FULafia Journal of Science & Technology
ISSN (print) 2449-0954 \ ISSN (online) 2636-4972
https://lafiascijournals.org.ng/index.php/fjst
Manuscript => Received: July, 2026; Accepted: Nov., 2026; Published: March, 2027 https://doi.org/10.62050/fjst2027.v11n1.889

A COMPARATIVE ANALYSIS OF MACHINE LEARNING AND DEEP LEARNING MODELS FOR TRANSPORTATION DATA ANALYTICS USING CRISP-ML METHODOLOGY AND POWER BI

Ahmed Kasimu Dahiru, Ifeanyi Christopher Eke and Muhammad Dahiru Liman
Department of Computer Science, Federal University of Lafia, Nasarawa State, Nigeria
*Corresponding email: ahmedqasim37@gmail.com

ABSTRACT
In the modern transport sector, the availability of digital devices and GPS sensors, which are attached to modern vehicles, has resulted in the easy accessibility of datasets. The Power BI dashboard provides an interactive means of visualizing and exploring the analytical results. This study investigates and compares the performance of some selected machine learning (ML), MLP, and deep learning (DL) models for transportation of data analytics within the framework of the Cross-Industry Standard Process for Machine Learning (CRISP-ML). The machine learning considered in this work comprises of Logistic regression (LR), Support vector machine (SVM), Decision Tree (DT) and Random forest (RF); the deep learning (DL) comprises of Long short term memory (LSTM), and Bi-Long Short-Term memory (Bi-LSTM), and Multi-layer perceptron (MLP). In order to carry out our analyses for the model comparison, an Uber dataset, which contains 1,156 records, was used. Handling of missing variables, feature engineering, and normalization are the data preparation steps used, while Synthetic minority oversampling techniques (SMOTE) were used for class balancing. The classification results showed that random forest, and decision tree achieved an accuracy of 93.2, and 92.6 % respectively while MLP achieved the highest accuracy among neural networks. In regression, linear regression achieved the lowest MAE of 2.167 and the highest R^2 of 0.63. All the models performed well in classifying trip category. The result also clearly indicated that ML provides more interpretability and efficiency over complex deep learning models.

Keywords: Transportation analytics, Machine learning, Deep learning, CRISP-ML, Power BI, Predictive modeling

INTRODUCTION
The transportation sector plays a vital role in economic development and urban growth. Transportation systems are now significant in data sources, thanks to the recent developments in digital technology. Information is constantly produced by GPS-enabled devices, mobile apps, ride-hailing services, and sensor-based infrastructures. These data streams provide important opportunities for understanding travel behavior, forecasting transportation demand, and supporting operational planning. Nevertheless, the increasing amount and variety of transportation data have created an analytical difficulty that is not always well handled by conventional statistical techniques; therefore, machine learning and deep learning techniques can identify complicated relationships in huge amounts of data, which are being used to address these issues (Behrooz & Hayeri, 2022).

Because of their great accuracy in prediction and relative ability to interpret, conventional machine learning algorithms like Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines have become widely used (Çelik *et al.,* 2025). Long Short-Term Memory, Bidirectional LSTM, and Multi-Layer Perceptron are examples of deep learning architectures that have demonstrated favorable results, particularly in situations involving sequential and time-based patterns(Alizadegan *et al.,* 2025). However, computational challenges, feature structure, and data size often influence a deep learning model's dominance (Ahmed *et al.,* 2023). The direct comparison between machine learning and deep learning approaches is frequently carried out using different datasets, evaluation standards, and experimental procedures, although many studies have used predictive models to solve problems related to transportation(Sinha & Kumar, 2023). Such variations make it difficult to determine which modeling approach is most appropriate for structured transportation datasets. For instance, decision tree algorithms, according to (Yan *et al.,* 2025), provide transparent and user friendly structures that enhance policy interpretation, but they suffer from overfitting when applied to complex datasets. Furthermore, (Kinasih *et al.,* 2024)stated that random forest models address the issue by combining multiple predictions across multiple expanded decision trees, which provide more reliability. The flexible nature of Support Vector Machines (SVMs) enables them to accurately model complex patterns in large variable datasets, this makes them particularly suitable for transportation analyses. Research results reveal that their excellent performances in transportation analyses achieved classification evaluation above 99%

and outperformed similar traditional machine learning approaches such as random forest and Naïve Bayes. Similarly,(Sharifi *et al.*, 2026) confirmed that decision tree-based ensemble models regularly outperformed deep learning for structured datasets related to travel pattern prediction. However, traditional ML algorithms have some setbacks when capturing an extended term temporal or ordered relationship of transport time series data. As noted by (Hinterleitner *et al.*, 2024), who explained variable selection such as Recursive Feature Elimination can enhance transparency but fail to resolve the static nature of these algorithms. DL models it is an advancement of ML which utilizes multilayered network frameworks that automatically derive from layered features of input variables. These models can capture complex, spatiotemporal, and space time relationships that traditional ML struggles to capture (Mienye & Swart, 2024). Furthermore, (Yin *et al.*, 2022) Outline that DL methods have demonstrated strong accuracy in traffic predictions, outlier detection, and multiple mobility prediction, where the relationship between temporal and space time features is critical.

Despite recent progress and the effectiveness of DL in modeling complex data, there limitation in using the models such as high computational power and a large volume of dataset is required (Ankita *et al.*, 2023). This observation underscores the importance of comparative experiments using standard conditions across ML and DL models, which is a methodological gap this research seeks to address. The integration of power Bi dashboard is also an integral part of the contribution of our study, where the accurate results of the models are displayed on the interactive dashboard for better planning and decision making by the stakeholders. As highlighted by (Zierz & Lordsleem Junior, 2026) stated that interactive dashboard are essential in interpreting complicated model results into practical decision support making output.

Therefore, our experiment utilized the same dataset to carry out these analyses by the use of CRISP-ML. CRISP-ML methodology, which provides a systematic process that promotes transparency, reproducibility, and quality assurance throughout model development, thereby offering a suitable foundation for comparative investigations(Sinha & Kumar, 2023).This study, therefore, presents a comparative analysis of ML and DL models for transportation analytics using the CRISP-ML framework and Power BI visualization.

The technique used integrates both interactive Power BI visualization and predictive modeling in order to evaluate both expected accuracy and the actual interpretability of the results. The purpose is to determine modeling techniques that provide reliable performance that are appropriate to use in real-world transportation analysis and decision-support applications.

MATERIALS AND METHODS

Research Design and CRISP-ML Framework

This study applied the methodology of the Cross-Industry Standard Process for Machine Learning. It consists of seven interrelated phases, which are business understanding, data acquisition and understanding, data preparation, modeling, model evaluation, model deployment, and monitoring. The framework offers a systematic and consistent approach for developing the ML and DL models throughout the study, as shown in Table 1 and Fig. 1.

Table 1: CRISP-ML framework

CRISP-ML phase	Implementation in this study
Business Understanding	The objective here is to classify trips into business or personal and to also predict trip miles
Data Acquisition & Understanding	Obtain and inspect 1,156 Uber trip records
Data Preparation	Missing-value handling, feature engineering, encoding, normalization
Modeling	Train LR, DT, RF, SVM, LSTM, Bi-LSTM and MLP
Evaluation	Accuracy, Precision, Recall, F1, MAE and R^2
Deployment	Present analytical outputs through Power BI
Monitoring	This is considered as component but is outside the scope of this work

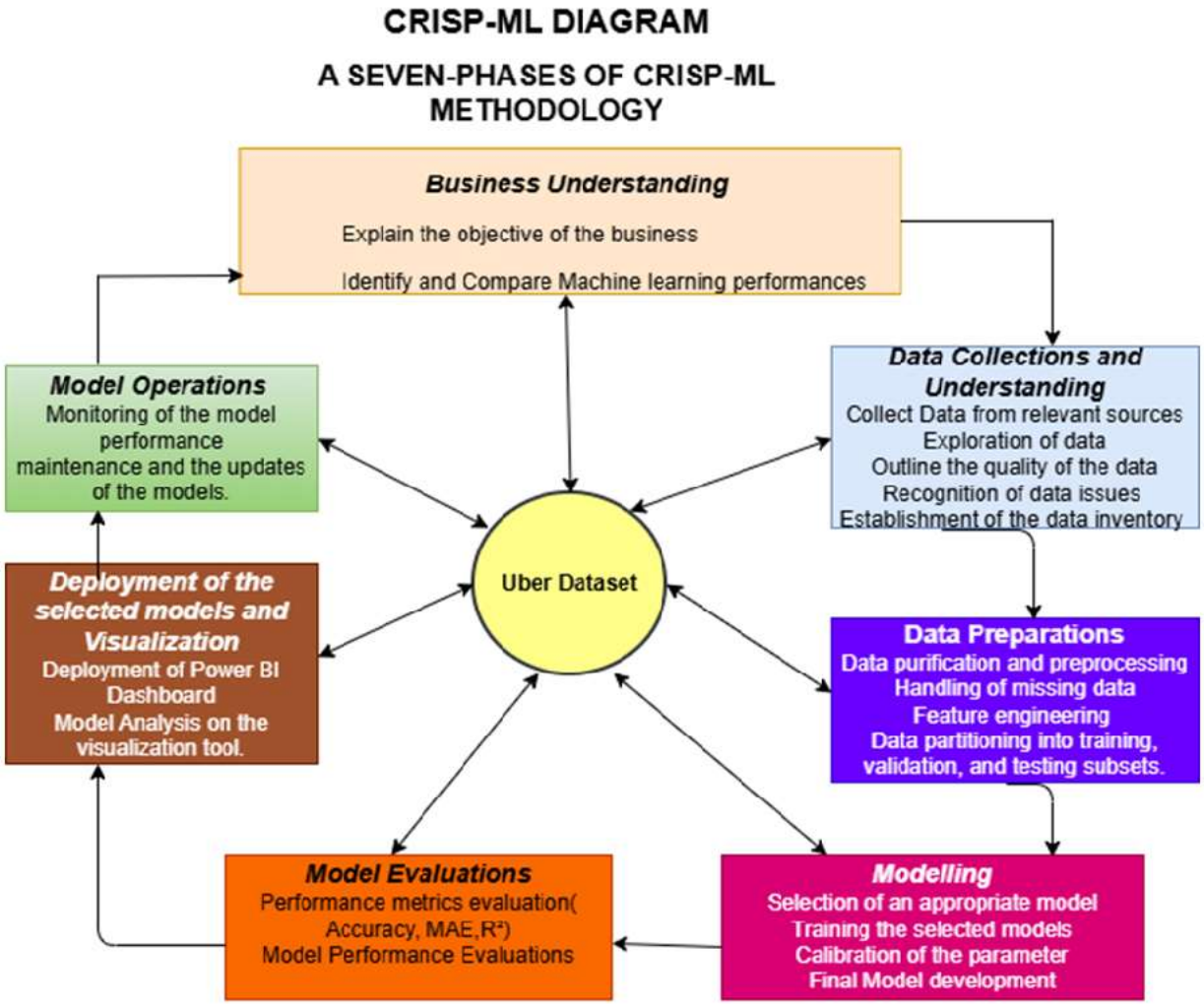


Figure 1: CRISP-ML framework

Dataset Description

The Uber trip dataset utilized in this study was obtained from Kaggle. The dataset contains 1,156 trip records with features including trip start time, trip end time, category, start location, stop location, miles covered, and trip purpose. An additional feature representing trip duration in minutes was computed from timestamp differences.

The feature representing trip duration was calculated from time interval differences. In the purpose attribute, the missing values were handled by the mode substitution method, while the class imbalance was handled by the synthetic minority oversampling technique (SMOTE). SMOTE was performed on the training data alone, after splitting. Before applying SMOTE Business and personal trip categories have 862, and 62 training data respectively. After applying SMOTE all the categories have 862 training data each.

Table 2: Dataset description

S/N	Feature	Description	Data Type
1	Trip Start Date	Trip commencement timestamp	Date & Time
2	End Date	Trip completion timestamp	Date & Time
3	Category	Business or Personal trip	Categorical
4	Start	Trip origin location	Text
5	Stop	Destination location	Text
6	Miles	Distance covered	Numeric
7	Purpose	Purpose of travel	Categorical
8	Duration (min)	Derive trip duration	Numeric

Table 2 shows the features in the dataset. For classification, the target variable is Category, while the target variable for regression is miles.

Data Pre-processing

In order to purify the dataset and make the model more suitable, data preprocessing was conducted. Missing data were recorded, unreliable data were corrected, and feature engineering was also applied in order to extract temporal variables. During our model training, mathematical features were adjusted in order to maintain uniform scaling. The images show how the features are distributed in Fig. 2.

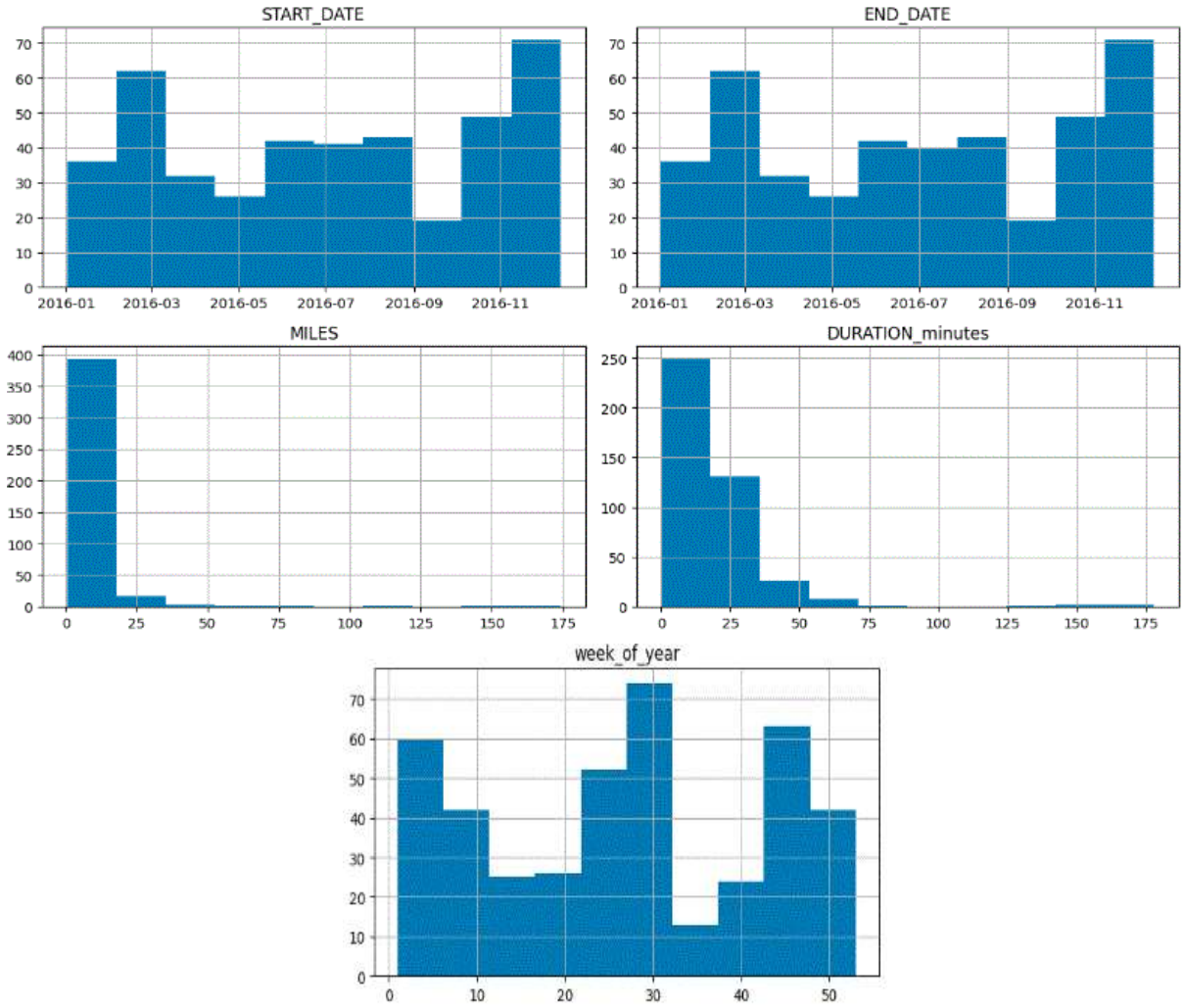


Figure 2: Distribution of selected features

Computational Environment

The computer system utilized for the experiment is a HP laptop with model number 14s-dq. It is an 11th-generation system with an Intel Core i3 which has a 12GB RAM and 500GB SSD operating on windows 11 Pro.

The Integrated Development Environment (IDE) was used in the model implementations is google Colab. A Python programming language was used which include the libraries such as Scikit-learn, Keras, Pandas, NumPy, and matplotlib. Keras was used to develop DL models, Matplotlib was used for data visualizations, Numpy and Pandas were applied to purify the dataset, while Scikit-learn was used to develop traditional ML.

Model Development

In this study, both machine learning and neural network models were adopted. The machine learning approaches utilized comprised of Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and Linear Regression, whereas the neural network approaches utilized include Long Short-Term Memory, Bidirectional Long Short-Term Memory, and Multi-Layer Perceptron (MLP) networks. Random seed was set to 42 for all the models. In logistic regression max_iter was set to 1000. For random forest, n_estimators was set at 100 with n_jobs set to -1. While SVM used rbf kernel with probability set to true.

The LSTM network was designed with two hidden layers, the first layer contains 64 neurons, while the second layer contains 32 neurons, respectively, and they are integrated with a dropout rate of 0.2 to prevent overfitting. The Bi-LSTM model utilized bidirectional processing in order to learn sequential dependencies from both previous and subsequent data points. The implementation of LSTM and BI-LSTM considered every trip as a sequence of length 1 since the trips are independent and there are so many data points with missing date. Meanwhile, the MLP model consists of three fully connected feed forward layers, which utilize the Rectified Linear Unit (ReLU) activation function in order to facilitate nonlinear learning.

Performance Evaluation Metrics

A standard evaluation metric was used in order to measure the performance of the model (Rainio *et al.*,

2024). The short explanations of the performance metrics with their mathematical expressions are as follows;

Accuracy (ACC): This measures the ratio of correct predictions to the total number of the given samples(Sujon *et al.,* 2025).

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision (PR): This evaluates the ratio of true positive observations among all positive predictions by the model(Rainio *et al.*, 2024).

$$PR = \frac{True\ Positive}{True\ positive+FalsePositive} \quad (2)$$

Recall (RE): This measures the ratio of true positive cases that are correctly identified out of all actual positives measured by the model (Pudjihartono *et al.*, 2022).

$$Recall = \frac{TP}{TP+FN} \quad . \quad (3)$$

The F1 Score: This can be defined as the harmonic mean of precision and recall by creating an accurate, unified measure that is suitable for a class in balance(Rainio *et al.*, 2024).

$$F1 = \frac{2TP}{2TP+FP+FN} \quad (4)$$

For regression evaluation, Mean Absolute Error (MAE) and R-Square were used.

Mean Absolute Error (MAE): This is used to quantify the average absolute deviation between predicted and actual observations (Robeson & Willmott, 2023).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \quad (5)$$

Smaller MAE values indicate that the model predictions are closer to observed values.

R-Squared: This is the coefficient of determination that measures the proportion of variance in the dependent variable that is explained by the independent features(Gao, 2023).

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (6)$$

Higher values of R^2 indicate greater explanatory power.

RESULTS AND DISCUSSION

This section explains the results of the analysis that was carried out in the research using the mean and standard deviation of accuracy, precision, recall, and F1-score. Precision, recall and f1-score are macro. The outcomes of the experiment are divided into classification and regression results, which are presented in the section below. K-fold cross validation was used to obtain the results.

Classification Performance

Under this section, our analysis shows that both ML and DL models achieved good prediction accuracy in the classification studies. However, traditional machine learning, such as decision trees, and random forests, achieved a higher accuracy of 93.2, and 92.6 % respectively. MLP achieved an accuracy of 90.70 %, which was also a slightly competitive result.

The feature importance analysis in this work indicated that trip destination, trip origin, travel distance, and trip duration are the coefficients that determined our classification analysis results.

Table 3: Classification accuracy comparison

S/N	Model	Accuracy Mean ± SD	Macro Precision Mean ± SD	Macro Recall Mean ± SD	Macro F1 Mean ± SD
0	Logistic Regression	0.8712 ± 0.0292	0.6366 ± 0.0369	0.7675 ± 0.0507	0.6675 ± 0.0435
1	Decision Tree	0.9264 ± 0.0137	0.6960 ± 0.0821	0.6588 ± 0.0755	0.6720 ± 0.0735
2	Random Forest	0.9329 ± 0.0179	0.7359 ± 0.1212	0.6635 ± 0.0735	0.6883 ± 0.0846
3	SVM	0.8809 ± 0.0367	0.6424 ± 0.0487	0.7430 ± 0.0652	0.6691 ± 0.0584
4	MLP	0.9070 ± 0.0380	0.6999 ± 0.0857	0.7563 ± 0.0265	0.7107 ± 0.0560
5	LSTM	0.8950 ± 0.0416	0.6785 ± 0.0896	0.7505 ± 0.0388	0.6923 ± 0.0613
6	Bi-LSTM	0.8842 ± 0.0412	0.6517 ± 0.0603	0.7370 ± 0.0419	0.6714 ± 0.0512

Regression Performance

Table 4 shows the performance of 5 models, in terms of Mean absolute Error and R-squared. The regression results indicated that linear regression achieved the lowest Mean Absolute Error (2.1660) and the highest coefficient of determination (0.63), followed by random forest which has an MAE of 2.190 and an R^2 value of 0.62 (Table 4). For the two neural network model, Bi-LSTM performed better in terms of R^2 while MLP performed better in terms of MAE.

Table 4: Regression Model Performance

S/N	Model	MAE	R^2
1	Linear Regression	2.1660	0.63
2	Random Forest	2.1910	0.62
3	Bi-LSTM	2.5390	0.56
4	MLP	2.6510	0.54

Linear regression despite being the best among the models tested on this dataset still has 37% of the variance in trip distance remains unexplained. Therefore, the model's suitability for transportation planning should be

interpreted cautiously, and its practical adequacy would require validation against a larger, more representative dataset and application-specific error tolerances.

The result obtained could depend on the size of the dataset. Traditional machine learning models tend to function well as a result of the size of the dataset, while neural network architectures usually need higher volumes of data to properly utilize their strengths. Therefore, rather than having a notion that more complex architectures will always produce better results, traditional machine learning can perform as well. The findings further highlight the significance of choosing models based on the dataset size available and features.

Deployment & Visualization

This study shows the usefulness of combining analytical models with visualization tools in addition to predicting accuracy. Power BI dashboards provide a clear visualization of trip distributions by categories, mobility developments, journey duration, and patterns of travel. The visual studies revealed that business travel has a large share of travel activity, and also transportation demand tends to increase over the course of the week, as shown in the dashboard. These results clearly indicate the need for the integration of interactive visualization and predictive modeling in order to improve transportation sector decision-support scenarios. The dashboards serve as a visual support for transport planners and policy makers in order to understand and communicate the results effectively for better decision making by the stakeholders (Sinha & Kumar, 2023).

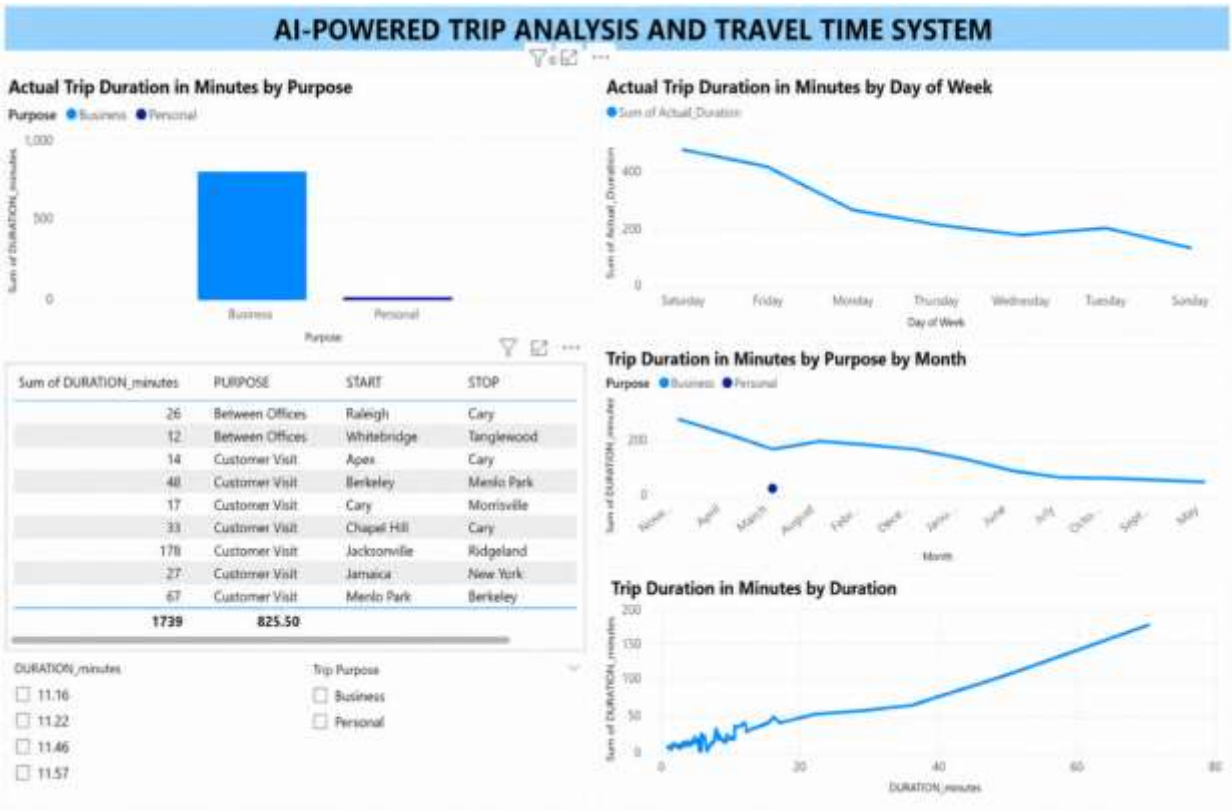


Figure 3: Power BI dashboard visualization

CONCLUSION

This study examined and compared the performance of selected machine learning and neural network models for transportation data analyses using the CRISP-ML methodology and Power BI visualization. From the experimental results, we can clearly indicate that machine learning models, particularly random forest, and decision tree, consistently delivered a strong predictive and classification performance on the structured Uber trip dataset. From the neural network models, it can be seen that MLP produced the highest accuracy of 0.90. In regression linear regression achieved the lowest MAE and highest R^2.

The results also highlighted to us that effective transportation analytics does not depend only on model sophistication but also on the size and the characteristics of the available dataset. As can be seen, a structured moderate dataset, when applying traditional machine learning approaches, can yield accurate, interpretable, and efficient results. The CRISP-ML approach utilized in the work provided a highly reproducible model development. The integration of the Power BI dashboard has provided a way of visualizing results for stakeholders' understanding.

From a practical perspective, the designed methodology can assist transport organizations in evaluating travel

patterns. This can be achieved by enhancing transportation routes, improving how resources are utilized for better decision-making. Lastly, this study contributes both educational and practical value for the advancement of modern transportation analytics.

Future Work

Future work may expand this study by utilizing larger and more diverse real-time transportation datasets, which can be obtained from sources such as GPS mobile devices, traffic monitoring systems, and smart city platforms. The integration of spatial and time series data can provide more understanding of travel behavior in order to enable a comprehensive evaluation of predictive and classification models.

Future research could also explore a hybrid framework that combines both machine learning and neural network models in order to take advantage of their supplementary strengths, with the possibility of enhancing model reliability and predictive performance across transportation analyses.

Furthermore, advanced machine learning models such as graph neural networks, attention mechanisms and relational neural networks can be explored for future work in order to capture more geographical-time-based data.

Lastly, explainable AI methods such as SHAP and LIME can be utilized in order to improve model transparency and interpretability. These advancements would lead to more improved decision-support systems for the next generation in smart transport applications.

Conflict of interest: The authors declare no conflict of interest.

References

Ahmed, S. F., Alam, M. S. Bin, Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Shawkat Ali, A. B. M. and Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artific. Intelligence Review*, 56(11), 13521–13617. https://doi.org/10.1007/s10462-023-10466-8

Alizadegan, H., Rashidi Malki, B., Radmehr, A., Karimi, H. and Ilani, M. A. (2025). Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction. *Energy Exploration and Exploit.*, 43(1), 281–301. https://doi.org/10.1177/01445987241269496

Ankita, Mittal, S., Sharma, I. and Kumar, A. (2023). Comparative analysis of shallow learning and deep learning. *2023 International Conference on Next Generation Electronics (NEleX)*, 1–8. https://api.semanticscholar.org/CorpusID:267702919

Behrooz, H. and Hayeri, Y. M. (2022). Machine learning applications in surface transportation systems: A literature review. In *Applied Sciences* 12(18), 9156. https://doi.org/10.3390/app12189156

Çelik, S., Doğanlı, B., Şaşmaz, M. Ü. and Akkucuk, U. (2025). Accuracy comparison of machine learning algorithms on world happiness index data. *Mathematics*, 13(7). https://doi.org/10.3390/math13071176

Gao, J. (2023). R-squared (R^2) – How much variation is explained? *Research Methods in Medicine & Health Sciences*, 5(4), 104–109. https://doi.org/10.1177/26320843231186398

Hinterleitner, A., Bartz-Beielstein, T., Schulz, R., Spengler, S., Winter, T. and Leitenmeier, C. (2024). Enhancing feature selection and interpretability in AI regression tasks through feature attribution. arXiv preprint arXiv:2409.16787. https://doi.org/10.48550/arXiv.2409.16787

Kinasih, A. N. S., Handayani, A. N., Ardiansah, J. T. and Damanhuri, N. S. (2024). Comparative analysis of decision tree and random forest classifiers for structured data classification in machine learning. *Science in Information Technology Letters*, 5(2), 13–24. https://doi.org/10.31763/sitech.v5i2.1746

Mienye, I. D. and Swart, T. G. (2024). A comprehensive review of deep learning: architectures, recent advances, and applications. *Information* (Switzerland), 15(12). https://doi.org/10.3390/info15120755

Pudjihartono, N., Fadason, T., Kempa-Liehr, A. W. and O'Sullivan, J. M. (2022). A review of feature selection methods for machine learning-based disease risk prediction. *Frontiers in Bioinformatics*, 2, 927312. https://doi.org/10.3389/fbinf.2022.927312

Rainio, O., Teuho, J. and Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, 14(1), 1–14. https://doi.org/10.1038/s41598-024-56706-x

Robeson, S. M. and Willmott, C. J. (2023). Decomposition of the mean absolute error (MAE) into systematic and unsystematic components. *PLoS ONE*, 18(2 February), 1–8. https://doi.org/10.1371/journal.pone.0279774

Sharifi, D., Mirzahossein, H., Kalantari, N. and Sedighi, F. (2026). Integrating hybrid recurrent neural networks and large language models for daily travel behavior prediction. *Transportation Research Interdisciplinary Perspectives*, 35. https://doi.org/10.1016/j.trip.2025.101793

Sinha, N. and Kumar, V. (2023). Comparing machine learning and deep learning models. *International Research Journal of Modernization in Engineering Technology and Science*, 11, 100–104. https://doi.org/10.56726/irjmets45900

Sujon, K. M., Hassan, R., Choi, K., & Samad, M. A. (2025). Accuracy, precision, recall, F1-score, or MCC? Empirical evidence from advanced statistics, ML, and XAI for evaluating business predictive models. *Journal of Big Data*, 12, Article 268. https://doi.org/10.1186/s40537-025-01313-4

Yan, Y., Liao, Y., Xu, G., Yao, R., Fan, H., & Sun, J. (2025). Large language models for traffic and transportation research: Methodologies, state of the art, and future opportunities. arXiv Preprint arXiv:2503.21330. https://doi.org/10.48550/arXiv.2503.21330

Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H. and Yin, B. (2022). Deep learning on traffic prediction: Methods, analysis, and future directions. *IEEE Transactions on Intelligent Transportation Systems*, 23, 4927–4943. https://doi.org/10.1109/TITS.2021.3054840

Zierz, O. I. S. and Lordsleem Junior, A. C. (2026). Business intelligence tools in organizations with a focus on power BI applications in civil construction: A systematic literature review. *Buildings*, 16(4). https://doi.org/10.3390/buildings16040869