SEMARAK ILMU PUBLISHING
202103268166(003316878-P)

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Public Transport Delay Prediction using Deep Learning

Siti Nurulain Mohd Rum[1,*], Meor Muhammad Nazmi Meor Yusoff[1], Amalia Mahdi[2]

[1] Faculty of Computer Science & Information Technology, Universiti Putra Malaysia, Malaysia
[2] Faculty of Computer Science & Information Technology, Universiti Utara Sumatera, Indonesia

ABSTRACT

A system of trains, buses, subways, ferries, and other vehicles that are accessible to the public is referred to as public transportation, also known as public transit or mass transit. These networks are usually run by private or public organisations are made to efficiently transport large numbers of people in urban and suburban areas. It usually operates on set schedules and routes and has a listed fare for each journey. Long wait periods and lengthy travel times resulting from delays are two problems that frequently plague public transportation services. There are several things that can cause a delay in public transport, such as an excess of passengers, heavy traffic, accidents, and other unforeseen circumstances. The availability of a more accurate delay prediction for public transportation might increase users' confidence and their willingness to pay more for transit services. Over the past two decades, a number of studies on prediction algorithms for transportation data have been proposed. Most of the work is on machine learning model development, focusing on delay prediction and taking into account several factors such as weather conditions and infrastructure issues. This paper proposes a deep learning model to predict public transportation delays using data from public transportation and the weather. The results obtained from this research work are compared with several other existing works. Our experiment has demonstrated that the deep neural network (DNN) is the best model to predict transit delay compared to several other machine learning and deep learning models.

Keywords:
Public transportation; delay prediction; machine learning; deep learning; deep neural network (DNN)

1. Introduction

Big Data is the term used to describe the enormous volume of data produced daily by people, companies, and organisations. This data encompasses everything from social media to sensor data from Internet of Things (IoT) devices and online transactions [1]. There are some key areas where big data is making significant impacts in research. Some recent examples of big data used in past studies are disease and epidemics [2-4], social media [5-7] and transportation [8-10]. Big data contains unknown and possibly significant knowledge and information that is frequently extracted using machine learning and data mining methods [11]. Increasing productivity and profit while raising living standards and the quality of work operations is the goal of many. Today, big data has also emerged

* *Corresponding author.*
E-mail address: snurulain@upm.edu.my

https://doi.org/10.37934/araset.59.2.168177

as a crucial instrument for studying and enhancing public transportation systems, as it is important for many people, whether in cities or rural areas. The quality and availability of public transport services in a country can greatly affect citizens' lives [12], as they may rely on them for transit to work, school, and other areas. Having a good and efficient public transport system is important for facilitating better travel and reducing road congestion. Public transport services commonly have several issues, including long waiting times and long travel times due to delays. Public transport delays can be caused by several issues, such as traffic congestion and bad weather conditions. Predicting bus delays is one example of a crucial element in an intelligent public transportation system (IPTS). The combination of several non-linear components, such as traffic conditions, dwell periods, incidents, etc. may affect public transport arrival times. Typically, it is challenging to represent these interactions using traditional modelling tools. The availability of a more accurate estimation of public transportation delays can increase user confidence to use these services daily. Precisely estimating journey time can lower transportation expenses by avoiding congested areas thus improve the punctuality and quality of services.

The transit and weather data include many features that are measured in numbers through sensors. While sensors and computers can precisely measure this data, it can be difficult for humans to understand the numbers. Therefore, there is a need for algorithms to improve prediction accuracy due to the complex nature of the urban environment and the various factors influencing bus arrival times. With machine learning algorithms, better results and robustness can be achieved. However, extrapolating large datasets over short time periods to produce precise forecasts is challenging. There are a number of studies have been developed to predict public transport delays using machine learning [13-15], however, there are no previous studies that make comparisons between these models. Therefore, this research paper presents Mean Absolute Error (MAE) comparison between several deep learning models as well as machine learning models for predicting public transport delays using a public transport and weather dataset.

2. Related

Over the past two decades, number of studies on prediction algorithms for transportation data have been proposed [16-18]. Most of them were focused on transportation and traffic flow analytics [19-21]. Some examples of these that use more complex approaches are the automatic personality categorization (APC) [22], artificial neural networks (ANN) [23], additive model [16], mean absolute percent error (MAPE), vehicle ad hoc networks (VANETs) [17] and support vector regression (SVR) [24]. In addition, multiple studies have used dynamic data from the advanced public transportation system [20], automatic fare collection (AFC) [23], and global location system (GPS) [25]. Furthermore, some related works use Kalman filters to include dynamic data in the aforementioned algorithms [19, 20, 22, 23]. Recently, there have been even more studies on developing prediction models, for public transportation delays. These studies have also used a wide variety of models, such as Bayesian networks [9, 10], decision trees [7], deep learning [26-30], Kalman filters [31], multivariate regression [32], and random forests [33]. Furthermore, Serin *et al.,* [34] have also tested and combined multiple models in a three-layer architecture. However, many of these recent studies have only used public transit data for their model. Many other factors, such as road [26] and weather conditions [26, 33] could have an impact on public transportation. By including other external data into the transit prediction models can provide more holistic model. Table 1 presents a summary of recent works on transportation delay prediction.

Table 1
Summary of previous studies on public transportation prediction models

Models	Authors	Dataset	Performance
Bayesian network	Corman and Kecman [10]	Train transit	MAE – ~0.7 – ~1.6 minutes
	Lessan *et al.,* [9]	Train transit	Accuracy – more than 80%
Decision trees	Audu *et al.,* [7]	Bus transit	Accuracy – 89%
Deep learning	Wu *et al.,* [28]	Train transit	MAE – 37.56 seconds
	Yu *et al.,* [30]	Flight transit	MAE – 8.41 minutes
	He *et al.,* [26]	Bus transit, roads, weather	MAE – 4.863 minutes
	Shoman *et al.,* [27]	Bus transit, location, probe	MAPE – 5.32%
	Zhang *et al.,* [29]	Train transit	MAE – 0.16 minutes
Kalman filters	Achar *et al.,* [31]	Bus transit, GPS	MAPE – 14% – 25%
Mixed	Serin *et al.,* [34]	Bus transit	MAPE – 2.552
Multivariate regression	Celan and Lep [32]	Bus transit, GPS	MAE – 55.7s – 63.7s
Regression	Evangeline *et al.,* [35]	Flight transit	MAE (0.2 and 0.1), MSE (0.1 and 0.04) RMSE (0.3 and 0.2) and Accuracy (99.7% and 99.8%)
Fuzzy logic	Leung *et al.,* [33]	Streetcar transit, weather	MAE – 3.70 minutes

The model presented in [10] computes the complexity of random-variable stochastic inference, enabling a real-time reduction of the uncertainty of future train delays as well as the update of probability distributions. The study focuses on the precision with which train delays are predicted, that is crucial for anticipating and acting proactively to control railway traffic in real time. The suggested model introduced a dynamic method for predicting train delays by taking into consideration the variability and unpredictable nature of railway traffic process times. The authors utilize a Bayesian network to model the uncertainty of train delays, using the conditional independences between events as a means of computing of their joint distribution. Lesson et al. [9] discuss the development and evaluation of a model for train operating delay prediction using a Bayesian network, specifically on a high speed railway line. The study addresses the challenges posed by uncertainties and disturbances in train operations, focusing on the significance of accurate delay predictions to improve the efficiency of train services. The proposed model aims to capture the complexities and dependencies in train delays by integrating historical data and domain knowledge to enhance prediction accuracy. The study evaluates three types of Bayesian network structures, including a heuristic hill-climbing approach, a primitive linear structure, and a hybrid model refined with domain knowledge and expertise. The evaluation results indicate that the hybrid Bayesian network model outperforms the other structures, attaining more than 80% prediction accuracy in a 60-minute time frame. The study further assesses the model's performance through various measures, like the root mean square error (RMSE), mean absolute error (MAE), and mean error (ME) and demonstrates its potential for real-world application in improving delay management decisions and the deployment of trains.

Yu *et al.,* [30] in their study used a high dimensional dataset provided by Beijing International Airport to create a practical model for flight delay prediction. They proposed using support vector regression (SVR) together with a deep belief network (DBN) approach to identify key determinants of delay. The study found important variables such as airport congestion, air route conditions, and previous flight delays, and demonstrated the DBN-SVR model's superior performance compared to benchmark techniques like support vector machines, linear regression, and k-nearest neighbors. The results indicated that the proposed model achieved high accuracy in predicting flight delays, meeting the industry's requirement of at least 98% of predicted delays being within 30 minutes deviation from actual delays [29].

The study's findings [31] revealed the data network selection has a major influence on how accurate bus arrival time estimates are, with the best performance observed for a data model that considers potential barriers affecting bus travel speed. The analysis also highlights the influence of time periods on prediction accuracy, showcasing the significance of segmenting travel times based on different time slots. The study's comprehensive evaluation, based on over 192,000 bus location data points, demonstrates that the proposed model, especially when considering the data model that determines travel speed barriers, yields accurate predictions with over 90% success rate for public bus service users, showcasing the practical applicability and reliability of the proposed approach.

Evangeline *et al.,* [35] in their study using different regression algorithms in machine learning explore the application of various regression algorithms in machine learning for predicting flight delays. It discusses the challenges associated with flight delays and the importance of accurate prediction for both airlines and passengers. The study compares the performance of different regression algorithms, such as gradient boosting regression, decision tree regression, random forest, and linear regression, in predicting flight delays. Through experimentation and evaluation, the article aims to identify the most effective algorithm for flight delay prediction, offering insights into potential improvements in air travel management and passenger experience. Through experimentation and evaluation, the article aims to identify the most effective algorithm for flight delay prediction, offering insights into potential improvements in air travel management and passenger experience.

Leung *et al.,* [33], in their study introduces an approach to predictive analytics in transportation using fuzzy logic-based machine learning. The method aims to address the challenges of analyzing large-scale transportation data by integrating fuzzy logic with machine learning techniques. It discusses the advantages of fuzzy logic in handling uncertainty and imprecision inherent in transportation data. The proposed algorithm is designed to predict various transportation-related outcomes, such as traffic flow patterns and travel times, to improve decision-making in urban transportation systems. The article highlights the potential of this innovative approach to enhance predictive analytics and support the development of smarter transportation systems [23].

3. Methods

This section provides the description of method to develop the proposed model that includes the environment, dataset used, data pre-processing, deep learning model specifications and model evaluation metric data description.

3.1 Environment

This study was conducted using a PC running on Windows 10 operating system with Intel Core i5-8400 CPU @ 2.80GHz, NVIDIA Geforce GTX 1060 6GB GPU, 16GB RAM 2,666MHz. The code was programmed using Keras, Tensorflow, sklearn, and pandas libraries using python programming language through Jupyter Notebook.

3.2 Data Description

The datasets used for this study are the Toronto Transit Commission (TTC) delay dataset from the Open Data Portal of the City of Toronto and weather data from the Weather Dashboard for Toronto. The data is selected for a six-year period, from January 1, 2014, to December 31, 2019. The streetcar delay dataset has 78,525 entries with 10 features describing streetcar delay incidents in Toronto, while the weather dataset has 2,191 entries with 70 features describing Toronto’s daily weather data. Table 2 provides a list of attributes of the dataset with descriptions.

Table 2
List of dataset attributes

Column Name	Column Description
A. 1 *B.* Time	*C.* Time of delay incident
D. 2 *E.* Temperature	*F.* Average of all hourly temperatures within the day
G. 3 *H.* Wind	*I.* Average of all hourly wind speeds within the day
J. 4 *K.* Visibility	*L.* Average of all hourly visibility within the day
M. 5 *N.* Rain	*O.* Amount of rainfall within the day
P. 6 *Q.* Snow	*R.* Amount of snowfall within the day
S. 7 *T.* Snow on Ground	*U.* Amount of snow on ground within the day
V. 8 *W.* Min Delay	*X.* Length of transit delay in minutes

Figure 1 shows the architecture of the deep neural network (DNN) model, which shows layers between the input and output with several hidden levels. These hidden layers enable DNNs to discover intricate patterns in data, making them powerful tools for reinforcement learning, natural language processing, and image recognition. In neural networks, activation functions specify how each node's output is created from the weighted sum of its inputs within the network’s layers. This model uses rectified linear units (ReLU) as the trigger function for the hidden layers. Here is how the ReLU function is defined:

$$f(x) = \max(0, x) \tag{1}$$

where x is the input data for the layer. The function returns 0 in the case of any negative input and returns x in the case of positive input. Rather, it gives back the value it received.

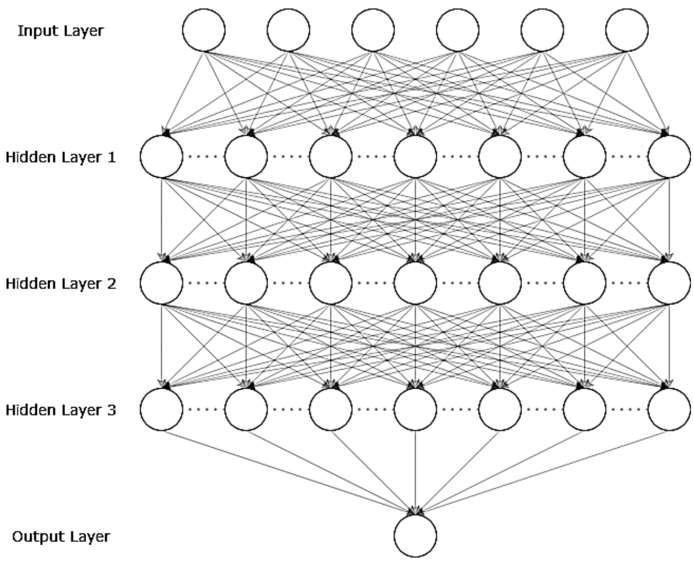

Fig. 1. Deep neural network model architecture

Futhermore, the output layer uses a linear activation function that returns the input value without any changes. The process of training neural networks is an optimization problem where we attempt to find the most optimal neural network model for our objective. Loss functions are functions that calculate the difference between the expected output and the predicted output of a model. The loss function that was used in this algorithm is mean squared error (MSE) which is defined as:

$$MSE = (y_{true} - y_{pred})^2 \quad (2)$$

where y_{true} is the true length of delay and y_{pred} is the predicted length of delay. During training, the algorithm tries to find the model with the minimum loss by using optimizers. Optimizers are algorithms or methods that modify neural network parameters to minimise loss. The optimizer chosen for this model is the Adam algorithm, a stochastic gradient descent technique predicated on adaptive first- and second-order moment estimation. Table 3 shows a list of the hyperparameters that were used for the DNN model.

Table 3
List of DNN model hyperparameters

Hyperparameter		Value of method	
Y.	Number of hidden layers	*Z.*	3
AA.	Number of hidden units	*BB.*	[100, 50, 25]
CC.	Weight initialization	*DD.*	Glorot uniform
EE.	Activation function	*FF.*	ReLU
GG.	Optimizer	*HH.*	Adam
II	Learning rate	*JJ.*	0.001
KK.	Number of epochs	*LL.*	100
MM.	Batch size	*NN.*	50

3.3 Cross-Validation

K-fold cross-validation is a frequently employed technique in machine learning for optimisation and to minimise overfitting. The data is divided into k folds of the same size using this procedure. Then, k-1 folds are used as the training set, while the remaining fold is used as the holdout set. The holdout set is used to assess the model after it has been fitted to the training set. This process is repeated k times, with a different fold used as the hold setting each time. For this study, a 10-fold cross-validation was used with an 80/20 train/test split.

3.4 Model Evaluation Metrics.

The corresponding percentage for mean absolute error (MAE) is mean absolute percentage error (MAPE), is the most widely used measures of model prediction accuracy. The MAPE providing a clear and interpretable measure of forecast error by a model. In other words, it expresses how much it deviates from the expectation measures. While knowing how to compute this measure and comprehending its significance are vital, it's equally crucial to comprehend its advantages and disadvantages before applying it in a production setting. Although the quality of a machine learning model may depend on the training data, the model's overall assessment is determined by the performance metrics employed in the real world. If the data exhibit irregularities such as skewing, an abundance of outliers, or a high percentage of zeros and nans, the performance metric will be responsible for assessing the efficacy of your model and identifying these problems. A performance metric must be select that properly fits the use case to compare the model's predictions to its ground truths, or actuals, and obtain a deeper understanding of how your model affects user behaviour, profitability, and other measurements. The mean absolute error (MAE) is used to assess the efficacy of the suggested model. These metrics are defined as follows:

$$MAE = |y_{true} - y_{pred}| \tag{3}$$

where y_{true} is the true length of delay and y_{pred} is the predicted length of delay.

4. Results

In this section, the accuracy of different models is presented. The results are measured in minutes. Table 4 displays the accuracy of the trained DNN model on the test set over 10 folds measured in MAE. The average MAE over the 10 folds is 3.21 minutes. Whereas the average of 10 folds of other models is presented in Table 5.

Table 4
DNN model performance over 10 folds

Model k	MAE
Model 1	3.15
Model 2	3.15
Model 3	3.26
Model 4	3.23
Model 5	3.26
Model 6	3.21
Model 7	3.21
Model 8	3.21
Model 9	3.20
Model 10	3.25
Mean	3.21

Table 5 presents MAE results for several model in machine learning together with deep learning models for predicting public transport delays using TTC dataset. Obviously, the results show the Deep Neural Network (DNN) outperformed other models with an MAE of 3.21 over 10 folds. In evaluating machine learning models, including deep learning models, 5-fold or 10-fold cross-validation is commonly used in practice. These values strike a balance between variance reduction and computational cost. The DNN model's results also show a promising improvement over the previous study. This result could suggest that, when predicting the length of delay in public transportation, a DNN model could be considered one of the best models. However, there are several things that need to be considered when predicting public transport delays, such as architectures, hyperparameters, and preprocessing techniques, to determine which model performs best. Another process that may have contributed to better model performance is feature selection. The suggested model does not include the rain and snow variables in its dataset because it is believed that they will negatively impact the model learning process.

Table 5
Comparison of model performance for transit delay prediction

Type	Model	MAE
Machine learning	Support vector machine (SVM)	3.60
Machine learning	Random forest	3.71
Machine learning	Bayesian network	3.59
Machine learning	Fuzzy logic	3.70
Deep learning	Long short-term memory (LSTM)	3.37
Deep earning	Deep neural network (DNN)	3.21
Deep learning	Feedforward neural network (FNN)	3.52

5. Conclusions

This paper presents the accuracy of different machine learning and deep learning models in predicting public transport delays. Our finding shows that the DNN model perform better than other models with Mean Absolute Error (MAE) 3.21 of over 10 folds. Based on the findings, deep learning appears to be a better method for predicting delays in public transport. However, to ascertain which model performs best, several factors, including preprocessing techniques, architectures, and hyperparameters, must be considered when predicting public transport delays. Determining the DNN model's ideal hyperparameters is one area where further development can be done. There are other methods for hyperparameter optimisation that can be investigated, and using the ideal hyperparameters should enhance the performance shown in this work. Future research could also benefit from looking into other deep learning methods for predicting public transit delays and comparing the results of these models on various transit datasets.

Acknowledgement
This research was funded by a grant from Universiti Putra Malaysia (Geran Putra Inisiatif (GPI) – Pusat kos 9777500).

References
[1] Katal, Avita, Mohammad Wazid, and Rayan H. Goudar. "Big data: issues, challenges, tools and good practices." In *2013 Sixth International Conference on Contemporary Computing (IC3), IEEE*, p. 404-409. 2013. https://doi.org/10.1109/IC3.2013.6612229

[2] Wang, C. Jason, Chun Y. Ng, and Robert H. Brook. "Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing." *Jama* 323, no. 14 (2020): 1341-1342. https://doi.org/10.1001/jama.2020.3151

[3] Bai, Yan, Lingsheng Yao, Tao Wei, Fei Tian, Dong-Yan Jin, Lijuan Chen, and Meiyun Wang. "Presumed asymptomatic carrier transmission of COVID-19." *Jama* 323, no. 14 (2020): 1406-1407. https://doi.org/10.1001/jama.2020.2565

[4] MacLaren, Graeme, Dale Fisher, and Daniel Brodie. "Preparing for the most critically ill patients with COVID-19: the potential role of extracorporeal membrane oxygenation." *Jama* 323, no. 13 (2020): 1245-1246. https://doi.org/10.1001/jama.2020.2342

[5] Leong, Lai-Ying, Teck-Soon Hew, Keng-Boon Ooi, Voon-Hsien Lee, and Jun-Jie Hew. "A hybrid SEM-neural network analysis of social media addiction." *Expert Systems with Applications* 133 (2019): 296-316. https://doi.org/10.1016/j.eswa.2019.05.024

[6] Jain, Gauri, Manisha Sharma, and Basant Agarwal. "Spam detection in social media using convolutional and long short term memory neural network." *Annals of Mathematics and Artificial Intelligence* 85, no. 1 (2019): 21-44. https://doi.org/10.1007/s10472-018-9612-z

[7] Audu, Abdul-Rasheed A., Alfredo Cuzzocrea, Carson K. Leung, Keaton A. MacLeod, Nibrasul I. Ohin, and Nadège C. Pulgar-Vidal. "An intelligent predictive analytics system for transportation analytics on open data towards the development of a smart city." In *Complex, Intelligent, and Software Intensive Systems: Proceedings of the 13th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS-2019)*, Springer International Publishing p. 224-236. 2020. https://doi.org/10.1007/978-3-030-22354-0_21

[8] Leung, Carson K., Jonathan D. Elias, Shael M. Minuk, A. Roy R. de Jesus, and Alfredo Cuzzocrea. "An innovative fuzzy logic-based machine learning algorithm for supporting predictive analytics on big transportation data." In *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, p. 1-8. 2020. https://doi.org/10.1109/FUZZ48607.2020.9177823
[9] Lessan, Javad, Liping Fu, and Chao Wen. "A hybrid Bayesian network model for predicting delays in train operations." *Computers & Industrial Engineering* 127 (2019): 1214-1222. https://doi.org/10.1016/j.cie.2018.03.017
[10] Corman, Francesco, and Pavle Kecman. "Stochastic prediction of train delays in real-time using Bayesian networks." *Transportation Research Part C: Emerging Technologies* 95 (2018): 599-615. https://doi.org/10.1016/j.trc.2018.08.003
[11] Bellatreche, Ladjel, Carson Leung, Yinglong Xia, and Didier El Baz. "Advances in cloud and big data computing." *Concurrency and Computation: Practice and Experience* 31, no. 2 (2019): e5053. https://doi.org/10.1002/cpe.5053
[12] Kabir, Afrida, Faiyaj Kabir, Saief Newaz Chowdhury, and AR M. Harunur Rashid. "Design, simulation and fabrication of an ergonomic handgrip for public transport in Bangladesh." *Malaysian Journal on Composites Science and Manufacturing* 13, no. 1 (2024): 98-111. https://doi.org/10.37934/mjcsm.13.1.98111
[13] Akhtar, Mahmuda, and Sara Moridpour. "A review of traffic congestion prediction using artificial intelligence." *Journal of Advanced Transportation* 2021, no. 1 (2021): 8878011. https://doi.org/10.1155/2021/8878011
[14] Boukerche, Azzedine, and Jiahao Wang. "Machine learning-based traffic prediction models for intelligent transportation systems." *Computer Networks* 181 (2020): 107530. https://doi.org/10.1016/j.comnet.2020.107530
[15] Majumdar, Sharmila, Moeez M. Subhani, Benjamin Roullier, Ashiq Anjum, and Rongbo Zhu. "Congestion prediction for smart sustainable cities using IoT and machine learning approaches." *Sustainable Cities and Society* 64 (2021): 102500. https://doi.org/10.1016/j.scs.2020.102500
[16] Kormáksson, Matthías, Luciano Barbosa, Marcos R. Vieira, and Bianca Zadrozny. "Bus travel time predictions using additive models." In *2014 IEEE international conference on data mining*, pp. 875-880. IEEE, 2014. https://doi.org/10.1109/ICDM.2014.107
[17] Kulla, Elis, Soushi Morita, Kengo Katayama, and Leonard Barolli. "Route lifetime prediction method in VANET by using AODV routing protocol (AODV-LP)." In *Complex, Intelligent, and Software Intensive Systems: Proceedings of the 12th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS-2018)*, Springer International Publishing, p. 3-11. 2019. https://doi.org/10.1007/978-3-319-93659-8_1
[18] Rajput, Prashant, Durga Toshniwal, and Apeksha Agggarwal. "Improving infrastructure for transportation systems using clustering." In *Big Data Analytics: 5th International Conference, BDA 2017, Hyderabad, India, Proceedings 5*, Springer International Publishing, p. 129-143. 2017. https://doi.org/10.1007/978-3-319-72413-3_9
[19] Chien, Steven I-Jy, and Chandra Mouly Kuchipudi. "Dynamic travel time prediction with real-time and historic data." *Journal of transportation engineering* 129, no. 6 (2003): 608-616. https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(608)
[20] Vanajakshi, Lelitha, Shankar C. Subramanian, and R. Sivanandan. "Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses." *IET intelligent transport systems* 3, no. 1 (2009): 1-9. https://doi.org/10.1049/iet-its:20080013
[21] Williams, Billy M., and Lester A. Hoel. "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results." *Journal of transportation engineering* 129, no. 6 (2003): 664-672. https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(664)
[22] Shalaby, Amer, and Ali Farhan. "Prediction model of bus arrival and departure times using AVL and APC data." *Journal of Public Transportation* 7, no. 1 (2004): 41-61. https://doi.org/10.5038/2375-0901.7.1.3
[23] Lin, Yongjie, Xianfeng Yang, Nan Zou, and Lei Jia. "Real-time bus arrival time prediction: Case study for Jinan, China." *Journal of Transportation Engineering* 139, no. 11 (2013): 1133-1140. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000589
[24] Khan, Mahnoor, Nadeem Javaid, Muhammad Nabeel Iqbal, Muhammad Bilal, Syed Farhan Ali Zaidi, and Rashid Ali Raza. "Load prediction based on multivariate time series forecasting for energy consumption and behavioral analytics." In *Conference on Complex, Intelligent, and Software Intensive Systems*, Cham: Springer International Publishing, p. 305-316. 2018. https://doi.org/10.1007/978-3-319-93659-8_27
[25] Majumdar, Sharmila, Moeez M. Subhani, Benjamin Roullier, Ashiq Anjum, and Rongbo Zhu. "Congestion prediction for smart sustainable cities using IoT and machine learning approaches." *Sustainable Cities and Society* 64 (2021): 102500. https://doi.org/10.1016/j.scs.2020.102500

[26] He, Peilan, Guiyuan Jiang, Siew-Kei Lam, and Yidan Sun. "Learning heterogeneous traffic patterns for travel time prediction of bus journeys." *Information Sciences* 512 (2020): 1394-1406. https://doi.org/10.1016/j.ins.2019.10.073

[27] Shoman, Maged, Armstrong Aboah, and Yaw Adu-Gyamfi. "Deep learning framework for predicting bus delays on multiple routes using heterogenous datasets." *Journal of Big Data Analytics in Transportation* 2 (2020): 275-290. https://doi.org/10.1007/s42421-020-00031-y

[28] Wu, Canrong, Yang Liu, Yueying Yang, Peng Zhang, Wu Zhong, Yali Wang, Qiqi Wang et al. "Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods." *Acta Pharmaceutica Sinica B* 10, no. 5 (2020): 766-788. https://doi.org/10.1016/j.apsb.2020.02.008

[29] Zhang, Dalin, Yunjuan Peng, Yumei Zhang, Daohua Wu, Hongwei Wang, and Hailong Zhang. "Train time delay prediction for high-speed train dispatching based on spatio-temporal graph convolutional network." *IEEE Transactions on Intelligent Transportation Systems* 23, no. 3 (2021): 2434-2444. https://doi.org/10.1109/TITS.2021.3097064

[30] Yu, Bin, Zhen Guo, Sobhan Asian, Huaizhu Wang, and Gang Chen. "Flight delay prediction for commercial air transport: A deep learning approach." *Transportation Research Part E: Logistics and Transportation Review* 125 (2019): 203-221. https://doi.org/10.1016/j.tre.2019.03.013

[31] Achar, Avinash, Dhivya Bharathi, Bachu Anil Kumar, and Lelitha Vanajakshi. "Bus arrival time prediction: A spatial Kalman filter approach." *IEEE Transactions on Intelligent Transportation Systems* 21, no. 3 (2019): 1298-1307. https://doi.org/10.1109/TITS.2019.2909314

[32] Čelan, Marko, and Marjan Lep. "Bus-arrival time prediction using bus network data model and time periods." *Future Generation Computer Systems* 110 (2020): 364-371. https://doi.org/10.1016/j.future.2018.04.077

[33] Čelan, Marko, and Marjan Lep. "Bus-arrival time prediction using bus network data model and time periods." *Future Generation Computer Systems* 110 (2020): 364-371. https://doi.org/10.1016/j.future.2018.04.077

[34] Serin, Faruk, Yigit Alisan, and Metin Erturkler. "Predicting bus travel time using machine learning methods with three-layer architecture." *Measurement* 198 (2022): 111403. https://doi.org/10.1016/j.measurement.2022.111403

[35] Evangeline, A., R. Catherine Joy, and A. Albert Rajan. "Flight delay prediction using different regression algorithms in machine learning." In *2023 4th International Conference on Signal Processing and Communication (ICSPC)*, IEEE, p. 262-266. 2023. https://doi.org/10.1109/ICSPC57692.2023.10125675