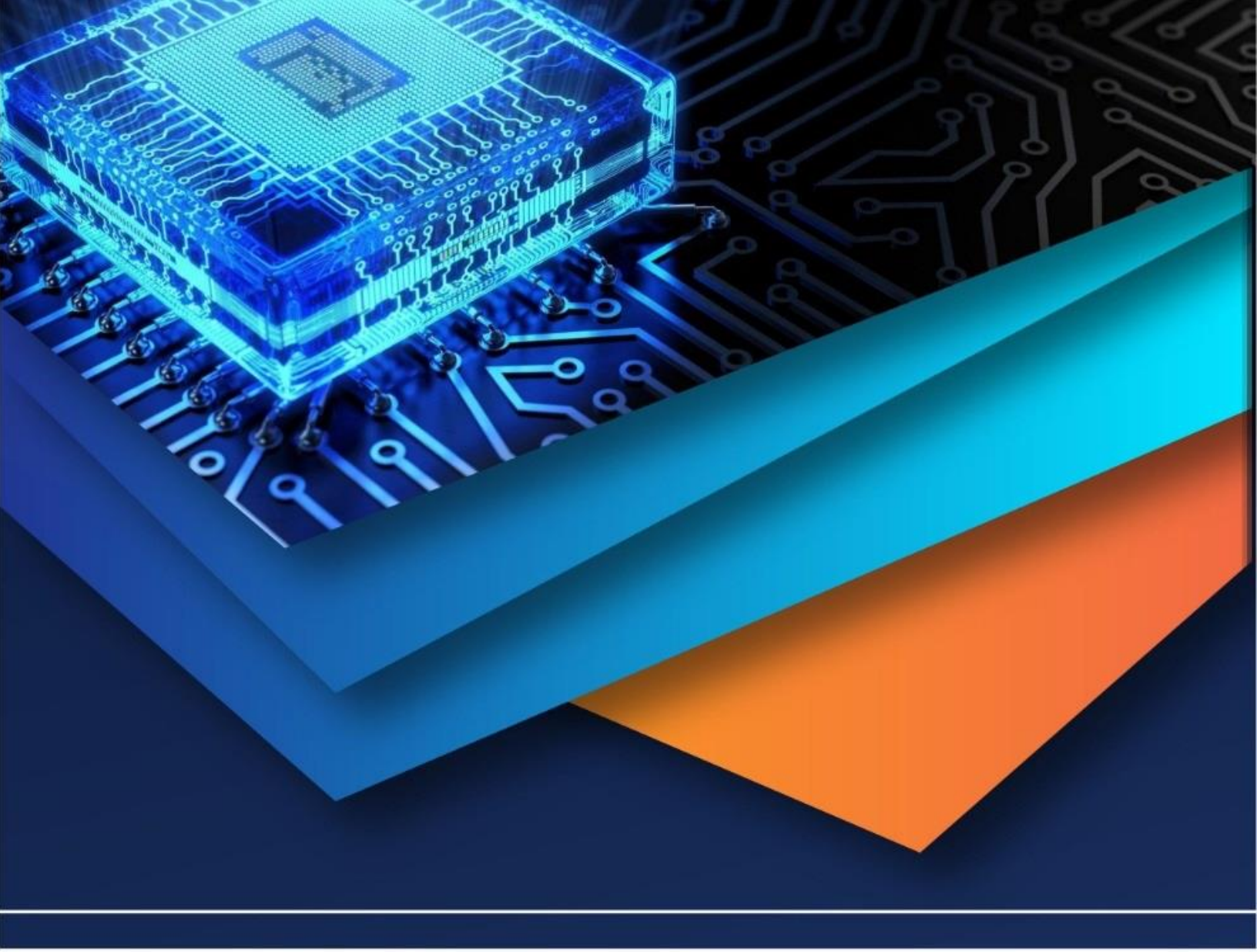
The authors thank the Department of Computer Science and Engineering, Adhiyamaan College of Engineering, Hosur, for institutional support.

©IJRASET: All Rights are Reserved

REFERENCES

[1] S. Deb, K. Tammi, K. Kalita, and P. Mahanta, "Review of recent trends in charging infrastructure planning for electric vehicles," IEEE Access, vol. 6, pp. 64145–64166, 2018.
[2] M. B. Arias and S. Bae, "Electric vehicle charging demand forecasting model based on big data technologies," IEEE Trans. Smart Grid, vol. 7, no. 3, pp. 1577–1586, May 2016.
[3] M. Majidpour, C. Qiu, P. Chu, H. R. Pota, and R. Gadh, "Fast charging station demand forecasting: A machine learning approach," IEEE Trans. Transp. Electrific., vol. 2, no. 1, pp. 25–34, Mar. 2016.
[4] Y. Zhang, L. Cai, and Z. Meng, "Queueing analysis and capacity sizing of fast-charging stations under heterogeneous arrival rates," Appl. Energy, vol. 268, p. 114976, 2020.
[5] B. Al-Hanahi, I. Ahmad, D. Habibi, and M. A. S. Masoum, "Charging infrastructure for commercial electric vehicles: Challenges and future works," IEEE Access, vol. 9, pp. 121476–121492, 2021.
[6] N. Sadeghianpourhamami, N. Refa, M. Strobbe, and C. Develder, "Quantifying flexibility in electric vehicle charging with real data," IEEE Trans. Smart Grid, vol. 9, no. 5, pp. 4734–4746, 2018.
[7] E. Xydas, C. Marmaras, L. M. Cipcigan, N. Jenkins, S. Carroll, and M. Barker, "A data-driven approach for characterising the charging demand of electric vehicles: A UK case study," Appl. Energy, vol. 162, pp. 763–771, Jan. 2016.
[8] H. Moon, S. Park, C. Jeong, and J. Lee, "Forecasting electric vehicle charging demand using machine learning and time-series decomposition," IEEE Trans. Veh. Technol., vol. 70, no. 9, pp. 8412–8424, Sep. 2021.
[9] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, Oct. 2001.
[10] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Ann. Statist., vol. 29, no. 5, pp. 1189–1232, Oct. 2001.
[11] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825–2830, Nov. 2011.
[12] S. Wang, Z. Y. Dong, C. Chen, and X. Fan, "Multi-factor electric vehicle charging load forecasting under adverse weather conditions," IEEE Trans. Ind. Informat., vol. 18, no. 11, pp. 7820–7831, Nov. 2022.
[13] A. M. Morro-Mello, A. Padilha-Feltrin, and J. D. Melo, "Dynamic tariff design and financial risk assessment for public EV fast-charging hubs," Renew. Sustain. Energy Rev., vol. 143, p. 110904, 2021.
[14] T. D. Chen, K. M. Kockelman, and M. Khan, "Locating electric vehicle charging stations with optimization modeling and queuing considerations," IEEE Trans. Intell. Transp. Syst., vol. 19, no. 10, pp. 3201–3210, Oct. 2018.
[15] P. Kumar, S. Bera, and A. Maiti, "Real-time web-based telemetry and analytics framework for distributed EV charging hubs," IEEE Access, vol. 11, pp. 45120–45131, 2023.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)