An ANN-based Prediction Model for Public Bus Journey Time

Yong Ting Lim*, Sye Loong Keoh*, Yung-Wey Chong†, Noor Farizah Ibrahim‡, Sharul Kamal Abdul Rahim§
*School of Computing Science, University of Glasgow, Glasgow, United Kingdom
†National Advanced IPv6 Centre, Universiti Sains Malaysia, Penang, Malaysia
‡School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia
§Wireless Communication Centre, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia
Email: 2609702L@student.gla.ac.uk, SyeLoong.Keoh@glasgow.ac.uk, {chong, nfarizah}@usm.my, sharulkamal@utm.my

***Abstract*—Traffic congestion is a major problem in cities worldwide, especially in developing countries. It has a significant impact on the local GDP, environment, and society. Public transport is used to ease congestion but it is not efficiently implemented in developing countries. Implementing an accurate bus arrival time prediction system is a necessity to improve the standard of public transport and consumer satisfaction. In this study, we have developed a basic ANN-based prediction model for bus journey duration to estimate the bus arrival time using a case study in Johor Malaysia. The model was trained using bus fleet GPS dataset and the results shows improved accuracy compared to baseline approaches by considering factors like bus stop, time of day, month, and travel distance. *Virtual bus stop* is introduced for a long stretch of road and this shows promise in addressing limitations and improving performance. The simplicity of the model allows its application on any route by breaking it down into smaller segments. The final model achieves an MAE of 0.0056 and RMSE of 0.0123.**

***Index Terms*—urban mobility, prediction of bus arrival time**

I. Introduction

Urban transport plays a pivotal role because it provides access to places of education, healthcare, entertainment and other services [1]. However, road traffic congestion remain a significant problem in most cities worldwide, especially in developing countries [2] such as Malaysia. Traffic congestion typically occurs when travel demand exceeds the road capacity and mainly attributed to the unavailability of efficient public transportation infrastructure. This is notable in Malaysia as traffic worsens post-pandemic [3], [4]. As public transportation provides group travel, it is an effectively way to mitigate road traffic congestion [5], [6].

In Malaysia, the government has subsidised a total of RM155 million for a month of free public transportation in the Klang Valley area to mitigate road congestion. Such a campaign only has its short-term effects, and the main issue lies with the quality and efficiency of its public transport system. Promoting public transport use by improving commuters' travel well-being is vital to mitigate traffic congestion. In addition, the commute satisfaction is an indicator of the utility of public transport and is affected by modal attributes, including travel time variability, predictability of travel, information, and other factors [7].

One of the major factors leading to the low utilisation of public transportation, i.e., public bus is the uncertain arrival and departure time, making it difficult for commuters to plan their journeys, despite the mobile apps and online tracking facilities provided by bus companies. In many cases, buses do not follow the published schedule and due to the shortage of bus fleet, the bus service frequency is rather low, thus causing delays in travel for commuters. This is further exacerbated by the traffic road congestion, resulting in longer journey time.

As most of the bus fleet's journeys are tracked using GPS, huge volume of their journey data can be analysed to learn about the journey duration, traffic conditions along the bus routes and deriving insights on the bus services. This paper proposes a deep learning approach using Artificial Neural Network (ANN) to train a journey duration model from the past historical GPS traces. The model can then be used to accurately predict the bus arrival time given the bus's current location in real-time. The experiment was based on the P-411 bus service route plying Kulai to Larkin Sentral in Johor, Malaysia. The GPS dataset was provided by the bus operator and we carried out data cleansing and developed a deep learning model to provide for an accurate analysis of the prediction of bus timings. Our results achieved close to 80% accuracy in the journey duration between bus stops and approximately two minutes difference from the expected arrival time. With this, the model can be generalised to be applied to other cities to provide commuters with real-time accurate information on the arrival times of buses, enabling them to plan their journeys more efficiently and reduce waiting times at bus stops.

The paper is organised as follows: Section II presents the related work in journey duration prediction using machine learning approaches. Section III describes the dataset, data engineering, features selection and model training of the proposed ANN-based approach. Section IV presents the results and we conclude the paper in Section V.

This work is the output of the ASEAN IVO project, 'An IoT-based Public Transport Data Collection and Analytic Framework using Bluetooth Proximity Beacons' (UTM Grant No: R.J130000.7606.4B856), and financially supported by NICT (http://www.nict.go.jp/en/index.html) and Universiti Teknologi Malaysia (UTM) Matching Grant (Q.J130000.3006.04M36).

II. Related Work

Machine Learning based approach have been used widely for urban transport. In [8], a bus travel time prediction model was developed using ANN and GPS data obtained for a 17.4 km route in Delhi India with 43 bus stops. The result showed that ANN outperformed the linear regression model. In [9], the bus arrival time was predicted using ANN and hierarchical ANN based on GPS data and automatic fare collection (AFC) system data for an 8.1 km route in Jinan, China with 15 bus stops. In a more recent study, [10] introduced the GPS position calibration method to help increase arrival accuracy using Long Short-Term Memory (LSTM) models. The results of their experiment demonstrated good accuracy in both peak-time and off-peak-time prediction compared to other traditional methods.

A study on predicting bus arrival time in urban and rural regions of China [11], employed a combination of Support Vector Regression (SVR) and K-Nearest Neighbours (KNN) in one model to compare against historical data-based prediction methods on a bus route that integrates both urban and rural areas. One of the factors this study accounts for is the stopping time at each stop as the number of passengers alighting and boarding would impact the timing. The results showed a significant difference in relative error between the SVR+KNN model and the historical data-based prediction method. The results showed that the use of machine learning models performed better than that of the predictive analysis.

Leveraging the IoT, [12] predicted the bus arrival time using the data gathered from the vehicles stored in a cloud server. The proposed method showed that accurate arrival time under different traffic conditions can be predicted using the bus and route with different parameters such as average speed, number of passengers, rush hour information, and number of bus stops. In a different study by [13] that analysed arrival time using historic and real-time route data, beacons were installed on all the busses and the best stops to estimate the arrival time. The captured data were then analysed using machine learning models to predict the arrival schedule and at the same time, allow the commuters to access the running status of the buses.

Furthermore, ANN has been used extensively in predicting transportation parameters because they are good at pattern recognition and classification, especially in complex nonlinear settings. In one study, ANN was able to identify the complex nonlinear relationship between travel time and the independent variables [14] and better results compared to historical data models and regression models. Additionally, ANN was able to map complicated relations for trips under different patterns, as seen in [15].

III. An ANN-based Journey Duration Prediction Model

This section presents the methodology in developing an ANN-based prediction model for the bus journey time using the P-411 route in Johor, Malaysia. We believe that the methodology can be applied to other bus routes in other cities.

A. Bus P-411 Service Information

The P-411 serves two bus terminals located in different cities, with Kulai being situated north of Larkin. The distance between them is 34.1 km and the bus journey duration typically averaging around 50-70 minutes each way, depending on the traffic. The routes for both directions follow very similar paths travelling in opposite directions. The bus service operates with a schedule that includes five journeys in each direction daily approximately every two to three hours.

The journey from Kulai to Larkin Sentral encompasses a total of 31 bus stops, denoted by the labels 6001 to 6031, with the Kulai bus terminal designated as bus stop 6001 and the Larkin terminal represented as bus stop 6031. Conversely, the journey from Larkin to Kulai consists of 34 bus stops, identified by the labels 7001 to 7035, with Larkin terminal designated as 7001 and Kulai terminal designated as 7035. The journey between each bus stop, i.e., 6002 $\rightarrow$ 6003, is denoted by 6002'.

The GPS traces of the bus journey were recorded every 4 seconds. The data fields are: date, time, location, speed, and bearing, with the location field recorded as latitude and longitude (lat, lng).

B. Dataset

The bus operator, Pengangkutan Awam Johor provided the GPS dataset from May 2021 to March 2022, and November 2022 to May 2023. It is important to note the limitations of this dataset as the earlier dataset was captured during the COVID-19 pandemic and traffic patterns might have changed since then due to lockdown and border closure. The earlier part of the dataset was mainly used for model training, while the more recent dataset was mainly used to test against the prediction model to validate our model's accuracy. Although this dataset only captured one route covering a small part of Johor Malaysia, it represents the main trunk of traffic congestion in Johor as it is the route taken from Johor Bahru to other cities and towns.

We observed that many locations logged by the GPS were not accurate or missing. There were occasions when the bus did not complete the full route of the journey or made detour halfway for reasons such as re-fueling, repair, etc. It is also observed from the dataset that the buses were parked at bus depots and had to be driven to the terminal at the start and end of their service daily.

The bus service operates with high punctuality, with buses typically departing on time. The standard deviation of approximately 3 minutes indicates a relatively low variability in departure times, suggesting that passengers can rely on the buses adhering closely to the scheduled departure times.

C. Data Engineering

The raw dataset was cleaned and pre-processed to filter out missing, incomplete, and irrelevant data to provide a more accurate, consistent, and quality data for a more accurate analysis outcome.

1) Removal of incorrect GPS data: Many of the data points are not on the route of the journey, such as buses travelling back and forth from bus depots at the start and end of each service daily, as well as the detour made by the bus. There were instances where buses detoured halfway e.g., started or ended from the middle of the journey instead of terminals. These data points had to be removed to provide better visualisation of the bus route.

2) Tag bus stop to GPS data: GPS data is then associated with its travel direction, and the nearest points to each bus stop are tagged with their relative bus stop codes to enhance the data's meaningfulness and utility. It is crucial to acknowledge that there are instances when buses travel too fast without stopping at certain bus stops; thus, a buffer of 60 meters was implemented to ensure that bus stops can be properly associated with the bus journey.

3) Duration of travel between two consecutive bus stops: The travelling time of each bus stop to its next bus stop was derived using (1) where t$_{n} isthetimeatthebusstopn$:

$$duration = t_{n+1} - t_n \tag{1}$$

4) Removal of Outliers: There are a few instances of extreme outliers in the dataset that were considered anomalies within the data. As a result, these outliers were removed from the dataset to ensure the reliability and robustness of the model.

5) Cyclic time and month: When dealing with time-series data, it has inherent periodicity, such as the daily cycle of 24 hours and the yearly cycle of 12 months. One-hot encoding should be avoided because it increases the data dimensionality and does not effectively capture the relationship of time. With cyclic features, we can effectively capture the periodic nature of time in a continuous representation, enabling machine learning models to understand and leverage the cyclical patterns in the data. It also allows us to preserve proximity, e.g., 11:59pm and 12:01am should be close to each other in the encoded space as they are adjacent in time. This is enabled by encoding with *sine* and *cosine* functions to transform them into continuous variables that can be easily used as inputs for the machine-learning models. In addition, this helps to avoid linearity assumptions by linearly encoding time with numerical values such as hours 1, 2, 3 and so on.

In order to accommodate a more granular analysis, we have opted to divide the daily cycle of 24 hours into intervals of half an hour each. This decision allows for a more detailed examination of patterns and trends within smaller time increments, providing a deeper understanding of the data.

D. Data Analysis, Visualisation and Model Training

The cleaned dataset was transformed into a set S = (a_1, b_1, c_1, d_1, e_1), (a_2, b_2, c_2, d_2, e_2), ... where a_i is the bus stop code, b_i and c_i is the *sine* and *cosine* value of half hour ranges, and d_i and e_i is the sine and cosine value of the month.

The training and test data for the bus prediction model encompassed the period from May 2021 to March 2022. The dataset was divided using a 70:30 split approach, where 70% of the data was allocated for training and 30% for testing. As for the November 2022 to May 2023 dataset, it was discovered that it was riddled with inaccuracies. After thorough cleaning and pre-processing, only the months of January, March, April, and May 2023 contain enough reliable data for utilisation. Considering that the '*month*' is one of the features used for prediction, it is important to note that the overlapping months from the new usable data are limited to January, March, and May. These months will be the focus for further analysis and modelling, as they provide the most reliable and comprehensive data points for predicting bus arrival times.

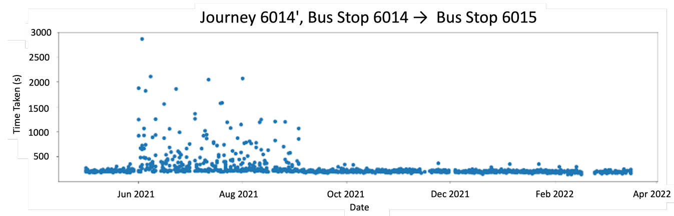

Fig. 1. Travel Duration for Journey 6014' (Bus Stop 6014 to 6015).

1) Visualisation of monthly trend: The dataset revealed valuable insights regarding monthly bus travel time variation trends. Upon analysing the data, it becomes evident that specific months exhibit distinctive patterns, where the time taken by the bus to complete its route fluctuates, as seen in Figure 1 for the journey of 6014'.

Recognising the significance of these monthly variations, it becomes crucial to incorporate the monthly variable as an independent feature in any subsequent analysis or machine learning modelling.

By considering the month as an independent variable, we can capture the inherent seasonal patterns or external factors that might influence bus travel time. This additional feature allows for a more comprehensive understanding of the data and enhances the predictive capabilities of the models.

While the analysis of the GPS data highlights noticeable fluctuations in the travel time across different months, it is important to note that the available dataset does not provide sufficient data of other yearly data. This limits our ability to generalise and assume the observed monthly trends are consistent throughout the years, and it is possible that the fluctuations witnessed are specific to the recorded year and may not necessarily represent an established yearly pattern. The monthly fluctuation trends must be further analysed when new data is available.

2) Standard deviation of travel duration: Standard deviation (SD) is a measure of spread around the mean. It is crucial to assess the variation in travel duration across the different bus stops to identify which exhibits the highest variability in their travelling duration. Bus stops with higher variability can have several implications and highlight potential areas where delays or unpredictable factors such as traffic congestion may be more prevalent. It is important to analyse these bus stops

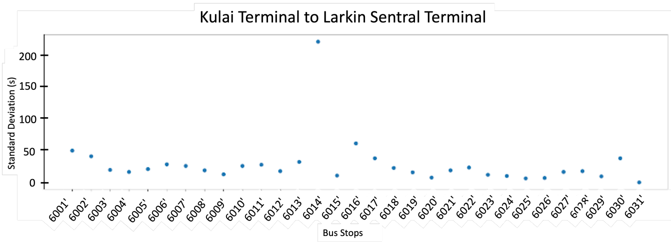

Fig. 2. Standard Deviation of Travel Duration from Kulai to Larkin Sentral.

that may require additional interventions to address a more accurate prediction timing.

Upon analysing the dataset, one notable observation is that journey 6014' exhibits a higher standard deviation of approximately two minutes as compared to the other journeys as seen in Figure 2. This variability may result from irregular traffic patterns, unforeseen circumstances, or specific conditions along the route that influence the journey's duration.

It was discovered that a distinct pattern along the journey of 6014', where occurrences of congestion (with speeds below 10 km/h) predominantly happen during the first half of the journey. Upon further investigation, we observed that the noticeable alleviation point of traffic congestion coincides with the presence of a U-turn. This finding suggests that the congestion at 6014' is primarily influenced by traffic leading up to the U-turn contributing to slower traffic speeds during the initial phase of the journey.

Thus, in this paper we introduce a *virtual bus stop* at this U-turn to help in predicting bus arrival timings from bus stop 6014 to bus stop 6015.

3) Correlation analysis: As bus stops are linked by connecting roads, it is important to analyse the relationship between the travelling times of consecutive bus stops. Pearson's correlation coefficient (Pearson's R), a widely used technique due to its ability to measure linear relationships, is adopted to analyse the relationship of consecutive bus stops.

Based on the observations, it is apparent that there is a varying degree of correlation between bus stop travelling times and their respective preceding bus stop travelling times. While some bus stops exhibit a moderate to strong correlation, this relationship is not consistently present across all bus stops. Therefore, careful consideration should be given to selecting this feature as a predictor, considering its strength of correlation for specific bus stops.

Similarly, a correlation analysis was done to analyse the correlation between the day of week, day of month and the travel duration. The analysis was done on every journey of consecutive bus stops and the results indicated a very weak relationship between these variables. This suggests that the day of the week and day of month have a very minimal influence on the travel duration, and that they are not a significant factor in determining the travel duration between bus stops.

E. Journey Duration Prediction Model

The model adopts a hop-to-hop approach to predict the journey duration between two consecutive bus stops, i.e., Kulai Terminal $\rightarrow$ Bus Stop 1, Bus Stop 1 $\rightarrow$ Bus Stop 2 and so on, as seen in Figure 3.

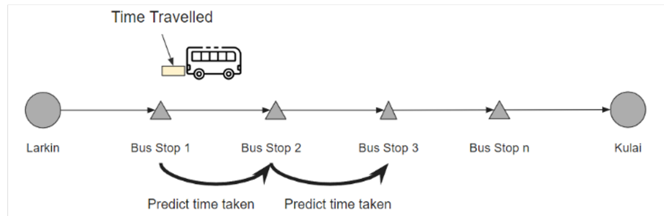

Fig. 3. A Hop-by-hop Journey Duration Prediction.

The arrival time will be predicted by adding the predicted time duration of preceding bus stops not yet travelled, subtracting the time travelled.

$$t_{k_{arrival}} = \sum_{t}^{k-1} t_{i_{predicted}} - t_{i_{travelled}} \tag{2}$$

The prediction framework used is an ANN model employing the Adam loss function, with a configuration consisting of two hidden layers, each with 100 neurons. The model utilises the Rectified Linear Unit (ReLU) activation function in the first hidden layer and the hyperbolic tangent activation function in the second hidden layer.

IV. Results and Analysis

This section evaluates the results of the proposed journey duration prediction model. We evaluated the model's accuracy and reliability using the MAE and RMSE.

The hop-by-hop journey prediction model was trained for 20 epochs at a learning rate of 0.00001, achieving an MAE of 0.6% (approx. 18 seconds) and an RMSE of 1.4% (approx. 40 seconds). The hop-to-hop logic takes into consideration the dynamic nature of bus travel, where delays can be accumulated at each stop, and the bus does not maintain a constant speed throughout the journey. By factoring in these delays and variations in travel time, the model provides more reliable predictions that align with the real-world conditions of bus operations.

In addition, the hop-to-hop approach allows for the model to identify the different patterns of bus travel by breaking down the journey into smaller segments. This helps the model to understand better the factors affecting bus travel time, such as traffic conditions of the different segments on the journey.

Figure 4 shows the results in predicting the duration from bus stop 6021 to bus stop 6031, it reveals that it accurately accounts for delays from previous bus stops, achieving an MAE of 2% (approx. 1 minute) and an RMSE of 2.7% (approx. 1 minute 15 seconds). Thus, it shows that the model is able to take into account previous delays on the preceding journey, making more precise predictions, resulting in reduced

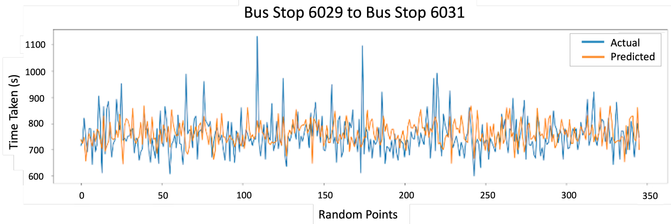

Fig. 4. Prediction of Travel Duration from 6021 to 6031.

TABLE I
PERFORMANCE OF PREDICTION MODEL WITH ADDITIONAL FEATURES

	MAE	**RSME**
Base Model	0.0067	0.0139
Additional features		
Travelling time of preceding journey	0.0066	0.0139
Day of the week	0.0066	0.0140
Day of the month	0.0067	0.0141
Distance of the journey	0.0056	0.0132

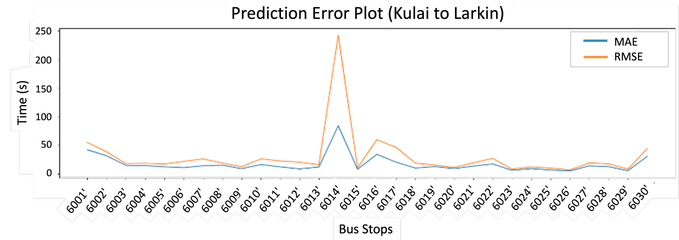

Fig. 5. Error Measurement of Each Journey from Kulai to Larkin Terminal.

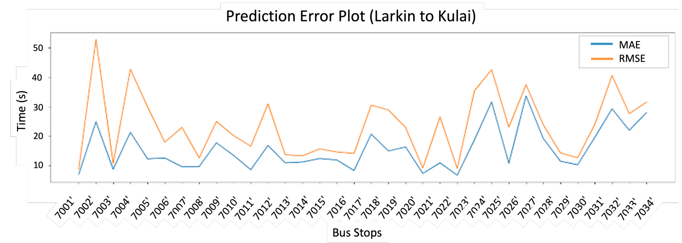

Fig. 6. Error Measurement of Each Journey from Larkin to Kulai Terminal.

errors and improved accuracy, showcasing its effectiveness in capturing the intricacies of bus travel.

A. Additional features

In this section, we discuss the augmentation of our bus prediction model by introducing additional features.

During the data engineering phase, it was noted that additional features such as previous journey travelling time, day of the week and day of the month did not correlate strongly with the travel duration. The lack of correlation suggests that these features may not capture the significant factors influencing travel duration between bus stops. These additional features were introduced into the bus prediction model to validate the findings. As shown in Table I, it is evident that the additional features besides distance of journey may not effectively capture the complexities and nuances of the bus travel time dynamics.

Thus, an additional feature, *distance of journey* has been added to enhance the feature set. The updated feature set S' will then become $S' = (a_1, b_1, c_1, d_1, e_1, f_1), (a_2, b_2, c_2, d_2, e_2, f_2), \ldots$ where f_i is the *distance of journey*.

B. Virtual bus stop

As shown in Figure 5, it has been identified that journeys 6001', 6016', and 6030' showed poor prediction accuracy. While for Figure 6, the prediction accuracy is quite accurate for the Larkin to Kulai journey. The observed error in journey 6001' can be attributed to the inherent variance in timing when the bus departs from the terminal. We attempted to add a virtual bus stops to journeys 6014', 6016', and 6030' to evaluate the effectiveness to improve the accuracy of travel time predictions for these specific bus stops.

The journey 6016' spans a straight road of 2 kilometres with no major roads joining it and experiences congestion (speeds less than 10km/h) throughout, it is proposed to place a virtual bus stop in the middle of the journey. As for 6030', the journey spans a distance of 4 kilometres and involves travelling across a highway. An analysis was conducted to assess the congested areas (speeds less than 10km/h) along the route of bus stop 6030. However, the findings revealed that the speed throughout the journey remained consistent without significant variations or congestion from the dataset.

By strategically placing virtual bus stops along the route, the model gains a finer granularity and a more detailed understanding of the variations in travel times. These virtual bus stops act as additional reference points for a more comprehensive analysis and a better representation of the underlying dynamics of the bus journey. The improved performance can be observed in the evaluation metrics showing more accurate predictions as seen in Table II and Table III.

C. Training and Testing with Year 2023 Dataset

To assess the performance and generalisation ability of the bus prediction model, the new data for the months of January,

TABLE II
PERFORMANCE OF PREDICTION MODEL USING VIRTUAL BUS STOPS

	MAE	**RSME**
Original Model	0.0056	0.0132
Model with Virtual Bus Stops	0.0056	0.0123

TABLE III
PERFORMANCE OF PREDICTION MODEL WITH VIRTUAL BUS STOPS

		6014'	**6016'**	**6030'**
Original Model	MAE	0.0294	0.0119	0.0109
	RMSE	0.0845	0.0209	0.0153
Model With Virtual Bus Stops	MAE	0.0294	0.0119	0.0109
	RMSE	0.0845	0.0209	0.0153

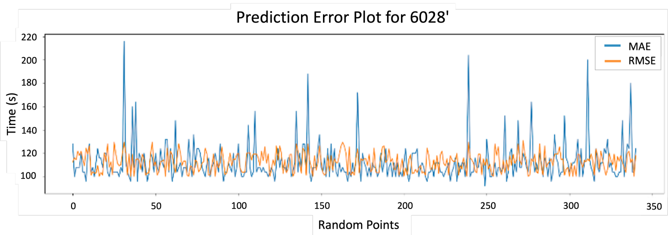

Fig. 7. Journey Prediction for 6028'.

March, April, and May 2023 was trained independently to compare with the current model trained with the data from 2021 to 2022. Both models were tested against the test dataset of the months of January, March, April, and May 2023.

As seen in Table IV, the performance between the two models are very similar, and this suggests that the model's effectiveness extends beyond the initial dataset and can generalise well to unseen instances. However, it is crucial to note that this evaluation is limited to the overlapping months of January, March, and April 2023 between 2021-2022 dataset. While the model demonstrates satisfactory performance during this period, it is essential to conduct further analysis and evaluation when more data becomes available.

TABLE IV
PERFORMANCE OF PREDICTION MODEL TRAINED WITH 2023 DATASET

	MAE	**RSME**
Original Model	25.7s	46.2s
Model Trained with 2023 Dataset	24.8s	42s

V. CONCLUSIONS

In conclusion, the development and evaluation of the bus prediction model have provided valuable insights into predicting bus stop timings. The model incorporates various features and techniques, including using artificial neural networks, early stopping, and introducing virtual bus stops.

The results obtained from the model show promising performance, with improved accuracy and predictions compared to baseline approaches. By considering factors such as bus stop, time of day, and month and incorporating additional features like distance of travel, the model demonstrates the ability to capture the complexities of bus travel and provide more accurate estimations.

However, it is important to acknowledge the limitations of the model. The unpredictable nature of traffic conditions, the influence of external factors such as weather, and the presence of outliers highlight the challenges in achieving perfect accuracy in arrival time predictions. While the model performs well overall, there will always be instances that are more difficult to predict accurately. Journey 6028' serves as an example of the inherent variability in traffic patterns and travel duration, as seen in Figure 7. Outliers at this bus stop may be caused by unique circumstances, such as unusually heavy traffic or specific events that disrupt the usual flow. These exceptional cases pose a challenge for the model, as they deviate significantly from the typical patterns observed usually.

In summary, the bus prediction model, with its incorporation of various features and techniques, offers a promising solution for estimating bus stop timings. By acknowledging its strengths, limitations, and the dynamic nature of traffic conditions, this work lays the foundation for further exploration and extension by considering a wider variety of dataset that encompass different conditions, such as weather and traffic variations.

REFERENCES

[1] D.Q. Nguyen-Phuoc, G. Currie, C. De Gruyter, W. Young, "Congestion relief and public transport: An enhanced method using disaggregate mode shift evidence", *Elsevier Case Studies on Transport Policy*, Vol. 6, No. 4, Pages 518-528, 2018.

[2] V. Jain, A. Sharma, and L. Subramanian, "Road traffic congestion in the developing world". In *Proceedings of the 2nd ACM Symposium on Computing for Development (ACM DEV '12)*, NY, USA, 2012.

[3] I. Zulkifli, "'Malaysia jam teruk' and how it's making working class Malaysians poorer", *CILISOS - Current Issues Tambah Pedas!*, 27-Sep-2022. [Online]. Available: https://cilisos.my/malaysians-lose-billions-in-rm-due-to-bad-traffic/. [Accessed: 02-Dec-2022].

[4] E. Teh, "Traffic congestions in Malaysia and the lessons we must learn", Heinrich-Böll-Stiftung. [Online]. Available: https://th.boell.org/en/2022/07/22/traffic-malaysia. [Accessed: 02-Dec-2022].

[5] I. Mulalic and J. Rouwendal, "Does improving public transport decrease car ownership? Evidence from a residential sorting model for the Copenhagen metropolitan area", *Regional Science and Urban Economics*, vol. 83, no. 103543, p. 103543, 2020.

[6] B. Younes, "The benefits of improving public transport: a myth or reality?", *Transport Reviews*, vol. 15, no. 4, pp. 333–356, 1995.

[7] M. Abou-Zeid and M. Ben-Akiva, "The effect of social comparisons on commute well-being," *Transportation Research Part A: Policy & Practice*, vol. 45, no. 4, pp. 345–361, 2011. [Accessed: 02-Dec-2022].

[8] J. Amita, S.S. Jain and P.K. Garg, "Prediction of Bus Travel Time Using ANN: A Case Study in Delhi", *Transportation Research Procedia*, Vol. Volume 17, Pages 263-272, 2016.

[9] Y. Lin and X. Yang and N. Zou and L. Jia, "Real-Time Bus Arrival Time Prediction: Case Study for Jinan, China", *Journal of Transportation Engineering*, Vol. 139, No. 11, pages 1133-1140, 2013.

[10] Q. Han, K. Liu, L. Zeng, G. He, L. Ye and F. Li. "A Bus Arrival Time Prediction Method Based on Position Calibration and LSTM." *IEEE Access*, Vol. 8, pages 42372-42383, 2020.

[11] Y. Wang, "Intellectualization of the urban and rural bus: The arrival time prediction method", *Journal of Intelligent Systems*, vol. 30, no. 1, pages 689–697, 2021.

[12] J. Jalaney and R. Ganesh, "Highly accurate bus arrival time prediction using k-nearest neighbor prediction in the internet of things (IoT) environment", *Journal of Green Engineering*, Vol. 10, No. 9, pages 4752–62, 2020.

[13] K.Desingu, D.Isaac, P. Mirunalini, B. Bharathi and C.Philipose, "LocateMyBus: IoT-Driven Smart Bus Transit", *Journal of Communications Software and Systems*, vol. 19, no. 2, pp. 136-146, 2023.

[14] M. As and T. Mine, "Dynamic bus travel time prediction using an ANN-based model". In *Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication (IMCOM '18)*, NY, USA, 2018.

[15] R. Jeong and R. Rilett, "Bus arrival time prediction using artificial neural network model", In *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*, Washington, WA, USA, pages 988-993, 2004.