LONTAR KOMPUTER VOL. 15, NO. 3 DECEMBER 2024 **p-ISSN 2088-1541**
DOI : 10.24843/LKJITI.2024.v15.i03.p03 **e-ISSN 2541-5832**
Accredited Sinta 2 by RISTEKDIKTI Decree No. 158/E/KPT/2021

Predicting the Number of Passengers in Public Transportation Areas Using the Deep Learning Model LSTM

Joko Siswanto[ab1], Sri Yulianto Joko Prasetyo[b2], Sutarto Wijono[b3], Evi Maria[b4], Untung Rahardja[c5]

[a]Road Transportation System Engineering, Politeknik Keselamatan Transportasi Jalan
Semeru Street, 3 Tegal, Indonesia
[1]siswanto@pktj.ac.id (Corresponding author)

[b]Faculty of Information Technology, Satya Wacana Christian University
Diponegoro Street, 52-60 Salatiga, Indonesia
[2]sri.yulianto@uksw.edu
[3]sutarto.wijono@uksw.edu
[4]evi.maria@uksw.edu

[c]Faculty of Science and Technology, University of Raharja
Jenderal Sudirman Street, 40 Tangerang, Indonesia
[5]untung@raharja.info

Abstract

Accurate predictions of the number of public transport passengers on buses in each region are crucial for operations. They are required by the planning and management authority for bus public transport. A deep learning-based LSTM prediction model is proposed to predict the number of passengers in 4 bus public transportation areas (central, north, south, and west), evaluated by MSLE, MAPE, and SMAPE with dropout, neuron, and train-test variations. The CSV dataset obtained from Auckland Transport(AT) New Zealand metro patronage report on bus performance(1/01/2019-31/07/2023) is used for evaluation. The best prediction model was obtained from the lowest evaluation value and relatively fast time with a dropout of 0.2, 32 neurons, and train-test 80-20. The prediction model on training and testing data improves with the suitability of tuning for four predictions for the next 12 months with mutual fluctuations. The strong negative correlation is central-south, while the strong positive correlation is north-west. Predictions are less closely interconnected and dependent, namely central-south. With its potential to significantly impact policy-making, this prediction model can increase public transport mobility in each region, leading to a more efficient and accessible public transport system and ultimately enhancing the public's daily lives. This research has practical implications for public transport authorities, as it can guide them in making informed decisions about service planning and resource allocation.

Keywords: *Prediction, LSTM, Deep Learning, Bus Transportation Area, Passenger*

1. Introduction

Transport planners and public transport agencies agree to achieve sustainable urban public transport development. Practical government policies significantly influence bus use and improve the quality of public transportation services[1]. Service efficiency and waiting time can increase the number of users of public transportation modes. Public transportation authorities will gain benefits and ease of activity[2]. Public transportation can provide a much larger transport capacity[3]. The public transport authorities of most countries are responsible for providing local and regional public transport. The reduction of private vehicle trips and disadvantaged groups can be achieved through public transportation[4]. Many cities have integrated a combination of bus and metro public transportation[5]. Bus public transport has been proven to be a relatively safe mode of transport in densely populated urban areas that plays a vital role in ensuring affordable and adequate mobility for most of the population[6]. Bus public transportation is a common means of transportation and is often subsidized in many countries and cities[7]. Public bus transportation passengers depend on transportation conditions in an area[8].

Operations and passenger satisfaction are enhanced by the potential development of an urban bus public transportation system[9]. Reliable passenger bus services in large cities need to be provided with efficient and prospective bus operational management[10]. The important role of public bus transportation is to guarantee the movement of passengers within the city or between cities. Time and space are variables that vary in the number of public transport passengers[11]. The decline in bus public transportation passengers occurred in several areas, causing inconsistencies in time, routes, and loads, which created problems in fleet capacity planning[12]. The issues of urban public bus transportation passengers in many developing countries include poor infrastructure, chaotic driving, and inadequate information systems[13]. Most bus public transportation passengers complain of dissatisfaction with fleet loading capacity services in each area, which results in disrupted trips[14]. Bus public transportation passengers dominate customer satisfaction[1] and an important operational aspect[15]. Bus public transportation passengers will choose according to the public transportation network that suits the destination area[5]. The number of public transportation passengers on buses in each area with a dense population cannot be ignored[6].

Prediction of the number of bus passengers needs to improve the accuracy of deep learning to optimize the subdivision matrix structure of hourly passenger flows[16]. The multitask deep learning-service level passenger flow prediction (MDL-SPFP) model is used to predict the number of passengers at the lane level with an increase in accuracy of 22.39%[17]. Deep Neural Network (DNN) is used to predict large-scale bus passenger flows throughout the city of Nanjing, China, with extraordinary performance results[18]. Bus line passenger flow prediction uses the point of interest data-extreme gradient boosting (PFP-XPOI) model with higher accuracy compared to others[19]. The Gaussian Process Regression model is used to predict the number of bus public transportation passengers with accuracy that can outperform the Student-t process model and Kernel Ridge Regression (KRR) process[20]. The Seasonal Autoregressive Integrated Moving Average (STSC-SARIMA) model, which groups time series, is used to predict passenger flows with high accuracy and predictive applicability[21]. Effectiveness of suitable predictions in scenic spots using a passenger flow prediction model with graph convolutional network–recurrent neural network (GCN–RNN)[22]. The bus passenger flow prediction model based on the combination of convolutional neural network (CNN) and gated recurrent unit (GRU) used has good prediction performance[23]. The model based on the Attention mechanism (TFMA) is used to predict bus passenger flows in real-time with very high accuracy (90%), which can outperform multiple linear regression, GRU, and LightGBM[24]. Prediction performance is 15% better and more accurate using the ISTL-LSTM method compared to single and hybrid models used to predict daily bus passenger flows in Beijing[25]. Prediction of bus public transport passenger demand using the BiLSTM model has an accuracy of more than 90%, which can outperform other models[9]. The LSTM network is better by achieving the coefficient of determination (R2) to predict sequential bus passenger time series[26]. The long short-term memory (LSTM) deep learning approach is used to predict the flow of public transport bus passengers for each region in the Karnataka State Road Transport Corporation (KSRTC) with better accuracy than recurrent neural network (RNN) and greedy layer-wise algorithm[27].

Estimating the number of passengers in each area by increasing the number of passengers is carried out by planning comfortable bus public transport transportation[12]. Accurate prediction of the number of public transport passengers for each region is very important for public transport operation, which will improve the quality and reliability of services and attract more public transport passengers[9]. An important aspect of the bus public transportation system is passenger prediction, an important operational problem faced in transportation planning[15]. Predicting the number of public transportation passengers on the bus is needed to reduce most of the time and effort and make travel comfortable[14]. Network planning and allocating public transport resources for buses are important in predicting the short-term number of urban public bus transport passengers[10]. The planning and management authority for public bus transportation requires estimates of the number of passengers for each region [5]. The proposed prediction

p-ISSN 2088-1541
e-ISSN 2541-5832

model differs from those carried out in that it considers the time required, the use of 3 model evaluation matrices, and the suitability of LSTM model tuning.

Deep learning-based LSTM is proposed to predict the time-series number of passengers in bus public transportation areas for the next 12 months. The prediction model that has been carried out focuses on accuracy without presenting the suitability of tuning, only produces 1 type of prediction, and does not cover relationships and patterns between predictions in depth. LSTM was chosen as the prediction model due to the model's suitability for prediction needs compared to other LSTM variants. Drop-out variations, neurons, and testing-training data division were carried out to find the best model optimization and evaluated using Mean-Squared Logarithmic Error (MSLE), Mean Absolute Percentage Error (MAPE), and Symmetric Mean Absolute Percentage Error (SMAPE) based on the values lowest and relatively fast time. The prediction model produces four types of bus public transportation area predictions in one model for the next 12 months. The use of 4 types of prediction of the number of passengers in bus public transportation areas has high complexity and various fluctuations. Four predictions are needed to display developments and relationships from the 1st to the 12th month. The proposed prediction model is something new that can be used as a reference for carrying out four types of predictions, relationships between predictions, and patterns between predictions in depth with abundant data about the number of passengers in bus public transportation areas. The prediction results can be used as a reference, capturing the dynamics of fluctuations in the number of passengers in public transportation areas of buses and optimizing prediction models for stakeholders in the transportation sector.

2. Research Methods

Data was 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[28]. The dataset is taken about the number of passengers using public bus transportation. Six public bus transportation areas are operating, but four bus areas are taken namely central, north, south, and west. The total data is 1,673 based on daily bus public transportation passenger report data. The data used in this dataset are the date and number of passengers in each bus public transportation area daily. 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. Minmax feature scaling is used for data processing first; then, the dataset is divided into two segments (training and testing). The most appropriate and best model results were obtained with a dataset run using an LSTM model based on different parameter tuning. The training data set and prediction accuracy are evaluated by comparison of the prediction data. The selection of LSTM model parameters can be seen from 4 model evaluations (MSLE, MAPE, and SMAPE). The experiment's lowest evaluation value and relatively fast time become the most optimal model. The system architecture produces prediction results for the number of passengers in bus public transportation areas with input from the dataset processed by the LSTM model with the best evaluation results as a model for predicting the number of passengers in 4 bus public transportation areas in the future (Figure 1).

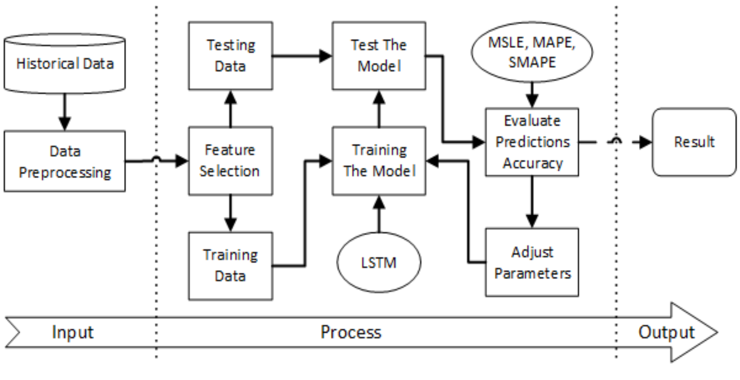

Figure 1. Bus Area Passenger Number Prediction Architecture

DOI : 10.24843/LKJITI.2024.v15.i03.p03
e-ISSN 2541-5832
Accredited Sinta 2 by RISTEKDIKTI Decree No. 158/E/KPT/2021

Prediction of passenger numbers in 4 bus public transportation areas adopting LSTM based on deep learning. The Recurrent Neural Network (RNN) was developed, and a special type is LSTM[29]. which is modified to address the weaknesses of RNN[30], [31], [32]. Neural network design gate units and memory cells were introduced to find a way to re-collect data over a certain period carried out by LSTM[33]. The four neural network layers are made up of the LSTM layer[29] on certain methods that interact[31], [34]. Memory blocks carry out the role of normal neurons in hidden layers, which are special LSTM units[32]. The main structure of LSTM consists of three gates (input gate, forget gate, and output gate)[29], [31] in the algorithm structure[35], which helps update and control the flow of information through memory blocks [32] during each activation function of the network layer nerves[36]. Information remembered for long periods rather than learned from struggle becomes the default LSTM behavior. The repetition module of the nervous system is likened to a chain with a basic structure covering a single ground layer[33]. LSTM contains a chain structure with repeating modules following different structures(
Figure ***2***). The LSTM layer that is built consists of 2 layers, each using hyperbolic tangent (tanh) activation. Drop out is always done at the end of each layer, after that you do a dense layer.

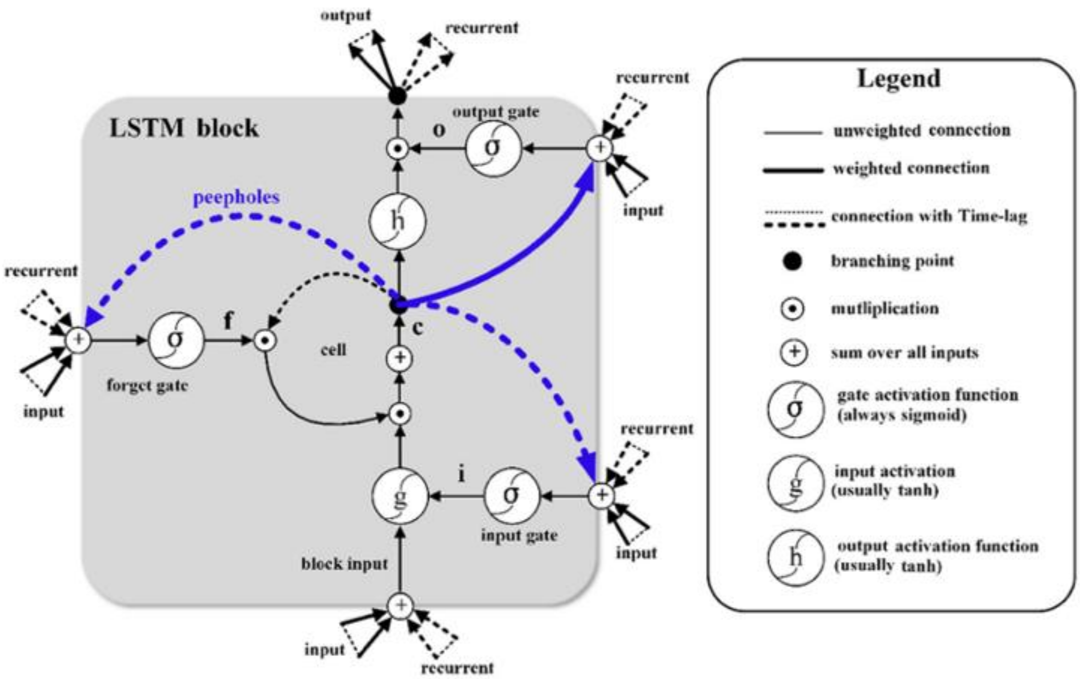

Figure 2. LSTM Cell Structure In Hidden Layer[37]

An output sequence that is known to be correct using multiple input sequences is trained by an LSTM model. Errors can be minimized by adjusting model parameters during training by passing actual and predicted output errors through the network (backpropagation)[30]. Newly experienced data is passed from one cell to another cell with different LSTM model gates. These are known as update, forget, and output gates. Two cell outputs (activation and candidate) are the outputs of the LSTM cell. The cell state is like a transport line throughout the chain that works together in a small linear way and passes data without any changes, which is a level line[33].

The performance of LSTM with the appropriate and best parameter values for predicting the number of road accidents is evaluated for accuracy. Mean-squared logarithmic Error (MSLE) is used to determine the accuracy of the prediction model with an evaluation matrix[38]. Mean Absolute Percentage Error (MAPE)[39], dan Symmetric Mean Absolute Percentage Error (SMAPE)[40]. MSLE uses logarithms to compensate for large outliers in the data set and treats them as if they were on the same scale as the target balanced model using similar error percentages. The model performance is considered by determining the accuracy proof of the MSLE loss function[41]. The target value and predicted value are denoted by y_i and $\hat{y}$ respectively, while n represents the total amount of data[42](1).

LONTAR KOMPUTER VOL. 15, NO. 3 DECEMBER 2024 p-ISSN 2088-1541
DOI : 10.24843/LKJITI.2024.v15.i03.p03 e-ISSN 2541-5832
Accredited Sinta 2 by RISTEKDIKTI Decree No. 158/E/KPT/2021

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

n is the number of observations in the dataset, y_i is the actual value at the ith observation, and $\hat{y}_i$ is the predicted value at the i-th observation.

Prediction accuracy uses key performance indicators commonly used by MAPE. Each request is based on each error shared by MAPE[43]. High errors during periods of low demand can significantly impact MAPE[39] (2). High and low prediction accuracy are in harmony with the MAPE value. A MAPE value that is getting closer to 0 reflects better model performance[44].

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

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

The difference in absolute value between the estimate and the actual is an absolute error[45]. SMAPE is a measure of accuracy based on the percentage of errors[46]. Evaluate the prediction performance of time series datasets using percentage-based and scale-independent SMAPE[40] (3). The accuracy value of a forecast is in line with the SMAPE value. SMAPE values that are closer to 0 represent better model performance[47].

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

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

3. Result and Discussion

The total number of public transport bus passengers daily is in numeric form from January 1, 2019, to July 31, 2023, into the data input format. The data is processed and formatted in the Comma-Separated Values (CSV) file type with 1,673 results. Five columns in the dataset are used: DATE, CENTRAL, NORTH, SOUTH, and WEST. DATE contains date data, CENTRAL contains data on the number of passengers in the central public bus transportation area, NORTH contains data on the number of passengers in the northern public bus transportation area, SOUTH contains data on the number of passengers in the southern public bus transportation area, and WEST contains data on the number of public transportation passengers western bus(Figure 3). Accurate and predicted data are subjected to appropriate scaling, training, and testing. The evaluation values from MSLE, MAPE, and SMAPE are observed using different layers and different units in 2 hidden layers and dense layers for prediction output.

	DATE	CENTRAL	NORTH	SOUTH	WEST
0	2019-01-01	24494	13060	7339	4868
1	2019-01-02	31920	16510	9711	7136
2	2019-01-03	46744	27126	15169	10188
3	2019-01-04	49467	27847	14807	10688
4	2019-01-05	37659	19822	12175	7750
...	...	...	...	...	...
1668	2023-07-27	83551	67003	26659	21099
1669	2023-07-28	81084	62788	26073	20312
1670	2023-07-29	45167	28496	13687	10059
1671	2023-07-30	37743	23597	9713	8122
1672	2023-07-31	79068	63316	25444	20987

Figure 3. Research Datasets

Analysis of different periods becomes the basis for training the LSTM model by selecting the best training period. 2 hidden layers, activation using hyperbolic tangent (tahn), epoch 80, and batch size 8 becomes the default LSTM model. The LSTM model used to train on training data with the optimizer used by Adam and Verbos is 1. Model variations The LSTM experiments carried out were the number of neurons, dropout, and training-testing data. The lowest evaluation scores given by MSLE, MAPE, and SMAPE, as well as the relatively fast time based on experimental variations of the LSTM model, are the most optimal models for predicting the number of passengers in public bus transportation areas. The quickest time occurred with a dropout variation of 0, neuron 32, and train-test 90-10, but the MSLE, MAPE, and SMAPE values were still relatively high compared to the others. The processing time depends on the tuning of the LSTM model parameters used. Increasing or decreasing the number of dropouts, neurons, and train tests 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. The most optimal deep learning-based LSTM model with a dropout of 0.2, 32 neurons, and train-test 80-20, because of the relatively fast time and the lowest MSLE, MAPE, and SMAPE values compared to the others(Table 1).

Table 1. LSTM Model Experiment Results

N	T	Drop Out: 0.0				Drop Out: 0.1				Drop Out: 0.2			
		E1	E2	E3	Time	E1	E2	E3	Time	E1	E2	E3	Time
8	50-50	0.283	55.98	37.21	00:01:09	0.350	67.04	38.38	00:01:27	0.341	65.74	38.15	00:00:45
	60-40	0.182	39.42	31.16	00:01:32	0.215	47.50	34.46	00:01:29	0.198	42.79	32.19	00:00:43
	70-30	0.147	31.20	27.98	00:01:26	0.152	32.76	28.96	00:01:29	0.149	32.70	28.76	00:01:25
	80-20	0.128	31.04	29.19	00:01:12	0.139	29.29	28.34	00:01:27	0.144	29.46	29.20	00:00:48
	90-10	0.134	28.65	28.79	00:01:09	0.150	29.10	28.87	00:01:26	0.151	31.17	31.07	00:00:42
16	50-50	0.313	60.43	37.27	00:01:27	0.280	54.68	36.06	00:01:18	0.347	67.05	38.68	00:01:28
	60-40	0.176	39.17	30.25	00:01:27	0.177	38.92	31.16	00:02:36	0.180	39.40	30.94	00:01:28
	70-30	0.145	30.58	29.17	00:02:29	0.145	31.23	29.64	00:01:28	0.144	31.25	29.21	00:01:17
	80-20	0.137	29.37	27.56	00:01:26	0.124	27.76	27.34	00:02:27	0.139	29.10	27.39	00:01:27
	90-10	0.153	30.18	31.90	00:01:23	0.166	28.69	29.81	00:01:23	0.150	28.59	28.06	00:01:26
32	50-50	0.300	58.62	35.63	00:01:28	0.299	58.52	36.18	00:01:28	0.246	47.95	34.01	00:01:32
	60-40	0.183	41.31	31.82	00:01:40	0.183	40.98	31.79	00:01:26	0.186	41.01	31.43	00:01:03
	70-30	0.145	31.28	25.01	00:02:29	0.142	32.58	27.51	00:01:08	0.143	31.36	28.17	00:01:07
	80-20	0.129	26.45	24.48	00:02:27	0.145	27.69	27.37	00:01:28	0.102	22.46	20.98	00:01:12
	90-10	0.167	28.74	29.63	00:00:41	0.152	26.78	26.10	00:01:27	0.146	28.89	26.13	00:01:05

Note: N=Neuron, T=Train-Test, E1=MSLE, E2=MAPE, E3=SMAPE.

Internet connection stability, appropriate data format, and adequate computer performance are the keys to conducting experiments on the proposed model. Combinations and variations greatly influence the time needed to carry out processing, which will be measured by the evaluation value. If there are more neurons, the LSTM model process will take longer. The experiment produces a large number of neurons, so, certainly, the time required is also relatively fast. In-depth investigations were also carried out on large datasets to achieve the best accuracy and suitability of the desired predictions. Appropriate combinations and variations are essential factors for training the model and design of the proposed framework to perform four passenger predictions of bus public transportation areas. Dataset normalization is the first task that must be considered when creating a time-series prediction model using LSTM in the preprocessing process. Empty data and data discrepancies are problems that must be resolved. LSTM can process time-series data; the suitability of time-series data is critical in the preprocessing stage.

The best LSTM model was evaluated from MSLE of 0.102, MAPE of 22.46, and SMAPE of 20.98. Optimal settings require a longer time because they require a more significant number of neurons, and the size of the train-test data distribution also influences this. Correspondence between the number of neurons, train-test data, and drop out is the key for the three evaluation matrices to produce the lowest values in a relatively fast time. The three evaluation models used have a value of less than 30, which means the LSTM model has good performance quality and is acceptable. This is because if the value is more than 30, then the proposed model needs improvement, and even a value of more than 50 is unacceptable. Comparison of loss and validation loss with dropout 0, 0.1, and 0.2 for epoch 80, batch size 8, neuron 32, and train-test 80-20. The train-test data ratio of 80-20 is the best model because the model is sufficient to train 80% of the dataset used and to test the dataset 20% of the model training results that have been carried out. This ratio is stable enough to train and test data with the lowest value results from the three evaluation matrices. In a comparison of graphs with neuron variations in general, the training and validation lines are almost the same, so dropout 0.2 with the lowest evaluation results is the most optimal model(

 p-ISSN 2088-1541
e-ISSN 2541-5832

Figure ***4***).

Figure 4. Output Epoch by LSTM Model

The loss and validation loss graphs display the model's learning performance during training and evaluation of validation data. The loss value generated during training is depicted at each epoch. The graphs displayed become a critical tool in evaluating model performance. The loss function measures how much the model can make accurate predictions on the training data. Loss graphs pay attention to the difference between loss on training data and validation data. Validation Loss graphs provide insight into how well a model can predict never-before-seen data and special attention is paid to potential overfitting or underfitting. The deviation between the Loss and Validation Loss graphs can provide important insight into the quality of the model's generalization to new data.

The proposed model learns on the dataset with the most optimal training-testing division, namely 80%-20%. The proposed model's prediction results are by 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 in 4 bus public transportation areas, which is getting better(Figure 5). The four movements fluctuate with the same rhythm, even though a few tangents exist between the data. This is due to a very sharp decline due to the complete cessation of public transportation operations. The highest passenger movement is in the central region, while the lowest is in the western region. The predictions made are greatly influenced by the data obtained as the primary material for making predictions. Prediction models have been carried out for various predictions of short-term public transport bus passengers[17], [18], [19], [20], [22], [23], [24], [25], [26]. The model focuses more on finding the best accuracy[18], [20], [21], model accuracy comparison[17], [19], [22], [23], [24], [25], and improve model accuracy[26], [27] by using various evaluation matrices. The best model for the proposed prediction model is based on an evaluation matrix that is different from previous prediction models in the form of settings for drop-outs, neurons, and the distribution of the prediction model's train-test data. The model also presents processing time, four types of predictions, relationships between predictions, and patterns between predictions. The results of the three evaluation metrics of the proposed prediction model have better values compared to several prediction models that have been carried out.

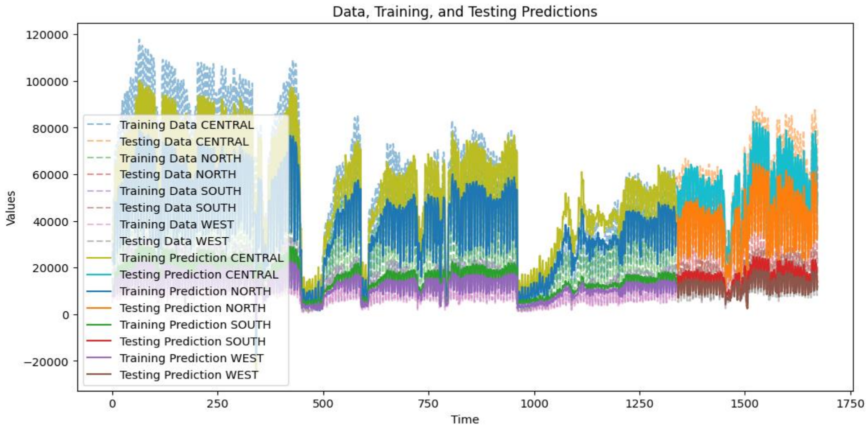

Figure 5. Data, Training, and Testing Prediction

Predictions are made for the next 12 months for the four types of predictions produced. The prediction results show a drastic decline starting from the 2nd month (October). Still, starting from the 4th month (February), the movement tends to stabilize and experience a slight increase until the 12th month (August). The predictions for the number of passengers do not overlap because all bus areas operate normally. Passengers between bus regions have long distances, passengers in the south and west regions being the closest. The distances between predictions do not intersect, which means that fluctuations in the number of passengers occur simultaneously. The proposed prediction model has stable accuracy when using three evaluation matrices to produce predictions dynamically according to the use of a dataset which can be used as a solution to see predictions of the number of passengers in bus public transportation areas in the next 12 months(Figure 6).

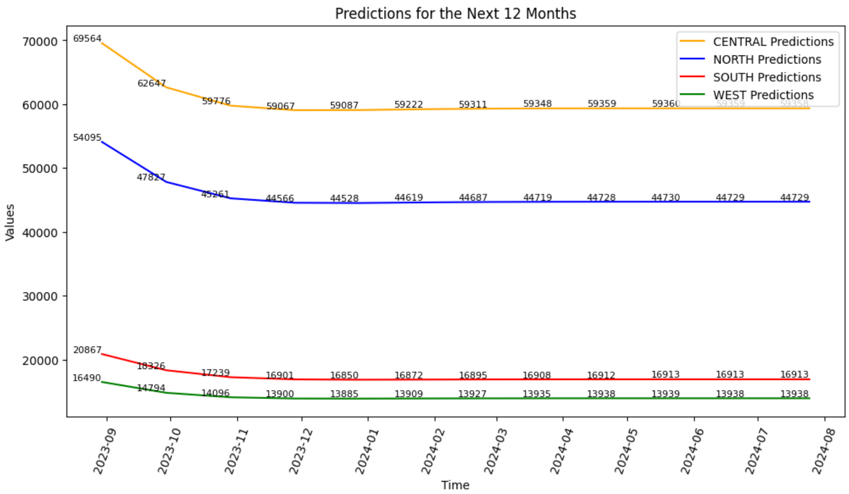

Figure 6. Predictions For The Next 12 Months

The four predictions for the number of passengers in public transportation areas on buses have a positive or negative relationship, which can be presented with a correlation heat map. In the correlation heatmap, blue indicates a negative correlation, and red indicates a positive correlation. A positive link means that the predictions are directly proportional, while a negative link implies that the predictions are inversely proportional. The predicted number of passengers in the bus area with a strong negative correlation is CENTRAL-SOUTH. The predicted number of passengers with a weak positive correlation is CENTRAL to WEST and NORTH. The predicted number of passengers that is strongly positively correlated is NORTH-WEST, while the one that is usually correlated is SOUTH to NORTH and WEST(Figure 7).

LONTAR KOMPUTER VOL. 15, NO. 3 DECEMBER 2024 p-ISSN 2088-1541
DOI : 10.24843/LKJITI.2024.v15.i03.p03 e-ISSN 2541-5832
Accredited Sinta 2 by RISTEKDIKTI Decree No. 158/E/KPT/2021

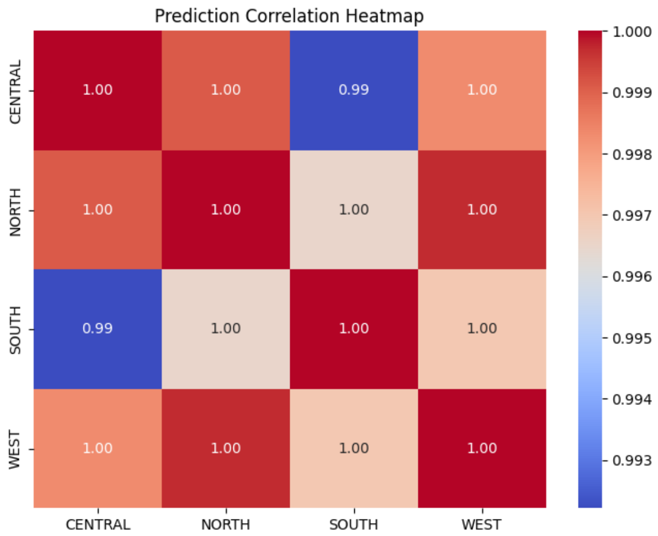

Figure 7. Prediction Correlation Heatmap

The relationship between prediction variables can be 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. If the points are concentrated in an area, it indicates that the variables are interdependent and can be used in predictions. As the diagonal scatter plot suggests, the four predictions have a solid relationship. Predictions closely related and dependent are CENTRAL-WEST, CENTRAL-NORTH, NORTH-WEST, and WEST-SOUTH. Predictions that are less closely interconnected and dependent are CENTRAL-SOUTH (Figure 8).

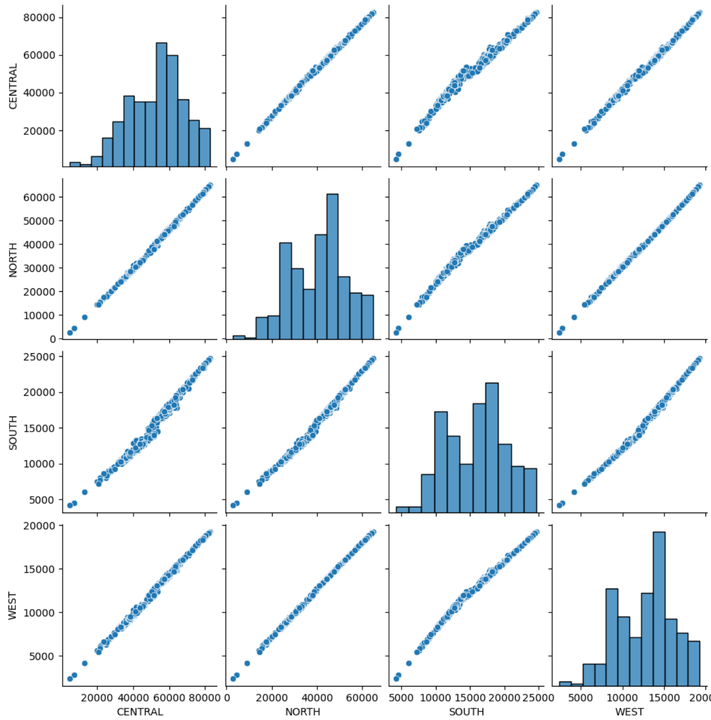

Figure 8. Prediction Scatter Matrix

The proposed prediction modeling can capture the dynamics of fluctuations in the number of passengers in 4 operational areas of public bus transportation. The presentation of passenger descriptions can be used as a reference for creating policies and strategies to deal with them. A solution for improving the quality of public bus transportation services is to implement appropriate policies and strategies based on time. Policy making and improving mobility using good public transportation in a new direction using predictive models to predict the number of bus passengers. Good bus operation management is the basis for decisions and plays an important role. Public bus transportation agencies can use this model and insight to improve the quality and breadth of information provided to users of each region's bus public transportation system.

4. Conclusion

Predicting the number of passengers in 4 bus public transportation areas (central, north, south, and west) using time-series data can be done using a deep learning-based LSTM model. The most optimal prediction model is epoch 80, batch size 8, neuron 32, dropout 0.2, and train-test 80-20 with the lowest MSLE, MAPE, and SMAPE evaluation values compared to other variations. Predictions on training and testing data improve with appropriate tuning. The proposed prediction model predicts 12 months later for four predictions of bus area passenger numbers with fluctuations occurring simultaneously. There is a strong negative correlation in central-south and a strong positive correlation in the north-west; apart from that, there is a weak negative and positive correlation. Predictions that are less closely interconnected and dependent are central-south. Policy making and increasing mobility using public transportation in each region get a new direction with a predictive model proposed to predict the number of bus passengers. The proposed model needs to be compared with other prediction models, and using short-term prediction models will be a step in future work.

References

[1] X. Shen and S. Feng, “How public transport subsidy policies in China affect the average passenger load factor of a bus line” *Research in Transportation Business & Management*, vol. 36, p. 100526, Sep. 2020, doi: 10.1016/j.rtbm.2020.100526.

[2] J. Jalaney and R. S. Ganesh, “Multiple Extreme Learning Machines Based Arrival Time Prediction for Public Bus Transport” *Intelligent Automation & Soft Computing*, vol. 36, no. 3, pp. 2819–2834, 2023, doi: 10.32604/iasc.2023.034844.

[3] X.-G. Luo, H.-B. Zhang, Z.-L. Zhang, Y. Yu, and K. Li, “A New Framework of Intelligent Public Transportation System Based on the Internet of Things” *IEEE Access*, vol. 7, pp. 55290–55304, 2019, doi: 10.1109/ACCESS.2019.2913288.

[4] A. Vigren and R. Pyddoke, “The impact on bus ridership of passenger incentive contracts in public transport,” *Transportation Research Part A: Policy and Practice*, vol. 135, pp. 144–159, May 2020, doi: 10.1016/j.tra.2020.03.003.

[5] W. Wang, Y. Wang, G. H. de A. Correia, and Y. Chen, “A Network-Based Model of Passenger Transfer Flow between Bus and Metro: An Application to the Public Transport System of Beijing” *Journal of Advanced Transportation*, vol. 2020, pp. 1–12, Dec. 2020, doi: 10.1155/2020/6659931.

[6] H. Zhou, C. Yuan, N. Dong, S. C. Wong, and P. Xu, “Severity of passenger injuries on public buses: A comparative analysis of collision injuries and non-collision injuries” *Journal of Safety Research*, vol. 74, pp. 55–69, Sep. 2020, doi: 10.1016/j.jsr.2020.04.003.

[7] N. B. Dhital, L.-C. Wang, H.-H. Yang, C.-H. Lee, W.-H. Shih, and C.-S. Wu, “Effects of the COVID-19 pandemic on public bus occupancy and real-world tailpipe emissions of gaseous pollutants per passenger kilometer traveled” *Sustainable Environment Research*, vol. 32, no. 1, p. 35, Aug. 2022, doi: 10.1186/s42834-022-00146-7.

[8] R. Burdzik, W. Chema, and I. Celiński, “A study on passenger flow model and simulation in aspect of COVID-19 spreading on public transport bus stops” *Journal of Public Transportation*, vol. 25, p. 100063, 2023, doi: 10.1016/j.jpubtr.2023.100063.

[9] 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.

[10] 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.

[11] I. Mariñas-Collado, A. E. Sipols, M. T. Santos-Martín, and E. Frutos-Bernal, "Clustering and Forecasting Urban Bus Passenger Demand with a Combination of Time Series Models" *Mathematics*, vol. 10, no. 15, p. 2670, Jul. 2022, doi: 10.3390/math10152670.
[12] F. M. A. Hassouna, "Sustainability assessment of public bus transportation sector in westbank, palestine" *Environmental Research Communications*, vol. 5, no. 1, p. 015001, Jan. 2023, doi: 10.1088/2515-7620/acb03f.
[13] R. Verma *et al.*, "A Smartphone-Based Passenger Assistant for Public Bus Commute in Developing Countries" *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 465–476, Apr. 2020, doi: 10.1109/TCSS.2019.2961390.
[14] B. Jain, "Intelligent Bus Tracking and Passenger Counting System" *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 6, no. 4, 2019.
[15] R. Thiagarajan and Dr. S. Prakashkumar, "Identification of Passenger Demand in Public Transport Using Machine Learning" *Webology*, vol. 18, no. Special Issue 02, pp. 223–236, Apr. 2021, doi: 10.14704/WEB/V18SI02/WEB18068.
[16] W. T. Wu, K. Zeng, W. Zhou, P. Li, and W. Z. Jin, "Deep learning method for bus passenger flow prediction based on multi-source data and surrogate-based optimization" *Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)*, vol. 53, no. 7, 2023, doi: 10.13229/j.cnki.jdxbgxb.20210935.
[17] D. Luo *et al.*, "Fine-Grained Service-Level Passenger Flow Prediction for Bus Transit Systems Based on Multitask Deep Learning" *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 7184–7199, Nov. 2021, doi: 10.1109/TITS.2020.3002772.
[18] 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.
[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] V. G S and H. V S, "Prediction of Bus Passenger Traffic using Gaussian Process Regression" *Journal of Signal Processing Systems*, vol. 95, no. 2–3, pp. 281–292, Mar. 2023, doi: 10.1007/s11265-022-01774-3.
[21] 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.
[22] 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.
[23] J. D. Zhao, J. Shen, and L. W. Liu, "Bus passenger flow classification prediction driven by CNN-GRU model and multi-source data" *Jiaotong Yunshu Gongcheng Xuebao/Journal of Traffic and Transportation Engineering*, vol. 21, no. 5, 2021, doi: 10.19818/j.cnki.1671-1637.2021.05.022.
[24] Z. Mei, W. Yu, W. Tang, J. Yu, and Z. Cai, "Attention mechanism-based model for short-term bus traffic passenger volume prediction" *IET Intelligent Transport Systems*, vol. 17, no. 4, pp. 767–779, Apr. 2023, doi: 10.1049/itr2.12302.
[25] 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.
[26] G. S. Vidya and V. S. Hari, "LSTM Network Integrated with Particle Filter for Predicting the Bus Passenger Traffic" *Journal of Signal Processing Systems*, vol. 95, no. 2–3, pp. 161–176, Mar. 2023, doi: 10.1007/s11265-022-01831-x.
[27] N. Nagaraj, H. L. Gururaj, B. H. Swathi, and Y.-C. Hu, "Passenger flow prediction in bus transportation system using deep learning" *Multimedia Tools and Applications*, vol. 81, no. 9, pp. 12519–12542, Apr. 2022, doi: 10.1007/s11042-022-12306-3.
[28] Auckland Transport, "AT Metro patronage report," Wellington, 2023. [Online]. Available: https://at.govt.nz/about-us/reports-publications/at-metro-patronage-report

[29] M. M. Patel, S. Tanwar, R. Gupta, and N. Kumar, “A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions” *Journal of Information Security and Applications*, vol. 55, no. August, p. 102583, 2020, doi: 10.1016/j.jisa.2020.102583.
[30] H. K. Dam, T. Tran, T. Pham, S. W. Ng, J. Grundy, and A. Ghose, “Automatic Feature Learning for Predicting Vulnerable Software Components” *IEEE Transactions on Software Engineering*, vol. 47, no. 1, pp. 67–85, 2021, doi: 10.1109/TSE.2018.2881961.
[31] A. Mokhtar *et al.*, “Estimation of SPEI Meteorological Drought Using Machine Learning Algorithms” *IEEE Access*, vol. 9, pp. 65503–65523, 2021, doi: 10.1109/ACCESS.2021.3074305.
[32] S. Ghimire, Z. M. Yaseen, A. A. Farooque, R. C. Deo, J. Zhang, and X. Tao, “Streamflow prediction using an integrated methodology based on convolutional neural network and long short-term memory networks” *Scientific Reports*, vol. 11, no. 1, p. 17497, Sep. 2021, doi: 10.1038/s41598-021-96751-4.
[33] Md. A. Istiake Sunny, M. M. S. Maswood, and A. G. Alharbi, “Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model” in *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, IEEE, Oct. 2020, pp. 87–92. doi: 10.1109/NILES50944.2020.9257950.
[34] M. Nabipour, P. Nayyeri, H. Jabani, S. S., and A. Mosavi, “Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis” *IEEE Access*, vol. 8, pp. 150199–150212, 2020, doi: 10.1109/ACCESS.2020.3015966.
[35] L. Mu, F. Zheng, R. Tao, Q. Zhang, and Z. Kapelan, “Hourly and Daily Urban Water Demand Predictions Using a Long Short-Term Memory Based Model” *Journal of Water Resources Planning and Management*, vol. 146, no. 9, Sep. 2020, doi: 10.1061/(ASCE)WR.1943-5452.0001276.
[36] Y. Jung, J. Jung, B. Kim, and S. Han, “Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea” *Journal of Cleaner Production*, vol. 250, p. 119476, Mar. 2020, doi: 10.1016/j.jclepro.2019.119476.
[37] D. Xu, Q. Zhang, Y. Ding, and D. Zhang, “Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting” *Environmental Science and Pollution Research*, vol. 29, no. 3, pp. 4128–4144, Jan. 2022, doi: 10.1007/s11356-021-15325-z.
[38] S. Basheer, R. M. Mathew, and M. S. Devi, “Ensembling Coalesce of Logistic Regression Classifier for Heart Disease Prediction using Machine Learning” *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 12, pp. 127–133, Oct. 2019, doi: 10.35940/ijitee.L3473.1081219.
[39] 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.
[40] 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.
[41] M. S. Devi, S. Basheer, and R. M. Mathew, “Exploration of Multiple Linear Regression with Ensembling Schemes for Roof Fall Assessment using Machine Learning” *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 12, pp. 134–139, Oct. 2019, doi: 10.35940/ijitee.L3474.1081219.
[42] N. Singh, P. Sharma, N. Kumar, and M. Sreejeth, “Short-Term Load Forecasting Using Artificial Neural Network and Time Series Model to Predict the Load Demand for Delhi and Greater Noida Cities” in *Lecture Notes in Networks and Systems*, vol. 177 LNNS, 2021, pp. 443–455. doi: 10.1007/978-981-33-4501-0_41.
[43] 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.
[44] 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.

[45] X. Xu and M. Yoneda, "Multitask Air-Quality Prediction Based on LSTM-Autoencoder Model" *IEEE Transactions on Cybernetics*, vol. 51, no. 5, pp. 2577–2586, May 2021, doi: 10.1109/TCYB.2019.2945999.
[46] 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.
[47] G. S. Vidya and V. S. Hari, "LSTM Network Integrated with Particle Filter for Predicting the Bus Passenger Traffic" *Journal of Signal Processing Systems*, vol. 95, no. 2–3, pp. 161–176, Mar. 2023, doi: 10.1007/s11265-022-01831-x.