Turkish Journal of Electromechanics and Energy
ISSN: 2547-975X
Vol:9, No:2, Page 49-58 (2024)
https://www.scienceliterature.com

Turkish Journal of Electromechanics & Energy
an international open access journal

Implementing deep learning techniques for travel time prediction

K. S. Jadaan[1], A. Najjar[1], S. Abu-Eisheh[2, *], A. Abuaisha[3]

[1]*Department of Civil Engineering, University of Jordan, Amman, Jordan*

[2]*Department of Civil and Architectural Engineering, An-Najah National University, Nablus, Palestine*

[3]*Department of Data Science and Artificial Intelligence, Monash University, Melbourne, Australia*

ARTICLE INFO

***Article Type*:**
Selected Research Article[c]

Article History:
Received: 20 March 2024
Revised: 6 July 2024
Accepted: 27 July 2024
Published: 15 October 2024

Editor of the Article:
M. E. Şahin

Keywords:
Travel time prediction, Deep learning, Artificial intelligence applications, Neural networks, Jordan Country

ABSTRACT

In recent years, deep learning (DL) has proved to be a powerful artificial intelligence prediction tool in solving many complex problems. Despite its demonstrated superiority over traditional approaches in various domains, the utilization of DL for traffic prediction remains limited. Nevertheless, prior studies have shown its potential and effectiveness in this area. As for travel time, DL can make accurate predictions for all segments in the transportation network with a single model structure. This study reports on the application of DL in simultaneously predicting the travel time over several segments of a network. Two DL models, namely multilayer perceptions (MLPs) and long-short term memory networks (LSTM), which can deal with high-dimensional input data, are proposed to tackle the problem. The outcome of the application of the two DL models in Amman city, the capital of Jordan, is presented in this study. Amman City Taxi Company provided a dataset containing information such as trip origin, destination, and travel time, for taxi trips over the years 2018 and 2019. The results demonstrate that DL models hold great promise in achieving precise and real-time prediction of travel time on a network-wide level.

Cite this article: K. Jadaan, A. Najjar, S. Abu-Eisheh, A. Abuaisha, "Implementing deep learning techniques for travel time prediction," *Turkish Journal of Electromechanics & Energy*, 9(2), pp.49-58, 2024.

1. INTRODUCTION

Accurate travel time predictions are crucial for cost management, planning, design, and decision making in transportation. However, since traffic is a dynamic phenomenon that varies with time, location, modes of transportation, and other factors, travel time is difficult to predict under different conditions [1]. The irregularity of travel time patterns and the difficulty of quantifying their characteristics have made developing high-performing and reliable models challenging. This necessitates the use of complex statistical methods to identify specific patterns and robust behaviors, and to achieve predictability. Artificial intelligence methods, including deep learning and machine learning, offer solutions to this problem.

The main objective of this study is to develop a predictive model, using deep learning, which provides sufficiently accurate predictions of the travel time in a road network. Two deep learning techniques utilized in this study comprised of the multilayer perceptron's (MLPs) and long short-term memory (LSTM) networks, which were selected due to their ability to effectively capture the complex nonlinear relationships and patterns present in travel time data [2]. Compared to conventional approaches, deep learning models can adaptively learn and generalize from the available historical data, without the need for explicit feature engineering or assumptions about the underlying data distributions. Furthermore, techniques like LSTM have the capability to capture long-term dependencies and temporal patterns in the data, enabling models to accurately predict travel time influenced by various factors, including traffic conditions, weather, events, and time-of-day effects [3].

This study reports on the application of deep learning techniques, specifically MLPs and LSTM networks, for simultaneously predicting travel time over several segments of a transportation network. The outcome of the application of these two deep learning models is presented for Amman City, Jordan. The results demonstrate that deep learning models hold great promise in achieving precise and real-time prediction of travel time on a network-wide level, outperforming traditional approaches. The key contributions of this research include (i) developing and evaluating the performance of MLP and LSTM models for travel time prediction, leveraging the inherent advantages of deep learning techniques; (ii) providing insights into the applicability and limitations of the proposed models using a real-world dataset from Amman; and (iii) highlighting the potential of deep learning in enhancing transportation planning,

[c]Initial version of this article was presented at the 5th International Conference of Materials and Engineering Technology (TICMET'23) held on November 13-16, 2023, in Trabzon, Türkiye. It was subjected to a peer-review process before its publication.

**Corresponding author's e-mail: sameeraa@najah.edu*

traffic management, and real-time decision-making for commuters and authorities.

The remainder of this research is organized as follows. Section two summarizes the most relevant works related to the problem at hand. A general overview of the architecture and types of artificial neural networks (ANN) is presented in section three. Next, section four outlines the methodology adopted in this research, followed by a description of the collected data in section five. The two models developed in this work are explained in section six, while section seven discusses the research results and model evaluation. Finally, section eight concludes with the main research findings and offers recommendations for future work.

2. LITERATURE REVIEW

A wide range of approaches has been proposed in the literature to predict travel time, including time series models, statistical regression, and traffic flow theory-based methods. Researchers tended to rely mostly on parametric statistics to forecast future travel times based on a fixed set of parameters. For example, Wu et al. proposed a spatiotemporal random effects model (STRE) which reduces computational complexity by reducing the number of mathematical dimensions [4]. While most of the modelling work tried to address the problem on a network scale rather than individual road segments, other researchers, such as Rice et al. [5], developed models to predict travel times on specific stretches of freeways.

Recently, deep learning (DL), a branch of machine learning, has proved to be successful in dealing with large amounts of data in a high-dimensional space [6]. However, its applications in traffic prediction are still limited. Artificial neural network (ANN), which forms the backbone of DL algorithms, consists of many intensive parallel distributed processors that perform mathematical operations among them. ANNs gain experience through learning from the environment and storing knowledge, represented by the strength of connections within the neurons, known as the neurons' weights.

In 2010, Park et al. used the ANN system to predict travel time on a highway in Houston, Texas, based on data collected from the Automated Vehicle Identification System as a test base [7]. The researchers found that the model could strongly predict travel time and that the mean absolute percentage error (MAPE) value ranged between 7.4%-17.9%. Wisitpongphan et al. designed an ANN model with three hidden layers to predict travel time. The model was based mainly on vehicles' speed data collected for one month by global positioning system GPS [8]. Travel time was expected during the morning and evening peak hours, off-peak hours, on a weekday, and a weekend. The predicted values of trip time obtained from the model were compared with those trip times obtained through Google Maps [9]. The model results were found superior with a mean square error (MSE) value of less than 3%.

The bus rapid transit (BRT) system is an important form of urban transportation in China, playing a crucial role in easing traffic congestion and improving service quality. Accurate travel time prediction is essential for BRT systems, as it enables optimal path planning and sufficient knowledge of passengers' journeys. In 2010, Huang et al. proposed a probabilistic method to predict travel times for BRT by analyzing historical traffic data, by collecting data from GPS-equipped BRT buses on a 17-stop route in Jinan, China over 13 days. The method first analyzes the expectations (average travel times) and MSEs of the historical data, split into consecutive time intervals, revealing that travel time variability differs across time intervals, with higher MSEs during rush hours. The probability distributions of travel times for each time interval are derived, providing detailed information on the likelihood of different travel times and capturing the uncertainties inherent in the data. Using the probability distributions, the method can generate confidence intervals corresponding to given confidence levels, such as when departing between 16:30-17:00, there is a 68.75% chance the travel time will be between 33-36 minutes. The study argues that this probabilistic approach is well-suited for BRT systems, where travel times tend to fluctuate less due to the dedicated bus lanes, and with more historical data, the method can be refined to provide even more precise predictions. Overall, the proposed probabilistic method leverages historical statistics to deliver informative travel time predictions for BRT, accounting for variability and uncertainty, which can support passenger planning and improve the overall quality of BRT services [10].

In 2016, Duan and Wang explored the application of LSTM neural networks, a type of recurrent neural network, for travel time prediction. Travel time is an important indicator of traffic conditions, but cannot be directly observed in real-time, so accurate travel time prediction is crucial for intelligent transportation systems. The study constructs a series prediction LSTM neural network model for each link in the travel time dataset, taking historical travel time data as input and predicting future travel times. The entries for this model were 66 roads in England on which travel time data were taken. The entire environment was divided into three sections, 80% for training 10% for verification, and the remaining 10% for testing. The used model achieved a good performance as the median of the mean relative error (MRE) was less than 7% for the model. During training, they tune the number of hidden units and optimize the model parameters to minimize prediction error on a validation set. Evaluating the trained models on a test set, they find the 1-step ahead travel time predictions have relatively small errors, with a median mean relative error of 7%, though accuracy decreases for longer forecast horizons, highlighting the importance of timely data acquisition. The findings conclude that deep learning models like LSTM neural networks are promising for travel time prediction, an essential component of intelligent transportation systems, providing a comprehensive methodology for applying LSTM networks to this domain [11].

In 2017, Wang et al. proposed a gaussian process regression (GPR) method for predicting urban road travel times, which outperforms traditional approaches like autoregressive integrated moving average (ARIMA) and supports vector machine regression. Compared to neural networks and support vector machine (SVM), GPR is easier to implement and has the advantage of self-adaptive hyper-parameter tuning. The GPR model considers factors like vehicle counts, average speeds, and historical travel times to forecast the next period's travel duration. Evaluated on both weekday and weekend data, the GPR method achieved a MAPE of just 2.8% and 97.2% prediction accuracy, significantly better than the benchmark models, especially during peak traffic periods on workdays when travel times fluctuate dramatically. The authors conclude that the GPR-based approach is a promising solution for urban road travel time prediction due to its strong performance and practical implementation advantages [12].

Rodriguez-Deniz et al. also explored the potential of Gaussian process-based methods for urban traffic prediction, but these two studies differ in their specific model formulations and evaluation approaches. Rodriguez-Deniz, H., et al presented a multi-output gaussian process regression (MOGPR) model, which is designed to capture the interdependencies between traffic speeds on different roads in the urban network. This was achieved through the use of a regionalized covariance structure that models the prior information about historical traffic correlations. In contrast to Wang, this study discusses a GPR model that focuses on forecasting travel times on individual roads, without explicitly considering the network effects. In terms of evaluation, Rodriguez-Deniz assessed the MOGPR model's performance using mean absolute error (MAE) in km/h, while the review reports mean absolute percentage error (MAPE) and prediction accuracy, providing a slightly different perspective on the models' strengths. The MOGPR model's ability to achieve low absolute errors is emphasized, whereas the review highlights the GPR model's high prediction accuracy, particularly during peak traffic periods on workdays. Additionally, the study emphasizes the computational efficiency of the MOGPR model, which can be trained in seconds by relying on the prior network dependencies. In contrast to Wang, this study discusses the GPR model's ease of implementation and self-adaptive hyperparameter tuning as key advantages [13].

Liu et al. presented a comprehensive investigation of short-term travel time prediction using various LSTM and deep neural network (DNN) models. The researchers collected 90 days of travel time data from the Caltrans performance measurement system (PeMS) on the Corridor I-880 freeway section in California, USA, and used this dataset to evaluate the performance of their proposed models. Specifically, the study explored 8 different LSTM-DNN model configurations by varying three key hyperparameters: the number of LSTM units (32, 64, 128), the number of LSTM layers (2, 3), and the number of DNN layers (1, 2). This resulted in a diverse set of deep learning models that were trained and tested multiple times to account for variability in the training process. The models were evaluated using MAPE and root mean squared error (RMSE) on the test data. In addition to the LSTM-DNN models, the researchers also tested several benchmark models, including linear regression, Ridge regression, Lasso regression, ARIMA, and a simple DNN model. Two sets of experiments were conducted, using sliding window lengths of 15 minutes and 30 minutes, and prediction horizons ranging from 1 to 6-time steps (5 to 30 minutes). The results of the study demonstrated the advantages of the LSTM-DNN models, while also revealing the different characteristics of these deep learning models with varying hyperparameter settings. This provided valuable insights for optimizing the model structures and architectures for short-term travel time prediction tasks [14].

Additional strategies have been created to deal with large-scale transport traffic forecasts. Recent research was carried out in 2018 by Hou et al. where an analytical model was designed using LSTM to predict travel time for a connected road network to be able to model spatial complexity to improve the model and make it stronger [15]. The network data and travel time data were collected for all the routes used in the model, and the time series was divided into a month of a year. A day from a week, then an hour from a day, and at the end, one hour was divided into four parts so that the measured period was 15 minutes to obtain a model with the most smoothing and fewer drop values as well as greater accuracy of the model. The researchers found a remarkable superiority of artificial intelligence methods as compared to other methods. The developed model produced MAPE values of =7.25% and 7.09% for the LSTM and the ANN models, respectively.

In 2019, Kankanamge et al. presented an approach to predict static travel time for taxi trip trajectories in New York City using XGBoost regression models. The study identified two categories of trips - "in-lier" trips that follow the shortest path, and "extreme-conditioned" trips that deviate from the shortest path. For the in-lier trips, the XGB-IN model outperforms other benchmark models like neural networks and support vector regression, achieving lower error metrics (MAE of 94.764s, RMSE of 130.627s, MAPE of 17.021%) and higher correlation (Spearman correlation of 0.93) with actual travel times, while also capturing time-varying features better. For extreme-conditioned trips, the XGB-Extreme model can provide reasonably accurate predictions, though it struggles for trips with very long actual travel times. The authors demonstrate the scalability and applicability of their approach to large-scale taxi trip trajectory datasets, highlighting the effectiveness of the XGBoost regression technique in static travel time prediction [16].

In 2020, Deng et al. presented a spatial and temporal analysis method for leveraging big data from public transport cards, using classification statistics and visualization techniques. The study first preprocesses the data to clean and categorize the cards into three main types - ordinary, discount, and free. They then analyze the distribution of active cards and swiping frequency for each type, as well as the peak passenger flow periods. Interestingly, the analysis reveals that while elderly discount card users tend to travel outside of the overall peak hours, student card users have peak usage times that largely overlap with the main congestion periods. Based on these insights, the authors propose two strategies to reduce student card usage during peaks: increasing fares for students during rush hours and introducing dedicated student-only bus lines. They evaluate the potential impact of these measures using real-world data from Xiamen City, China, finding the dedicated student line approach to be more effective at directly removing student passengers from the main congestion [17].

In 2023, Xu et al. explored the use of machine learning techniques, specifically MLP and MLP Regressor, to predict bus arrival times using historical GPS data collected from a bus route in Malaysia. The methodology included performing data cleaning and feature engineering on the dataset to prepare it for training the machine learning models. The results showed that the MLP model outperformed the MLP Regressor in terms of various error metrics (R-squared, MAE, MSE, and RMSE) on the small dataset. This aligns with findings from prior studies that have used different machine learning techniques, such as support vector regression (SVR) and k-nearest neighbors (KNN) combined model, ANN and linear regression, and dual-stage attention-based recurrent neural network (DA-RNN), long short-term memory recurrent neural network (LSTM RNN), Kalman Filter, and SVR, which have generally demonstrated that machine learning-based approaches outperform traditional historical data-based prediction methods. This study compared the performance of MLP and MLP regressors on a small dataset, which is a common scenario in real-world applications. The study acknowledges the limitations comprised in the need for further research with larger and more diverse datasets, as well as exploring alternative machine learning models and techniques to improve the prediction accuracy, in order to enhance the performance of such bus arrival time prediction systems [18].

In 2024, Nampalli and Gudla proposed a novel methodology for predicting travel time in complex road structures using machine learning techniques, specifically linear regression and LSTM networks. The approach utilizes real-time traffic data collected through Bluetooth sensors deployed at traffic intersections, which provides valuable information such as vehicle speed, travel time, and other relevant parameters. Compared to traditional approaches, the study highlights the limitations of using average travel time or GPS-based methods, which may not accurately capture the impact of traffic signals and other factors that influence travel time. The proposed machine learning models, on the other hand, are designed to learn from the complex patterns and dependencies within the Bluetooth sensor data, enabling more accurate and reliable travel time predictions. The results of the study demonstrate the superior performance of the LSTM model, which achieved significantly lower RMSE values and higher accuracy compared to the linear regression baseline. The LSTM model's ability to capture long-term dependencies and effectively handle sequential data makes it well-suited for the travel time prediction task. The study also emphasizes the potential of the proposed methodology to contribute to efficient and reliable travel time prediction systems, which can assist commuters in making informed decisions and improve traffic management strategies [19].

The applications of deep learning approaches in traffic and travel time prediction have been explored, but only in a limited number of previous studies. Polson and Sokolov introduced a deep learning architecture that combines a linear model with a sequence of layers for predicting traffic flow [6]. This structure addresses the challenge of sharp spatio-temporal nonlinearities in traffic flow transitions, such as those from free flow to congested conditions caused by factors like work zones, weather, special events, and incidents, all of which contribute to travel time uncertainty. The developed predictor effectively models these spatiotemporal relations, enabling forecasts of congestion propagation and providing forty-minute speed predictions. Additionally, it can integrate various data sources, such as weather forecasts and police reports, to enhance the accuracy of its forecasts.

Huang et al. proposed a deep learning architecture for traffic flow prediction, which combines a deep belief network (DBN) for unsupervised feature learning and a multitask regression layer for supervised prediction [2]. Their results suggest that deep learning techniques can be highly effective in complex systems, such as transportation research, beyond traditional neural-related areas. Ammoura et al. applied deep learning to predict travel time over several segments in a network simultaneously using two deep learning models, namely convolutional neural networks (CNN) and long-short term memory networks (LSTMN) [20]. They employed a refined representation of the data where travel time estimates for each hour of the day were provided along some roads in New York City. They concluded that deep learning can provide accurate real-time prediction of travel times on a network scale.

This study aims to build upon and extend the existing body of knowledge in several ways. Previous studies have demonstrated the potential of deep learning models, such as LSTM and MLP, in achieving accurate travel time predictions. For instance, Duan and Wang [11] developed an LSTM-based model for travel time prediction on 66 road segments in England, achieving a median mean relative error of less than 7%. Liu et al. [13] also explored the use of various LSTM-DNN configurations for short-term travel time prediction on the I-880 freeway in California, outperforming traditional regression and time series models. However, this study distinguishes itself from previous investigations in two key aspects. First, it compares the predictive capabilities of both MLP and LSTM architectures, offering insights into their relative strengths and weaknesses in travel time prediction. Second, it applies these models to a fast-growing city, Amman, in a developing country, Jordan.

By building upon the foundations laid by previous studies and addressing their limitations, this research aims to contribute to the growing body of knowledge on the application of deep learning for travel time prediction, providing valuable insights and methodological advancements that can inform future research and practical implementations in predicting travel time.

3. NETWORK ARCHITECTURE AND DESCRIPTION

The results, graphs and tables if there are any, should be expressed in this section clearly. The subject can be discussed with similar studies by giving references.

3.1. Neural Networks Components

In general, ANNs consist of the following components, as shown in Figure 1 [21].

- Neurons: The main component of the neural network is artificial neurons.
- Neuron Weights: The magnitude of the neurons' influence representing the strength of the neurons' relationship with the function.
- Activation Function: The mathematical calculation inside the neurons that works to calculate the neurons' output by measuring the weighted input.

- Input Layer: This is the first layer in the ANN from which the analysis process begins. This layer consists of a group of neurons containing the inputs from the project's data group, and it is called the visual layer.
- Hidden Layers: These are the second layers in the ANN that come after the input layer. The neurons' values are taken from the inputs to perform the calculations based on the activated function and the distribution of the neuron's weights according to the amount of its participation, which then are delivered to the third layer.
- Output Layer: The final layer is called the output layer. It is responsible for presenting the output values and/or vectors of values that correspond to the format required for the problem.

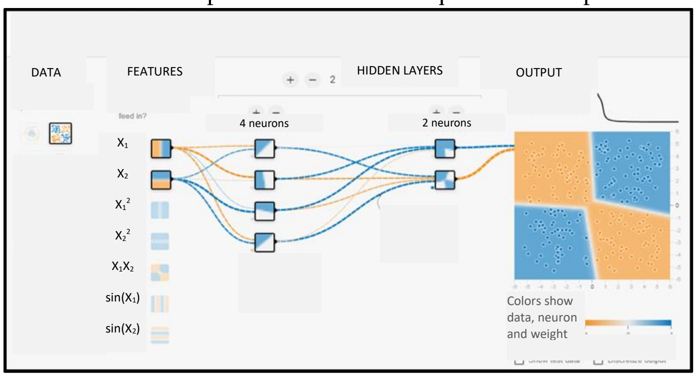


Fig. 1. Neural network components.

3.2. Types of ANN

In this research, concentration is made on two specific types of ANN, namely multilayer perceptrons (MLPs) and long short-term memory networks (LSTM). MLPs may be considered traditional networks. They consist of one or more layers of neurons. Information is supplied from the input layer to one or more hidden layers giving different levels of thinking, while forecasts are made on the output layer. The elements of MLP networks and the sequence of their calculation process are shown in Figure 2 [22]. Equations (1) and (2) explain the forward propagation step of MLPs.

$$Y = f(v) \tag{1}$$
$$v = \sum(x \times w + b) \tag{2}$$

Where w is the weight matrices of the neuron, x is the input neuron, v is the net input neuron, b is biased (threshold value), Y is the output neuron, and f(v) is the activation function.

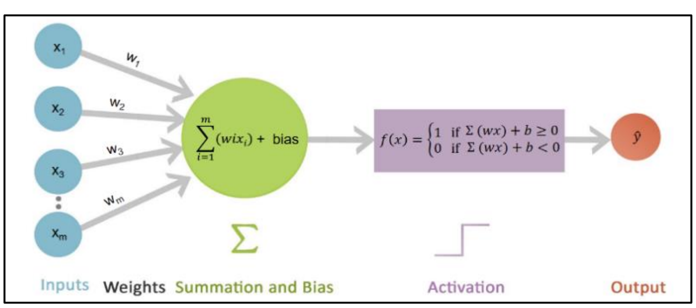


Fig. 2. MLPs network model.

The MLP is a fully connected and feed-forward neural network, composed of multiple layers of interconnected nodes, with at least three layers: input, hidden, and output layers. The MLP model is known for its flexibility in modelling and generalization potential, which has made it popular in various transportation engineering fields such as predicting travel time. MLP uses a supervised learning method called back-propagation for model training, making it suitable for regression problems where historical data is available.

On the other hand, LSTMs are a branch of the recurrent neural networks (RNNs), which are capable of learning long-term dependencies. LSTM has four layers interacting specially. The layers and the interactions within the cell are shown in Figure 3 [23].

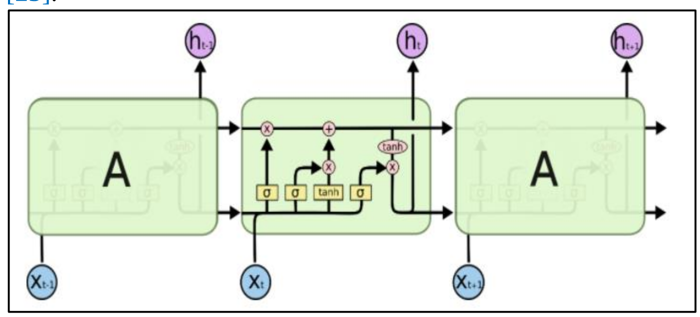


Fig. 3. LSTMs network model.

LSTM is a specialized type of RNN designed to overcome the limitations of traditional RNNs in capturing long-term dependencies in sequential data. The key innovation of LSTM is its gating mechanisms, which include the input gate, forget gate and output gate. These gates regulate the flow of information into and out of the LSTM's memory cells, allowing the model to selectively retain or forget relevant details over time. This enables LSTM to effectively learn and leverage complex temporal patterns in the data, making it well-suited for time series forecasting tasks like travel time prediction.

The LSTM networks were selected in this study for the travel time prediction task due to the inherent advantages of LSTM over traditional RNNs. LSTM is designed to address the limitations of RNNs, particularly the vanishing gradient problem, which can hinder the ability to capture long-term dependencies in sequential data. The LSTM architecture incorporates gating mechanisms that regulate the flow of information, allowing the model to selectively retain or forget relevant details over time. This capability is crucial for accurately predicting travel time, as it enables the model to effectively learn and capture the complex temporal patterns and dependencies in the dataset.

4. METHODOLOGY

The process of developing the deep learning models for travel time prediction can be systematically represented in the following steps:

(a) Data collection and preprocessing:

- Gather the necessary travel time data from the respective sources (e.g., taxi company, Uber) for the study locations.

- Perform data cleaning and handling of any missing or inconsistent data points.
- Explore the data characteristics, such as the relationship between travel time, distance, and time of day.

(b) Model selection and hyper parameter tuning:

- Choose the deep learning architectures to be investigated (in this case, MLPs and LSTM networks).
- Determine the appropriate hyperparameters for each model, such as the number of layers, neurons, activation functions, learning rates, and optimization algorithms.
- Split the dataset into training and testing sets to enable model evaluation.

(c) Model training and validation:

- Train the MLP and LSTM models using the respective hyperparameters and the training dataset.
- Monitor the model performance during the training process, using metrics such as RMSE.
- Validate the trained models using the testing dataset to assess their generalization capabilities.

(d) Model evaluation and comparison:

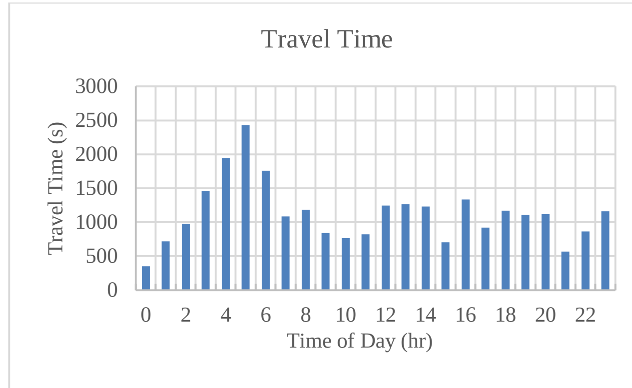
- Compare the performance of the MLP and LSTM models on the testing dataset.
- Analyze the results, including the RMSE values, learning curves, and the relationship between actual and predicted travel times.
- Assess the strengths and limitations of each model based on the specific characteristics of the dataset.

(e) Model deployment and future improvements:

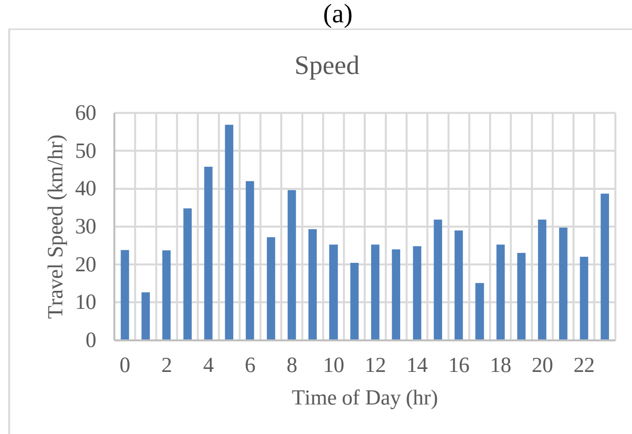
- Identify the best-performing model(s) for each study location based on the evaluation results.
- Discuss the potential for deploying the developed models in real-world applications for travel time prediction.
- Provide recommendations for future research, such as incorporating additional data sources, exploring alternative deep learning architectures, or addressing the limitations of the current study.

5. DATASET DESCRIPTION

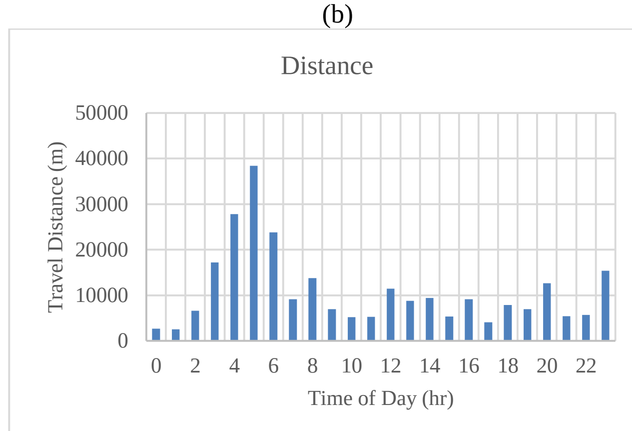
The dataset used in this research is described here. The dataset used for the purpose of this research was obtained from Al-Moumayaz Taxi in Amman, Jordan, which covers 1300 trips over the years 2018 and 2019. The collected trip data include origin and destination locations, date and time of the trip, distance of the trip, and travel time. The time series study was conducted during a full week from Monday to Sunday. Figure 4 (a-c) shows the traffic characteristics for all segments for one random day, illustrating travel time, speed, and distance, respectively, while Figure 5 displays the relationship between the time of day for all segments and the speed of trips. It can be seen that the relationship is not simple or clear due to the variation in the volume of vehicles on the road, which may lead to congestion and long travel times in some peak hours.



(a)



(b)



(c)

Fig. 4. Traffic characteristics for all segments on Monday. (a) time of day with travel time, (b) time of day with travel speed, and (c) time of day with travel distance.

6. MODEL DEVELOPMENT

This section describes the data pre-processing applied to the dataset, followed by the method used to construct each of the MLPs and LSTM models.

6.1. Data Restraints

Data pre-processing is often considered an essential step in deep learning to be done before starting to build the models for the traffic forecasting task. A few challenges were encountered while developing the model. Some data were missing, which was solved through linear interpolation. The data used were provided by Al-Moumayaz Taxi in Amman, Jordan, without an explanation for such gaps. Moreover, due to the changing traffic conditions throughout the day, records should ideally be provided for all days, hours of the day, and every sequence of five minutes (which was not the case in the dataset), to build a robust model. Given these limitations, the analysis and comparison of the models are interpreted with the primary focus of this research, which is demonstrating the potential of deep learning techniques, specifically MLPs and LSTM, in addressing the travel time prediction problem.

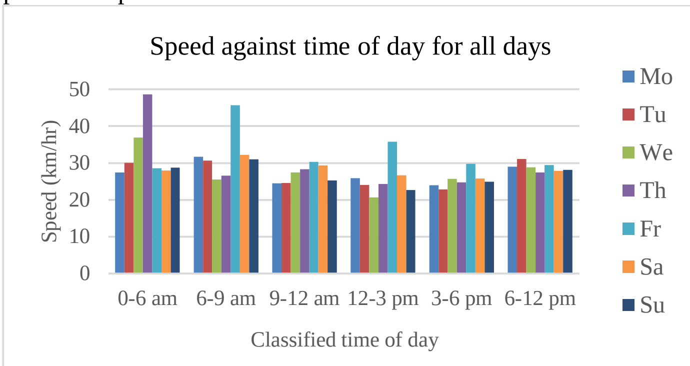


Fig. 5. Travel speed on all segments on all days.

6.2. Data Distribution

The distribution of travel data representing the relationship between travel time and travel distance in the Amman dataset is shown in Figure 6. The data illustrates that there is a specific pattern between time and distance, which helps to get good output from learning methods.

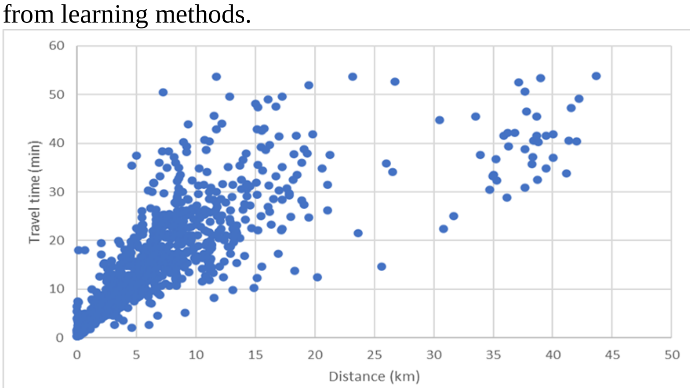


Fig. 6. Distribution of travel distance and travel time.

6.3. MLPs Model

Table 1 shows the parameters of the MLPs model that is used for Amman. In this model, the dataset was divided into two subsets: training and test subsets with an 80% and 20% split, respectively.

The choice of three layers is a common practice in MLP modelling. This specific choice of several layers ensures the balance between the model complexity and the training efficiency. The choice of the specific numbers of neurons in each layer is based on experimentation and tuning of the model in the training process. The activation of the rectified linear unit (ReLU), which is computationally efficient compared to other activation functions, in the model developing process introduces non-linearity to enable the model to capture the complex nature of the datasets. This complexity is compelled on the dataset likely due to the changing surrounding environment such as changing traffic conditions. The activation of ReLU also helps the model to avoid slow learning, by the simple assumption of outputting the input if it is positive and zero otherwise, which helps gradients to flow more easily during backpropagation. Additionally, the ReLU function helps prevent model overfitting by allowing the model to predict sparse output. The utilization of Normal and Glorot_uniform initializer is also a common practice in developing MLP models, aimed at preventing the vanishing or exploding of the gradient in the training process [24]. The utilization of Adam optimizer is a common practice which allows the model to adapt to a wide range of problems. The number of epochs (160) sets the number of times the entire dataset is passed forward and backward through the neural network in the training process. The specific number of epochs is often determined through experimentation and validation performance. A higher number of epochs can lead to better convergence and performance, but it also increases the risk of overfitting if the model learns noise in the training data. Whereas the batch size sets the number of samples propagated through the network before the weights are updated. A batch size of 32 is a common practice in developing MLP models that ensures the balance in training speed and model performance.

Table 1. MLPs parameters.

Parameter	*Value*
Number of layers	3
Number of neurons	100, 80, 10
Activation function	ReLU
Initializer	Normal, Glorot_uniform
Optimizer	Adam
Number of the epoch	160
Batch size	32

6.4. LSTM Model

Table 2 lists the choice of hyperparameters for the developed LSTM model. The LSTM model was enhanced by incorporating more parameters in the new model while keeping the division of the dataset into two subsets: training and testing with an 80% and 20% split, respectively. The additional parameters were the activation function, learning rate, and the optimizer.

The choice of two layers in the LSTM model against three layers compared to MLP modelling, allows the model to capture patterns in the data without introducing unnecessary complexity. The number of neurons is configured with 50 and 1 neurons, with the first layer of 50 neurons to take in the input data, and the second layer to give one output. The learning rate of 0.003 is specified through experimentation and tuning of the model. The rationale behind the activation of the ReLU, utilization of the Adam optimizer, specifying the number of epochs, and setting the batch size in the LSTM model is similar to that of the MLPs model (refer to subsection 6.3).

Table 2. LSTM parameters.

Parameter	*Value*
Number of layers	2
Number of neurons	50, 1
Activation function	ReLU
Learning rate	0.003
Optimizer	Adam
Number of the epoch	160
Batch size	32

7. RESULTS

This section summarizes and discusses the main results obtained from applying MLPs and LSTM to the test dataset. In addition, the models are evaluated using a routing engine.

7.1. Root Mean Square Error (RMSE)

The mean squared error function was chosen as the baseline for our prediction. The RMSE values obtained from the MLPs and LSTM models for the Amman case were 3.62 and 8.73, respectively. Compared to the cut-off value, which is considered 15 similar to previous models [11], it is clear that the error rate in Amman models is lower than the acceptable error. This means that deep learning in the study case produces a powerful model that can make accurate travel time predictions in the road network. The RMSE is widely accepted as an appropriate measure for travel time prediction tasks since it directly assesses the average magnitude of the errors, which is crucial for evaluating the accuracy of regression models. The choice of RMSE is further supported by its extensive use in previous studies on deep learning-based travel time prediction. It has demonstrated effectiveness in evaluating the performance of models in the literature, including MLPs and LSTM models.

7.2. Learning Curves

Figures 7 and 8 depict the learning curves for MLPs and LSTM models, respectively. The models exhibit different learning curves. The figures show a strong relationship between the testing and training values across the number of epochs. The close alignment between the training and testing curves in these figures suggests that neither model is overfitting, and both models generalize well to unseen data. The consistent decrease in loss values further indicates good model performance. Overall, the learning curves highlight the effectiveness of both models, demonstrating low training and validation losses and suggesting strong generalization.

7.3. Actual and Predicted Value Relationship

The actual and predicted values for all segments in the model are illustrated in Figure 9. The x-axis represents the number of segments in the sample, while the y-axis represents the speed in km/hr. It can be noticed that the MLP model shows predicted values that closely match the actual values, which confirms the strength of this model. Additionally, the consistency between predicted and actual values across the entire range of segments shows that the model generalizes well to different segments, maintaining accuracy throughout. The relatively low variance in the difference between predicted and actual values further supports the robustness of the model.

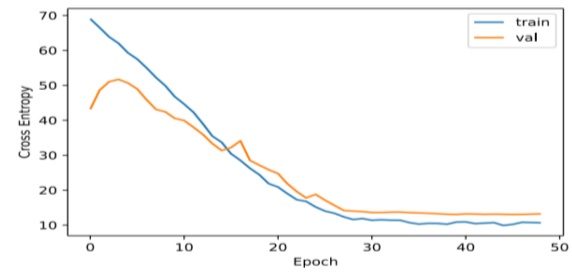


Fig. 7. Learning curves for the MLPs model.

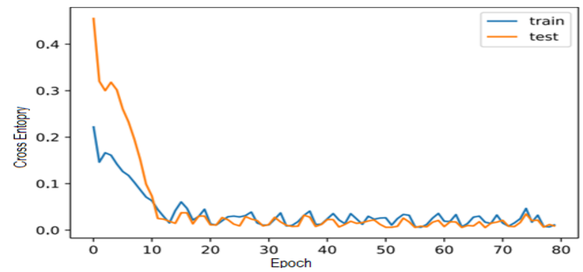


Fig. 8. Learning curves for the LSTM model.

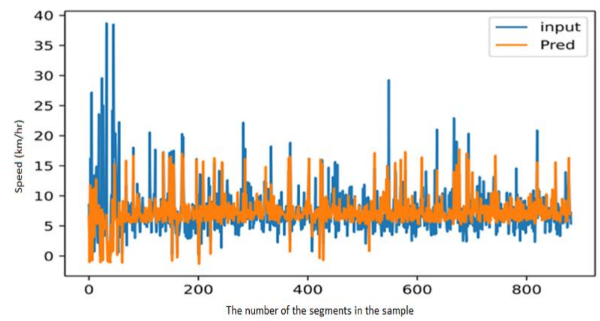


Fig. 9. Actual and predicted values for the MLPs model.

7.4. Model Evaluation

To evaluate the developed models, we compared their output with travel data from an open-source routing machine (OSRM). Figure 10 shows the relationship between the predicted value and the value obtained from OSRM, where the x-axis represents the number of the segments in the sample while the y-axis represents the speed in km/hr. It can be observed that there is a difference between the two curves. This difference arises because OSRM and the website mapping tools determine the travel time as a fixed value between any two points, without taking into account the effect of temporal factors on travel times. This discrepancy underscores the need for a more robust model that incorporates the time of travel through deep learning.

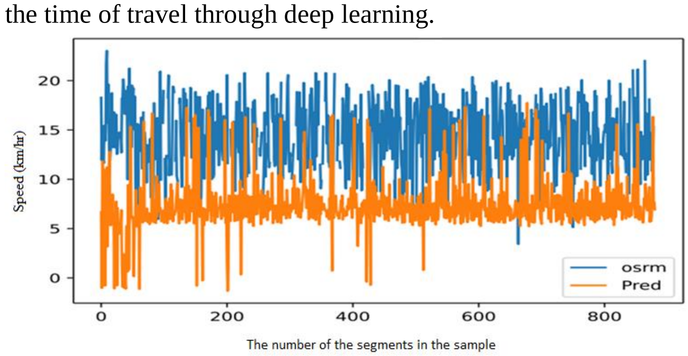


Fig. 10. Predicted value vs. OSRM value.

8. CONCLUSION AND RECOMMENDATIONS

8.1. Conclusion

Travel time prediction is a complex and challenging task. This study has tackled this issue using multiple deep learning techniques, specifically MLPs and LSTM Networks. The developed networks performed well, competing with state-of-the-art methods and models. The results indicate that the MLPs outperform the LSTM model, with RMSE values of 3.62 and 8.73, respectively. These results can be attributed to the nature of the collected data, which were not fully sequenced through the travel time, leading to higher errors in the LSTM model due to its reliance on data arrangement. Additionally, the use of one-hour interval data with coarse information and the presence of missing data points impacted the robustness of the model and its ability to generalize.

8.2. Recommendations for Future Work

Deep Learning has proven to be a promising method for travel time prediction. However, the coarse nature of the data (one-hour intervals) produced sub-par performances from the models. The predictive power of the developed models could be further improved by introducing augmented data, such as dividing each hour into four fifteen-minute intervals and determining travel times for each interval. Future research should employ advanced computational technology such as graphic processing units (GPUs). Additionally, to provide a robust and scientifically sound comparison of deep learning model performance in travel time prediction across different datasets, future studies should aim to collect and analyze travel time data from a diverse set of locations and periods. Finally, further research is necessary to thoroughly validate and evaluate the proposed models against predefined evaluation metrics.

References

[1] E.I. Vlahogianni, M.G. Karlaftis, and J.C. Golias, “Short-term traffic forecasting: Where we are and where we’re going,” *Transportation Research Part C: Emerging Technologies*, 43(1), pp. 3-19, 2014.

[2] W. Huang, G. Song, H. Hong, and K. Xie, “Deep architecture for traffic flow prediction: Deep belief networks with multitask learning,” *IEEE Transactions on Intelligent Transportation Systems*, 15(5), pp. 2191–2201, 2014.

[3] R. Fu, Z. Zhang, and L. Li, “Using LSTM and GRU neural network methods for traffic flow prediction,” *in Proccedings of the 31st Youth academic annual conference of Chinese association of automation (YAC),* Wuhan, China, 11-13 November 2016, pp. 324-328.

[4] Y.J. Wu, F. Chen, C.T. Lu, and S. Yang, "Urban traffic flow prediction using a spatiotemporal random effects model,” *Journal of Intelligent Transportation Systems*, 20(3), pp. 282–293, 2015.

[5] J. Rice and E. vanZwet, “A simple and effective method for predicting travel times on freeways,” *IEEE Transactions on Intelligent Transportation Systems*, 5(3), pp. 200–207, 2004.

[6] N. Polson and V. Sokolov, “Deep learning predictors for traffic flows,” *arXiv preprint arXiv:1604.04527*, 2016.

[7] D. Park and L. R. Rilett, “Forecasting freeway link travel times with a multilayer feedforward neural network,” *Computer-Aided Civil and Infrastructure Engineering*, 14(5), pp. 357–367, 1999.

[8] N. Wisitpongphan, W. Jitsakul, and D. Jieamumporn, “Travel time prediction using multi-layer feed-forward artificial neural network,” in *Proceedings of the 2012 Fourth International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN 2012), Phuket, Thailand*, *July 24-26, 2012*, pp. 326-330.

[9] Google Maps, [Online], Available: https://www.google.com/maps [Accessed: Sept. 12, 2024].

[10] G. Huang, J. Xing, L. Meng, F. Li, and L. Ma, “Travel time prediction for bus rapid transit using a statistics-based probabilistic method,” in *Proceedings of the 2010 2nd International Conference on Signal Processing Systems (ICSPS)*, Dalian, China, 05-07 July 2010, pp. V1-457–V1-460,

[11] Y. Duan and F. Wang, “Travel time prediction with LSTM neural network,” in *Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil, 2016*, pp. 1053-1058.

[12] D. Wang, Y. Wu, and Z. Xiao, “A Gaussian process regression method for urban road travel time prediction,” in *Proceedings of the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD),* Guilin, China, 29-31 July 2017, pp. 890-895,

[13] H. Rodriguez-Deniz, E. Jenelius, and M. Villani, “Urban network travel time prediction via online multi-output Gaussian process regression,” in *Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC),* Yokohama, Japan, 16-19 October 2017, pp. 1-6.

[14] Y. Liu, Y. Wang, X. Yang, and L. Zhang, “Short-term travel time prediction by deep learning: A comparison of different LSTM-DNN models,” in *Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC),* Yokohama, Japan, 16-19 October 2017, pp. 1-8.

[15] Y. Hou and P. Edara, “Network scale travel time prediction using deep learning,” *Transportation Research Record*, 2672(1), pp. 1–9, 2018.

[16] K. D. Kankanamge, Y. R. Witharanage, C. S. Withanage, M. Hansini, D. Lakmal, and U. Thayasivam, “Taxi trip travel time prediction with isolated XGBoost regression,” in Proceedings of the *2019 Moratuwa Engineering Research Conference (MERCon),* Moratuwa, Sri Lanka, 3-5 July 2019, pp. 54–59.

[17] C. Ding, C. Wang, Y. Cao, and J. Chen, “A spatial and time analysis method of big data of public transport card based on classification statistics and visualization,” in *Proceedings of the 2020 Eighth International Conference on Advanced Cloud and Big Data*, Beijing, China, 05-06 December 2020, pp. 62–67.

[18] X. Xu, S. L. Keoh, C. K. Seow, Q. Cao, and S. K. Abdul Rahim, “Towards prediction of bus arrival time using multi-layer perceptron and MLP regressor,” in *Proceedings of the 2023 8th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand*, 18-19 May *2023*, pp. 669-674.

[19] V. V. Nampalli, C. Gudla, and Md. S. Rana, “Predicting travel time in complex road structures using deep learning,” in *Proceedings of the 2024 IEEE 14th Annual Computing and Communication Workshop and Conference*, January 2024.

[20] R. Ammoura, K. Jadaan, and N. Sobh, "Developing deep learning models for network-scale travel time prediction," *Unpublished Report*, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, USA, 2022.
[21] F. Hakem, "Data Scientist | Machine Learning Engineer," [Online]. Available: https://www.faroukhakem.com/index.html# (Accessed: May 10, 2024).
[22] F. Filiz, "4.1.1 Artificial Neural Networks," *Medium*, Aug. 19, 2017, [Online]. Available:https://medium.com/@fahrettinf/4-1-1-artificial-neural-networks-6257a7a54bb3 [Accessed: May 10, 2024].
[23] F. Dijkstra, "Breakpoint detection through neural nets: A feasibility study," *Semantic Scholar*, Master Thesis, TU Delft, Naderland, 2020.
[24] X. Glorot, and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *in Proceedings of the thirteenth international conference on artificial intelligence and statistics*, JMLR Workshop and Conference Proceedings, March 2010. pp. 249-256.

Biographies

Khair Said Jadaan was born on March 27 1948, and attended the high school in Jordan. In 1975, he earned his Ph.D. degree in traffic engineering and planning from the University of Bradford, UK His working experience covers a variety of positions in both private and public sectors in various developed and developing countries including New Zealand, Germany, UK, USA, Kuwait, Iraq and Jordan. He is currently an Emeritus Professor of transportation engineering at the University of Jordan. He has worked as a consultant for several international organizations including the World Bank, UN organizations (UNECE, UNECEF, ESCWA) and the Arab Fund for Economic and Social Development (AFESD). He is a member of the Editorial Board of four International Scientific journals. He has published over 300 papers in international journals and conferences. He is a Fellow of IHTE (UK), member of IPENZ (New Zealand), member of ASCE (USA), and Jordan Engineers Association, and has been awarded several honoraries.
E-mail: kjadaan@gmail.com

Abdullah Ahmad Najjar born in 1995 and attended high school in Jordan. He earned his Bachelor's degree in Civil Engineering from the University of Jordan in 2017 with an excellent grade, then earned his Master's degree in Transportation Engineering in 2020 from the same university with an excellent grade. He is a member of the junior research team at the Civil Engineering Department, University of Jordan, and has co-authored a number of papers and presentations. His main research interests include traffic safety and Artificial Intelligence.
E-mail: abd.ah.najjar@gmail.com

Sameer Abdallah Abu-Eisheh was born on August 25 1960, and attended the high school in Palestine. In 1987, he earned his Ph.D. degree in Civil Engineering (Transportation) from the Pennsylvania State University, USA, the M.Sc. from the same university in 1894, and the B.Sc. in Civil Engineering from University of Jordan in 1982. He served as the Head of the Department of Civil Engineering, the Dean of Engineering, and the President's Assistant for Planning and Development. He was a visiting Professor at French, Germany and US universities. He served as an advisor to the Minister of Planning in Palestine since its establishment, and appointed later as the Minister of Planning. He served also as the Acting Minister of Finance, and Acting Minister of Education and Higher Education. His fields of experience include as well road and traffic engineering, management of transportation systems, strategic planning, and national and regional planning. He authored more than 120 published or conference papers. He has offered consulting services to various international, Arab, and local entities. He is member of the ASCE (USA), and Jordan Engineers Association, and has been awarded several honoraria.
E-mail: sameeraa@najah.edu

Abdallah Sameer Abuaisha was born on September 23 1992, and attended the high school in Palestine. He is currently a Ph.D. candidate in Data Science and Artificial Intelligence at Monash University in Australia, where he has been focusing on optimized routing for personalized public transportation systems since 2021. He has also been serving as a teaching associate at Monash University since 2022, contributing to both the Civil Engineering Department and the Department of Data Science and Artificial Intelligence. Before pursuing his Ph.D., he worked as a Transportation and Traffic Engineer in Palestine between 2015 and 2021. He earned his Master's degree in Transportation Engineering at the Technical University of Munich in Germany in 2019, following his Bachelor's degree in Civil Engineering from An-Najah National University in Palestine in 2015. He also completed internships in the Sultanate of Oman in 2014 and at Siemens AG in Germany in 2017. He has been a member of the ASCE (USA) since 2013 and was a co-founding member of the first ASCE student chapter in Palestine. He has been a member of the Engineers Association of Palestine and Jordan since 2015. He has published several papers on travel behavior, public transportation planning, and transportation optimization.
E-mail: abdallah.abuaisha@monash.edu