JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 9 Number 1 | June 2024: 18-28
DOI: 10.15575/join.v9i1.1245

Deep Learning Based LSTM Model for Predicting the Number of Passengers for Public Transport Bus Operators

Joko Siswanto[1], Danny Manongga[2], Irwan Sembiring[3], Sutarto Wiyono[4]
[1]Politeknik Keselamatan Transportasi Jalan, Indonesia
[1,2,3,4]Faculty of Information Technology, Satya Wacana Christian University, Indonesia

Article Info

Article history:

Received December 09, 2023
Revised March 03, 2024
Accepted March 05, 2024
Available Online April 23, 2024

Keywords:

Bus operator
Deep learning
LSTM
Number of passenger

ABSTRACT

Predicting the number of passengers is the main demand for a modern public transportation system and the focus of solving problems for bus operating companies. Time-series prediction models with high accuracy, high complexity, and large historical data to produce predictions of passenger numbers, correlations, and relationships between prediction results are problems that need to be solved. The stages of proposing a prediction model consist of input(historical data and data preprocessing), process(feature selection, testing and training model, and evaluating prediction accuracy), and output (prediction model, correlation and relationship between prediction results). Historical data from Auckland Transport(AT) New Zealand metro patronage reports. Preprocessing data from the number of passengers of 4 operators(Go Bus, New Zealand Bus, Pavlovich, and Ritchies). The proposed prediction model process is LSTM based on deep learning with ajust parameters epoch 60, batch size 16, and neurons 32 based on the lowest and fastest evaluation value(00:00:37) from MSLE(0.132), MAPE(17.30), and SMAPE(17.68). The output of the prediction model makes predictions 12 months later for 4 operator passenger numbers. The prediction of a strong negative correlation is New Zealand Bus-Pavlovich, while predictions that are less closely related and dependent are New Zealand Bus against Go Bus, Pavlovich, and Ritchies. Prediction results, correlations and dependency relationships can be used by operators or stakeholders to support operational policies and strategies for public bus transportation.

Corresponding Author:

Joko Siswanto,
Politeknik Keselamatan Transportasi Jalan
Semeru Street, 3 Tegal, Indonesia. 52125
Email: siswanto@pktj.ac.id

1. INTRODUCTION

Bus public transportation is one of the most important modes of travel for society[1] which is expected to experience rapid growth in the future[2]. Bus public transportation plays a big role in the short and long distance trips that people rely on every day to go to work, continue their studies, and many other important purposes[3]. The erratic accessibility of public bus transportation has become a significant issue[4]. The development of fast and comfortable public transportation is a serious challenge[5]. More efficient and prospective management of public bus transportation operations is needed in big cities to provide reliable bus services[6]. The quality of bus public transportation has a serious impact on people's happiness, sense of belonging and brand[7]. Public views can be utilized by companies as a decision support system to improve and evaluate company services[8]. Bus public transportation services are influenced by traffic congestion, weather conditions, number of passengers, and traffic signals[9]. Passengers are one of the fundamental inputs in a rational system of recommendation and service for public bus transportation[10]. Bus transportation passenger demand plays an important role in network planning, resource allocation[6], and determinants of the development of public transportation [11]. Many governments are pushing for better bus public transportation so that it can serve more passengers[12].

Predicting the number of passengers is the main demand for a modern public transportation system which has experienced a revolution in recent years[13] and becomes the focus for solving various problems of bus operating companies[14]. Accurate predictions of the number of bus passengers can help bus operating companies make reasonable plans and encourage the development of good bus public transportation[11]. Increasing the accuracy of passenger flow predictions for public bus transportation can help managers make better decisions regarding operations[15]. Bus passenger prediction algorithms can improve the reliability and quality of services and reduce operating costs[6] by extending artificial intelligence to bus public transportation systems[16]. Prediction algorithms need to be developed to find solutions that are accurate and can be trusted by passengers[17].

Predicting the flow of public transportation bus passengers is very important for operator companies in operating buses[18] and provides a better basis for decisions[19]. A secondary decomposition integration method was proposed by combining Empirical Modal Decomposition (EMD), Sample Entropy (SE), and Kernel Extreme Learning Machine (KELM) to predict bus route passenger flows in a short time with high accuracy and good robustness[11]. Passenger flow prediction for several lines is proposed using a model based on the Point-Of-Interest (POI) data and eXtreme Gradient Boosting (XGBoost) or PFP-XPOI with higher prediction accuracy results when compared to other models[19]. The XGBoost model was proposed to predict bus public transportation passenger flows from various routes with the results of achieving advanced prediction accuracy and computational efficiency[20]. The ARIMA model compared with Bocks-Young, normalized BIC, and stationary R-square is used to predict short-term bus passenger flows with better results[21]. The Seasonal Auto-Regressive Integrated Moving Average (SARIMA) model is used to predict the frequency of passenger flows[22]. The Short Time Series Clustering-Seasonal Autoregressive Integrated Moving Average (STSC-SARIMA) model was built to predict bus public transportation passenger flows with high prediction accuracy and applicability[23]. An integrated model to accurately predict short-term passenger flows for public bus transportation using Multivariable Linear Regression (MLR), K-Nearest Neighbor (KNN), XGBoost, and Gated Recurrent Unit (GRU) with results that can enrich the short-term passenger flow prediction system and provide effective data support[24]. A prediction model combining Principal Component Analysis (PCA) and error Back Propagation (BP) neural networks to predict short-term bus passenger flows with high accuracy and prediction performance[25]. The recurrent neural network (RNN) model is used to predict bus public transportation passenger flows in the Karnataka State Road Transport Corporation Bus Rapid Transit (KSRTCBRT) with high accuracy results[18]. The Graph Convolutional Network–Recurrent Neural Network (GCN–RNN) model was proposed to predict passenger flow with the results of the model being able to predict effectively[26]. The Diffusion Convolution Recurrent Neural Network (DCRNN) architecture was adopted to predict the number of bus public transportation passengers on each line with more accurate prediction results compared to classical RNN[27]. A long short-term memory (LSTM) network model was proposed for predicting new passenger flows for public bus transportation with better accuracy results compared to other baseline methods[28]. The Long Short Term Memory (LSTM) model is used to accurately predict passenger flows for the next thirty days[29]. The BiLSTM model is used to predict bus passenger demand based on actual patronage data obtained from the smart card ticket system in Melbourne with accuracy results of more than 90% which can outperform the LSTM model[30]. The CNN-BiLSTM model is used for short-term bus passenger flow predictions with better accuracy results compared to LSTM, CNN, and CNN-LSTM[31]. Predictions using artificial intelligence have become an option to produce high accuracy, data flexibility, and fast time.

Predictions for bus public transportation that explore passengers are still few[32] with low reliability and capability[33]. Low prediction accuracy results in poor performance of the prediction model which greatly affects the performance of public transportation buses[34]. The accuracy of predicting the number of passengers using deep learning[35] directly influences decisions in the operation and management of public bus transportation[36]. Predictions of the number of bus public transportation passengers are still weak in capturing and measuring passenger distribution[37] with still low accuracy[38]. Improving the passenger number prediction model algorithm pays less attention to deep learning, data mining[39] and big data capacity[40], thus producing a prediction model with low accuracy which has an impact on the failure of implementing the solution[41]. The time-series prediction model uses deep learning with high accuracy, high complexity, and the use of large historical data to produce predictions of the number of passengers for bus public transportation operators, the correlation of prediction results, and the relationship of prediction results are problems that need to be resolved.

The reliability of the deep learning-based LSTM model by achieving maximum accuracy, handling time-series data, and handling complex prediction problems is the basis for using LSTM to propose a time-series prediction model for the number of passengers for bus public transport operators. The prediction model uses a time-series dataset of the number of passengers by public transport bus operators every day for at least 3 years. Variations in neurons, epochs, and batch sizes were carried out to find the best model optimization and were evaluated using Mean-Squared Logarithmic Error (MSLE), Mean Absolute Percentage Error (MAPE), and Symmetric Mean Absolute Percentage Error (SMAPE) based on the lowest value and fastest time. The prediction model produces 4 types of bus public transportation operator predictions in one model for the next 12 months. 4 types of predictions are needed to display developments and relationships respectively from the 1st to the 12th month. The proposed prediction model is something new that can be used as a reference for carrying out 4 types of predictions, relationships between predictions, and patterns between predictions in depth with abundant data about the number of passengers of bus transportation operators. The proposed prediction model is the first in the development and application of deep learning using LSTM to predict the number of passengers by public transport bus operators. The prediction results in the form of the number of passengers of bus public transportation operators, the correlation of prediction results, and the relationship between prediction results can be used to support stakeholders and operators in making operational strategies and policies for bus public transportation.

2. METHOD

The stages in proposing a model for predicting the number of passengers for bus public transport operators consist of input, process, and output. Input consists of historical data and preprocessing data. The process consists of feature selection, training and testing the model, and evaluating prediction accuracy. The output will produce the best prediction model with the highest accuracy values, correlation of prediction results, relationship of prediction results(Figure 1). The prediction simulation environment uses the Python programming language running on Google Colaboratory with the macOS Venture 13.5 Operating System and 8 GB RAM. The deep learning framework used is Tensor Flow.

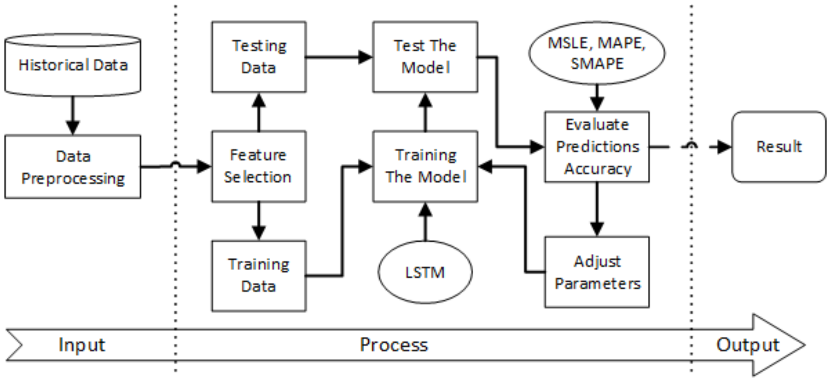

Figure 1. Bus Operator Passenger Number Prediction Architecture

Historical data collected from a CSV dataset containing Auckland Transport (AT) metro patronage reports regarding public transportation in New Zealand. The report used is a bus performance report every day starting from January 1 2019 to July 31 2023[42]. Data preprocessing from the dataset of the number of passengers using public bus transportation. There are 8 bus public transportation operators operating, but 4 bus operators were taken, namely Go Bus, New Zealand Bus, Pavlovich, and Ritchies. The 4 bus operators were chosen because they have the most complex fleet, routes and passengers. The total data is 1,673 data based on daily bus public transportation passenger report data. The data used in this dataset are the date and number of passengers on each bus public transportation operator each day.

Feature selection uses minmax feature scaling, then the dataset is divided into two segments (training and testing). The dataset is run using a deep learning-based LSTM model. Very robust dynamical systems with vanishing/exploding gradient problems can be solved with LSTM structures[43], [44]. LSTM has a distinct recurrent chain structure[45] with recurrent hidden units forming self-contained loop cycles to implicitly store information about the history of all past chain elements[43]. The three-gate structure of the interaction layer (input gate, output gate, and forget gate) and the tanh unit of the unit cell[43], [45], [46]. Forget gate and input gate to control the input and output of each cell state[47]. The sequence input data is input to the model at the input layer and then a number of LSTM layers use the output signal from the input layer to prepare the input signal for the output layer. In the output layer, the output signal of the last LSTM layer is connected through multiplication by weight and bios addition, and the output estimate is calculated[46](Figure 2). LSTM can be applied to detect various types of time series data which basically uses multilayer neural networks to learn the time series relationship between input and output parameters[47]. LSTM models have great learning capabilities for time series problems built on various structures in the network to achieve the most accurate prediction model for the turning process[46].

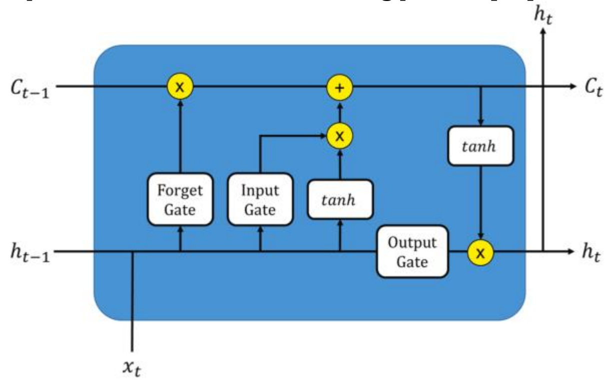

Figure 2. Schematic of An LSTM Unit[46]

Adjust LSTM parameters to find the most suitable and best model. Comparison of training datasets and prediction accuracy will be evaluated. The lowest evaluation value and the fastest time of the experiment become the most optimal model. The selection of LSTM model parameters can be seen from Mean-Squared Logarithmic Error (MSLE), Mean Absolute Percentage Error (MAPE), and Symmetric Mean Absolute Percentage Error (SMAPE). The consistency and speed of performance of the proposed model will be seen from these 3 types of evaluation matrices. MSLE calculates the average of the logarithm of the squared errors between predicted and actual values[48] (1). An MSLE value that is closer to 0 is a reflection of better model performance[49]. MAPE is a key performance indicator commonly used for prediction accuracy. MAPE divides each error based on each request[50](2). The smaller the MAPE value, the higher the prediction accuracy[51]. SMAPE calculates accuracy based on error percentage[52] that can be used to evaluate the prediction performance of time series data sets[53] (3). The lower the SMAPE value of a forecast, the higher the accuracy[29].

$$MSLE = \frac{1}{n}\sum_{i-1}^{n}(log(y_i + 1) - log(\hat{y} + 1))^2 \quad (1)$$

n is the number of observations in the dataset, y_i is the actual value at the ith observation, and ŷi is the predicted value at the i-th observation.

$$MAPE = \frac{1}{n}\sum_{i-1}^{n}\left|\frac{\hat{p}_i - p_i}{p_i}\right| \quad (2)$$

$\hat{p}_i$ represents predicted prices, p_i represents actual prices and n is the total number of observations.

$$SMAPE = \frac{1}{n}\sum_{j-1}^{n}\left|\frac{f_j - y_j}{|y_j| + |f_j|}\right| \quad (3)$$

f_j represents the predicted value, y_j represents the actual value, n is the size of the prediction horizon.

The architectural output produces a prediction model for the number of passengers for bus public transportation operators using LSTM with the best and fastest accuracy values based on the experiments carried out. The prediction model is used to predict the number of passengers for 4 bus public transportation operators in the following year every month. Correlations and relationships between predictions can be presented in the proposed prediction model.

3. RESULT AND DISCUSSION

Historical data with input format in the form of numeric results of the number of passengers of public transport bus operators every day from January 1 2019 to July 31 2023. Data preprocessing was processed and formatted in the Comma-Separated Values (CSV) file type with 1,673 results. There are 5 columns in the dataset used, namely DATE, GO_BUS, NEW_ZEALAND_BUS, PAVLOVICH, and RITCHIES. DATE contains date data, GO_BUS contains data on the number of passengers of the public transport operator Go Bus, NEW_ZEALAND_BUS contains data on the number of passengers of the public transport operator New Zealand Bus, PAVLOVICH contains data on the number of passengers of the public transport operator Pavlovich bus, and RITCHIES contains data on the number of passengers of public transport operators Ritchies bus(Figure 3).

	DATE	GO_BUS	NEW_ZEALAND_BUS	PAVLOVICH	RITCHIES
0	2019-01-01	6441	26955	2591	12789
1	2019-01-02	8860	34838	3819	16478
2	2019-01-03	13216	50557	5776	26955
3	2019-01-04	13215	53483	6074	27455
4	2019-01-05	11055	40968	4345	19519
...	...	...	...	...	...
1668	2023-07-27	25261	89432	12317	58350
1669	2023-07-28	25242	86696	11903	55506
1670	2023-07-29	13994	48314	5994	25818
1671	2023-07-30	10471	40522	4649	21014
1672	2023-07-31	24000	84652	11973	56062

Figure 3. Data Preprocessing

Feature selection with minmax feature scaling to divide training data by 80% and testing data by 20%. Training the model uses LSTM with 2 hidden layers, activation uses hyperbolic tangent (tahn), and dropout is 0.20. There are 1 LSTM models used to train the training data with the optimizer used by adam and verbos=1. Adjust parameters of the LSTM model for the experiment from the number of neurons (8, 16, and 32), epochs (20, 40, 60, 80, and 100), and batch size (4, 8, and 6). Evaluate predictions accuracy is carried out by MSLE, MAPE, and SMAPE with the lowest values and fastest times based on experiments to find the most optimal model for predicting the number of passengers for bus public transport operators. MSLE, MAPE, and SMAPE evaluation values are observed using different layers, and different units in 2 hidden layers and dense layers for prediction output. The fastest time occurred with a variation of 8 neurons, epoch 20, and batch size 16, but the MSLE, MAPE, and SMAPE values were still quite large compared to the others. The most optimal deep learning-based LSTM model with 32 neurons, epoch 60, and batch size 16, because the time is relatively fast and the MSLE, MAPE, and SMAPE values are lowest compared to the others(Table 1).

Table 1. LSTM Model Experiment Results

E	B	8 Neuron				16 Neuron				32 Neuron			
		E1	E2	E3	Time	E1	E2	E3	Time	E1	E2	E3	Time
20	4	0.143	21.11	19.22	00:01:27	0.142	19.97	19.62	00:00:52	0.138	19.06	18.38	00:00:52
	8	0.147	20.37	20.11	00:00:48	0.146	20.19	18.93	00:00:30	0.147	20.18	19.77	00:00:38
	16	0.145	21.69	18.67	00:00:25	0.148	20.46	19.10	00:00:16	0.147	21.52	18.78	00:00:22
40	4	0.137	20.98	18.72	00:02:47	0.137	17.35	18.54	00:02:36	0.140	17.50	18.54	00:01:39
	8	0.139	19.53	18.92	00:01:28	0.137	19.89	18.69	00:01:27	0.145	18.81	20.24	00:01:27
	16	0.143	20.79	18.51	00:00:32	0.140	19.61	18.60	00:00:47	0.139	19.51	18.31	00:00:33
60	4	0.136	19.32	18.46	00:02:37	0.162	18.94	22.36	00:02:27	0.148	18.09	21.07	00:03:27
	8	0.140	20.69	18.69	00:01:32	0.150	18.08	20.45	00:01:17	0.196	20.15	25.23	00:01:28
	16	0.139	20.74	18.61	00:00:46	0.143	18.64	19.35	00:00:46	0.132	17.30	17.68	00:00:37
80	4	0.143	19.88	20.40	00:03:29	0.153	17.43	19.06	00:03:29	0.153	17.93	20.85	00:03:28
	8	0.140	19.55	18.47	00:01:20	0.145	18.01	20.49	00:01:46	0.139	17.63	19.64	00:01:38
	16	0.138	20.25	18.93	00:00:48	0.150	18.08	20.41	00:00:54	0.165	18.27	21.56	00:00:57
100	4	0.250	23.60	31.16	00:04:27	0.141	17.86	19.24	00:03:07	0.233	20.04	24.92	00:03:57
	8	0.186	19.69	24.33	00:02:28	0.175	19.76	24.26	00:02:03	0.202	20.36	25.33	00:02:12
	16	0.138	19.35	20.11	00:01:07	0.174	19.19	23.18	00:01:31	0.155	18.01	20.97	00:01:10

Note: E=Epoch, B=Batch Size, E1=MSLE, E2=MAPE, E3=SMAPE.

The best LSTM model with an evaluation value from MSLE of 0.132, MAPE of 17.30, and SMAPE of 17.68. The three evaluation models used have a value of less than 20, which means the LSTM model has good performance quality and is acceptable. This is because if the value is more than 20, then the proposed model needs improvement and even a value of more than 50 is unacceptable. Comparison of loss and validation loss with neurons 8, 16, and 32 for epoch 60 and batch size 16. Comparison of graphs with variations in neurons. In general, the training and validation lines are almost the same, so that neuron 32 with the lowest evaluation results is the most optimal model(Figure 4).

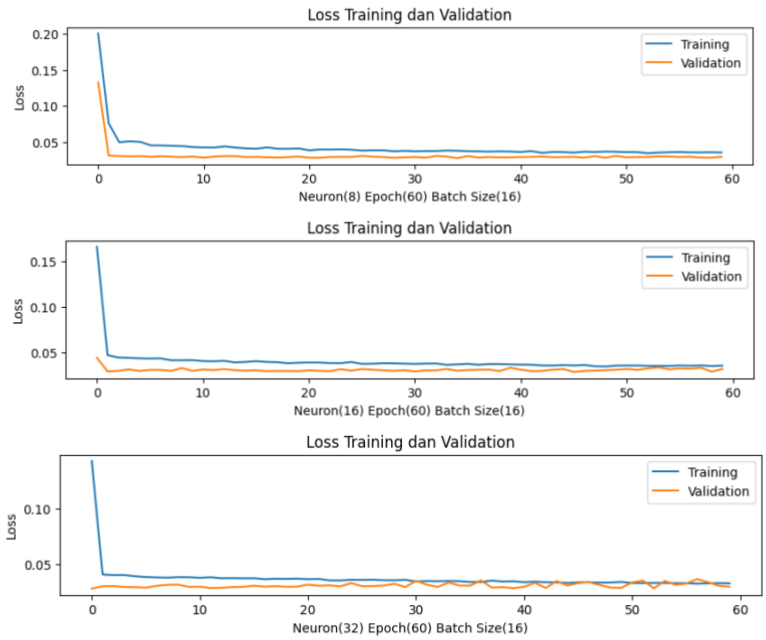

Figure 4. Output Epoch by LSTM Model

The prediction results using the proposed model are in accordance with the movement of testing and training data. Prediction results on training and testing data improve with more training carried out. Deep learning carries out deeper learning based on long and short term time by looking at the movement of the number of passengers of 4 bus public transportation operators which is getting better. The four movements fluctuate with the same rhythm, although there are several times there are allusions between Ritchies operator passengers and Go Bus and New Zealand Bus(Figure 5).

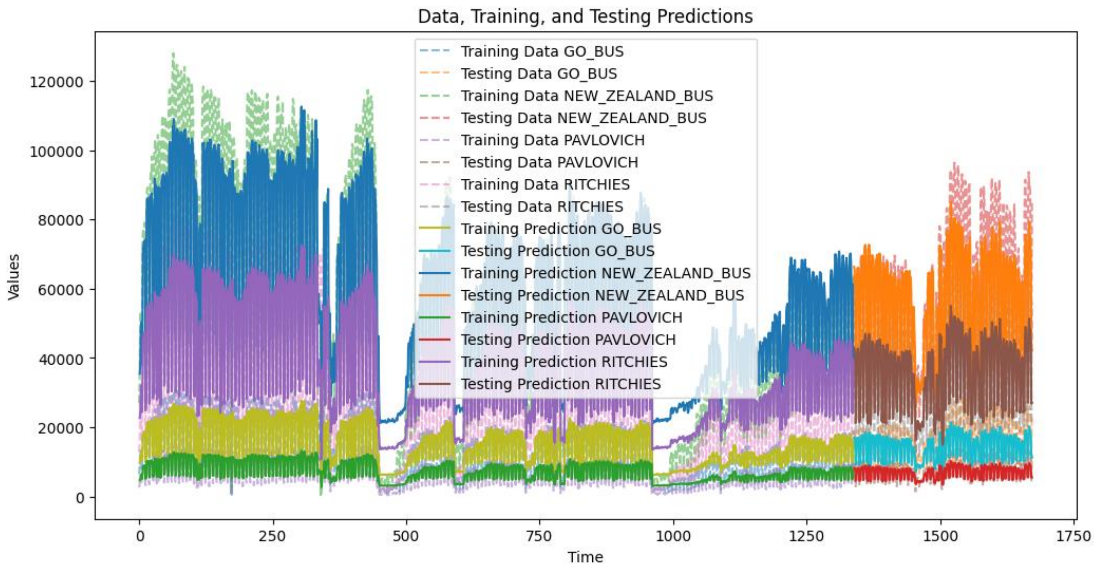

Figure 5. Data, Training, and Testing Prediction

The number of passengers for bus public transportation operators consisting of Go Bus, New Zealand Bus, Pavlovich, and Ritchies is predicted at the same time using a deep learning-based LSTM model. Prediction output results are carried out for the following 12 months for the 4 types of predictions produced. The prediction results show that there is a gentle decline starting from the 2nd month (October), but starting from the 5th month (February) the movement tends to stabilize until the 12th month (August). The predictions for the number of passengers do not overlap because all bus operators are operating normally(Figure 6).

p-ISSN: 2528-1682
e-ISSN: 2527-9165

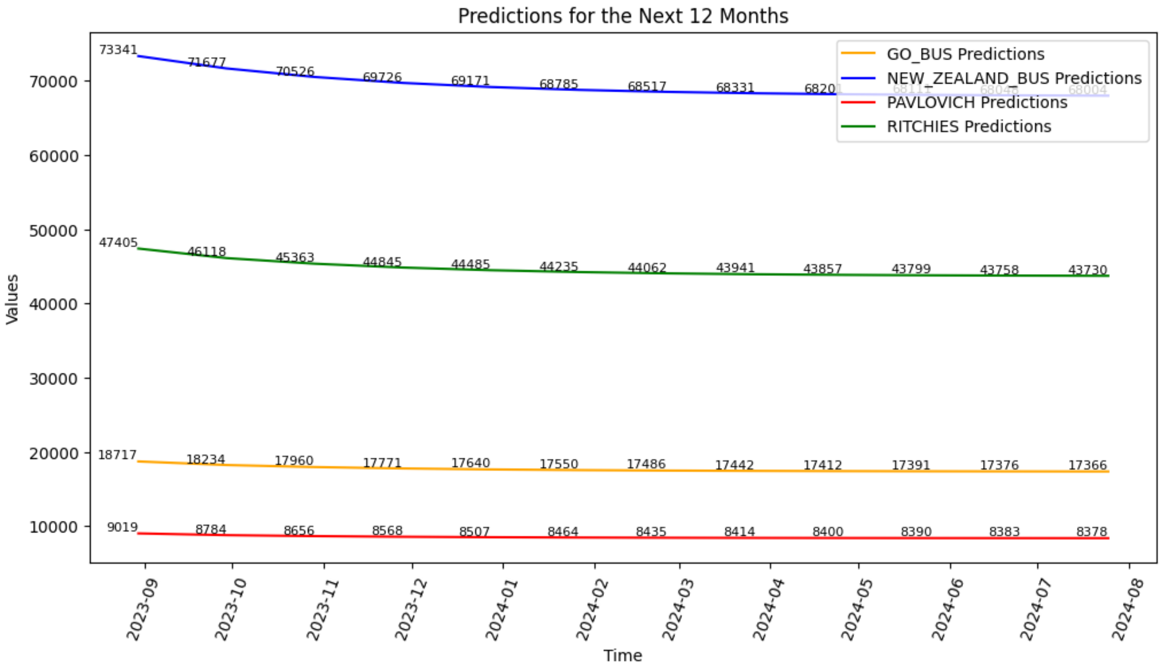

Figure 6. Predictions For The Next 12 Months

The 4 predictions for the number of passengers by public transport bus operators have a correlation with the positive (red) or negative (blue) characteristics presented in the correlation heatmap. The predicted number of bus operators with a strong negative correlation is 1. The predicted number of passengers with a weak negative correlation is 1. The predicted number of passengers with a weak positive correlation is 2. The predicted number of passengers with a strong positive correlation is 2(Figure 7).

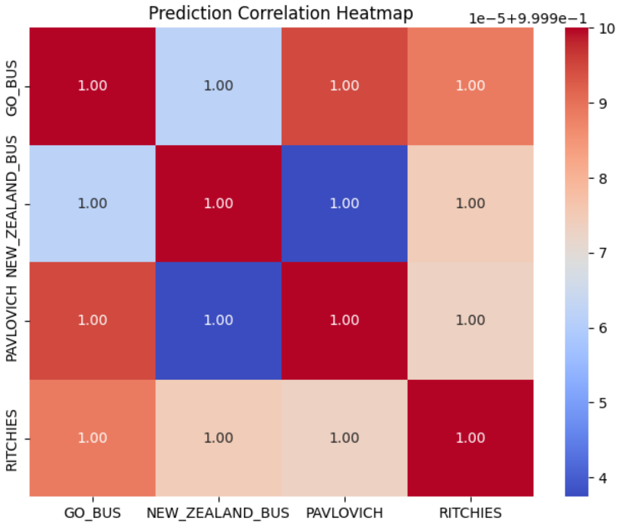

Figure 7. Prediction Correlation Heatmap

Prediction dependencies are visualized with a prediction scatter matrix. The points on a scatter plot that move together or form a line become a consistent positive or negative pattern in the prediction model. The four predictions have very strong dependencies as indicated by the diagonal scatter plot. There are 3 predictions that are closely interconnected and dependent. There are 3 predictions that are less closely interconnected and dependent (Figure 8).

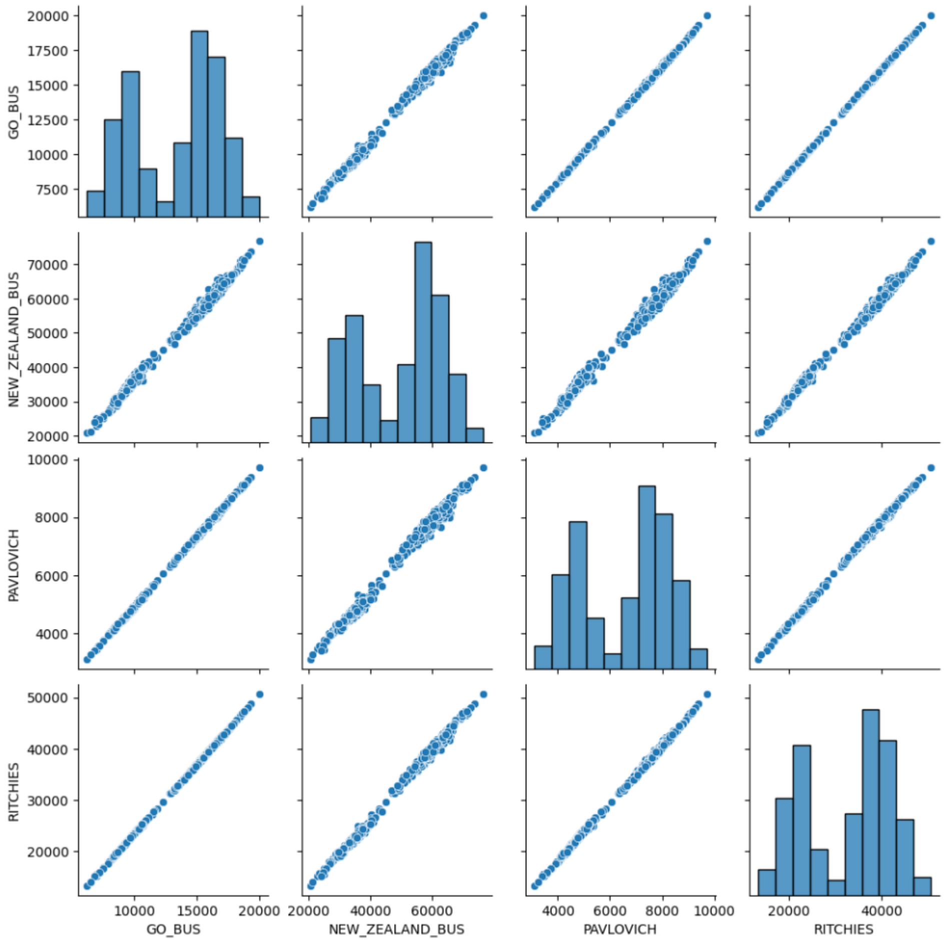

Figure 8. Prediction Scatter Matrix

Internet connection stability, appropriate data format, and adequate computer performance are the keys to conducting experiments on the proposed model. In-depth investigations are carried out on large historical data to achieve the best accuracy and suitability of the desired predictions. Empty data and data discrepancies are problems that must be resolved. LSTM can process time-series data, so the suitability of time-series data is key in the preprocessing stage. Dataset normalization is the first task that must be considered in creating a time-series prediction model using LSTM in the preprocessing process.

Adjusting the parameters of the deep learning-based LSTM model appropriately is an important factor for training the model and designing the proposed framework to carry out 4 passenger predictions for bus public transportation operators. The experimental results produce a smaller number of epochs, neurons, batch size, which requires faster time. Increasing or decreasing the number of neurons, epochs, and batch size does not always make the model better, the suitability of the data and the prediction goal is accuracy in selecting the appropriate model in a relatively fast time. Combinations and variations greatly influence the time needed to carry out processing which will be measured by the evaluation value. Evaluate prediction accuracy is carried out by MSLE, MAPE, SMAPE with the lowest value as proof of the stability of optimal prediction model performance and high accuracy.

Training epochs are selected in the best way to train the model according to the analysis of different epochs for LSTM models. Loss and validation loss graphs are used to see the performance of the learning model during training and evaluation of validation data. The deviation between the Loss and Validation Loss graphs can provide important insight into the quality of the model's generalization to new data. Validation Loss begins to increase to optimize the LSTM model to improve prediction performance.

The highest passenger movement is owned by the bus operator New Zealand Bus, while the lowest passenger movement is owned by the Pavlovich bus operator. The use of the LSTM model provides more reliable support in long-term and short-term prediction modeling. The predictions made are greatly influenced by the data obtained as the main material for making predictions. A very significant decline occurred in 2 periods in 2020 and 2021, the number of passengers for 4 bus public transportation operators began to increase starting in 2022, although several times there were insignificant declines.

The output results of the prediction model are predictions of the number of passengers of public transport bus operators, correlation of prediction results, and dependency of prediction results. The prediction results show that passengers between bus operators are very far apart, passengers from Go Bus and Pavlovich operators are the closest. The distances between predictions do not intersect, which means that fluctuations in the number of passengers occur simultaneously. The prediction for the number of passengers by bus operators that has a strong negative correlation is New Zealand Bus-Pavlovich. The predicted number of passengers with a weak negative correlation is New Zealand Bus-Go Bus. Predictions on the number of passengers that have a weak positive correlation are Ritchies against New Zealand Bus and Pavlovich. The prediction of the number of passengers that has a strong positive correlation is Go Bus with Pavlovich and Ritchies. Predictions that are closely related and dependent are Go Bus-Ritchies, Go Bus-Pavlovich, and Pavlovich-Ritchies. Predictions that are less closely related and dependent are New Zealand Bus against Go Bus, Pavlovich, and Ritchies. Updating the dataset can be a solution to see predictions that are more appropriate and accurate with developments that are occurring.

Deep learning using the LSTM model with high accuracy is proposed as a prediction model for the number of passengers for bus public transportation operators. High complexity, large capacity historical data, and time-series types of predictions can be carried out by the proposed prediction model. Modeling can be used as an illustration of fluctuations in the number of bus operator passengers in the following year every month. The correlation and dependency relationships between the prediction results presented can be used to support operators or stakeholders in making policies and operational strategies for public bus transportation. Policies and strategies can be adjusted based on developmental flexibility and predicted results with high accuracy.

4. CONCLUSION

The prediction model uses deep learning-based LSTM to predict the number of passengers of bus public transportation operators to produce predictions of the number of bus public transportation operators, correlation of prediction results, and correlation of prediction results. Predicting the number of passengers for 4 bus public transportation operators (Go Bus, New Zealand Bus, Pavlovich, and Ritchies) using time-series data can be done using a deep learning-based LSTM model. The prediction model with the best and fastest accuracy (00:00:37) uses adjust parameters LSTM epoch 60, batch size 16, and neurons 32 based on the lowest MSLE (0.132), MAPE (17.30), and SMAPE (17.68) values. The proposed prediction model predicts 12 months later for 4 bus operator passenger number predictions simultaneously with the resulting distance between predictions of fluctuations occurring simultaneously. Strong negative correlation in New Zealand Bus-Pavlovich, strong positive correlation in Go Bus-Ritchies and Go Bus-Pavlovich, apart from that there is a weak negative and positive correlation. Predictions that are less closely related and dependent are New Zealand Bus against Go Bus, Pavlovich, and Ritchies. Modeling can be used as an illustration of fluctuations in the number of bus operator passengers in the following year every month. The correlation and dependency relationships between the prediction results presented can be used to support operators or stakeholders in making policies and operational strategies for public bus transportation. Policies and strategies can be adjusted based on developmental flexibility and predicted results with high accuracy. Comparison of drop outs and training-testing percentages of LSTM models for predicting time series data is a step to optimize the model in future work.

ACKNOWLEDGEMENTS

This work was supported by Politeknik Keselamatan Transportasi Jalan and Satya Wacana Christian University.

REFERENCES

[1] R. Yao and X. Du, "Modelling lane changing behaviors for bus exiting at bus bay stops considering driving styles: A game theoretical approach," *Travel Behav Soc*, vol. 29, pp. 319–329, Oct. 2022, doi: 10.1016/j.tbs.2022.07.008.

[2] Z. Zhang, H. He, J. Guo, and R. Han, "Velocity prediction and profile optimization based real-time energy management strategy for Plug-in hybrid electric buses," *Appl Energy*, vol. 280, 2020, doi: 10.1016/j.apenergy.2020.116001.

[3] S. M, G. GS, J. M, A. Suresh, and V. A, "Commutator Bus Transit Recommendation by Predicting Dynamic Bus Flow Occupancy Using Machine Learning," in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, IEEE, Nov. 2021, pp. 538–541. doi: 10.1109/I-SMAC52330.2021.9640869.
[4] W. Yu, H. Sun, J. Wu, Y. Lv, X. Shang, and X. Wang, "Mapping multimodal random accessibility using smart card data: a case study of bus and subway stations in Beijing," *Transportation Planning and Technology*, vol. 45, no. 1, pp. 76–99, Jan. 2022, doi: 10.1080/03081060.2021.2017206.
[5] 刘铭哲, "Design Method of Express Lines at Bus Stations—Case Study in Guilin," *Open Journal of Transportation Technologies*, vol. 09, no. 03, pp. 197–206, 2020, doi: 10.12677/OJTT.2020.93024.
[6] A. M. Nayak, A. Ladha, and N. K. Chaubey, "A Comprehensive Comparative Analysis of Passenger Demand Prediction for Improving the Urban Bus Transportation System (UBTS)," *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 269–279, Oct. 2022, doi: 10.14445/22315381/IJETT-V70I9P227.
[7] G. Duan, X. Ma, J. Wang, Z. Wang, and Y. Wang, "Optimization of Urban Bus Stops Setting Based on Data Mining," *Intern J Pattern Recognit Artif Intell*, vol. 35, no. 08, p. 2159028, Jun. 2021, doi: 10.1142/S021800142159028X.
[8] Z. Nurthohari, D. I. Sensuse, and S. Lusa, "Sentiment Analysis of Jakarta Bus Rapid Transportation Services using Support Vector Machine," in *2022 International Conference on Data Science and Its Applications (ICoDSA)*, IEEE, Jul. 2022, pp. 171–176. doi: 10.1109/ICoDSA55874.2022.9862903.
[9] K. Hiroi, K. Hananouchi, H. Imai, and N. Kawaguchi, "An Analysis of Bus Traffic Delay with Geographical Characteristics and an Application to Arrival Time Prediction in Snowfall Situation," *IEEJ Transactions on Electronics, Information and Systems*, vol. 140, no. 2, pp. 257–266, Feb. 2020, doi: 10.1541/ieejeiss.140.257.
[10] S. Chen, X. Liu, C. Lyu, L. Vlacic, T. Tang, and Z. Liu, "A holistic data-driven framework for developing a complete profile of bus passengers," *Transp Res Part A Policy Pract*, vol. 173, p. 103692, Jul. 2023, doi: 10.1016/j.tra.2023.103692.
[11] Y. Li and C. Ma, "Short-Time Bus Route Passenger Flow Prediction Based on a Secondary Decomposition Integration Method," *J Transp Eng A Syst*, vol. 149, no. 2, Feb. 2023, doi: 10.1061/JTEPBS.TEENG-7496.
[12] J.-M. Clairand, M. González-Rodríguez, I. Cedeño, and G. Escrivá-Escrivá, "A charging station planning model considering electric bus aggregators," *Sustainable Energy, Grids and Networks*, vol. 30, p. 100638, Jun. 2022, doi: 10.1016/j.segan.2022.100638.
[13] A. M. Nayak and N. K. Chaubey, "Intelligent passenger demand prediction-based rerouting for comfort perception in public bus transport systems," *International Journal of Communication Systems*, vol. 35, no. 18, Dec. 2022, doi: 10.1002/dac.5351.
[14] T. Alexandre, F. Bernardini, J. Viterbo, and C. E. Pantoja, "Machine Learning Applied to Public Transportation by Bus: A Systematic Literature Review," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2677, no. 7, pp. 639–660, Jul. 2023, doi: 10.1177/03611981231155189.
[15] F. Jiao, L. Huang, R. Song, and H. Huang, "An Improved STL-LSTM Model for Daily Bus Passenger Flow Prediction during the COVID-19 Pandemic," *Sensors*, vol. 21, no. 17, p. 5950, Sep. 2021, doi: 10.3390/s21175950.
[16] T. A. Kumar, R. Rajmohan, M. Pavithra, S. A. Ajagbe, R. Hodhod, and T. Gaber, "Automatic Face Mask Detection System in Public Transportation in Smart Cities Using IoT and Deep Learning," *Electronics (Basel)*, vol. 11, no. 6, p. 904, Mar. 2022, doi: 10.3390/electronics11060904.
[17] R. Gracious, B. A. Kumar, and L. Vanajakshi, "Performance Evaluation of Passenger Information Systems," *Transportation in Developing Economies*, vol. 8, no. 1, p. 4, Apr. 2022, doi: 10.1007/s40890-021-00140-5.
[18] N. Nagaraj, H. L. Gururaj, B. H. Swathi, and Y.-C. Hu, "Passenger flow prediction in bus transportation system using deep learning," *Multimed Tools Appl*, vol. 81, no. 9, pp. 12519–12542, Apr. 2022, doi: 10.1007/s11042-022-12306-3.
[19] W. Lv, Y. Lv, Q. Ouyang, and Y. Ren, "A Bus Passenger Flow Prediction Model Fused with Point-of-Interest Data Based on Extreme Gradient Boosting," *Applied Sciences*, vol. 12, no. 3, p. 940, Jan. 2022, doi: 10.3390/app12030940.
[20] L. Zou, S. Shu, X. Lin, K. Lin, J. Zhu, and L. Li, "Passenger Flow Prediction Using Smart Card Data from Connected Bus System Based on Interpretable XGBoost," *Wirel Commun Mob Comput*, vol. 2022, pp. 1–13, Jan. 2022, doi: 10.1155/2022/5872225.
[21] T. Wu, H. Xu, and J. Lu, "Short-term Prediction of Up and Down Passenger Flow at Bus Station Based on Arrival Spacing," *Wuhan Ligong Daxue Xuebao (Jiaotong Kexue Yu Gongcheng Ban)/Journal of Wuhan University of Technology (Transportation Science and Engineering)*, vol. 44, no. 6, 2020, doi: 10.3963/j.issn.2095-3844.2020.06.030.
[22] X. Fang, M. Lin, W. Chen, and X. Pan, "Short-Time Bus Passenger Flow Prediction by Identifying Features of Incomplete Data," *Huanan Ligong Daxue Xuebao/Journal of South China University of Technology (Natural Science)*, vol. 48, no. 4, 2020, doi: 10.12141/j.issn.1000-565X.190465.
[23] G. Cheng and C. He, "Analysis of bus travel characteristics and predictions of elderly passenger flow based on smart card data," *Electronic Research Archive*, vol. 30, no. 12, pp. 4256–4276, 2022, doi: 10.3934/era.2022217.
[24] X. Wang, L. Huang, H. Huang, B. Li, Z. Xia, and J. Li, "An Ensemble Learning Model for Short-Term Passenger Flow Prediction," *Complexity*, vol. 2020, pp. 1–13, Dec. 2020, doi: 10.1155/2020/6694186.
[25] S. Zhang, Z. Liu, F. Shen, S. Wang, and X. Yang, "A prediction model of buses passenger flow based on neural networks," *J Phys Conf Ser*, vol. 1656, no. 1, p. 012002, Sep. 2020, doi: 10.1088/1742-6596/1656/1/012002.
[26] Z. Xu, L. Hou, Y. Zhang, and J. Zhang, "Passenger Flow Prediction of Scenic Spot Using a GCN–RNN Model," *Sustainability*, vol. 14, no. 6, p. 3295, Mar. 2022, doi: 10.3390/su14063295.
[27] X. Zhai and Y. Shen, "Short-Term Bus Passenger Flow Prediction Based on Graph Diffusion Convolutional Recurrent Neural Network," *Applied Sciences*, vol. 13, no. 8, p. 4910, Apr. 2023, doi: 10.3390/app13084910.
[28] Q. Ouyang, Y. Lv, J. Ma, and J. Li, "An LSTM-Based Method Considering History and Real-Time Data for Passenger Flow Prediction," *Applied Sciences*, vol. 10, no. 11, p. 3788, May 2020, doi: 10.3390/app10113788.
[29] G. S. Vidya and V. S. Hari, "LSTM Network Integrated with Particle Filter for Predicting the Bus Passenger Traffic," *J Signal Process Syst*, vol. 95, no. 2–3, pp. 161–176, Mar. 2023, doi: 10.1007/s11265-022-01831-x.
[30] S. Liyanage, R. Abduljabbar, H. Dia, and P.-W. Tsai, "AI-based neural network models for bus passenger demand forecasting using smart card data," *Journal of Urban Management*, vol. 11, no. 3, pp. 365–380, Sep. 2022, doi: 10.1016/j.jum.2022.05.002.
[31] C. Wu and X. Qi, "Short-term Bus Passenger Flow Forecast Based on CNN-BiLSTM," *Advances in Engineering Technology Research*, vol. 5, no. 1, p. 448, May 2023, doi: 10.56028/aetr.5.1.448.2023.
[32] T. Alexandre, F. Bernardini, J. Viterbo, and C. E. Pantoja, "Machine Learning Applied to Public Transportation by Bus: A Systematic Literature Review," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2677, no. 7, pp. 639–660, Jul. 2023, doi: 10.1177/03611981231155189.

[33] A. M. Nayak, A. Ladha, and N. K. Chaubey, "A Comprehensive Comparative Analysis of Passenger Demand Prediction for Improving the Urban Bus Transportation System (UBTS)," *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 269–279, Oct. 2022, doi: 10.14445/22315381/IJETT-V70I9P227.

[34] L. Zeng *et al.*, "An LSTM-based driving operation suggestion method for riding comfort-oriented critical zone," *J Ambient Intell Humaniz Comput*, vol. 14, no. 2, pp. 755–771, Feb. 2023, doi: 10.1007/s12652-021-03327-1.

[35] G. S. Vidya and V. S. Hari, "LSTM Network Integrated with Particle Filter for Predicting the Bus Passenger Traffic," *J Signal Process Syst*, vol. 95, no. 2–3, pp. 161–176, Mar. 2023, doi: 10.1007/s11265-022-01831-x.

[36] X. Wang, L. Huang, H. Huang, B. Li, Z. Xia, and J. Li, "An Ensemble Learning Model for Short-Term Passenger Flow Prediction," *Complexity*, vol. 2020, pp. 1–13, Dec. 2020, doi: 10.1155/2020/6694186.

[37] T. Tang, A. Fonzone, R. Liu, and C. Choudhury, "Multi-stage deep learning approaches to predict boarding behaviour of bus passengers," *Sustain Cities Soc*, vol. 73, p. 103111, Oct. 2021, doi: 10.1016/j.scs.2021.103111.

[38] J.-L. Wu, M. Lu, and C.-Y. Wang, "Forecasting metro rail transit passenger flow with multiple-attention deep neural networks and surrounding vehicle detection devices," *Applied Intelligence*, vol. 53, no. 15, pp. 18531–18546, Aug. 2023, doi: 10.1007/s10489-023-04483-x.

[39] P. Ye and Y. Ma, "Clustering-Based Travel Pattern for Individual Travel Prediction of Frequent Passengers by Using Transit Smart Card," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2677, no. 2, pp. 1278–1287, Feb. 2023, doi: 10.1177/03611981221111355.

[40] S. Zhang, Z. Liu, F. Shen, S. Wang, and X. Yang, "A prediction model of buses passenger flow based on neural networks," *J Phys Conf Ser*, vol. 1656, no. 1, p. 012002, Sep. 2020, doi: 10.1088/1742-6596/1656/1/012002.

[41] R. Gracious, B. A. Kumar, and L. Vanajakshi, "Performance Evaluation of Passenger Information Systems," *Transportation in Developing Economies*, vol. 8, no. 1, p. 4, Apr. 2022, doi: 10.1007/s40890-021-00140-5.

[42] Auckland Transport, "AT Metro patronage report," Wellington, 2023. [Online]. Available: https://at.govt.nz/about-us/reports-publications/at-metro-patronage-report

[43] L. Ni *et al.*, "Streamflow and rainfall forecasting by two long short-term memory-based models," *J Hydrol (Amst)*, vol. 583, p. 124296, Apr. 2020, doi: 10.1016/j.jhydrol.2019.124296.

[44] H. Liu and Z. Long, "An improved deep learning model for predicting stock market price time series," *Digit Signal Process*, vol. 102, p. 102741, Jul. 2020, doi: 10.1016/j.dsp.2020.102741.

[45] W. Chen, L. Yang, B. Zha, M. Zhang, and Y. Chen, "Deep learning reservoir porosity prediction based on multilayer long short-term memory network," *GEOPHYSICS*, vol. 85, no. 4, pp. WA213–WA225, Jul. 2020, doi: 10.1190/geo2019-0261.1.

[46] M. Marani, M. Zeinali, V. Songmene, and C. K. Mechefske, "Tool wear prediction in high-speed turning of a steel alloy using long short-term memory modelling," *Measurement*, vol. 177, p. 109329, Jun. 2021, doi: 10.1016/j.measurement.2021.109329.

[47] M. Dehghan Manshadi, M. Ghassemi, S. M. Mousavi, A. H. Mosavi, and L. Kovacs, "Predicting the Parameters of Vortex Bladeless Wind Turbine Using Deep Learning Method of Long Short-Term Memory," *Energies (Basel)*, vol. 14, no. 16, p. 4867, Aug. 2021, doi: 10.3390/en14164867.

[48] A. Poulose and D. S. Han, "UWB indoor localization using deep learning LSTM networks," *Applied Sciences (Switzerland)*, vol. 10, no. 18, 2020, doi: 10.3390/APP10186290.

[49] B. Sadhukhan, S. Chakraborty, and S. Mukherjee, "Predicting the magnitude of an impending earthquake using deep learning techniques," *Earth Sci Inform*, vol. 16, no. 1, pp. 803–823, Mar. 2023, doi: 10.1007/s12145-022-00916-2.

[50] H. Zheng, F. Lin, X. Feng, and Y. Chen, "A Hybrid Deep Learning Model With Attention-Based Conv-LSTM Networks for Short-Term Traffic Flow Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 6910–6920, Nov. 2021, doi: 10.1109/TITS.2020.2997352.

[51] L. Liu *et al.*, "Transient rotor angle stability prediction method based on SVM and LSTM network," *Dianli Zidonghua Shebei/Electric Power Automation Equipment*, vol. 40, no. 2, 2020, doi: 10.16081/j.epae.202001009.

[52] Y. Liu, C. Lyu, X. Liu, and Z. Liu, "Automatic Feature Engineering for Bus Passenger Flow Prediction Based on Modular Convolutional Neural Network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2349–2358, Apr. 2021, doi: 10.1109/TITS.2020.3004254.

[53] Z. Wang, X. Su, and Z. Ding, "Long-Term Traffic Prediction Based on LSTM Encoder-Decoder Architecture," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6561–6571, Oct. 2021, doi: 10.1109/TITS.2020.2995546.