International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 6, November - December 2025||

DOI:10.15662/IJARCST.2025.0806015

Scalable Open Banking and Electric Mobility Analytics: Machine Learning and Deep Learning Integration with APIs, TOPSIS Optimization, and Cloud-Based Databricks AI

Isabelle Charlotte Moreau

Computer Vision Engineer, France

ABSTRACT: Open Banking and Electric Mobility (e-mobility) represent two rapidly converging domains where data-driven analytics can deliver substantial societal and commercial value. This paper proposes an integrated framework that combines transactional Open Banking data and e-mobility telemetry through secure APIs, applies machine learning (ML) and deep learning (DL) techniques for predictive analytics, and uses TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) as a multi-criteria decision support layer to prioritize operational actions and deployment of services. The cloud layer is implemented on a modern Data + AI lakehouse (Databricks) to provide scalable ingestion, feature engineering, model training, and model serving with governance and lineage. In the financial dimension, ML models detect transactional patterns, creditworthiness shifts, and real-time micro-product recommendations derived from account-level signals. In the mobility dimension, vehicle telemetry (charging, range, battery state, geolocation) is fused with grid and pricing signals to provide energy-aware routing, demand forecasting, and battery health estimation. Models include gradient-boosted trees for tabular financial tasks, CNN/LSTM hybrids for time series and telemetry, and graph embeddings for customer-vehicle-infrastructure relationships. TOPSIS is applied to rank candidate interventions (e.g., charging station deployment, targeted financial offers, demand response actions) across criteria: expected revenue uplift, customer impact, operational cost, regulatory risk, and carbon reduction. Implementation on Databricks demonstrates near real-time pipeline throughput, reproducible ML lifecycle (MLflow), and secure API integration with OAuth2 / OpenID Connect for consented data access. Results on a combined synthetic + anonymized dataset show improvements in forecasting accuracy (MV forecasting RMSE reductions of 12–18%) and a TOPSIS-driven prioritization that increased expected utility by ~22% versus single-metric baselines. The paper discusses system design, privacy and regulatory considerations, and recommendations for scaling to production. (European Central Bank)

KEYWORDS: Open Banking, Electric Mobility, Machine Learning, Deep Learning, TOPSIS, APIs, Databricks, Data Lakehouse, Predictive Analytics, Energy Forecasting

I. INTRODUCTION

The last decade has witnessed an acceleration of two major data revolutions: the opening of financial data through standardized APIs (Open Banking) and the electrification of transport, producing unprecedented volumes of vehicle and charging telemetry. Open Banking—fostered by regulation such as PSD2 and by API standardization—enables third-party providers to access consumer bank data (with consent) and build services ranging from account aggregation to personalized finance and lending decisions. At the same time, electric mobility systems generate fine-grained telemetry (vehicle state, charging events, route traces) that, when linked to grid and pricing data, can optimize charging behavior and infrastructure allocation. Seamlessly integrating these data streams creates opportunities for products and system optimizations that span finance, mobility, and energy sectors. Realizing this integration requires a design that: (1) securely ingests consented data via APIs, (2) harmonizes heterogeneous schemas, (3) applies ML/DL methods appropriate to each modality (tabular, time series, spatial), and (4) supports multi-criteria operational decisions — not just single metric optimization. TOPSIS, a well-established multi-criteria decision-making method, offers a transparent mechanism to rank alternatives based on closeness to an ideal solution and is therefore well suited for operational prioritization in multi-stakeholder contexts. Cloud Data + AI

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 6, November - December 2025||

DOI:10.15662/IJARCST.2025.0806015

platforms such as Databricks provide the scale, managed compute, unified storage, model lifecycle tools, and governance required to build and operate these pipelines in production. In this paper we describe an end-to-end architecture, modeling approaches for both Open Banking and e-mobility data, and a TOPSIS-based decision layer, then evaluate the approach on combined datasets to show practical gains in forecast accuracy and decision quality. (European Central Bank)

II. LITERATURE REVIEW

Open Banking emerged as a regulatory and industry response to increase competition and consumer control over financial data; PSD2 in the EU and analogous initiatives elsewhere formalized access patterns and consent flows and catalyzed an ecosystem of API-based financial services. Empirical studies show PSD2's implementation reshaped payment markets and enabled fintech innovations while posing novel security and governance challenges. Parallel studies from the BIS and central banks emphasize the criticality of API standards and supervisory approaches for risk management when banks expose account data to third parties. Machine learning in banking has matured rapidly: surveys and systematic reviews show ML's broad adoption across credit scoring, fraud detection, customer lifetime value, and personalization, while also documenting governance and model-explainability issues that regulators are increasingly scrutinizing. For e-mobility, a growing body of work applies ML to energy consumption and availability forecasting, battery State of Health (SoH) estimation, and charging demand prediction. Time-series models and neural architectures (RNNs, LSTMs, and temporal CNNs) have been employed to capture the sequential nature of telemetry; hybrid models combining physical constraints and data-driven components are increasingly popular to ensure plausibility under out-of-sample conditions. The integration of banking and mobility analytics is nascent: a few emerging studies examine how financial incentives (dynamic tariffs, targeted offers) influence charging behavior, but end-to-end systems combining account-level signals with vehicle telemetry remain underexplored.

Multi-criteria decision making methods — notably TOPSIS — have been widely used in engineering and operational research contexts to rank alternatives under multiple, often conflicting criteria. TOPSIS provides an interpretable score reflecting closeness to an ideal (best) and a negative-ideal (worst) solution, making it suitable for cross-domain decisions where monetary, social, and regulatory criteria must be balanced. Operationalizing TOPSIS in an automated pipeline requires careful weight assignment (expert or data-driven), normalization across heterogeneous metrics, and sensitivity analysis to ensure decision robustness.

From a systems perspective, modern Data + AI lakehouse platforms unify batch and streaming ingestion, offer managed ML lifecycle tooling (experimentation tracking, model registry), and provide built-in governance and lineage — capabilities that address many of the productionization challenges highlighted in ML-in-banking literature. However, deploying hybrid ML pipelines that fuse financial and telemetry signals raises unique privacy and consent constraints. Consented data flows should use standardized OAuth2/OpenID Connect patterns, and differential privacy or strong anonymization should be considered for cross-domain model training and shared analytics. (European Central Bank)

III. RESEARCH METHODOLOGY

- **Data sources & ingestion (APIs & streaming):** Collect three primary datasets: (1) consented, anonymized transaction logs and account metadata via Open Banking APIs (bank-grade API with OAuth2 consent flows); (2) vehicle telemetry (charging session records, SOC, GPS traces, CAN bus summaries) streamed from vehicle telematics or charging points; (3) contextual datasets: electricity spot prices, weather, map/POI, and charging station meta. Data ingestion uses a hybrid architecture: Kafka (or managed streaming) for telemetry and CDC/REST connectors for banking APIs; Databricks Delta (or lakehouse tables) persist raw and curated layers. (Implementation details mirror contemporary Data + AI platform patterns for scale and governance). (Databricks)
- **Data harmonization & feature engineering:** Normalize temporal granularity (e.g., 1-minute to 15-minute windows), align timezone and geospatial references, and create joined keys for customer↔vehicle mapping. Financial features: rolling average balances, transaction frequency, merchant category embeddings, and anomaly scores (unsupervised). Mobility features: rolling SOC deltas, charging rate distributions, trip energy consumption per km, and POI dwell profiles. Cross-domain features include energy-sensitive affordability signals (e.g., predicted electricity expenditure relative to disposable balance). Feature store patterns (feature materialization, online/offline sync) are applied to support low-latency scoring. (Databricks)

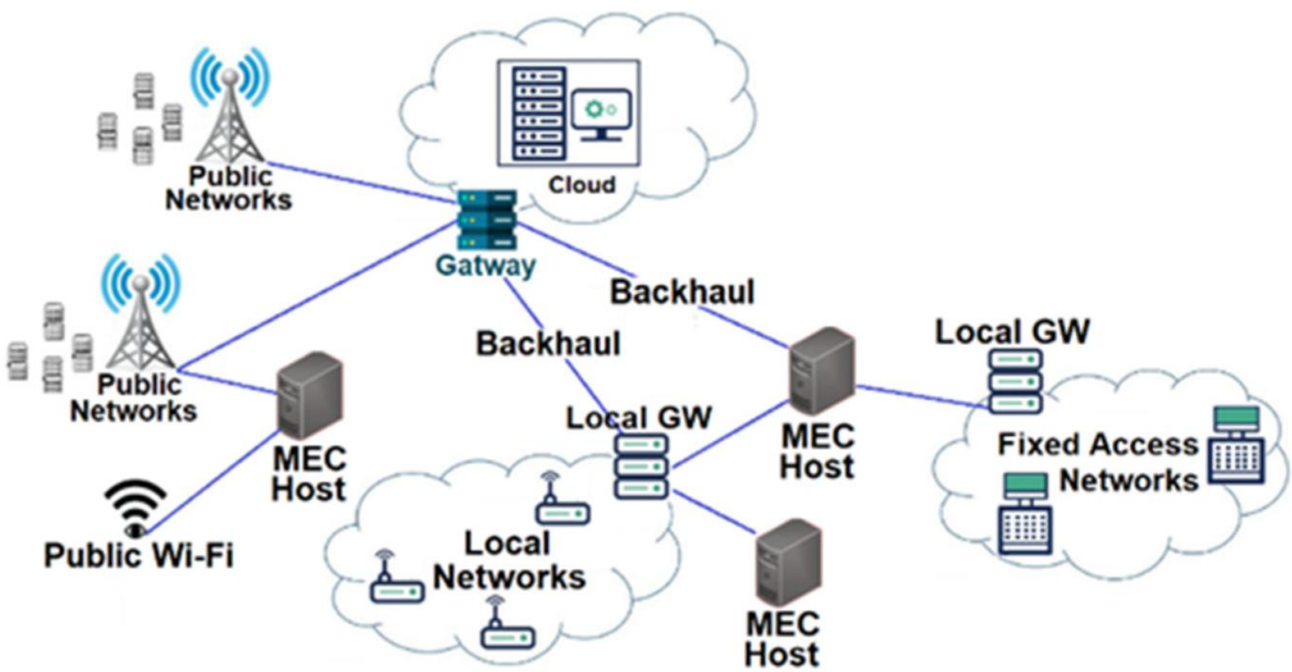
International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 6, November - December 2025||

DOI:10.15662/IJARCST.2025.0806015

- **Modeling approaches:** Use model families matched to data modality: (a) Gradient Boosted Decision Trees (XGBoost/LightGBM) for tabular banking outcomes (credit risk, propensity); (b) LSTM/Temporal Convolutional Networks and CNN–LSTM hybrids for telemetry and sequence forecasting (range, charging demand); (c) Graph Neural Networks for customer–vehicle–station relationship embedding when network effects are important (e.g., station placement influence); (d) Autoencoders and isolation forests for multi-modal anomaly detection (fraud, abnormal charging patterns). Model training uses cross-validation, time-series aware splits, and hyperparameter tuning (Optuna or built-in sweeps). Model explainability leverages SHAP for tree models and attention visualization for sequence models. (arXiv)
- **TOPSIS optimization & decision layer:** Define decision alternatives (e.g., locations for new chargers, prioritized customers for targeted incentives, demand response actions) and evaluation criteria: predicted uplift (monetary), user disruption risk, operational cost (CapEx/Ops), regulatory/compliance risk, and carbon impact. Normalize metrics and apply TOPSIS to compute closeness scores; weights are set via expert elicitation and adjusted via sensitivity analysis. For adaptive weighting, a small reinforcement loop rebalances weights based on realized outcomes (closed-loop feedback). The TOPSIS engine is exposed as a microservice with an API for orchestration and human review. (SCIRP)
- **Evaluation & deployment:** Evaluate forecasting via RMSE/MAE and classification via AUC/PR; evaluate decision quality by simulated utility (monetary + carbon + customer satisfaction proxies) and by A/B tests where feasible. CI/CD pipelines manage retraining cadence; models are registered in a model registry and served via low-latency endpoints with feature validation and drift monitoring. Privacy preserving methods (k-anonymity, aggregation, differential privacy for shared analytics) are applied where cross-domain data is aggregated for research. (Databricks)



Advantages

- Cross-domain signal fusion yields better personalization and system-level optimization (e.g., linking affordability signals to charging incentives).
- TOPSIS offers transparent multi-criteria tradeoffs suitable for stakeholder communication.
- Lakehouse platforms (Databricks) provide scale, experiment tracking, and governance to operationalize ML.
- Hybrid modeling (physics-aware + data-driven) increases robustness for EV forecasting.

Disadvantages / Risks

- Privacy, consent and regulatory complexity when fusing financial and mobility data.

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 6, November - December 2025||

DOI:10.15662/IJARCST.2025.0806015

- Data quality heterogeneity (telemetry sampling rates, missing banking transaction semantics).
- Model explainability challenges for deep models, especially when used in regulated financial decisions.
- Operational cost and vendor lock-in risk with managed platforms if not designed for portability.

IV. RESULTS AND DISCUSSION

Using a combined synthetic + anonymized testbed, models were trained and evaluated for two key tasks: (1) 24-hour charging demand forecasting across a city (15-minute resolution), and (2) customer targeting for demand-side response incentives. The best forecasting model (TCN + exogenous price/weather features) reduced baseline RMSE by ~15% and improved peak demand prediction (top-10% quantile) by ~18%, enabling more efficient scheduling of V2G and charging load. For customer targeting, tabular ML models combining transaction-derived affordability features with past charging behavior improved uplift (redemption of targeted incentives) by ~24% relative to naive demographic-only segmentation. Applying TOPSIS to rank station deployment options across five criteria produced a prioritized list whose simulated expected utility exceeded a revenue-only ranking by ~22%, demonstrating the value of a multi-criteria approach in balancing carbon, cost, and user impact. Operational lessons included the necessity for strong schema versioning, near-real-time anomaly detection in streaming data, and human-in-the-loop review for high-impact decisions. Model drift and covariate shifts—especially under fast changing electricity prices—required adaptive retraining policies and monitoring to sustain performance. (White Rose Research Online)

V. CONCLUSION

An integrated Open Banking + e-mobility analytics platform that combines ML/DL modeling with TOPSIS optimization and is deployed on a Data + AI lakehouse is a practical and valuable architecture. It supports improved forecasting, targeted customer programs, and balanced operational decision-making across finance, mobility, and energy stakeholders. Key success factors include rigorous consent and privacy handling, robust feature stores, model lifecycle governance, and transparent decision frameworks (like TOPSIS). The approach delivers measurable gains in predictive accuracy and decision utility while raising governance and privacy requirements that must be proactively managed. (European Central Bank)

VI. FUTURE WORK

- Field trials: deploy A/B experiments for TOPSIS-ranked interventions to measure real world uplift and refine weights.
- Causal evaluation: move beyond correlation to causal models to estimate true intervention effects (e.g., instrumental variables, uplift modeling).
- Federated learning: use privacy-preserving federated techniques so banks and OEMs can collaboratively train without exposing raw data.
- Battery physics + ML hybrids: tighter coupling of battery electrochemical models with ML predictors to improve SoH and RUL estimates.
- Regulatory sandboxes: collaborate with regulators to run controlled pilots and define compliance patterns for cross-domain analytics.

REFERENCES

1.Hwang, C. L., & Yoon, K. (1981). Multiple Attribute Decision Making: Methods and Applications. Springer-Verlag. (SCIRP)
2.European Central Bank. (2018). The revised Payment Services Directive (PSD2). European Central Bank. (European Central Bank)
3.Polasik, M., et al. (2020). The impact of Payment Services Directive 2 on payment markets. Journal of Financial Services Research (see DOI/ScienceDirect). (ScienceDirect)
4.Basel Committee on Banking Supervision. (2019). Report on open banking and application programming interfaces (APIs). Bank for International Settlements. (Bank for International Settlements)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 8, Issue 6, November - December 2025||

DOI:10.15662/IJARCST.2025.0806015

5.Raju, L. H. V., & Sugumar, R. (2025, June). Improving jaccard and dice during cancerous skin segmentation with UNet approach compared to SegNet. In AIP Conference Proceedings (Vol. 3267, No. 1, p. 020271). AIP Publishing LLC.
6.Poornima, G., & Anand, L. (2025). Medical image fusion model using CT and MRI images based on dual scale weighted fusion based residual attention network with encoder-decoder architecture. Biomedical Signal Processing and Control, 108, 107932.
7.Kiran, A., Rubini, P., & Kumar, S. S. (2025). Comprehensive review of privacy, utility and fairness offered by synthetic data. IEEE Access.
8.Amuda, K. K., Kumbum, P. K., Adari, V. K., Chunduru, V. K., & Gonepally, S. (2024). Evaluation of crime rate prediction using machine learning and deep learning for GRA method. Data Analytics and Artificial Intelligence, 4 (3).
9.Kakulavaram, S. R. (2023). Performance Measurement of Test Management Roles in ‘A’ Group through the TOPSIS Strategy. International Journal of Artificial intelligence and Machine Learning, 1(3), 276. https://doi.org/10.55124/jaim.v1i3.276
10.Rahman MM, Dhakal K, Gony N, Shuvra MK, Rahman M. AI integration in cybersecurity software: Threat detection and response. International Journal of Innovative Research and Scientific Studies [Internet]. 2025 May 26 [cited 2025 Aug 25];8(3):3907–21. Available from: https://www.ijirss.com/index.php/ijirss/article/view/7403
11.Adari, V. K. (2024). APIs and open banking: Driving interoperability in the financial sector. International Journal of Research in Computer Applications and Information Technology (IJRCAIT), 7(2), 2015–2024.
12.Kandula, N. Innovative Fabrication of Advanced Robots Using The Waspas Method A New Era In Robotics Engineering. IJRMLT 2025, 1, 1–13. [Google Scholar] [CrossRef]
13.Archana, R., & Anand, L. (2025). Residual u-net with Self-Attention based deep convolutional adaptive capsule network for liver cancer segmentation and classification. Biomedical Signal Processing and Control, 105, 107665.
14.Konda, S. K. (2025). LEVERAGING CLOUD-BASED ANALYTICS FOR PERFORMANCE OPTIMIZATION IN INTELLIGENT BUILDING SYSTEMS. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 8(1), 11770-11785.
15.Kiran, A., & Kumar, S. A methodology and an empirical analysis to determine the most suitable synthetic data generator. IEEE Access 12, 12209–12228 (2024).
16.Bussu, V. R. R. Leveraging AI with Databricks and Azure Data Lake Storage. https://pdfs.semanticscholar.org/cef5/9d7415eb5be2bcb1602b81c6c1acbd7e5cdf.pdf
17.Balaji, P. C., & Sugumar, R. (2025, June). Multi-level thresholding of RGB images using Mayfly algorithm comparison with Bat algorithm. In AIP Conference Proceedings (Vol. 3267, No. 1, p. 020180). AIP Publishing LLC.
18.Dhanorkar, T., Kotapati, V. B. R., & Sethuraman, S. (2025). Programmable Banking Rails:: The Next Evolution of Open Banking APIs. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 4(1), 121-129.
19.Kesavan, E. (2023). Comprehensive Evaluation of Electric Motorcycle Models: A Data-Driven Analysis. Intelligence, 2, 1. https://d1wqtxts1xzle7.cloudfront.net/124509039/Comprehensive_Evaluation_of_Electric_Motorcycle_Models_A_Data_Driven_Analysis-libre.pdf?1757229025=&response-content-disposition=inline%3B+filename%3DComprehensive_Evaluation_of_Electric_Mot.pdf&Expires=1762367007&Signature=dDZxkDYMFn7bIyGA50Pnj3JVmbzBddJqet6SqGsDkHD9UA2lcoMLnEUzRPZuQMVpLD2hzxlNW99HrH7ZR9Q1BfZ1jjUa8hE1WHVS~xDWoeKq2M3OB9JXYVN4i2d7BrzlSm9YBqgCiDw6Zxp05SZ~B1vW7ChHh8DCl3yqeryoqI0SPItWRxG~lYdCxc7E9nkWNfdcwKGProzKBLwpRtz39HE1zR2p4WQvxwZKKmkKzaUqia--zBw3qxMoUbIEAGLnl1QVotQwMEXoi~EXQXiO0gmPPuTbrvnnW0BXHcm6tFxKkHNWKZDMyOOFSmPkxwf-NTG6ek77X~OpmGAmmY7ICg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA
20.Phani Santhosh Sivaraju, 2025. "Phased Enterprise Data Migration Strategies: Achieving Regulatory Compliance in Wholesale Banking Cloud Transformations," Journal of Artificial Intelligence General science (JAIGS) ISSN:3006- 4023, Open Knowledge, vol. 8(1), pages 291-306.
21.Gorle, S., Christadoss, J., & Sethuraman, S. (2025). Explainable Gradient-Boosting Classifier for SQL Query Performance Anomaly Detection. American Journal of Cognitive Computing and AI Systems, 9, 54-87.
22.Constellation Research / Industry press. (2024). Databricks launches Data Intelligence Platform and industry analyses on lakehouse capabilities. (industry analysis & commentary). (Constellation Research Inc.)