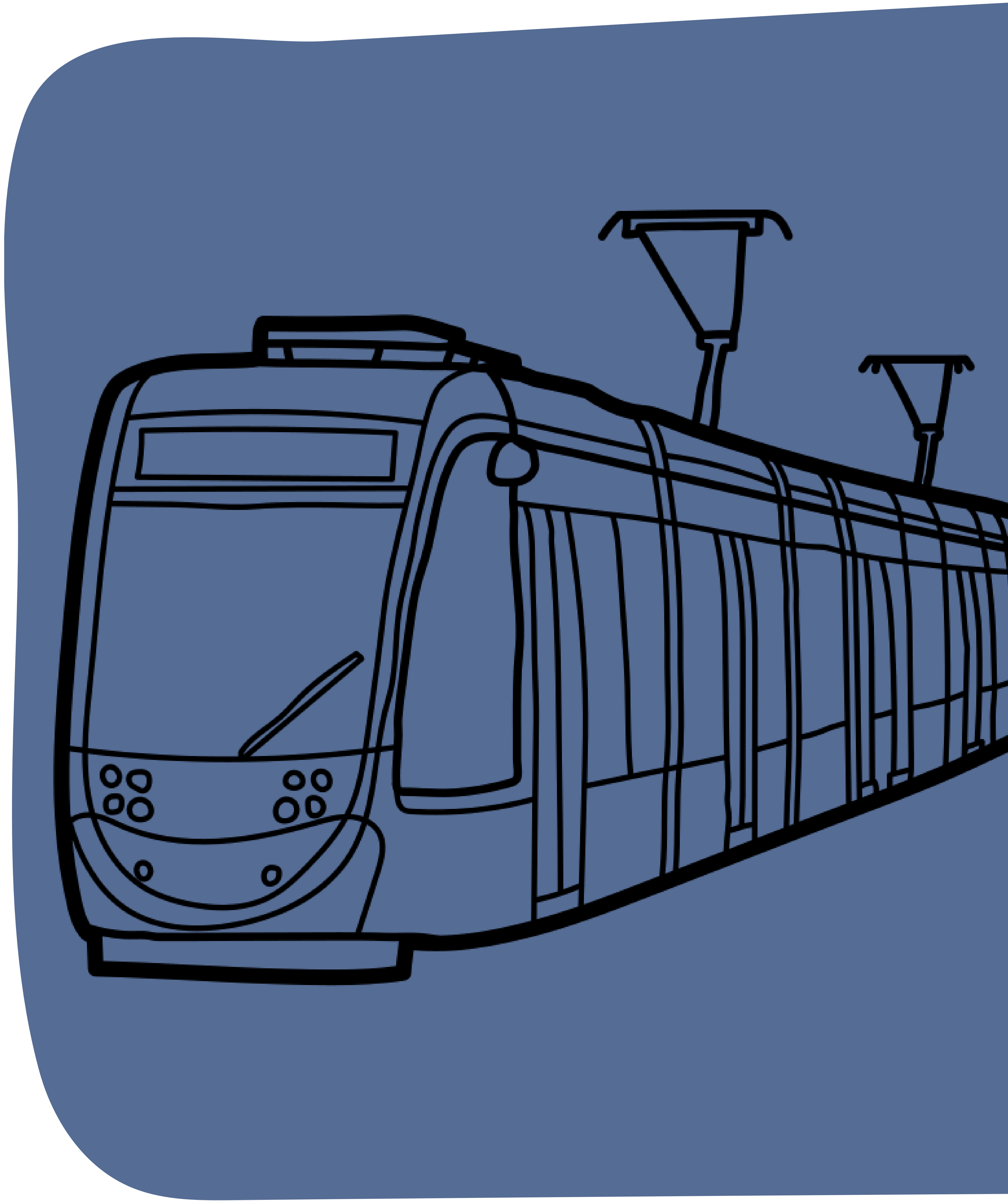
FAULT PREDICTION AND CALCULATING REMAINING USEFUL LIFE IN APU SYSTEM WITHIN METRO TRAIN

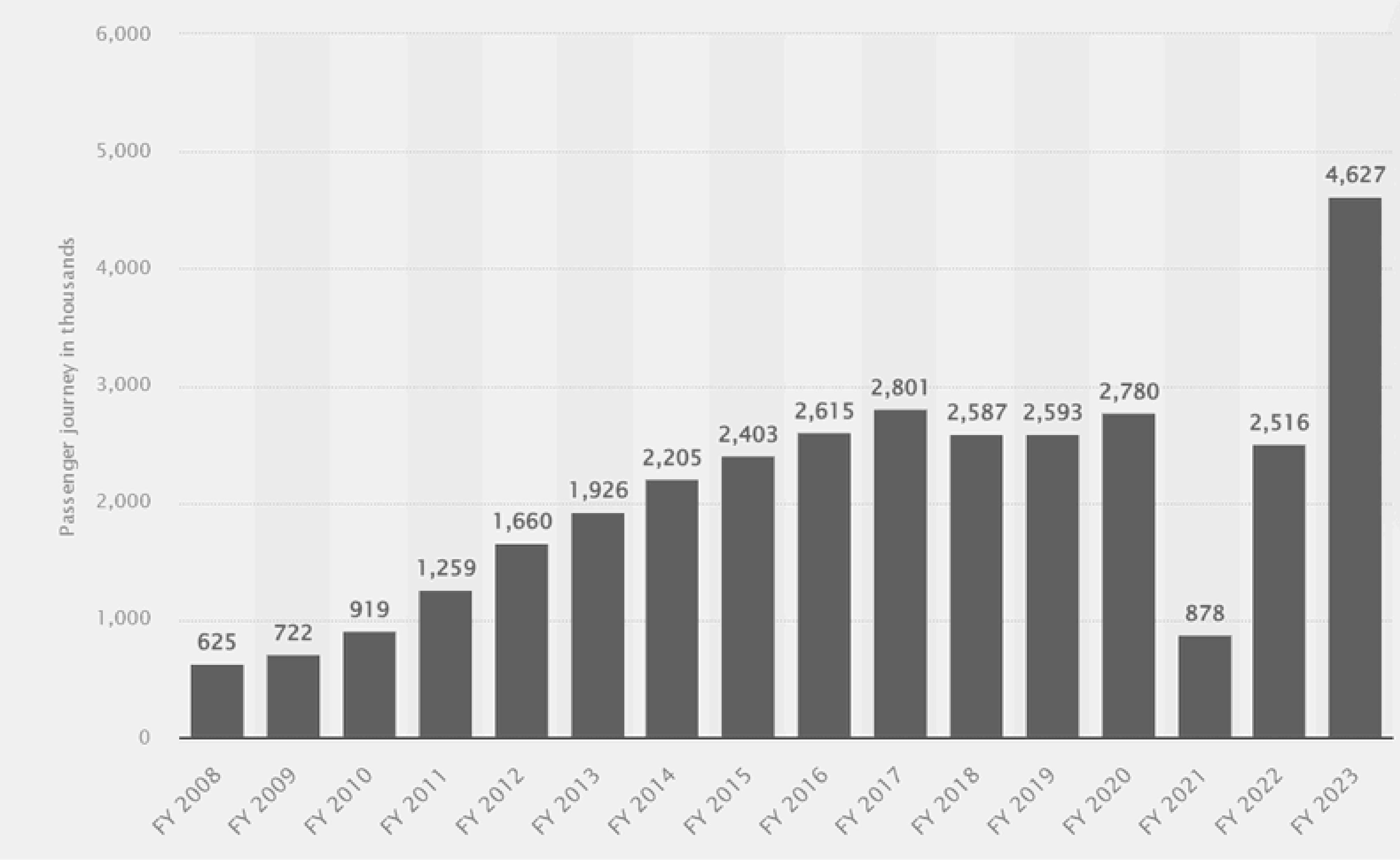
MEHER SIDHU AND UMA GABA

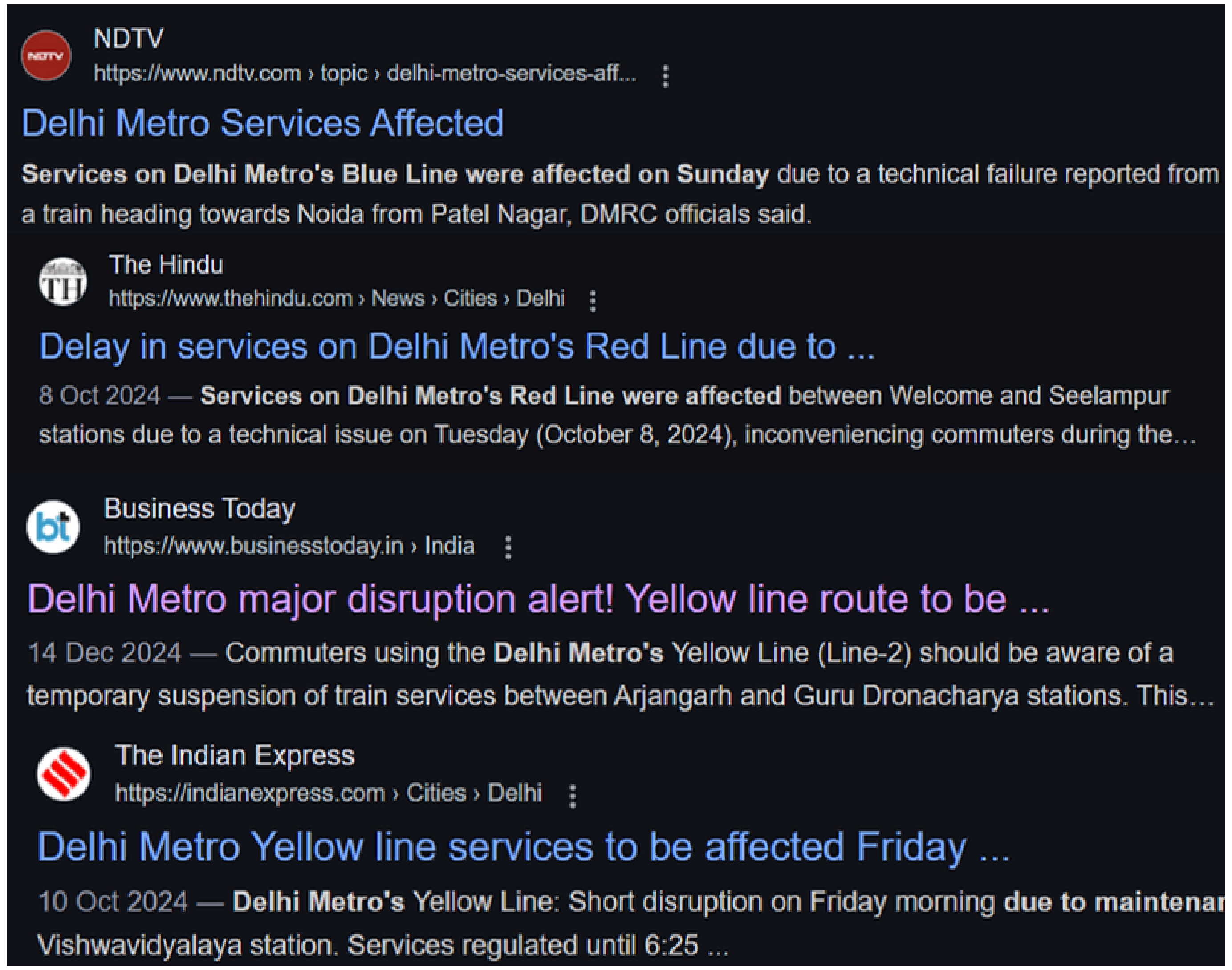


VALIDITY OF OUR PROBLEM STATEMENT

The Air Production Unit (APU) plays a vital role in feeding various units, including the secondary suspension responsible for maintaining vehicle height.
The failure of APU necessitates immediate train removal for repair.

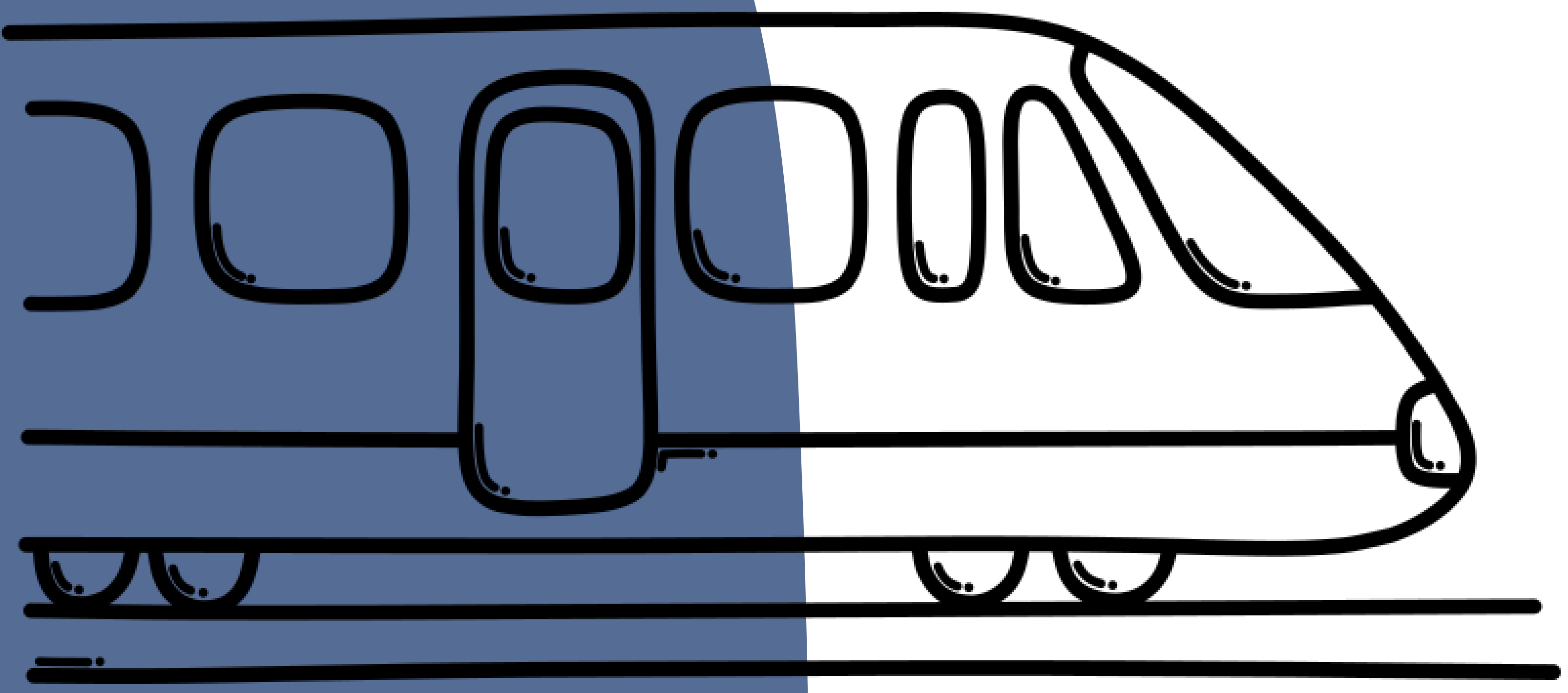
Average daily ridership of metro in Delhi, India from financial year 2008 to 2023(in 1,000 passenger journeys)





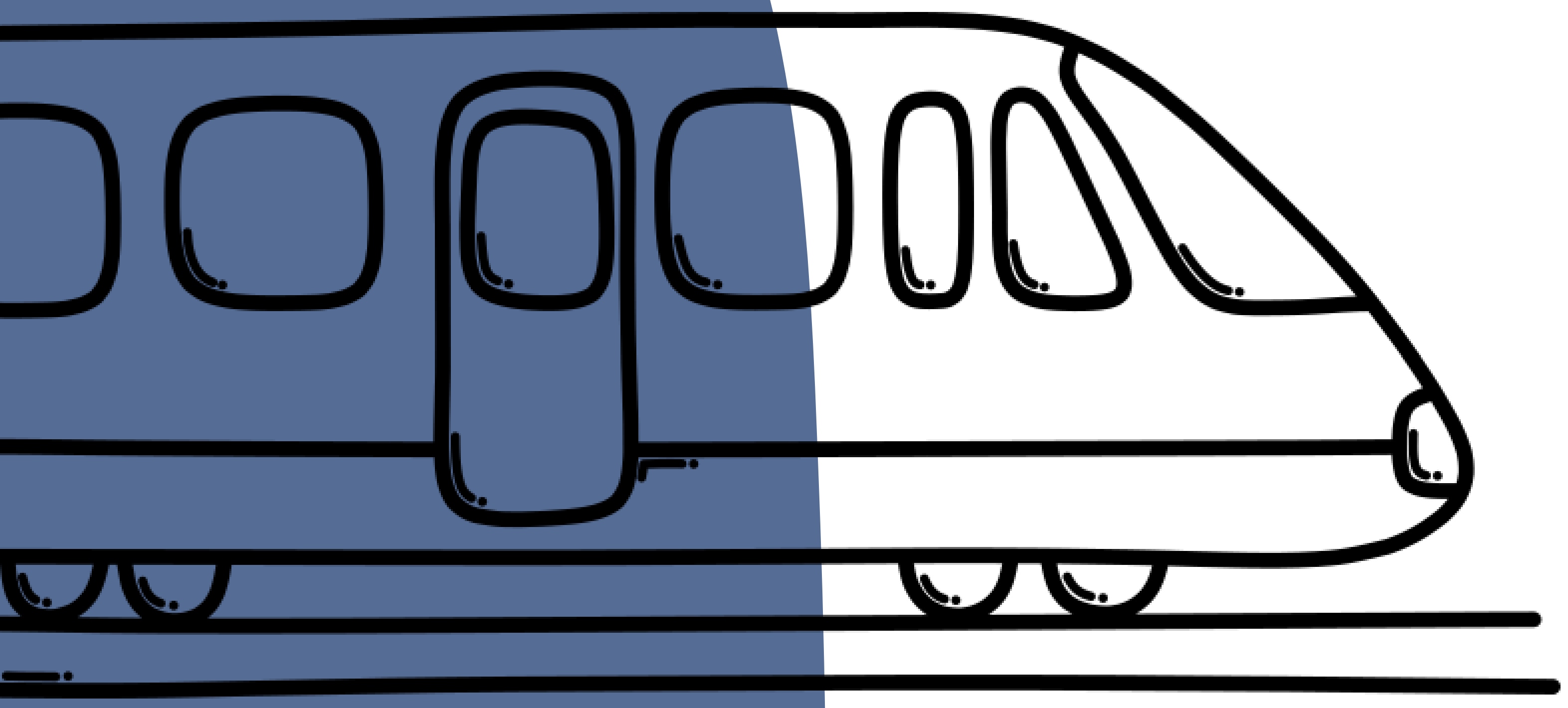
In financial year 2023, more than 4.6 million passenger journeys took place in metro in Delhi daily.

"India: Average Daily Ridership of Metro in Delhi 2023." n.d. Statista. Accessed March 11, 2025. https://www.statista.com/statistics/1240001/india-average-daily-ridership-of-public-transport-in-delhi/.

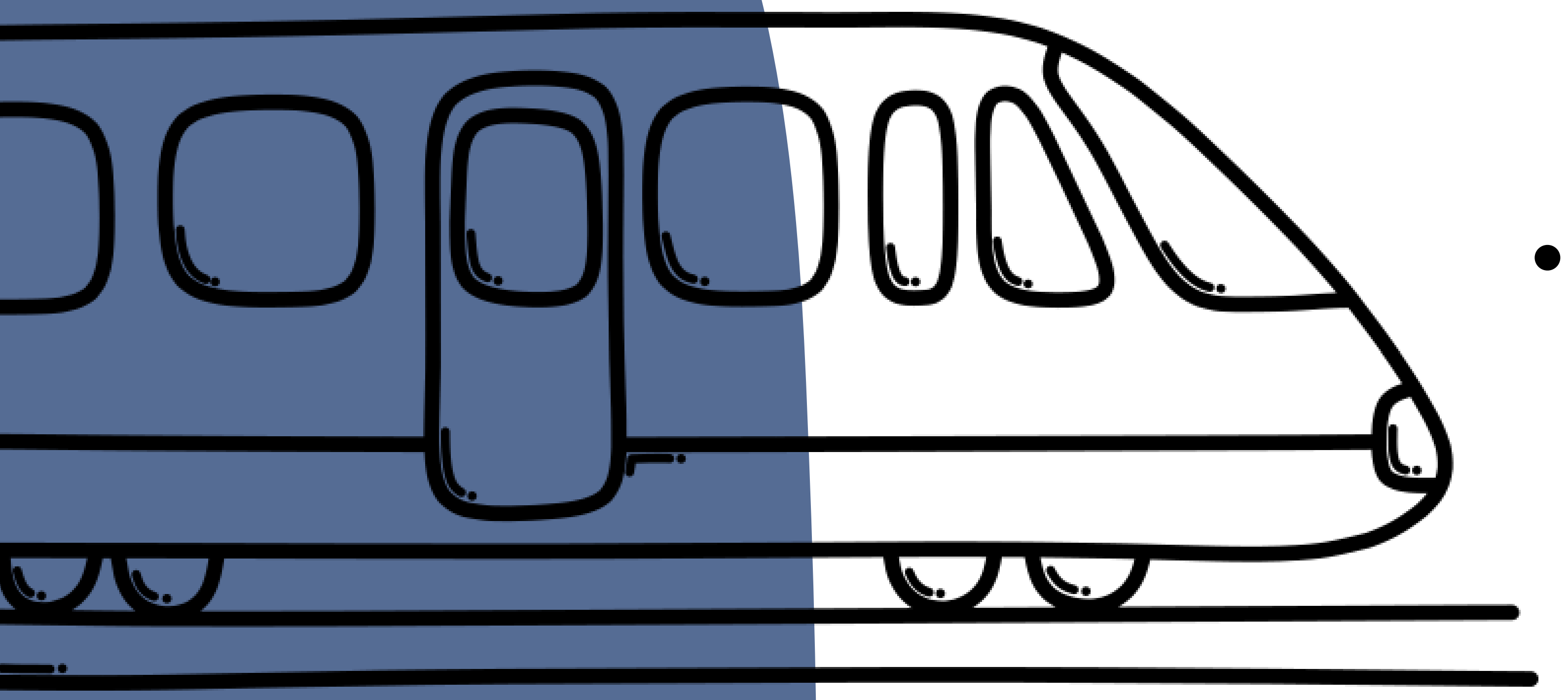


LITERATURE SURVEY

- **Earlier Techniques for fault identification**
  - Vibration Analysis using FFT & Time-Domain Features (TDF)
    - Attoui, Issam & Meradi, Hazem & Boulkroune, Ramzi & Saidi, Riad & Grid, Azzeddine. (2013). Fault detection and diagnosis in rotating machinery by vibration monitoring using FFT and Wavelet techniques. 401-406. 10.1109/WoSSPA.2013.6602399.
  - Rule-Based Threshold Monitoring
    - Najjar, Ayat & Ashqar, Huthaifa & Hasasneh, Ahmad. (2023). Predictive Maintenance of Urban Metro Vehicles: Classification of Air Production Unit Failures Using Machine Learning.



- The significance of predictive maintenance
  - Cost Savings
    - Zhu, T., Ran, Y., Zhou, X., & Wen, Y. (2019). A survey of predictive maintenance: Systems, purposes and approaches. arXiv preprint arXiv:1912.07383.
  - Extended Equipment Lifespan and Sustainability
    - Kane, A. P., Kore, A. S., Khandale, A. N., Nigade, S. S., & Joshi, P. P. (2022). Predictive maintenance using machine learning. arXiv preprint arXiv:2205.09402.
- Predictive Maintenance of Metro APU
  - Classify Whether a Failure is Imminent or Not
    - Nair, Vishak & M, Premalatha & Ramalingam, Srinivasa Perumal & M, Braveen. (2024). Enhancing Metro Rail Efficiency: A Predictive Maintenance Approach Leveraging Machine Learning and Deep Learning Technologies. 10.21203/rs.3.rs-4319916/v1.
  - Classification of Faults
    - Najjar, Ayat & Ashqar, Huthaifa & Hasasneh, Ahmad. (2023). Predictive Maintenance of Urban Metro Vehicles: Classification of Air Production Unit Failures Using Machine Learning.



- Significance of Remaining Useful Life (RUL) prediction and transparency in predictive maintenance
  - RUL helps optimize maintenance schedules, reduce costs, and prevent catastrophic system failures.
  - Accurate RUL estimation allows timely intervention before component breakdowns.
  - Lack of transparency leads to reduced trust and limited deployment.
  - Incorporating XAI techniques, like SHAP , provides insights into feature contributions,
    - Youness, G., & Aalah, A. (2023). An Explainable Artificial Intelligence Approach for Remaining Useful Life Prediction. *Aerospace*, *10*(5), 474. https://doi.org/10.3390/aerospace10050474

DATASET

MetroPT dataset, collected in 2022 from Porto, Portugal, as part of the eXplainable Predictive Maintenance (XPM) project. Multivariate time-series sensor data from Air Production Units (APUs) in metro trains.

Veloso, B., Ribeiro, R. P., Gama, J., & Pereira, P. M. (2022). The MetroPT dataset for predictive maintenance. Scientific data, 9(1), 764. https://doi.org/10.1038/s41597-022-01877-3

Size & Structure:

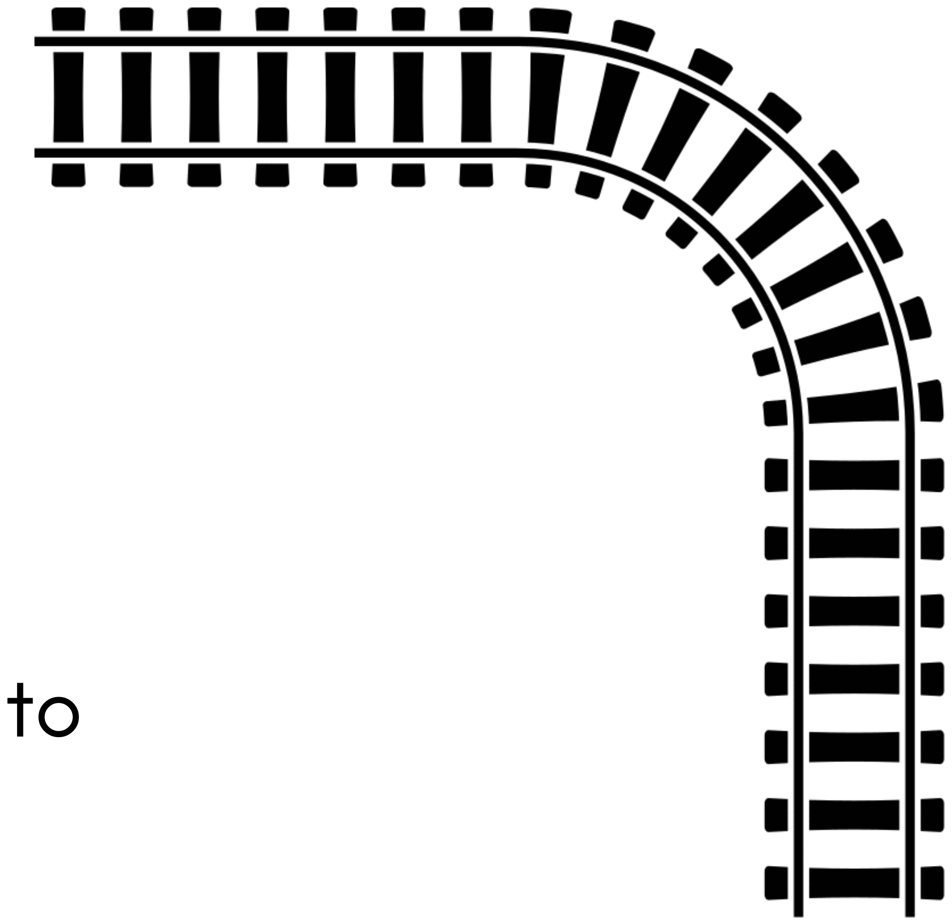
- 10,979,547 data points
- 21 features (analog and digital sensors, plus GPS)
- No missing values
- Three main fault types:
  - Air leak in the dryer
  - Air leak in the clients
  - Oil leak in the compressor

DATASET

Features:

- Analog: TP2, TP3, H1, DV_pressure, Reservoirs, Oil_Temperature, Flowmeter, Motor_current
- Digital: COMP, DV_electric, TOWERS, MPG, LPS, Oil Level, CAudal_impulses

FEATURES PREPROCESSING



- Timestamp Handling: Converted 'timestamp' from string to DateTime for temporal analysis.
- Feature Selection: Dropped GPS features.
- Label Engineering: Added ‘fault’ and ‘fault type’ columns to create ground truth for supervised learning.
- Sequence Creation: Created time sequences for temporal modeling (essential for LSTM).
- Train/Test Split: Dataset split into 80% training and 20% testing

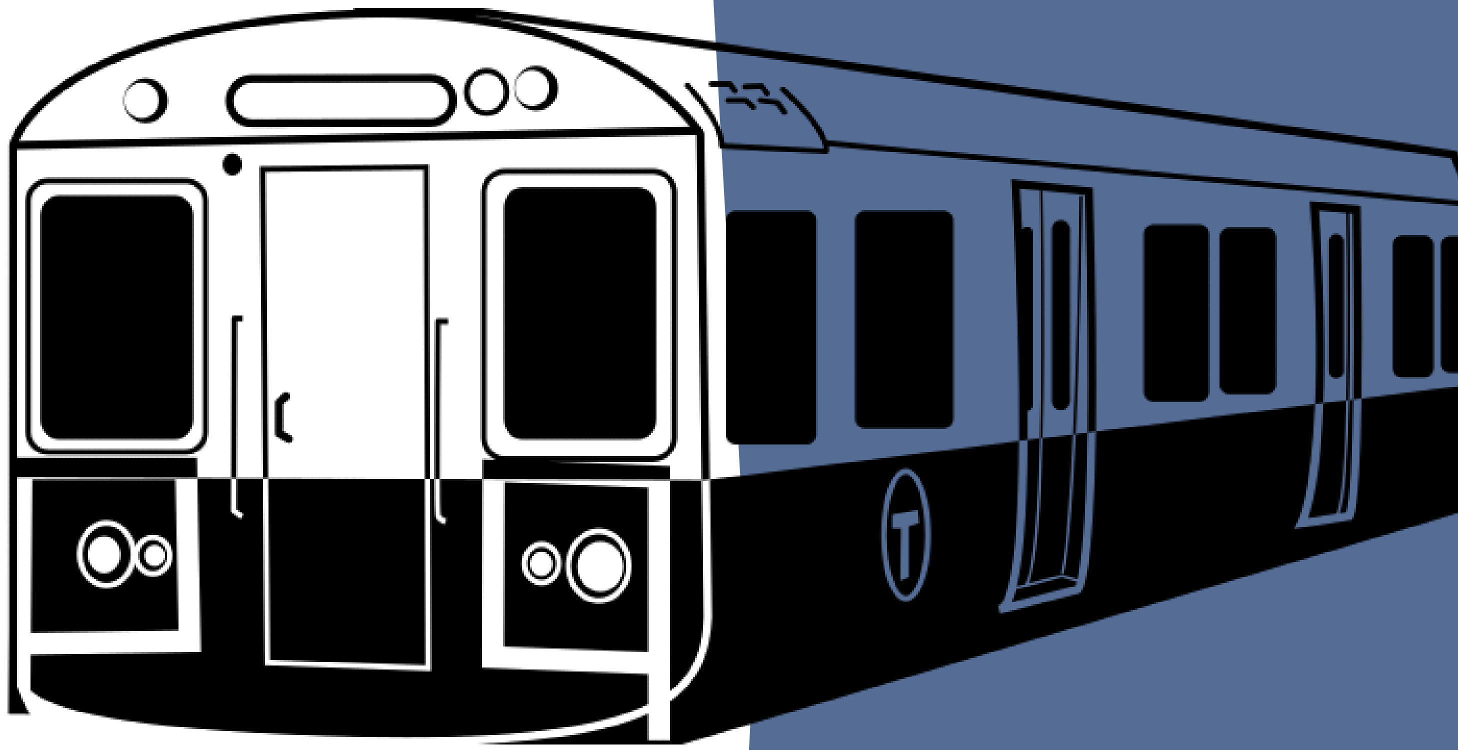
ML METHODOLOGY

Prediction of Fault - Random Forest

- High Precision(98.2) Recall(98.7) and Accuracy(99.4)
- Ensemble approach - reduces overfitting
- Computationally light
    - Nair, Vishak & M, Premalatha & Ramalingam, Srinivasa Perumal & M, Braveen. (2024). Enhancing Metro Rail Efficiency: A Predictive Maintenance Approach Leveraging Machine Learning and Deep Learning Technologies. 10.21203/rs.3.rs-4319916/v1.

Classification of Fault - Random Forest

- Multiclass classification - 97% F1 Score
- Ensemble approach - reduces overfitting
- Computationally light
    - Najjar, Ayat & Ashqar, Huthaifa & Hasasneh, Ahmad. (2023). Predictive Maintenance of Urban Metro Vehicles: Classification of Air Production Unit Failures Using Machine Learning.

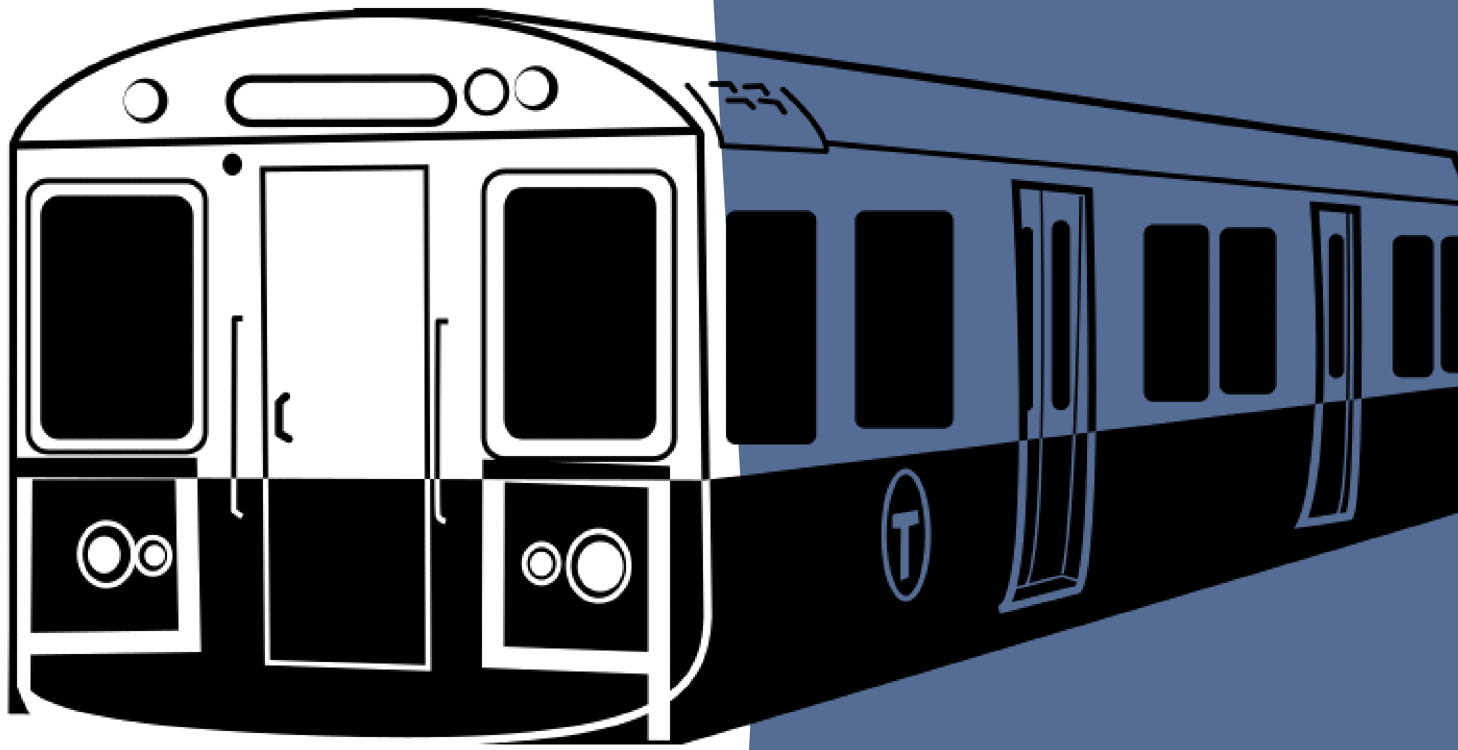


Remaining Useful Life Estimation - LSTM

- Captures Temporal Dependencies
- Handles Nonlinear and Complex Patterns
- Outperforms traditional ML models (e.g., Random Forest, SVM) in early RUL prediction
    - Safavi, V., Mohammadi Vaniar, A., Bazmohammadi, N., Vasquez, J. C., & Guerrero, J. M. (2024). Battery Remaining Useful Life Prediction Using Machine Learning Models: A Comparative Study. Information, 15(3), 124. https://doi.org/10.3390/info15030124

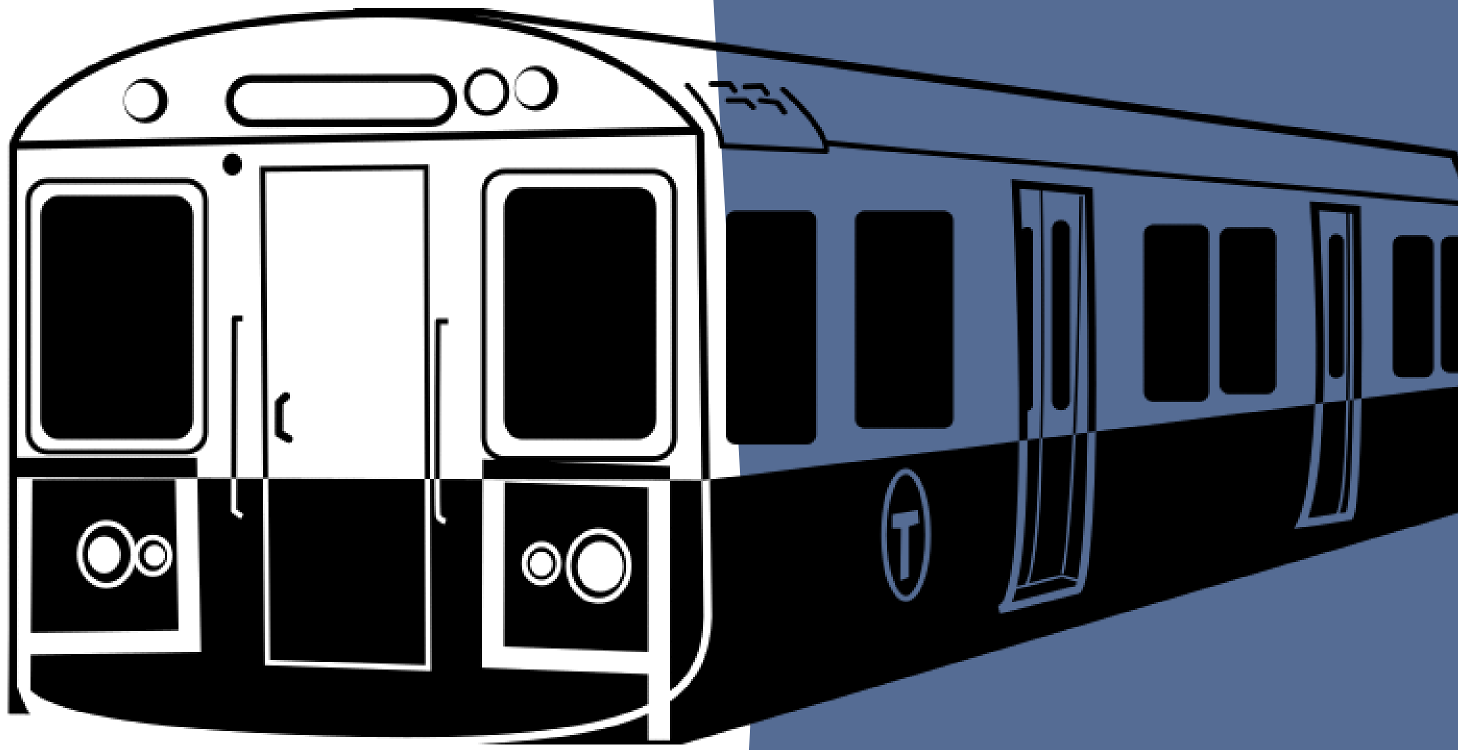
Remaining Useful Life Estimation - XGBoost

- High accuracy in RUL prediction
- Handles Nonlinear and Complex Patterns
- Highly optimized for structured/tabular data
    - Zhang, W., Wang, Y., & Li, X. (2021). Remaining Useful Life Prediction of Lithium-Ion Batteries Using XGBoost. IEEE Access, 9, 123456–123467. Safavi, V., Mohammadi Vaniar, A., Bazmohammadi, N., Vasquez, J. C., & Guerrero, J. M. (2024). Battery Remaining Useful Life Prediction Using Machine Learning Models: A Comparative Study. Information, 15(3), 124. https://doi.org/10.3390/info15030124



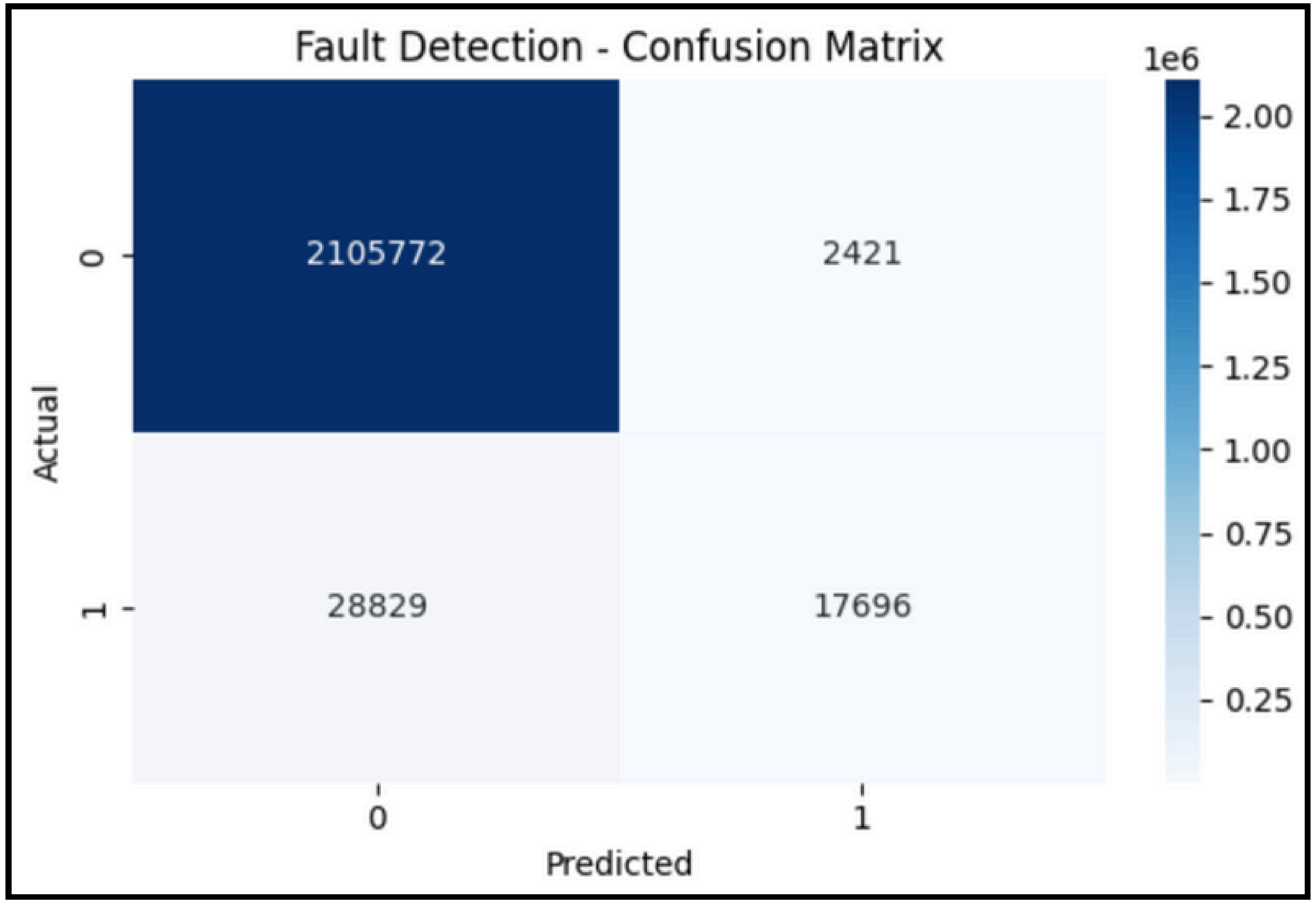
CHALLENGES

- LSTM model training caused RAM issues due to large batch sizes.
    - Reduced batch size to fit within memory constraints and ensure model convergence.
- SHAP (SHapley Additive exPlanations) was incompatible with Kaggle environment.
    - Shifted to Google Colab, which provided better support for SHAP and custom dependencies.
- Kaggle had numerous dependency conflicts when using advanced explainability libraries.
- Difficulty implementing prediction-to-classification cascading logic.



PERFORMANCE METRICS

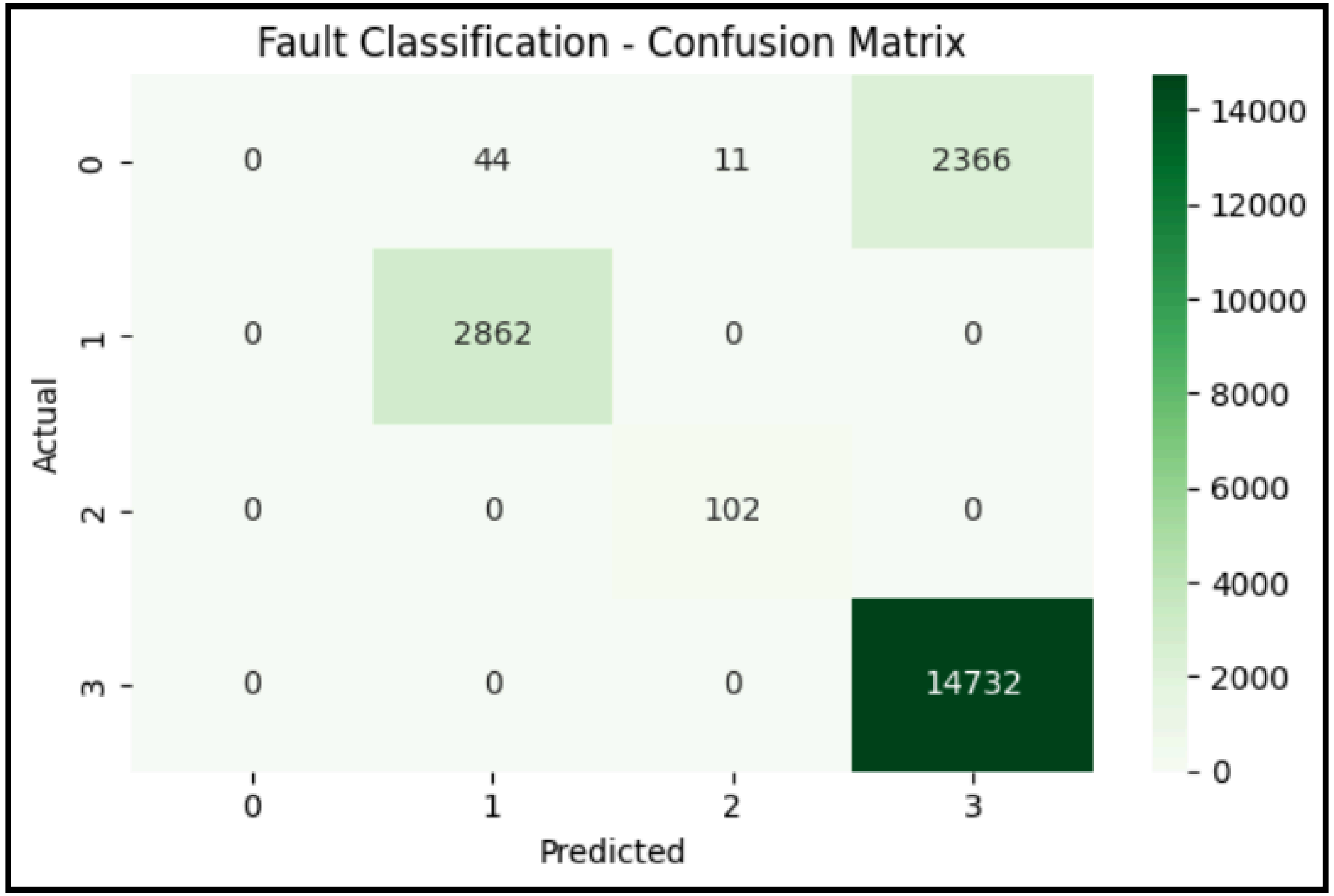
```
Fault Detection - Classification Report:
              precision    recall  f1-score   support

           0       0.99      1.00      0.99   2108193
           1       0.88      0.38      0.53     46525

    accuracy                           0.99   2154718
```



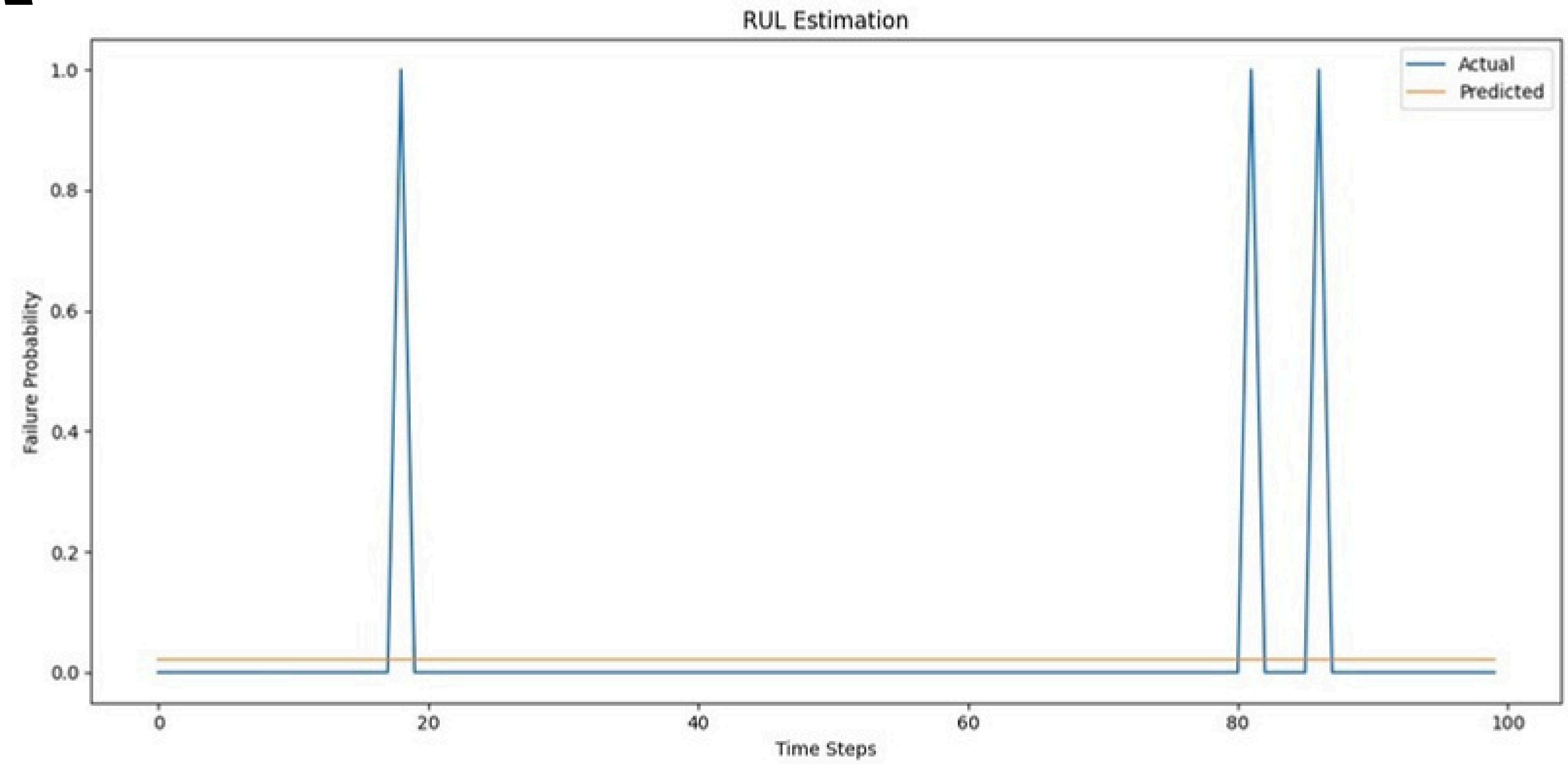
PERFORMANCE METRICS

Fault Classification - Classification Report:

	precision	recall	f1-score	support
0	0.00	0.00	0.00	2421
1	0.98	1.00	0.99	2862
2	0.90	1.00	0.95	102
3	0.86	1.00	0.93	14732
accuracy			0.88	20117

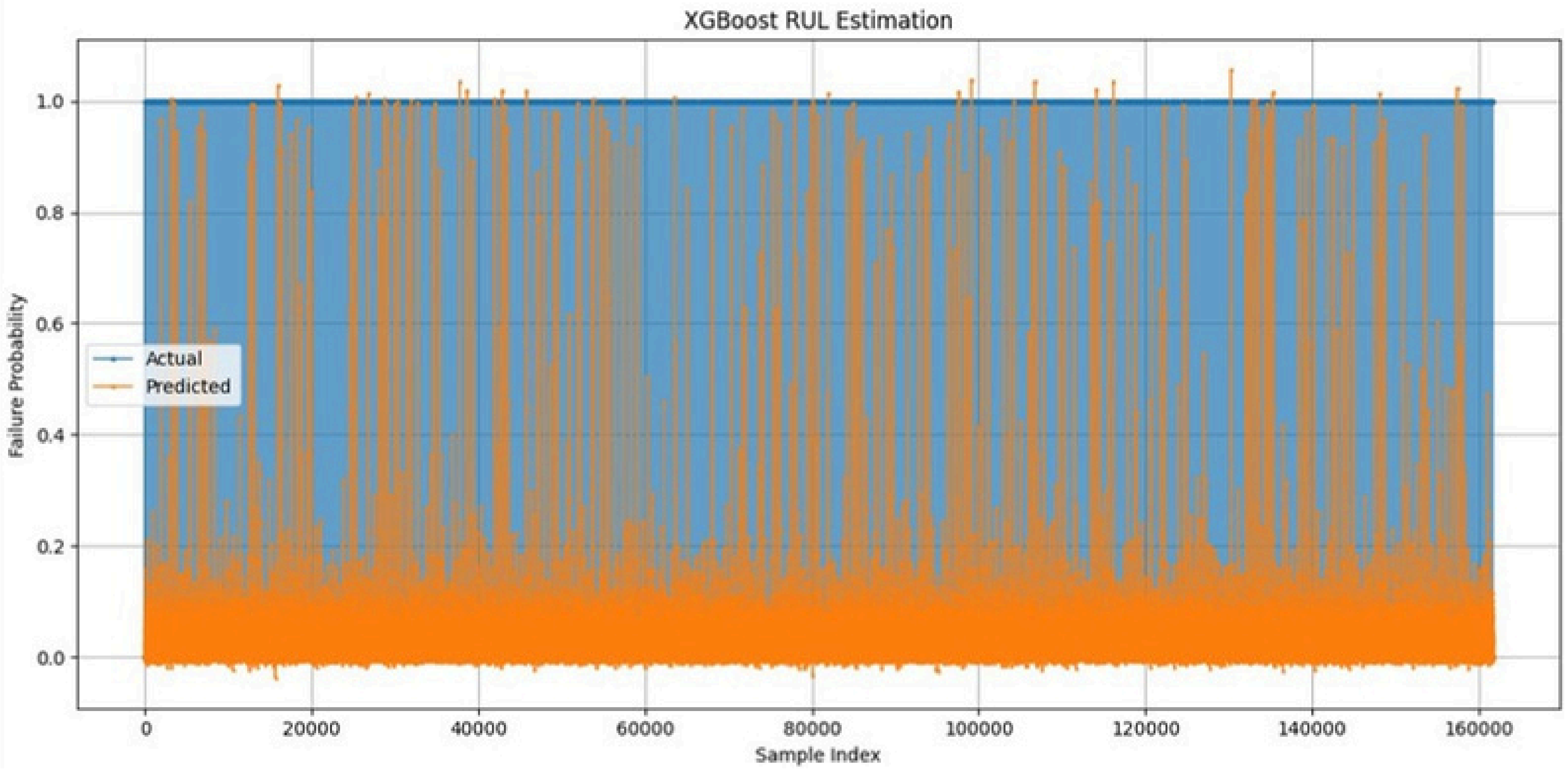


PERFORMANCE METRICS - LSTM



PERFORMANCE METRICS XGBOOST

MAE: 0.0397
RMSE: 0.1398



- Zhang, W., Wang, Y., & Li, X. (2021). Remaining Useful Life Prediction of Lithium-Ion Batteries Using XGBoost. IEEE Access, 9, 123456–123467. Safavi, V., Mohammadi Vaniar, A., Bazmohammadi, N., Vasquez, J. C., & Guerrero, J. M. (2024). Battery Remaining Useful Life Prediction Using Machine Learning Models: A Comparative Study. Information, 15(3), 124. https://doi.org/10.3390/info15030124

DEPLOYABILITY

Our model can be effectively integrated into existing Supervisory Control and Data Acquisition (SCADA) systems and sensor infrastructures. This integration enables real-time data acquisition and analysis, facilitating timely maintenance decisions.

- Hyderabad Metro Rail has implemented AI and IoT technologies for predictive maintenance, utilizing their existing CBTC (Communications-Based Train Control) systems to monitor and reduce maintenance costs .
- Metro de Madrid adopted machine learning models for predictive maintenance in tunnels, integrating data from various sensors to anticipate and optimize maintenance actions .

DEPLOYABILITY

Enhanced Model Transparency - SHapley Additive exPlanations

- Offers clear insights into how each input feature influences model predictions
- Transparency is crucial for understanding model behavior for Predictive maintenance
  - He, Z., Yang, Y., Fang, R., Zhou, S., Zhao, W., Bai, Y., Li, J., & Wang, B. (2023). Integration of Shapley additive explanations with random forest model for quantitative precipitation estimation of mesoscale convective systems. Frontiers in Environmental Science, 10, 1057081. https://doi.org/10.3389/fenvs.2022.1057081