ISSN: 2513-2504

# TRAVEL TIME MODEL USING GPS DATA AND MACHINE LEARNING FOR BUS INFORMATION SYSTEMS

Panchali Samarasinghe, Amal Kumarage, Asoka Perera, Samudaya Nanayakkara
*Department of Transport Management and Logistics Engineering, University of Moratuwa.*
*samarasinghepanchali@gmail.com, amal.kumarage58@gmail.com, asoka.uom@gmail.com, samudaya@uom.lk*

**ABSTRACT** - Public Transportation modes are prevalent and extensively utilised means of transportation for commuters. Road congestion, bus crew issues, malfunctions, and miscellaneous factors impede buses from adhering to schedules. As a result, it is becoming problematic for commuters to arrange their travel plans confidently. Intelligent transportation systems use Global Positioning System (GPS) technology and data analytics to accurately predict real-time travel information and improve traveller and operator experience. The research gap is the unavailability of standardised techniques for mass travel time predictions using standardised analytical methods. However, recent research has focused on developing accurate travel-time models employing machine learning algorithms. Predictive models rely on past data gathered through GPS systems. The study uses the GPS data of public buses in Central Province, Sri Lanka, one thousand buses have been fitted with GPS units since 2019 (7 million of data). Realtime and historical data that were gathered through GPS units can be used to develop machine learning-based models to predict bus or passenger transport information accurately. The study analysed available data using Microsoft Azure, Statistical, Time Series and Machine algorithms for performance accuracy with lower error rates on predictions used for comparison purposes.

**Keywords:** Travel Time; Prediction; GPS Data; Machine Learning

## 1. INTRODUCTION

The study explores the importance of travel time prediction in improving public transportation systems. Accurately predicting travel times can help passengers better plan their journeys, reduce waiting times at bus stations or train depots, and provide operators with real-time information for more efficient network management. The study explores various predictive methodologies while identifying areas that need improvements. Additionally, artificial intelligence (AI) and machine learning (ML) can enhance public transportation systems by improving efficiency, predicting demand, and reducing delays by effectively and efficiently analyzing data.

Implementing advanced analytics tools with machine learning algorithms can track vehicle locations and evaluate passenger flow to optimize routes or schedules instantaneously while minimizing disruptions. Integrating GPS data analysis with machine learning algorithms through a travel time model development offers transit operators a valuable opportunity to increase service delivery while controlling operational costs. This research utilized Microsoft Azure platform, a comprehensive suite of tools and services tailored precisely for this purpose. It bolsters predictive analytics across various industries through many use cases, such as healthcare, finance, and retail.

Furthermore, incorporating Microsoft Azure provides benefits such as minimum development time and cost-effectiveness over conventional solutions running on-premises. Moreover, it provides reliable and effective outcomes by providing insights into technological advancements. GPS-enabled data tracking of bus locations aids in the accurate analysis of massive datasets through historical patterns and real-time data. Hence, machine learning algorithms ultimately give companies a competitive advantage within their industries. Align with using Microsoft Azure platform to analyze

ISSN: 2513-2504

GPS data is used to determine the accuracy of existing travel time models using machine learning for the prediction of travel times of buses and other vehicles.

## 2. MATERIALS AND METHODS

The estimation of travel time is a critical factor that plays a vital role in enhancing reliability, efficiency, and overall user experience in public transport. Predicting travel times accurately and in real-time offers valuable insights to public transport users about bus arrival times at specific stops. (Ranhee Jeong, n.d.) and (Tran et al., 2020) emphasized the relevance of forecasting travel time for improving public transportation by using Artificial Neural Network (ANN) models based on six months' worth of Automatic Vehicle Location (AVL) data and Automatic Passenger Counting (APC). Their research covered variables such as bus stops' number, route length between them, average passengers per journey, and dwell time, along with intersections between stops. Inference to literature, predicting travel time is essential for improving public transportation systems.

Recent research by (Servos et al., 2020) created an artificial neural network model that could accurately predict real-time bus travel times using only GPS data. According to (Ranhee Jeong, n.d.), real time and machine learning based models could be more efficient in improving transportation system management than traditional data analytics solutions such as statistical and time series models that developed using limited number of historical data. On the other hand, implementing intelligent concepts such as machine learning cloud provide more competent ways of managing transport infrastructures than traditional data analytics (Wu et al., 2004). According to (Wu et al., 2004) and (Chen et al., 2023), devising a neural network model based on GPS data can predict real-time bus travel times precisely. These findings indicate the potential benefits of machine learning models while applying an Advanced Public Transportation System (APTS). Additionally, (Mehmet Altinkaya, 2013) accentuates that integrating GPS data with machine learning tactics can significantly upgrade travel time models by supporting more effective and responsible public transportation services. Moreover, (Tran et al., 2020) advise utilizing machine learning algorithms such as Artificial Neural Networks (ANN) and Support Vector Machines (SVM) to augur travel time with lower error rates and propose suitable routes for shuttle buses.

This research utilized the Microsoft Azure platform, which is a cloud-based service that helps businesses speed up their predictive analytics and machine learning tasks. As (Gal et al., 2017) highlights, the advantages of using Microsoft Azure are numerous compared to traditional analytical tools). Rapidly changing present technology environment, Microsoft Azure has become essential for companies to gain valuable insights from data and make data-driven decisions.

## 3. RESULTS AND DISCUSSION

### 3.1. Data Analysis

The study used GPS data collected from buses in Central Province, Sri Lanka. The selected bus route connects Kandy and Kadugannawa, which is 15.2 km long. The raw dataset includes multiple attributes, and the data cleaning process involves removing data records with default errors. The second phase involved determining the validity of the extracted data using the day of the week, average time, and corrected GPS locations. Variations in travel time collected route travel times are analysed through the mean median, standard deviation, and coefficient of variation (COV) are used to analyse the data. Model development was conducted through predesigned algorithms, and model validation can be performed through Table 1.

**Table 1.** Classification Matrix

| Factor | Accuracy | Precision | Recall | FI Score | AUC |
|---|---|---|---|---|---|
| Linear Regression | 67.40% | 0.81 | 0.59 | 0.31 | 0.61 |
| Logistics Regression | 70.79% | 0.79 | 0.53 | 0.37 | 0.65 |
| Support Vector Machine | 78.10% | 0.72 | 0.44 | 0.42 | 0.71 |
| Neural Network Regression | 82.40% | 0.64 | 0.32 | 0.49 | 0.75 |

The neural network regression-based model performed well in the analysis.

## 4. CONCLUSION

In conclusion, this research provides direction to policymakers and stakeholders to make appropriate decisions to improve public transportation networks. The neural network regression-based model performed well in the analysis. Microsoft Azure is a powerful tool for advanced predictive analytics, allowing users to design, deploy and manage machine learning models. Machine learning models can revolutionise bus travel time prediction, improving public perception and efficiency.

## REFERENCES

1. Chen, B., Bai, R., Li, J., Liu, Y., Xue, N., & Ren, J. (2023). A multiobjective single bus corridor scheduling using machine learning-based predictive models. *International Journal of Production Research*, *61*(1), 131–145. https://doi.org/10.1080/00207543.2020.1766716
2. Gal, A., Mandelbaum, A., Schnitzler, F., Senderovich, A., & Weidlich, M. (2017). Traveling time prediction in scheduled transportation with journey segments. *Information Systems*, *64*, 266–280. https://doi.org/10.1016/j.is.2015.12.001
3. Mehmet Altinkaya, M. Z. (2013). *Urban Bus Arrival Time Prediction: A Review of Computational Models*.
4. Ranhee Jeong, L. R. R. (n.d.). *Prediction Models of Bus Arrival Time for Real-Time Applications*.
5. Servos, N., Liu, X., Teucke, M., & Freitag, M. (2020). Travel Time Prediction in a Multimodal Freight Transport Relation Using Machine Learning Algorithms. *Logistics*, *4*(1). https://doi.org/10.3390/logistics4010001
6. Tran, L., Mun, M. Y., Lim, M., Yamato, J., Huh, N., & Shahabi, C. (2020). DeepTRANS: A Deep Learning System for Public Bus Travel Time Estimation using Traffic Forecasting. *Proceedings of the VLDB Endowment*, *13*(12), 2957–2960. https://doi.org/10.14778/3415478.3415518
7. Wu, C. H., Ho, J. M., & Lee, D. T. (2004). Travel-time prediction with support vector regression. *IEEE Transactions on Intelligent Transportation Systems*, *5*(4), 276–281. https://doi.org/10.1109/TITS.2004.837813