J. Intell. Comput. Health Inform. Vol. 3, No. 2, 2022
ISSN: 2715-6923 e-ISSN: 2721-9186, pp.23~29 | DOI: 10.26714/jichi.v3i2.11168.

Comparison of Deep Neural Network Architectural Models for Predicting Tourist Visits to Bali during the Pandemic Period

I Nyoman Purnama

Department of Information System, STMIK PRIMAKARA, Jl.Tukad Badung No. 135 Denpasar Bali
*Corresponding author: purnama@primakara.ac.id

Article histroy :
Received: 29 DES 2022
Accepted: 14 JAN 2023
Avalaible online: 22 FEB 2023

Research article

Abstract: Tourism is a vital component of any economy, providing a source of income for the community. Bali, one of the provinces in Indonesia, has tremendous potential in the tourism industry, with most people employed in this sector. However, fluctuations in the number of tourist visits can pose challenges when devising policies to overcome issues in the field of tourism. Hence, forecasting is necessary to predict post-pandemic tourist arrival patterns to ensure a smooth tourism recovery process. Forecasting is an essential tool that aids in making sound decisions. In this study, we utilized three forecasting methods: Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). We adopted a comparative approach using these three deep neural network architectures to predict tourist visits to Bali during a pandemic. We tested the architectural models with datasets sourced from Badan Pusat Statistik (BPS) and evaluated the model's performance using RMSE and MAE. The results indicated that the LSTM model outperformed the CNN and GRU models, with an RMSE value of 329036 and MAE value of 285874. Based on the study, we can conclude that the LSTM model has better performance and can predict tourist arrivals in Bali with reasonable accuracy.

Keywords: CNN; GRU; LSTM; PREDICTIONS; TOURISM

1. Introduction

Tourism has been designated as the primary sector contributing to the second-largest foreign exchange in the country, after the palm oil industry (Surtiningsih et al., 2018). Therefore, tourism plays a vital role in providing income to the people. Bali, one of the provinces in Indonesia, has great potential in tourism due to its various attractions, including natural and cultural sites. It is not surprising that Bali has been chosen as one of the most beautiful tourist destinations globally, as the number of tourist visits keeps increasing each year. However, along with the economic benefits of tourism, come problems such as uneven development of culture and tourism, lack of infrastructure, transportation, and accommodation. Additionally, the COVID-19 pandemic has caused a significant decline in tourist visits, leading to a paralyzed economy that had previously grown rapidly. According to data on tourist arrivals in May 2021 from the Badan Pusat Statistik website, only eight people passed through Bali's Ngurah Rai Airport, whereas in 2022, it increased to 115,553. This significant increase resulted in a substantial multiplier effect on the economy of Bali, which is highly dependent on tourism.

When making policies to address tourism-related problems, fluctuations in tourist visits can often pose constraints. To ensure that the tourism recovery process runs smoothly, it is essential to have an overview of post-pandemic tourist arrival patterns. Forecasting is an important planning tool that helps make informed decisions (Sukraini, 2017). It enables one to determine when an event will occur and take necessary action to prevent it. Forecasting heavily relies on past data that is collected regularly, such as daily, monthly, or yearly. Accurate and precise results can be obtained after analyzing the data with the right forecasting technique. Time series data, including light period data, can be used for forecasting (Oxaichiko Arissinta et al., 2022). Time series forecasting predicts future events based on past data and machine learning algorithms.

Several studies have been conducted on forecasting tourist visits, such as forecasting the number of tourist visits to Bali using Support Vector Regression and genetic algorithms (Surtiningsih et al., 2018). The predictions are made on time series data. Optimization of the important parameter values of the SVR results in good prediction outcomes. The GA-SVR method is suitable for short-term predictions, with a MAPE value obtained of 2.513% for foreign tourist visits to Bali. Another study by Wenjie Lu et al. focused on predicting stock prices using the CNN and LSTM algorithms on time series data. The study used MLP, CNN, RNN, LSTM, and CNN-RNN methods, and the results showed that the CNN-LSTM method provided price predictions at a very high level of accuracy, with MAE values obtained of 27,564 and RMSE 39,688.

Based on the aforementioned problems, we will predict the number of tourists visiting Bali during the pandemic using the CNN, LSTM, and GRU methods. In this study, we will compare the three methods by predicting test data using RMSE and MAE. Furthermore, based on previous research on small time series data, we anticipate that the resulting error rate will be minimal.

2. Methods

2.1 Dataset

In this study, we utilized a dataset of tourist arrivals in Bali that was sourced from BPS. The data can be accessed by the public through the following link: https://bali.bps.go.id/subject/16/pariwisata.html. The dataset includes 54 time series data from 2018 to 2022 and contains two attributes: the date attribute and the number of tourists attribute. The date attribute represents the cumulative number of tourist visits per month, while the tourist attribute denotes the number of tourist visits per month. Table 1 displays detailed tourist visit data in Bali.

This data is separated into training data and test data in a ratio of 80:20. Before being used in each method, the data is processed by normalization to improve prediction accuracy (Chen et al., 2015). Normalization aims to prevent too many differences in values. However, due to the pandemic, the data has significantly different values. For this study, the Min-Max scale method was used for the normalization process (Khalis Sofi et al., 2021).

Theorem 1: Equations of Min-Max scale:

$$Xn = \frac{Xo - Xn}{Xmax - Xmin} \tag{1}$$

Theorem 1 outlines the min-max scaling method, in which the value of xn is the result of normalization. x0 represents the data that needs to be normalized, while Xmin and Xmax indicate the minimum and maximum values of the time series data being used. This data will serve as the input feature for the CNN, GRU, and LSTM methods, with four data points as input and one additional data point as output.

2.2 Architecture of a deep neural network

A Convolutional Neural Network (CNN) is a type of deep learning that is effective for predicting time series applications (Hamoudi & Elseifi, 2018). In the built model, the input for Long Short Term Memory (LSTM) comes from the output of the CNN. CNN is a neural network typically used for identifying patterns in images and is widely used in feature processing (Lu et al., 2020). Meanwhile, LSTM has the ability to expand based on the time sequence and is widely used in time series data. The architecture of the CNN consists of the input layer, 1D convolutional layer, pooling layer, hidden layer, and fully connected layer. CNN is comprised of two parts, the convolution layer and the pooling layer, and has shown good performance in image processing and Natural Language Processing.

Table 1.Dataset.

DATE	NUMBER OF VISITS
1-2018	743456
2-2018	655719
3-2018	762622
4-2018	777287
5-2018	682521
6-2018	1156151
7-2018	906347
8-2018	770364
9-2018	774144
10-2018	762124
11-2018	806397
12-2018	960859
1-2019	793527
2-2019	692113
3-2019	787616
4-2019	795997
5-2019	656082
6-2019	1287877
7-2019	935930
8-2019	925360
9-2019	812003
10-2019	853007
11-2019	852626
12-2019	1152901
1-2020	879702
2-2020	721105
3-2020	567452
4-2020	175120
5-2020	101948
6-2020	137395
7-2020	229112
8-2020	355732
9-2020	283349
10-2020	337304
11-2020	425097
12-2020	382841
1-2021	282248
2-2021	240608
3-2021	305579
4-2021	330593
5-2021	363959
6-2021	498852
7-2021	166718
8-2021	202187
9-2021	298950
10-2021	468826
11-2021	513482
12-2021	629590
1-2022	527447
2-2022	389690
3-2022	547726
4-2022	500740
5-2022	960692
6-2022	753907

CNN, LSTM, and GRU are hybrid algorithms capable of extracting existing knowledge from time series data, with LSTM being particularly effective for identifying short-term and long-term dependencies (Lu et al., 2020). The Adaptive Moment Estimation (ADAM) optimization method is used for CNN because of its ease of implementation, good computational efficiency, and low memory consumption. The effectiveness of the ADAM method has been proven to produce low costs during the training process (Kingma & Ba, 2014). Similarly, ADAM can also be used in LSTM and GRU to optimize loss and assist in determining the epoch, which is typically done manually.

2.3 LSTM architecture

Long Short-Term Memory (LSTM) is one of the developments of the RNN (Recurrent Neural Network) algorithm. The LSTM algorithm (see in Fig. 1) is a solution to the vanishing gradient problem found in conventional RNNs. In LSTM there are three gates that control the use and updating of previous text information, namely the input gate, forget gate, and output gate The memory cell and three gates are designed to be able to read, store, and update past information (Ashari & Sadikin, 2020).

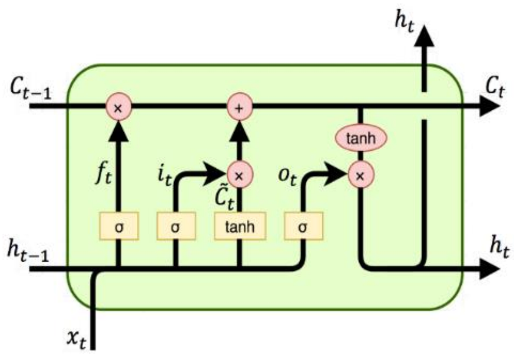

Fig 1. LSTM architecture (Zhang et al., n.d.).

2.4 GRU architecture

The Gated Recurrent Unit (GRU) is a development of Long Short Term Memory (LSTM). GRU has two information control components, namely the reset gate and the update gate, can be seen in Fig. 2. The reset gate determines how to combine new input with past information, while the update gate regulates how much information will be stored (Muliani Harahap & Fitrie, 2021).

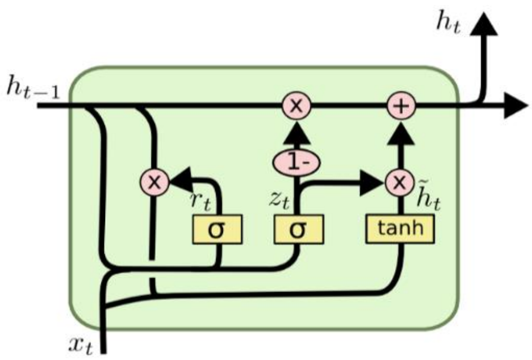

Fig 2. GRU architecture (Arissintaa et al., 2022).

2.5 CNN architecture

The architecture of the Convolutional Neural Network (CNN) comprises two Feature Extraction Layers and a Fully Connected Layer (Juliansyah et al., 2021). The feature extraction section involves two processes: convolution and pooling. During convolution, a filter is used to shift the input feature, which is a dataset of tourist visits. Following that, a pooling process is implemented to continuously reduce dimensions and minimize the number of parameters and computations in the network. This study employs max pooling, which selects the maximum value in each block. The classification section includes a fully connected layer that is fully connected to the previous layer. The CNN architecture can be seen in Fig. 3.

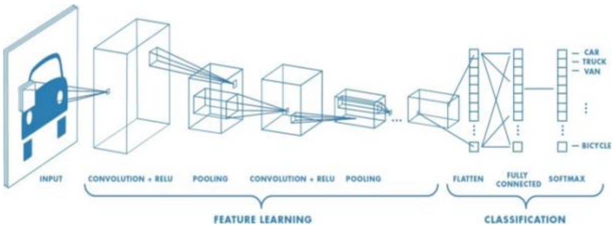

Fig 3. CNN architecture (Septiana et al., 2020).

2.6 Training and testing

The training process involves the use of three algorithms, namely CNN, LSTM, and GRU, each of which employs ReLu activation. Optimization is performed using ADAM. The training process is conducted over 100 epochs, and the results of this process will be utilized in the subsequent testing phase. The testing phase utilizes data that has been partitioned between the training and testing phases, and the outcome of this process is the prediction of the number of tourist visits. A flowchart outlining all the training and testing processes for predicting tourist visits is provided below.

2.7 Evaluation process

The developed system is evaluated using the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) methods. MAE and RMSE are commonly used for prediction models. A smaller value for these methods indicates better performance of the prediction model. The formulas for RMSE and MAE are stated in Eq. (2) and (3).

Theorem 2: Equations RMSE:

$$RMSE = \sqrt{\sum \frac{(Y' - Y)^2}{n}} \tag{2}$$

Theorem 3: Equations MAE:

$$MAE = \sum \frac{|Y' - Y|}{n} \tag{3}$$

From Eq. (2) and (3), Y' represents the prediction value, Y represents the actual value, and n is the number of datasets. Overall, the prediction process utilizing CNN, GRU, and LSTM methods can be described as follows:

- Preprocess the data to be used as input for CNN, GRU, and LSTM training using the min-max scale method.
- Change the composition of the training and test data with an 80:20 ratio.
- Initialize the CNN, LSTM, and GRU networks with optimal parameters.
- Conduct the training process with the CNN, LSTM, and GRU methods.
- Evaluate the prediction results using the MAE and RMSE methods.
- Save the training result model for later use in the testing process.
- Conduct the testing process using the previously trained models and test data

2.8 Proposed model

In this study, we propose a comparative prediction model approach using three deep neural network architectures to determine the best model for predicting tourist visits to Bali. The proposed model is illustrated in Fig. 4."

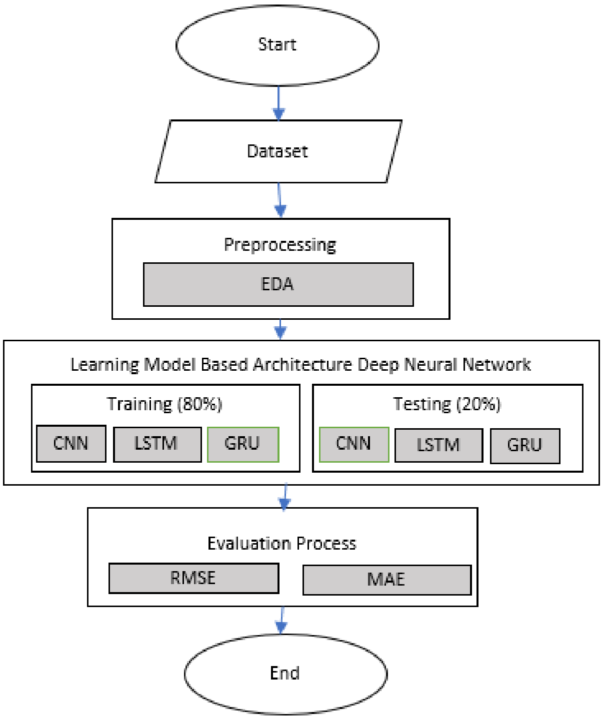

Fig 4. Block diagram of the porposed model.

3. Results and Discussions

After completing the design process, the next step is implementation. The implementation process is divided into several stages, which are as follows:

3.1 Exploratory Data Analysis (EDA)

The data used in this study is in the form of time series data, consisting of columns for dates and the number of tourist visits to Bali. The data is sourced from the Badan Pusat Statistik (BPS) website, based on tourist visits through I Gusti Ngurah Rai Airport from 2018-2022, and includes data on tourist visits during the Covid-19 period when movements were uncertain. The development environment used for this study was a single PC with an Intel Core i7 processor, 8GB RAM, and a VGA GeForce 1080. Python was used as the programming language.

During the exploratory data analysis (EDA) process, the csv file containing the data was uploaded to the Colabs environment. The dataset was then split into training data and test data, using an 80:20 composition. To aid in the prediction process, several libraries were used, including Pandas, Numpy, Sklearn, and Keras.

After processing the data, it was obtained in the form of a single column containing a total of 54 rows. Using the mathplotlib library, the data was then plotted in graphical form, can be shown in Fig. 5.

There was a drastic decrease in tourist arrivals when the Covid-19 cases reached their peak. Before using the data in the training process with CNN, GRU, and LSTM, it is first processed to obtain time-series data with a total of 54 points. To use this data in the supervised learning process, it must first be converted into sequential data. To do this, first determine the length of the sequence to be used for the conversion. For example, if a sequential length of 3 is used, then the sequence would be t−1, t−2, and t−3, and would be used to predict the value of t. Table 2 shows an example of the initial data that will be used as input for each algorithm.

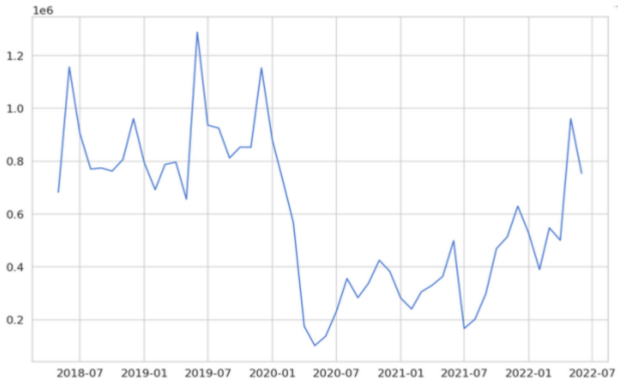

Fig 5. Graphical tourist visit.

Table 2.Example of sequential data.

SEQUENCE DATA	VALUE
t-1	743456
t-2	655719
t-3	762622
t	777287

With a sequence length of 4, 53 pieces of data were obtained, which will be used in the learning process. Afterward, we split this data using an 80:20 combination of training and testing data, resulting in 34 training data and 17 test data, as previously explained. Before being inputted into the CNN-LSTM network, this data is first normalized using the MinMax scaling method, which transforms the data to have values between 0 and 1. All data, including both the training and test data, is normalized to ensure that it has a wide range of values that will be used in the evaluation process later. An example of the normalized data results can be seen in Table 3.

Table 3. Example of sequential normalized data.

SEQUENCE DATA	VALUES
t-1	0.42285784
t-2	0.37828767
t-3	0.78107537
t	0.55982102

To be used as input for the CNN, LSTM, and GRU processes, this data needs to be processed into a 3D format comprising the number of samples, number of steps, and number of features. In this study, the number of steps is 4, indicating that the model will make predictions based on the last four months.

3.2 Learning model-based architecture deep neural network

After normalizing the data to be used in the training process, proceed with entering the data to be used in the training processes for the CNN, LSTM, and GRU. During the first iteration, the network parameter initialization process is carried out to determine the weights and bias values at the beginning of the CNN, GRU, and LSTM layers. The stopping condition for training is determined by the epoch number. If the stopping condition has not been reached, weight and bias adjustments will be made until the stopping condition is met. The resulting model from the training process is used in the testing process with test data. The test results are then compared with the actual visit data to produce predictive results. The accuracy value, calculated using RMSE and MAE values, is the deviation result.

Each of these algorithm models needs to be configured to produce an optimal output value. The configuration used in Table 4 is provided below.

Table 4. Algorithm configuration.

PARAMETERS	VALUES
Convolutional activation function	Relu
Hidden layer on LSTM, GRU, CNN	64
Activation function layer LSTM, GRU, CNN	tanh
Learning rate	0.01
Optimization	Adam
Epoch	100

As previously mentioned, the CNN, GRU, and LSTM models were executed using the Python programming development environment and its accompanying libraries. This study comprises of three main modules: the prediction module, the CNN-LSTM-GRU model development module, and the evaluation module. In the model building module, the input is designated as "x," which represents the dataset that will be utilized for predictive input, while "y" represents the target predictive value. The CNN, LSTM, and GRU models are the output of the model building process and will be employed for the prediction process using test data. While the CNN model can process input in 1D, 2D, and 3D sizes, it is better suited for processing time series data using the 1D CNN model. For the prediction module implementation, "X" is designated as the collection of data that will be used as input for the prediction process along with the trained CNN, LSTM, and GRU models. The predicted result, designated as "y_pred," can be observed in Fig. 5.

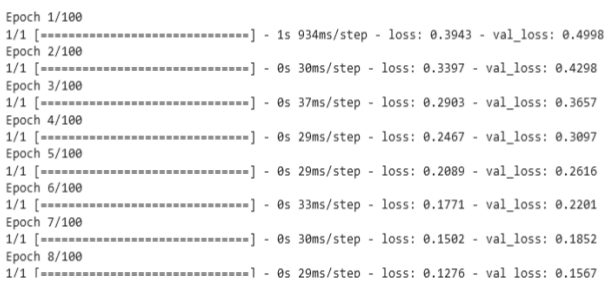

Fig 5. Training process.

The modeling task in this research was performed using the TensorFlow library. The training process consisted of several epochs of 100 for each method used. The prediction process will utilize input test data that accounts for 20% of the total data. After training the CNN, LSTM, and GRU models with the training data (which accounts for 80% of the total 54 tourist visit data between January 1, 2018, and June 6, 2022), performance evaluation was conducted using the test data. The remaining 20% of the data was utilized for this purpose.

Given that the time series data processed falls within the Covid-19 pandemic period, during which tourist visits decreased significantly in March, the CNN, LSTM, and GRU models still have the capability to predict the time series data of tourist visits with good performance. The following presents a comparison of the validation loss and training loss for each method, with the validation and training loss using CNN, LSTM, and GRU methods can be seen in Fig. 6.

In GRU and LSTM-based models, there is overfitting when the predictions on the training data are very accurate, but the predictions on the testing data are poor. Overfitting does not generalize well, so if the model is tested using different data, it can lead to reduced accuracy. The following presents a comparison of the predicted and actual results using the CNN, LSTM, and GRU methods can be seen in Fig. 7.

After the experimental process has been carried out, the RMSE and MAE values generated for each method can be seen in Table 5.

Table 5. Performance results.

MEASUREMENT	LSTM	GRU	CNN
RMSE	329036	338026	337403
MAE	285874	290803	285889

Based on the results presented in the table above, it can be concluded that the LSTM method performs better than the other two methods. With a minimal amount of data, LSTM is able to make accurate predictions. However, as shown in the graph of the loss, all three methods still experience overfitting. Therefore, future research should focus on generating more training data to improve accuracy.

4. Conclusion

The training model used in this study employed the CNN, LSTM, and GRU algorithms to compare their performance in predicting tourist visits to Bali from 2018-2022, represented as a time series data. A comparative approach was used, employing three deep neural network architectures to predict tourist visits during a pandemic. The model's architecture was tested with datasets obtained from Badan Pusat Statistik (BPS), and its performance was evaluated using RMSE and MAE. The results of the study revealed that the LSTM model outperformed the CNN and GRU models, with an RMSE value of 329036 and an MAE value of 285874. Consequently, the study concludes that the LSTM model offers better performance and accurately predicts tourist arrivals in Bali.

For future research, we aim to explore other relevant approaches for data analysis and data extraction, as these have been shown to significantly improve model performance.

ISSN: 2715-6923 e-ISSN: 2721-9186, pp.23~29 | DOI: 10.26714/jichi.v3i2.11168.

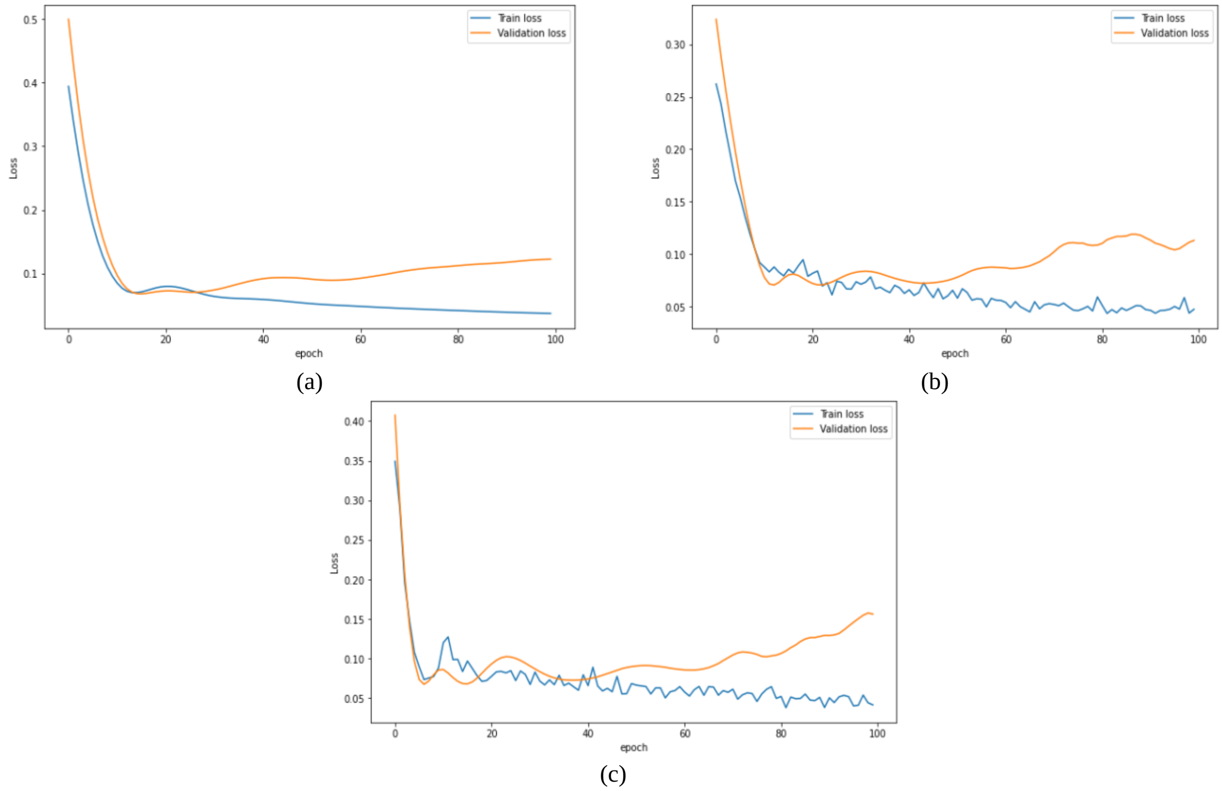

Fig 6. (a) CNN, (b) LSTM, and (c) GRU train and validation loss.

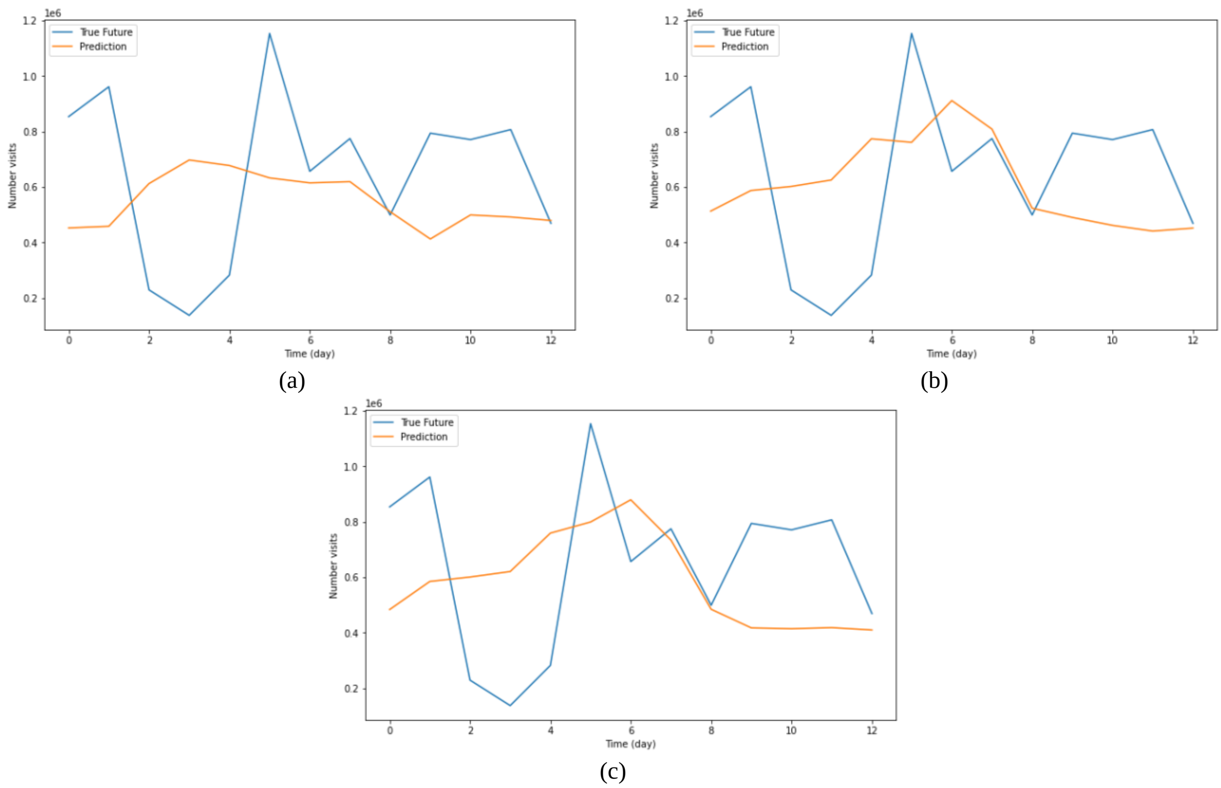

Fig 7. (a) CNN, (b) LSTM and (c) GRU prediction result.

References

Arissintaa, I. O., Sulistiyawatib, I. D., & Kharisudin, D. K. I. (2022). Pemodelan Time Series untuk Peramalan Web Traffic Menggunakan Algoritma Arima, LSTM, dan GRU. 5, 693–700.

Ashari, M. L., & Sadikin, M. (2020). Prediksi Data Transaksi Penjualan Time Series Menggunakan Regresi Lstm. Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI), 9(1), 1. https://doi.org/10.23887/janapati.v9i1.19140

Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. 2015 IEEE International Conference on Big Data (Big Data), 2823–2824. https://doi.org/10.1109/BigData.2015.7364089

Hamoudi, H., & Elseifi, M. A. (2018). Stock Market Prediction using CNN and LSTM.

Juliansyah, S., Laksito, A. D., 허원호, Lionel, D., Adipranata, R., Setyati, E., Naufal, M. F., Shania, S., Millenia, J., Axel, S., Soebroto, J. T., Febrina, P., Mercifia, M., Putra, A. I., Santika, R. R., Hsu, K.-C., Chou, S.-Y., Yang, Y.-H., Chi, T.-S., … Rizal, A. (2021). Klasifikasi Citra Buah Pir Menggunakan Convolutional Neural Networks. Jurnal Infra Petra, 7(1), 489–495. https://doi.org/10.22441/incomtech.v11i1.10185

Khalis Sofi, Aswan Supriyadi Sunge, Sasmitoh Rahmad Riady, & Antika Zahrotul Kamalia. (2021). Perbandingan Algoritma Linear Regression, LSTM, dan GRU dalam Memprediksi Harga Saham Dengan Model Time Series. Seminastika, 3(1), 39–46. https://doi.org/10.47002/seminastika.v3i1.275

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. http://arxiv.org/abs/1412.6980

Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN-LSTM-based model to forecast stock prices. Complexity, 2020. https://doi.org/10.1155/2020/6622927

Muliani Harahap, A., & Fitrie, S. (2021). Implementasi Gated Recurrent Unit (GRU) Untuk Prediksi Harga Saham Bank Konvensional di Indonesia. JISTech (Journal of Islamic Science and Technology) JISTech, 6(2), 42–49. http://jurnal.uinsu.ac.id/index.php/jistech

Oxaichiko Arissinta, I., Dwi Sulistiyawati, I., & Kurnianto Iqbal Kharisudin, D. (2022). Pemodelan Time Series untuk Peramalan Web Traffic Menggunakan Algoritma Arima. Prosiding Seminar Nasional Matematika, 5, 693–700. https://journal.unnes.ac.id/sju/index.php/prisma/

Septiana, T., Puspita, N., Fikih, M. al, & Setyawan, N. (2020). Face Mask Detection Covid-19 Using Convolutional Neural Network (CNN). 27–32.

Sukraini, T. T. (2017). Peramalan Kunjungan Wisatawan Ke Uluwatu Dengan Menggunakan Model Autoregressive Integrated Moving Average. Matrix Jurnal Manajemen Teknologi Dan Informatika, 6(1), 47.

Surtiningsih, L., Furqon, M. T., & Adinugroho, S. (2018). Prediksi Jumlah Kunjungan Wisatawan Mancanegara Ke Bali Menggunakan Support Vector Regression dengan Algoritma Genetika. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 2(8), 2578–2586. http://j-ptiik.ub.ac.id

Zhang, Y., Yuan, H., Wang, J., & Zhang, X. (n.d.). YNU-HPCC at EmoInt-2017: Using a CNN-LSTM Model for Sentiment Intensity Prediction. http://ie.com