Published online on: https://journal.uii.ac.id/jurnalsnati/
DOI: 10.20885/snati.v3.i3.37

Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi
Vol. 3 No. 3 (2024) 95 - 103 ISSN Media Electronic: 2807-5935

Deep Learning Based LSTM Model Hyperparameter Testing to Predict the Number of Road Accidents

Joko Siswanto[1*], Benny Daniawan[2], Haryani Haryani[3], Pipit Rusmandani[4]
[1,4]Road Transportation Systems Engineering, Politeknik Keselamatan Transportasi Jalan, Tegal, Indonesia
[2]Faculty of Science and Technology, Universitas Buddhi Dharma, Tangerang, Indonesia
[3]Department of Informatics Engineering, UIN Alauddin Makassar, Makassar, Indonesia
[1]siswanto@pktj.ac.id, [2]b3n2y.miracle@gmail.com, [3]hariani.kasim@uin-alauddin.ac.id, [4]pipit@pktj.ac.id

Abstract

Many have used the prediction of the number of road accidents, but it is still rare to find those who use and test prediction models that are not suitable. Predictive models that have been used to predict road accidents have proven successful, but have not provided model testing with data that is different from the deep learning approach. The LSTM model test is proposed to be tested with 5 different datasets from Kaggle and 3 hidden layer variations. The test results of the LSTM model are that with variations of 4 hidden layers it can achieve higher accuracy results than those without hidden layers and 2 hidden layers. The results are obtained from stability with the lowest average MSLE value and relatively balanced average time. Deep learning-based LSTM model testing was carried out to ensure and prove the stability of the model for predicting the number of road accidents in the future. Stakeholders can predict the number of road accidents using the resulting prediction model.

Keywords: *testing model; road accident; deep learning, LSTM*

1. Introduction

Stakeholders always monitor and store road accident data in large and reliable public databases [1]. Road accidents are documented as a major source of human and financial disaster as they create global, social and economic problems, and millions of people are killed each year in these accidents [2]. Road accidents pose a serious socio-economic challenge and cause loss of life and property [3]. The increase in the possibility of road accidents is contributed by the parameters of low levels of poverty, vulnerable age, road length and traffic density [4]. Road accidents include various elements and conditions [5]. Several methods can be used to help predict road accidents in the future [6]. Many predictive model tests are still found to be inappropriate and it is still rare to discuss prediction models for the number of road accidents.

The Long Short Term Memory (LSTM) model is used to predict truck traffic flows with superior performance and improve prediction accuracy [7]. The patient volume is much closer to the actual one in predicting the volume of COVID-19 patients in Pakistan using LSTM [8]. The accuracy advantage obtained from predicting software defects by utilizing the LSTM model compared to others [9]. Fine dust prediction for the area where the user is present has reliable accuracy using LSTM [10]. LSTM for predicting nonlinear stress has strong applicability and higher accuracy than others [11]. LSTM achieves increased accuracy above average for predicting traffic flow on East Beijing Sanlihe Road in China [12]. Prediction of tool wear during machining of alloy steel derived from LSTM modeling with more accurate results [13]. Short-term urban water demand prediction in Hefei Cine city using LSTM with the best performance and accuracy compared to other prediction models [14]. The LSTM model has been widely used to make predictions in various fields and needs. The model is most appropriate for making predictions with time-series data.

Predicting the number of road accidents per year in certain time intervals using deep learning called the auto encoder procedure can provide better accuracy [15]. Deep Learning (DL) Multi-Layer Perceptron (MLP) has the highest increase in accuracy than other algorithms used to predict road accidents in India [16]. Road accident prediction using Stacked Sparse AutoEncoder (SSAE) can achieve the best performance compared to other baseline models [17]. Road accident predictions based on social media platform data can be superior to Gated Recurrent Units (GRU) and Convolutional Neural Networks (CNN) compared to other basic deep learning algorithms [18]. Road accident prediction models with deep learning from tweet messages can improve and outperform the accuracy of other sophisticated algorithms, so that the

Received: 30-05-2024 | Revised: 25-06-2024 | Accepted: 29-06-2024 | Published: 25-07-2024

use of advanced features will help predict road accidents better [19]. Predictive models that have been used to predict road accidents have proven successful, but have not provided model testing with data that is different from the deep learning approach.

Deep learning can be used to create a road accident prediction model which is considered a standard paradigm approach that is different from the suitability of a logical review and targets [20]. The incidence of road accidents is difficult to predict precisely with the number in Serbia, but its development can use machine learning or deep learning [21]. The prediction model for describing the timing of the number of road accidents in Poland was selected based on the results of testing the predictive ability of each model [22]. The best results from testing prediction models by comparing performance based on correction and classification techniques were carried out for prediction models for the number of road accidents in South Korea [23]. Testing of the prediction model for the number of road accidents in China is used as a good reference in preventing and controlling road accidents [24]. The road accident prediction model is useful for predicting losses caused by first testing the proposed prediction model [25]. A precise prediction model using a deep learning approach is used to predict the number of road accidents. A deep learning approach using the LSTM model is used to predict the number of road accidents using time-series data. Testing the proposed LSTM model will be tested with different datasets and hidden layer variations. The best results will be obtained based on the average value of Mean-Squared Logarithmic Error (MSLE) and time for each hidden layer variation. Testing the prediction of the number of road accidents according to hidden layer variations using LSTM provides proof and reference for the use of appropriate and appropriate prediction models.

2. Research Methods

The data collected is obtained from Kaggle which can be accessed by the public. There are 5 datasets used in the form of CSV consisting of Great Britain Road Accidents [26], 1.6 Million UK Traffic Accidents [27], Road Accidents in UK [28], UK Car Accidents 2005-2015 [29], and Thailand Fatal Road Accidents 2011-2022 [30]. The formatting, filtering and summing of road accident events based on date was adjusted to the 5 datasets used. The date and number of road accident incidents are columns in each dataset. Google Colab was used to run the Python programming language on macOS Sonoma 14.1.1 with a MacBook Air Apple M2 8GB memory. Tensor Flow is used for deep learning framework. Minmax feature scaling and dividing training and testing segments become the first data processing. Predictions on the number of road accidents are produced from an LSTM model that runs on a training dataset with a variety of different hidden layers. The prediction and training datasets are compared and the prediction accuracy is evaluated. The prediction model with the best evaluation was produced by LSTM from the system architecture based on the input dataset for predicting the number of road accidents in the coming year (Figure 1). The results of the accuracy and time values for each dataset are averaged, so that it will display the average accuracy and time required for each hidden layer variant. The variations of the hidden layer used are non-hidden layer, 2 hidden layers, and 4 hidden layers.

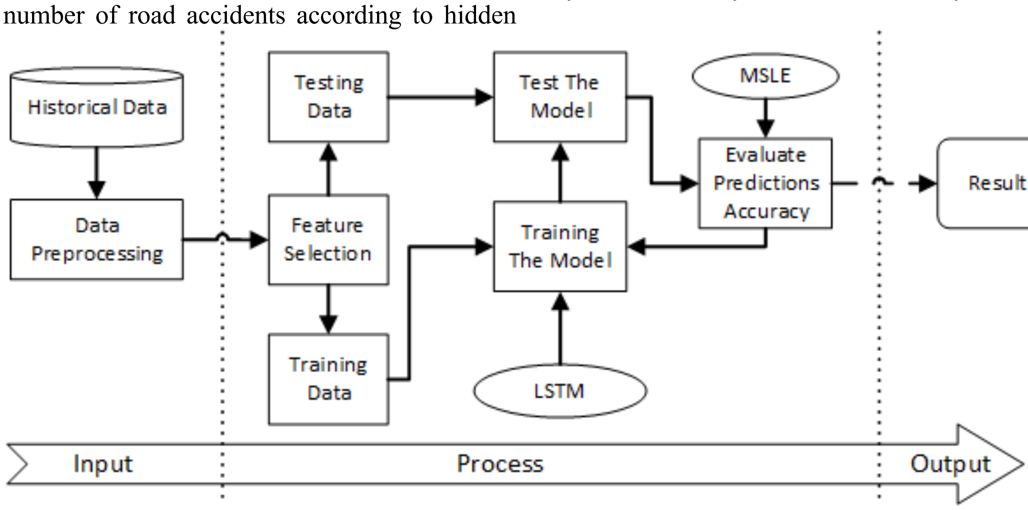


Figure 1. LSTM Model Testing Architecture for Predicting the Number of Road Accidents

LSTM is adopted based on deep learning to predict the number of road accidents. LSTM is a development and a special type of Recurrent Neural Network (RNN) [31] which is modified to overcome the weaknesses of RNN [32], [33], [34], [35], [36]. The gate unit and memory cells in neural network design are presented by LSTM to find out how to collect data again over a certain period [37]. LSTM neural networks are made of 4 layers [31] with interactions in certain methods [33], [34]. The role of normal neurons is performed in the memory blocks in the hidden layer of the special LSTM dependency unit [36]. LSTM consists of a main structure of three gates (input gate, forget gate, and output gate) [31], [33] in the structure of the algorithm

[14]. The flow of information via memory blocks is updated and controlled with the help of these gates [36] during each activation function of the neural network layer [35]. Long-term remembered information that is not based on lessons learned from the struggle becomes the default behavior of LSTM. The repetition module of the nervous system is likened to a chain that has the basic structure of a single layer of soil [37]. The LSTM contains a chain structure with repeating modules following different structures (Figure 2). The built LSTM layer consists of 2 layers, each of which uses an activation hyperbolic tangent (tanh). Drop out is always done at the end of the layer, after that do dense layers.

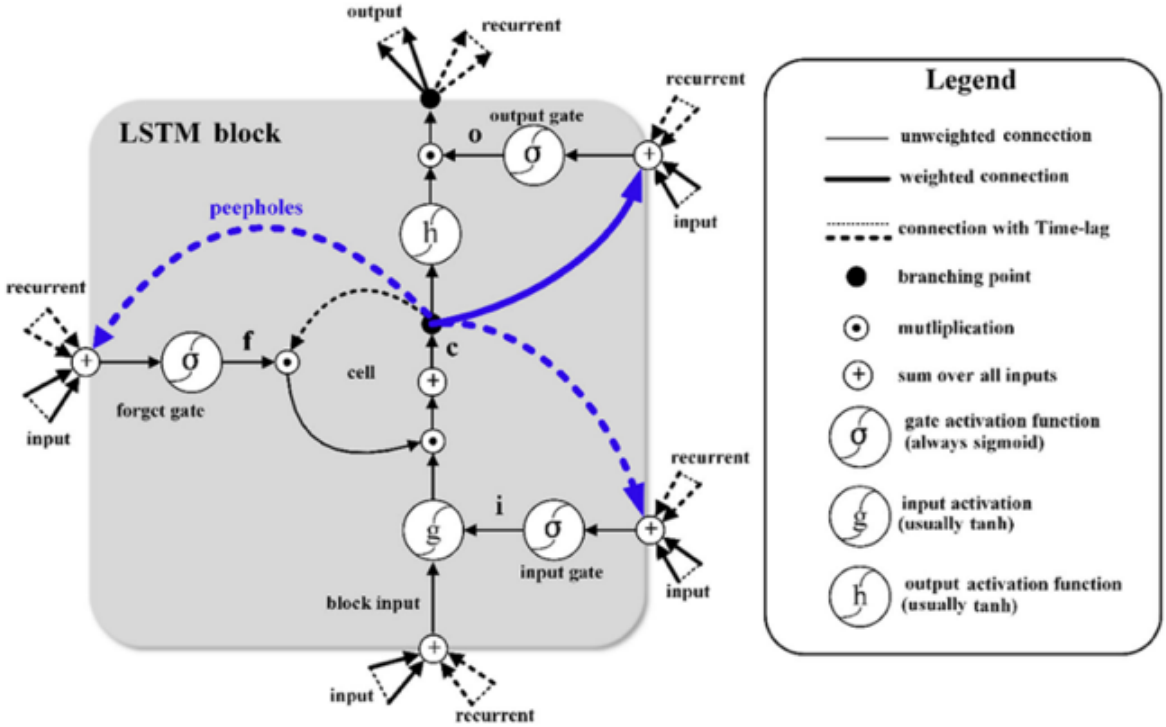


Figure 2. LSTM Cell Structure In Hidden Layer[38]

Many input sequences with output sequences that are known to be correct are used to train the LSTM model. Errors can be minimized by adjusting the model parameters during training in such a way that the actual output errors and predictions are passed back through the network (backpropagation) [32]. Data to or from cell status can be included or excluded by LSTMThe certainty of the cell's existence is gated to manage the data deliberately. Gate is a combination of the hyperbolic tangent (tanh) function and the dot product process. Data that has just been experienced from one cell to another cell is passed through different gates by the LSTM model known as update gates (Formula 1), forgetting gates (Formula 2), and output gates (Formula 3). Two outputs from the cell in the form of activation and candidate values will be produced by LSTM (Formula 4). The highest point that passes through the level line will be passed through the data. The level line is referred to as the state of the cell which is somewhat similar to the transport line by walking directly through the chain only to a few small linear cooperation and can only pass the data without any changes [37].

Update Gate: $\Gamma_u = \sigma(W_u[h^{<t-1>}, x^t] + b_u)$ (1)

Forget Gate: $\Gamma_f = \sigma(W_f[h^{<t-1>}, x^t] + b_f)$ (2)

Output Gate: $\Gamma_o = \sigma(W_o[h^{<t-1>}, x^t] + b_o)$ (3)

Output: $\begin{cases} c^{<t>} = \Gamma_u * c^{N<t>} + \Gamma_f * c^{<t-1>} \\ a^{<t>} = \Gamma_o * c^{<t>} \end{cases}$ (4)

The accuracy of the LSTM performance with the appropriate and best parameter values for predicting the number of road accidents is evaluated. Estimating the accuracy of the prediction model can use various types of evaluation matrices, one of which is Mean-Squared Logarithmic Error (MSLE) [39], [40], [41]. MSLE uses logarithms to offset large outliers in the data set and treats them as if they were on the same scale as the target balanced model using a similar percentage of error. The performance of the model under consideration is determined by a more accurate loss function as evidence of MSLE [41]. The target value and the predicted value are denoted by yi and ŷ, while n represents the total amount of data [42] (Formula 5).

$$MSLE = \frac{1}{n}\sum_{i-1}^{n}(log(y_i + 1) - log(\hat{y} + 1))^2 \quad (5)$$

3. Results and Discussions

The number of accidents that occur every day with the data input format in numeric form in the 5 datasets used. 80% of the training data and 20% of the testing data are the data division carried out. Appropriate scaling, training and testing are carried out on real data and predictions that will be observed MSLE evaluation values with variations without hidden layers, 2 hidden layers and 4 hidden layers of the LSTM model. LSTM without hidden layers only has one LSTM layer without additional hidden layers which may have a relatively high MSLE. This is due to the lack of complexity in capturing data patterns. LSTM models without hidden layers are not sufficient to handle complex datasets or have many variables, so they may fail to effectively capture non-linear relationships in the data. LSTM 2 hidden layers has two hidden layers besides the main LSTM layer which tends to have a lower MSLE compared to models without hidden layers. The addition of hidden layers allows the model to capture more complex patterns and deep relationships in the data. The 2 hidden layer LSTM model is better able to learn relevant features and capture temporal dynamics in the data, thereby increasing prediction accuracy. The LSTM 4 hidden layer has four hidden layers in addition to the main LSTM layer which may have an even lower MSLE. If the data is not complex enough or the dataset size is not large enough, diminishing returns or even overfitting may occur. Too many hidden layers can cause the model to become too complex and risk memorizing training data rather than learning common patterns. The LSTM 4 hidden layer model can be very good at capturing very complex relationships in the data, but the risk of overfitting must be taken into account. Increasing the number of hidden layers typically improves the model's ability to capture complex patterns and decreases MSLE. The right balance needs to be found for overly complex models that can suffer from overfitting and require more computational resources.

Activation hyperbolic tangent (tanh), dropout 0.20, epochs 100, batch size 32, and neurons 128 are the default LSTM models. The hyperbolic tangent (tanh) activation function is used by LSTM because it has an output range of −1 to 1 which can help to overcome the vanishing gradient problem that often arises in deep neural networks. Activation of tanh allows the model to learn richer representations due to its non-linear nature. Dropout is used to help prevent overfitting by randomly eliminating units (neurons) during training. A dropout rate of 20% (0.20) means that 20% of neurons will drop out in each epoch. It is often considered a good compromise between preserving model capacity and preventing overfitting. The number of epochs determines how many times the entire dataset will be processed through the network during training. 100 epochs is usually sufficient to ensure that the model has enough opportunity to learn data patterns which can reduce the risk of overfitting, but depends on the complexity of the data and the size of the dataset. Batch size refers to the number of samples processed before the model is updated. A batch size of 32 is often used which can offer a balance between computing speed and training stability. Larger batches can take advantage of parallelism in the computation, while smaller batches can provide more frequent and finer-grained updates to the model. The number of neurons in an LSTM layer determines the model's capacity to learn representations from the data. 128 neurons is often chosen as a standard size that is large enough to capture complex patterns, but not too large to avoid overfitting or extremely high computational requirements. Hyperparameter selection is based on experimentation and understanding of the data used. Hyperparameter tuning is often carried out to find the optimal combination.

There are 1 Adam and Verbos optimizers used by LSTM to train training data. The average the fastest time to complete the process is 00:00:48 with no hidden layer. LSTM with 4 hidden layers has the best prediction accuracy with the lowest average MSLE value, but the longest average time compared to the others (Table 1). These results are obtained by making an average of 5 datasets on hidden layer variations. Analysis was carried out on all datasets for changes in hidden layer variations. The more hidden layers, the more time it takes, this happens evenly for the 5 datasets. The suitability of hidden layer variations was obtained from testing the LSTM model with 5 different datasets.

Table 1. Comparison Of MSLE On Hidden Layer LSTM Variations

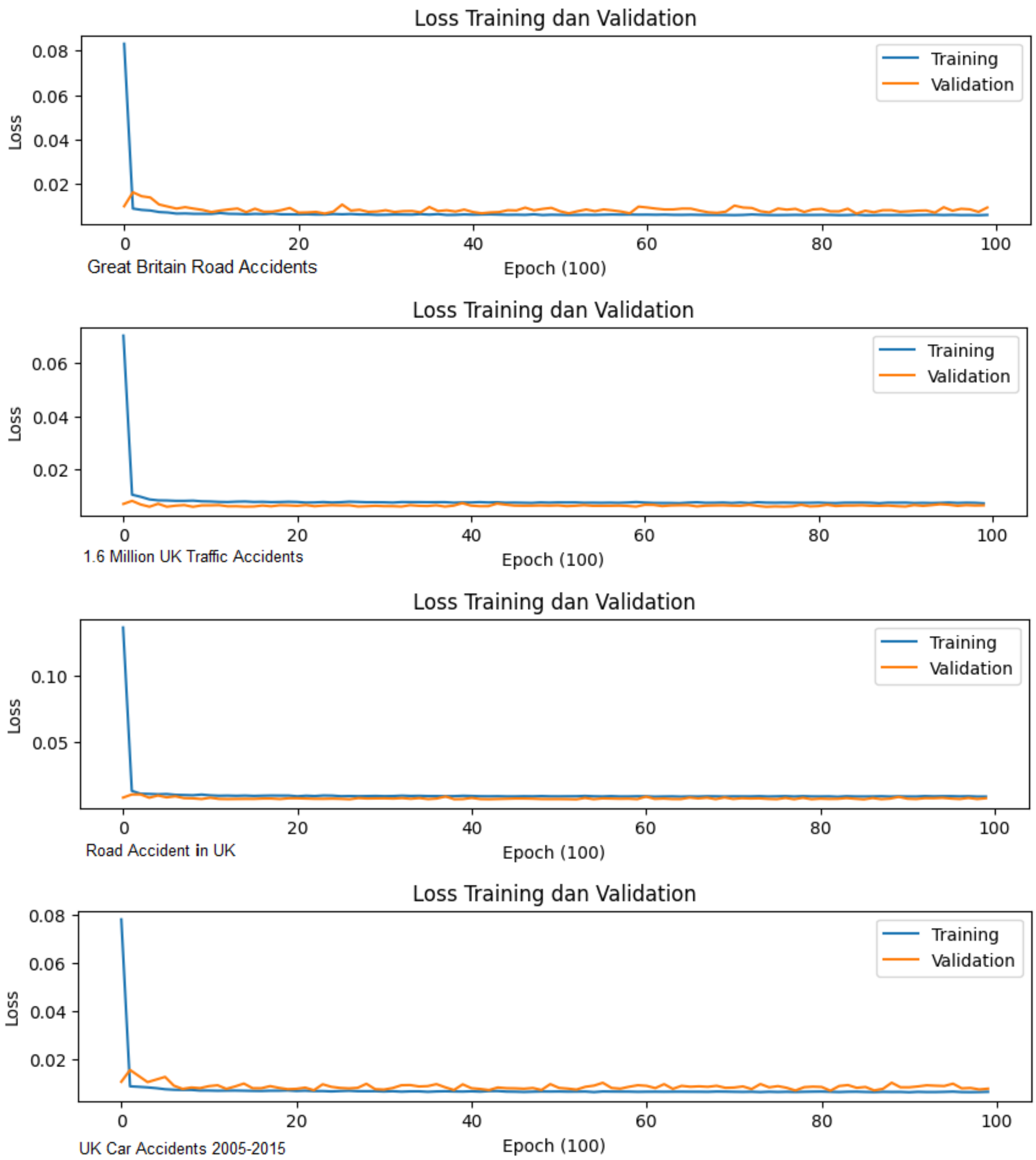
No	Dataset	Non hidden layer		2 hidden layer		4 hidden layer	
		MSLE	Time	MSLE	Time	MSLE	Time
1	Great Britain Road Accidents[26]	0.036361	00:00:44	0.036238	00:01:29	0.036598	00:01:34
2	1.6 Million UK Traffic Accidents[27]	0.036530	00:00:43	0.037271	00:01:08	0.035779	00:01:34
3	Road Accident in UK[28]	0.029293	00:00:31	0.029439	00:00:52	0.029099	00:01:34
4	UK Car Accidents 2005-2015[29]	0.034891	00:00:45	0.035170	00:01:09	0.034363	00:01:34
5	Thailand Fatal Road Accident 2011-2022[30]	0.038671	00:01:19	0.040422	00:02:30	0.038582	00:02:35
	Average	0.035149	00:00:48	0.035708	00:01:26	0.034884	00:01:58

The time required will increase as the number of hidden layers increases, but that does not apply to the MSLE value. Comparison of MSLE 5 dataset values for hidden layer variations greatly influences the evaluation results. The 4 hidden layer variation has the longest time requirement, but has a better value than the other hidden layer variations in the 5 datasets. Further investigation to achieve the best accuracy with hidden layer variations to train the model. The selection of the hidden layer variation is an important factor for training the model. The suitability of the data, the need for predictions, variations in the hidden layer are important for the design of the right LSTM model. The best MSLE is with an average value of 0.034884 with an average time of 00:01:58 with 4 hidden layers, overall the 5 datasets have improved. Table 1). Loss and loss validation with a total of 4 hidden layers in 5 datasets are presented and analyzed. The best MSLE is in the LSTM model with 4 hidden layers.

Loss graphs and loss validation are used for model performance to do deep learning during training and evaluation of validation data. Loss and loss validation on 5 datasets show the same movement. The resulting loss value is reflected in each epoch during training. Loss in validation data measures the performance of the model making accurate predictions on data that has never been seen before. Loss at the beginning of training will differ greatly because the model does not yet understand complex patterns in the data. As the

The LSTM model with 2 hidden layers is the worst because it produces the highest average MSLE value and the time is not far from the 4 hidden layers (

epoch goes on, we hope that the loss on these 5 datasets will decrease. The model gradually learns the patterns in the data and is able to make better predictions. Loss and loss validation on 5 datasets for predicting the number of road accidents provide important information for monitoring model training, optimizing performance, and preventing overfitting. Loss and loss validation show a steady decreasing trend in 4 hidden layers in 5 datasets (Figure 3). The number of accidents every day varies greatly, which is influenced by working days, holidays, and various events that are held. The validation data shows various variations in each model. Training is consistently always close to validation to be an accurate picture of the performance of the model built.



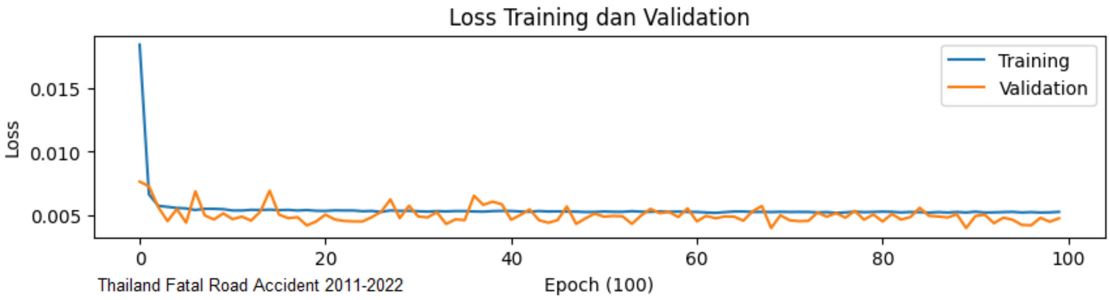
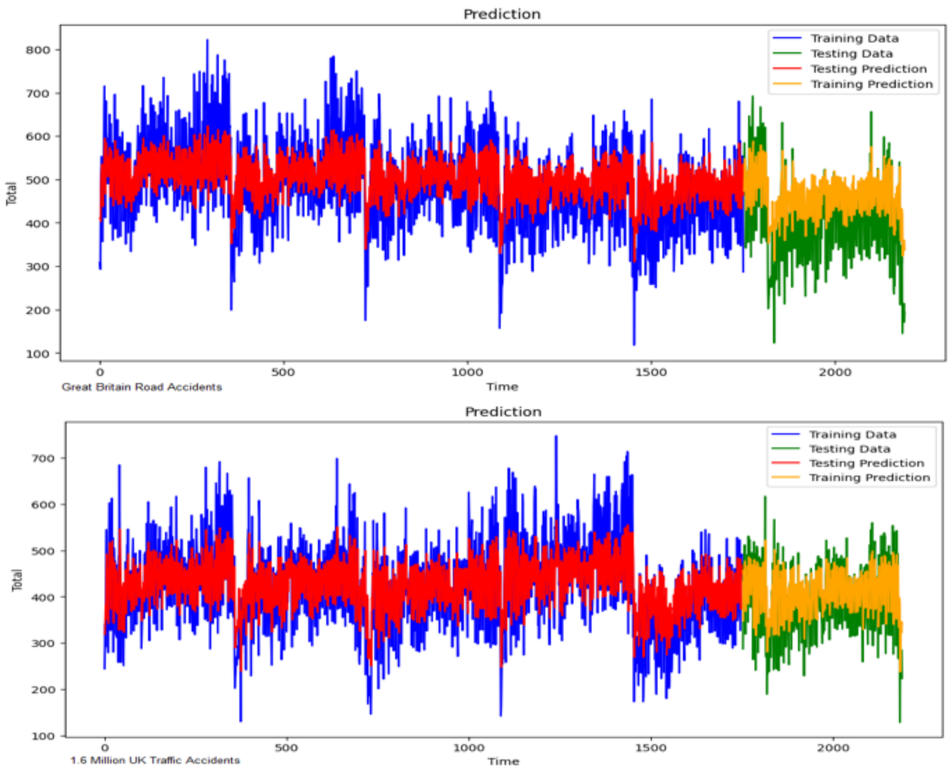


Figure 3. Output Epochs 5 Dataset With LSTM Model

The hidden layer in LSTM is enabled to retrieve information from the previous input and produce output that can be used for prediction or input for the next time step. The LSTM hidden layer consists of a sequentially connected collection of LSTM cells. Each LSTM cell in the hidden layer receives input from the LSTM cell in the previous time step and produces output in the current time step. The LSTM hidden layer produces output that comes from each LSTM cell at a certain time step. The hidden layer in the LSTM plays an important role in understanding and processing information along the input sequence. The LSTM deep learning model with 4 hidden layers that has been built learns from the 5 datasets used. Learning is done with training and testing data taken from the dataset with the distribution of 80% training data and 20% testing data. This division also occurs to make predictions on testing and training. In general, the movement of testing and training data is in accordance with the predicted results using the model that has been built. The more training that is done, the better the prediction results on the training and testing data (Figure 4). The movement in the number of road accident predictions is getting better as an indication that the deep learning algorithm is carrying out deeper learning based on long and short term time. This learning supports the adoption of a more reliable LSTM model in long and short term modeling.



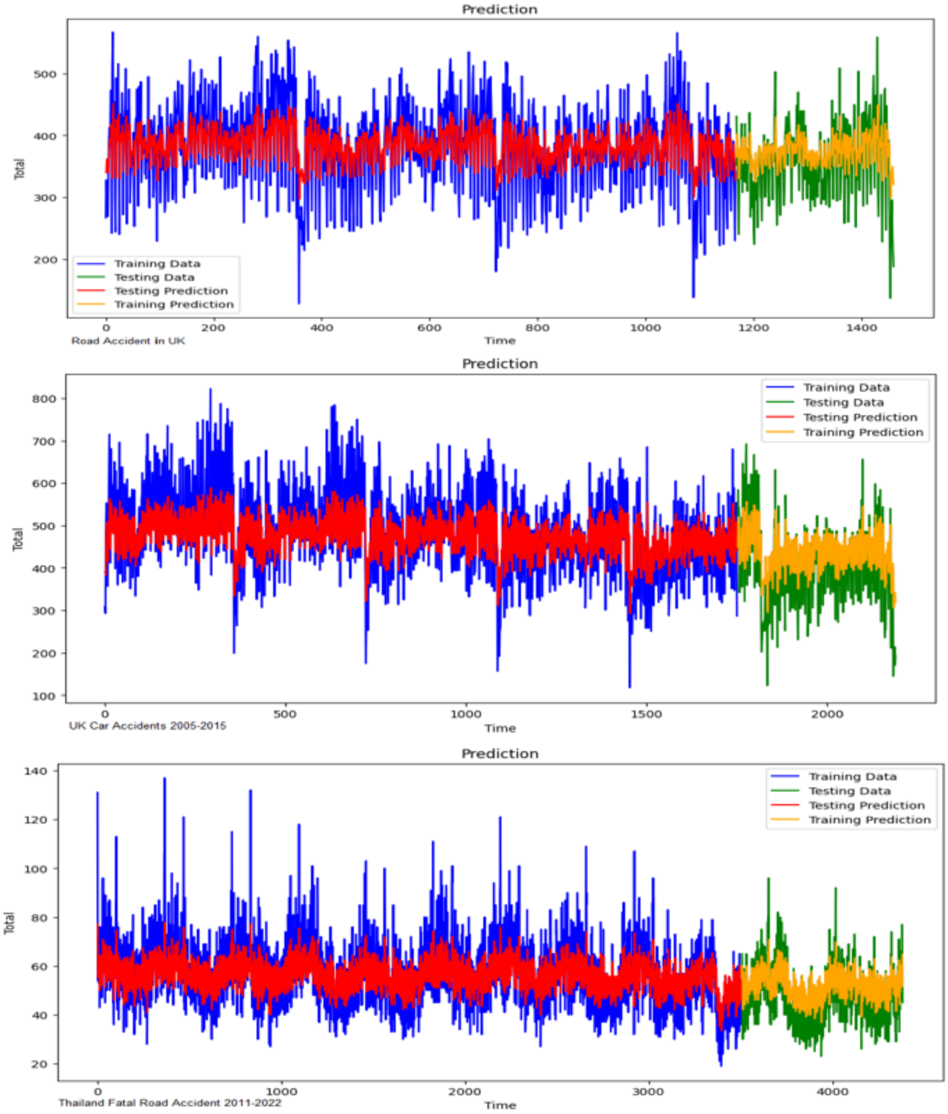


Figure 4. LSTM Output With 5 Datasets

4. Conclusion

Predictions with deep learning-based time series data for the number of road accidents need to be selected according to the data and needs. The LSTM model with 4 hidden layers has an accuracy that can outperform the LSTM model with variants without hidden layers and 2 hidden layers. The selection of the model is based on the variation of the hidden layer which is searched for on average and tested to produce the appropriate accuracy value. The measure of accuracy is obtained from the lowest MSLE results with 3 hidden layer variations. The LSTM model with 4 hidden layers built can be used to predict the number of road accidents according to the desired time requirements in a stable manner. The results of the prediction model can be used as a reference for predicting the number of road accidents by stakeholders. Stakeholders can have an overview of the number of road accidents from the predicted results, so they can plan alternative ways to reduce the number of accidents. Comparison of several types of appropriate evaluation methods can be carried out to test the stability of the proposed model in further work.

Reference

[1] S. Szénási, "Analysis of historical road accident data supporting autonomous vehicle control strategies," PeerJ Computer Science, vol. 7, p. e399, Feb. 2021, doi: 10.7717/peerj-cs.399.

[2] K. Khan, S. B. Zaidi, and A. Ali, "Evaluating the Nature of Distractive Driving Factors towards Road Traffic Accident," Civil Engineering Journal, vol. 6, no. 8, pp. 1555–1580, Aug. 2020, doi: 10.28991/cej-2020-03091567.

[3] T. Mphela, "Causes of road accidents in Botswana: An econometric model," Journal of Transport and Supply Chain Management, vol. 14, Sep. 2020, doi: 10.4102/jtscm.v14i0.509.

[4] J. Mesquitela, L. B. Elvas, J. C. Ferreira, and L. Nunes, "Data Analytics Process over Road Accidents Data—A Case Study of Lisbon City," ISPRS International Journal of Geo-Information, vol. 11, no. 2, p. 143, Feb. 2022, doi: 10.3390/ijgi11020143.

[5] S. Shahsavari et al., "Analysis of injuries and deaths from road traffic accidents in Iran: bivariate regression approach," BMC Emergency Medicine, vol. 22, no. 1, p. 130, Dec. 2022, doi: 10.1186/s12873-022-00686-6.

[6] H. Himanshi, “AN ANALYSIS OF ROAD ACCIDENTS IN INDIA,” INDIAN JOURNAL OF APPLIED RESEARCH, 2020, doi: 10.36106/ijar/4219296.

[7] S. Wang, J. Zhao, C. Shao, C. Dong, and C. Yin, “Truck Traffic Flow Prediction Based on LSTM and GRU Methods With Sampled GPS Data,” IEEE Access, vol. 8, pp. 208158–208169, 2020, doi: 10.1109/ACCESS.2020.3038788.

[8] M. Iqbal et al., “COVID-19 Patient Count Prediction Using LSTM,” IEEE Transactions on Computational Social Systems, vol. 8, no. 4, 2021, doi: 10.1109/TCSS.2021.3056769.

[9] J. Deng, L. Lu, and S. Qiu, “Software defect prediction via LSTM,” IET Software, vol. 14, no. 4, pp. 443–450, Aug. 2020, doi: 10.1049/iet-sen.2019.0149.

[10] T. Xayasouk, H. Lee, and G. Lee, “Air Pollution Prediction Using Long Short-Term Memory (LSTM) and Deep Autoencoder (DAE) Models,” Sustainability, vol. 12, no. 6, p. 2570, Mar. 2020, doi: 10.3390/su12062570.

[11] Y. Chen, “Voltages prediction algorithm based on LSTM recurrent neural network,” Optik, vol. 220, p. 164869, Oct. 2020, doi: 10.1016/j.ijleo.2020.164869.

[12] D. Xia et al., “A distributed WND-LSTM model on MapReduce for short-term traffic flow prediction,” Neural Computing and Applications, vol. 33, no. 7, pp. 2393–2410, 2021, doi: 10.1007/s00521-020-05076-2.

[13] 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.

[14] 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,” J Water Resour Plan Manag, vol. 146, no. 9, Sep. 2020, doi: 10.1061/(ASCE)WR.1943-5452.0001276.

[15] A. Behura and A. Behura, “Road Accident Prediction and Feature Analysis By Using Deep Learning,” in 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, Jul. 2020, pp. 1–7. doi: 10.1109/ICCCNT49239.2020.9225336.

[16] T. Vaiyapuri and M. Gupta, “Traffic accident severity prediction and cognitive analysis using deep learning,” Soft Computing, Nov. 2021, doi: 10.1007/s00500-021-06515-5.

[17] Z. Ma, G. Mei, and S. Cuomo, “An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors,” Accident Analysis & Prevention, vol. 160, p. 106322, Sep. 2021, doi: 10.1016/j.aap.2021.106322.

[18] C. Gutierrez-Osorio, F. A. González, and C. A. Pedraza, “Deep Learning Ensemble Model for the Prediction of Traffic Accidents Using Social Media Data,” Computers, vol. 11, no. 9, p. 126, Aug. 2022, doi: 10.3390/computers11090126.

[19] A. Azhar et al., “Detection and prediction of traffic accidents using deep learning techniques,” Cluster Computing, vol. 26, no. 1, pp. 477–493, Feb. 2023, doi: 10.1007/s10586-021-03502-1.

[20] K. R. Ballamudi, “Road Accident Analysis and Prediction using Machine Learning Algorithmic Approaches,” Asian Journal of Humanity, Art and Literature, vol. 6, no. 2, pp. 185–192, Dec. 2019, doi: 10.18034/ajhal.v6i2.529.

[21] S. Mićić, R. Vujadinović, G. Amidžić, M. Damjanović, and B. Matović, “Accident Frequency Prediction Model for Flat Rural Roads in Serbia,” Sustainability, vol. 14, no. 13, p. 7704, Jun. 2022, doi: 10.3390/su14137704.

[22] A. Borucka, E. Kozłowski, P. Oleszczuk, and A. Świderski, “Predictive analysis of the impact of the time of day on road accidents in Poland,” Open Engineering, vol. 11, no. 1, pp. 142–150, Dec. 2020, doi: 10.1515/eng-2021-0017.

[23] Y. Boo and Y. Choi, “Comparison of mortality prediction models for road traffic accidents: an ensemble technique for imbalanced data,” BMC Public Health, vol. 22, no. 1, p. 1476, Dec. 2022, doi: 10.1186/s12889-022-13719-3.

[24] X. Jin, J. Zheng, and X. Geng, “Prediction of Road Traffic Accidents Based on Grey System Theory and Grey Markov Model,” International Journal of Safety and Security Engineering, vol. 10, no. 2, pp. 263–268, Apr. 2020, doi: 10.18280/ijsse.100214.

[25] R. PEČELIŪNAS, V. ŽURAULIS, P. DROŹDZIEL, and S. PUKALSKAS, “Prediction of Road Accident Risk for Vehicle Fleet Based on Statistically Processed Tire Wear Model,” Promet, vol. 34, no. 4, Jul. 2022, doi: 10.7307/ptt.v34i4.3997.

[26] T. Lei, “Great Britain Road Accidents 2005_2016.” Accessed: Jun. 20, 2023. [Online]. Available: https://www.kaggle.com/datasets/nichaoku/gbaccident0516

[27] D. Fisher-Hickey, “1.6 million UK traffic accidents.” Accessed: Jun. 20, 2023. [Online]. Available: https://www.kaggle.com/datasets/daveianhickey/2000-16-traffic-flow-england-scotland-wales

[28] S. Jalalian, “Road Accident in UK.” Accessed: Jun. 20, 2023. [Online]. Available: https://www.kaggle.com/datasets/sadeghjalalian/road-accident-in-uk

[29] silicon99, “UK Car Accidents 2005-2015.” Accessed: Jun. 20, 2023. [Online]. Available: https://www.kaggle.com/datasets/silicon99/dft-accident-data

[30] T. R, “Thailand Fatal Road Accident [2011-2022].” Accessed: May 10, 2023. [Online]. Available: https://www.kaggle.com/datasets/thaweewatboy/thailand-fatal-road-accident

[31] 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.

[32] 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.

[33] 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.

[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] 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,” J Clean Prod, vol. 250, p. 119476, Mar. 2020, doi: 10.1016/j.jclepro.2019.119476.

[36] 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,” Sci Rep, vol. 11, no. 1, p. 17497, Sep. 2021, doi: 10.1038/s41598-021-96751-4.

[37] 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.

[38] 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.

[39] 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.

[40] M. S. Devi, R. M. Mathew, and R. Suguna, “Regressor Fitting Of Feature Importance For Customer Segment Prediction With Ensembling Schemes Using Machine Learning,” Int J

Eng Adv Technol, vol. 8, no. 6, pp. 952–956, Aug. 2019, doi: 10.35940/ijeat.F8255.088619.

[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.