International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-1, October, 2019

Time Series Data Prediction using Elman Recurrent Neural Network on Tourist Visits in Tanah Lot Tourism Object

Putu Sugiartawan, Sri Hartati

***Abstract**: The prediction of time series data is a forecast using the analysis of a relationship pattern between what will be predicted (prediction) and the time variable. The prediction process using the recurrent neural network (RNN) model could recognize and learn the data pattern of time series, but the presence of fluctuations in data makes the introduction of data patterns difficult to be learned. The data used for forecasting are tourist visits to Tanah Lot Bali tourist attraction for 10 years (2008-2017). The training process uses the RNN method on high fluctuating data, which requires a relatively long time in recognizing and studying the data patterns. Modification of the RNN method on learning rate and momentum by using dynamic values, can shorten learning time. The results showed the learning time using the RNN dynamic value, smaller than the variants of the RNN method such as the RNN Elman, Jordan RNN, Fully RNN, LSTM and the feedforward method (Backpropagation). The resulting error value is 0,05105 MSE. This value is smaller than the Fully RNN, Jordan RNN, LSTM and Feedforward methods. The elman method has the shortest training time among other models. The purpose of this research is to make a prediction design consisting of sliding windows techniques, training with neural network models and validation of results with k-fold cross-validation.*

***Keywords**: Time series, recurrent neural network, k-fold cross validation, sliding windows, prediction.*

I. INTRODUCTION

Bali Island is one of the tourist destinations which is known globally, one of them is Tanah Lot tourist destination. With the various events held in Bali, it has a positive impact on tourist visits especially Tanah Lot. The Central Statistics Agency noted that tourist arrivals visiting Tanah Lot in January-December 2017 period reached 3.5 million tourists [1], and it has increased from the previous year [2]. However, in the few previous months there was a significant decrease in the number of tourist visits [1], it makes the determination of tourist visits can't be predicted [3]–[7]. One solution that can be used to determine the pattern of tourist visits is to perform a prediction process [8]–[10]. By knowing the prediction pattern of tourist visits, it is expected to increase the knowledge of tourism object management, travel agents and the government to manage a better management of tourism objects [7] - [9]. Prediction of tourist visits is one of the predictions of time series data, there are some sequent variables depend on time [3]. Time series is a collection of sorted data observations in time[7], [11]–[13]. The method of time series is also a forecast method by using an analysis of relationship pattern between variables to be estimated with time variables [14]. There are various approaches in forecasting techniques, including statistical models [15]–[17], neural networks [18] and multivariates [19], [20]. Some approaches above have several advantages and disadvantages, including statistical models have a relatively large error rate, because the model applies the previous data to determine subsequent results[3], [21]. While the neural network model is able to study the overall data pattern and produce knowledge that can predict next some data [20]. The obtained results using neural networks have a relative small error rate when it is compared to other approaches [13], [22], [23]. The neural network method has several different architectures, generally the neural network architecture is divided into 2 namely: feed forward and recurrent neural network (RNN) [19]. The RNN method produces a smaller error value than feed forward, although the learning time is slightly longer than the feed forward nueral network [3], [8], [24]. The RNN network is a modification of feed forward with the main difference on the addition of associated neuron layer which provides a network output pattern to give feedback itself as input in order to produce the next network output. RNN is considered as a partial recurrent network because the connection is generally only in the form of feed forward [25]. This recurrent network has two inputs, namely real input and contextual input. This recurrent network has two inputs, namely: real input and contextual input. There is Feedback that may cause the iteration process faster, thus making the speed of parameter updating and convergence faster. RNN model can recognize fluctuating patterns for a fairly high level of data fluctuation. There were several studies related to the optimization of RNN model in prediction case, including the comparison of meurowavelet model to predict the short-term of stock returns from high-prequency of financial data [26], [27], the use of RNN method in network performance is to measure the river flows [20].

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

Putu Sugiartawan, Informatics Program, STMIK STIKOM Indonesia, Bali, Indonesia. Email: putu.sugiartawan@mail.ugm.ac.id

Sri Hartati*, Department of Computer Science and Electronics Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia. Email: shartati@ugm.ac.id

The combination of ERNN method with Cooperative coevolution method produces a more stable predictive value comparing to the neuroevolution method [28].

The modification of elman neural network produces more accurate prediction results, and the abnormal habits of internet use can be found [29]. This research applied Recurrent Neural Network approach to predict time series data, as well as the need to obtain an appropriate sample rate and optimize the size of the input windows (Windows Size).

The present study is going to predict tourist visits by using sliding windows time series techniques and K-fold validation techniques toward the validation of training results, and using RNN model for the training process. This study also compared several neural network methods, so the comparison of each model toward predicted results can be known.

II. METHODOLOGY

Reducing the learning time of data patterns and increasing the accuracy of the predicted results of tourist visits, it took several approaches and methods. This study applied the recurrent network approach and its optimization in predicting tourist visits. There were several stages in conducting the time series prediction process in this study, including ;

- Step 1 The Normalization of data with vulnerable values between 0 and 1, aiming to reduce the target value (Y).
- Step 2 Validation of the training and testing process by using the k-fold cross-validation technique
- Step 3 Optimize the training process by determining the value of an optimal neural network architecture
- Step 4 The prediction of time series data using sliding windows techniques.
- Step 5 The comparison of the RNN Optimization model with different neural network architectural models.
- Step 6 The normalization of prediction results, to produce true values.

The optimization model testing is done by comparing the results of predictions with different neural network architectures, such as feed forward / back propagation, Elman, Jordan RNN, LSTM and fully RNN. Each subsection described the detailed process of calculating predictions.

This study used Python program with tensorflow, hard and neurolab functions. The research data were the data of tourist visits in Tanah Lot attractions every month. There were 120 vector data used in the prediction calculation process.

A. Architectural prediction model

In Figure 1 described the process of predicting tourist visits using the RNN method, as well as comparing methods with other neural network methods. The initial process begins with preprocessing data through the normalization process. Value x_i which is the initial data consisting of 120 data vectors, the data consists of 2 values, namely: time (time of occurrence) and value (number of tourist visits). The value of x_i after normalizing became x'. The value of x' was then broken down into several by the k-fold cross validation technique, with a value of k = 5.

The training process using the RNN method begins with determining the value of each parameter, through several trials. These parameters include learning rate (α), momentum (μ), number of units in hidden layer and number of hidden layer (*n*). Determination of the optimal RNN architecture was carried out by conducting repeated experiments, to obtain the best value for training.

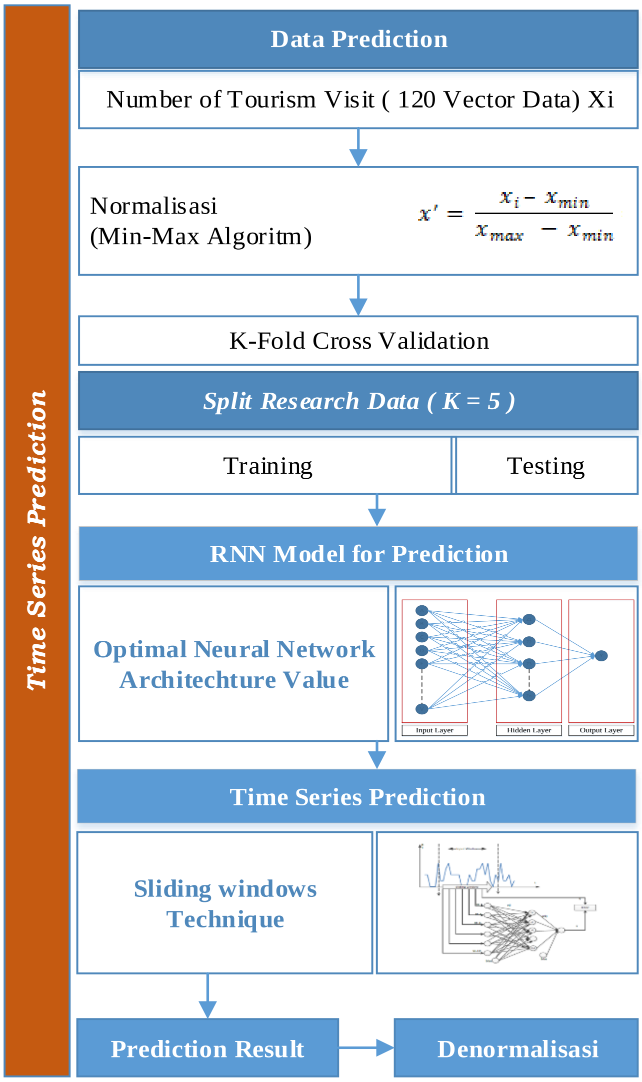

Fig. 1.Architecture of the proposed model.

The next process was to predict using the sliding windows technique, the value of the window size (*t*) was determined based on the number of predictions going forward. Sliding windows technique on univariate data, the movement of the data was on the coherent time (windows size). If the value (t = 3) then the predicted value or target of xt is determined from x_{t-3}, x_{t-2}, x_{t-1}. The final result of the prediction process was the prediction of the normalized visit value (x_t) and the error rate value calculated using the MSE method. The final process was to change the predicted value (x_t') to the actual value (x_t), using de-normalization techniques.

B. Data normalization and denormalization

Data used in the prediction process were first normalized using the min-max normalization method [30]–[32]. The range value used in the normalization process is $0 \le n \le 1$, Normalization also aimed to include the range of original data to be the same as the range used in the Artificial Neural Network, namely sigmoid binner. The value that has not been normalized was represented by (x), then the normalization process was carried out using the mathematical formula (1), becoming (x').

$$x' = \frac{(x_i - x_{\min})}{(x_{max} - x_{min})}(D - C) + C \quad (1)$$

The Min-max normalization method is a normalization method by linearly transforming the original data [31], [33], [34], while the calculation of the Min-max method is described in equation 2. The normalization value is symbolized by x', the value of x_i is the data to be normalized, D is the maximum given value, C is the minimum given value.

$$x_i = \frac{x'(x_{max} - x_{\min})}{(D - C) + C} + x_{\min} \quad (2)$$

x_{min} is the smallest value of the data, x_{max} is the largest value of the data, and (x_{min}, x_{max}} is the range of values of the data. As shown in Table I, the tourist visit data was normalized by using equation 1. Returning the results of normalization to its original condition can be done with denormalisation techniques. The denormalization calculation formula is described as follows, x_i is determined by equation 2.

C. K-Fold Cross Validation

Cross-validation is a statistical model that aims to evaluate learning algorithms by dividing data into two sets [8], [35]–[37], the first data set is used for training data or training the model and the second dataset is used to test the model [8].

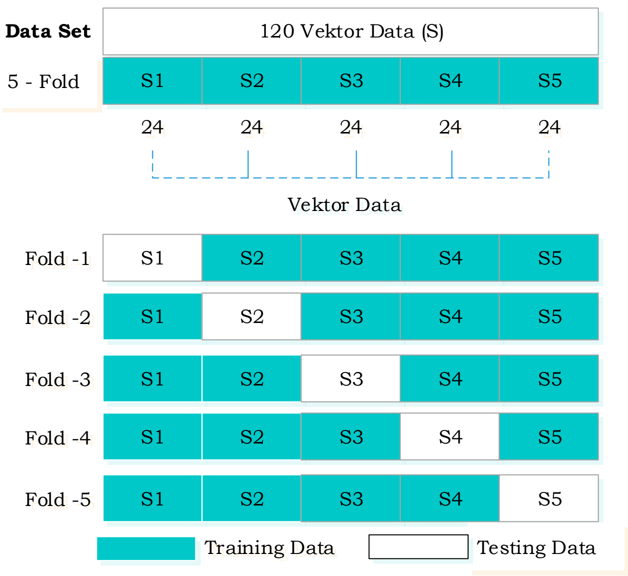

Fig. 2.Cross validation with 5 fold

Cross-validation has functions to assess / validate the accuracy of a model that is built based on a specific dataset [37]. Modeling usually aims to make predictions and classifications of new data that have never appeared in the dataset. The data used in the model development process is called as training data, while the data used to invariably model is called test data. Models that are too accurate for trained data will be models exclusively and produce good output for trained data, but cannot produce good output for validation data (data not included in the frame training), this event is called overfiting.

To estimate the error rate of model that has been developed in the field of pattern recognition, K-Fold Cross Validation can be used as one of techniques of cross validation methods.. The use of K-fold cross-validation aims to eliminate bias on data. The process method divides the data into k subsets or folds, in this study using a K of 5 namely S_1, S_2, S_3, S_4, S_5 with the size of each subset of 24, as shown in Figure 2.

D. Sliding windows technique

The data forecasting of time series technique using the RNN algorithm, could also be done with the sliding windows method [38], [39] . This method allowed more accurate periodic data predictions, and the prediction process could be determined for several periods x_i [27], [40], [41]. he RNN architecture design with the sliding windows method was shown in Figure 3.

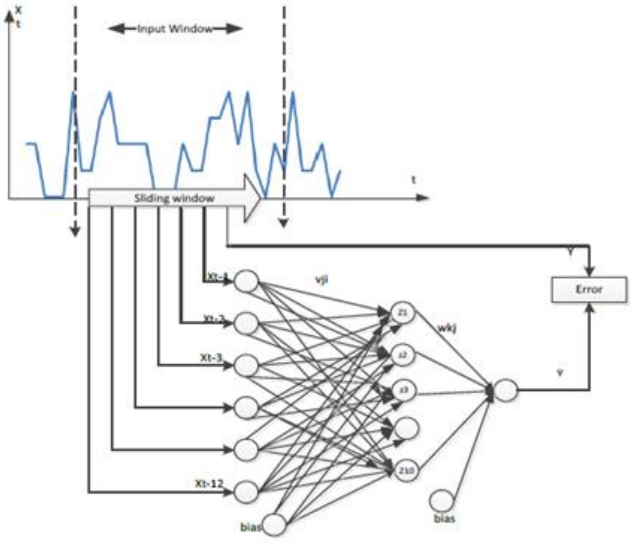

Fig. 3.Time Series Prediction with RNN

The sliding windows technique was based on the predetermined size of the windows, the sample data can be converted to $x_0, x_1, x_2, \ldots, x_{n-1}, x_n, x_{n+1}$ which is a series of data. If the time series data is set to k, the interval data will change to $x_{i-k}, x_{i-k+1}, \ldots, x_i, x_{i+1}$. As shown in Figure 4, the data interval is $x_{t-2}, x_{t-1}, x_t, x_{t+1}$ with a windows size of 3. the value of x_{i+1} is obtained from the value of x_{i-2}, x_{i-1}, x_i .

In the proposed model, the size of windows on windows size depended on the amount of data, where the window determined the prediction accuracy. Based on several experiments, the window size chosen in this study was 3, because it provided the best value and minimizes data reduction. The amount of data was 120 after going through the sliding windows process produces 3 data input variables, 1 output variable, and there were 21 training vector data. The amount of data was reduced because windows size removed the last data value.

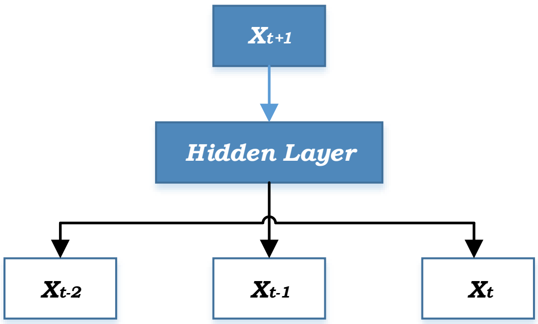

Fig. 4.Sliding windows with 3 windows size

E. Data training with RNN model

Recurrent Neural Network (RNN) is a neural network that has a Feedback connection[3], [6], [7], [42]. There are two types of Recurrent Network models, namely Elman Network and Hopfield Network. Elman Network is two-layer Back propagation network with the addition of a Feedback connection from output to input. This feedback makes Elman Network is able to study, recognize, and make temporary patterns such as spatial patterns [43]–[45].

Hopfield Network is used to store one or more target vectors. Recurrent architecture is almost the same as feed forward back propagation, but it is added with the context layer to accommodate the output of the hidden layer [8]. Feedback can cause the iteration process to be much faster, thus making the parameter update speed and convergence faster. The recurrent neural network structure consists of input, hidden layers, context layers and the output layer as shown in Figure 5.

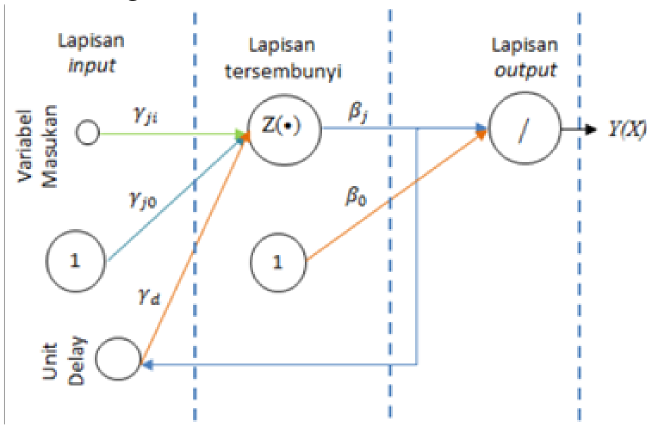

Fig. 5.Elman RNN Architecture

The The Elman RNN architecture is almost the same as the Back propagation feed forward architecture, but coupled with the context layer to accommodate the output results from the hidden layer, as shown in Figure 5. The stages in the Elman RNN algorithm are [45], [46];

A simple RNN network, the input vector is equally distributed through layers weighting (w), but also combined with the activation of the previous state through an additional weight layer, U, where U is the delay weight. Each unit will calculate its activation as in the feed forward network in Figure 6. Each layer will have its own index variables (*k*) for output nodes, (*j*) for hidden nodes, (*h*) for node context , and (*i*) for the input node. An advanced feed network has an input vector of (*x*), which is propagated through a weighted layer.

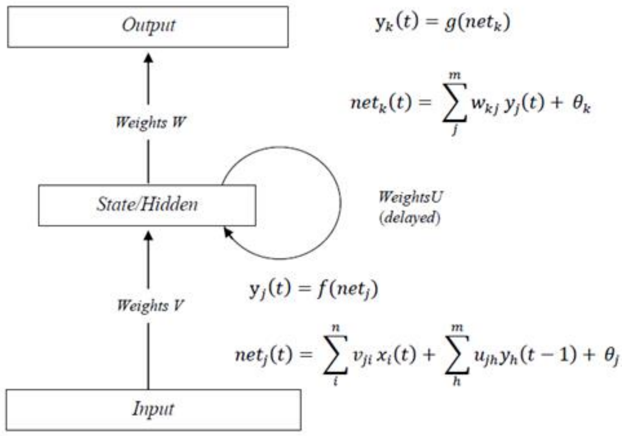

Fig. 6.Koneksi Antar Layer Elman RNN

F. Prediction accuracy

Accurate projection results are predictions that can minimize forecast errors [20], [47]. Estimated error of magnitude is calculated by reducing the real data with forecast [11], [48]. In addition, to determine the estimated error value, there are several techniques in calculating prediction errors, namely: the Mean Squared Error (MSE).

Mean Squared Error (MSE) is a technique for evaluating forecasting methods [47], [49], [50]. Each error or remainder is squared. Then added up and added to the number of observations. This approach manages significant forecasting errors because they are squared. The method produces moderate errors which are probably better for small mistakes, but sometimes produce substantial differences [8], [51]. As shown in the following equation 3 ;

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - f_i) \qquad (3)$$

The value of y_i is the value of the time measure *i* or minimum target on a neural network. f_i is the predicted value at time *i*. The value of n is the number of data samples.

III. RESULT AND DISCUSSION

The prediction process starts from the preparation of the data, the used data was the data of tourist visits to Tanah Lot attractions. The range data was for 10 years, starting from 2009 until 2018. It was about the number of visits each month, so there were 120 data records used for the prediction process. The data were then normalized, to match the data range with the activation function. The results of x' normalization were changed into a time series data format with the size of the windows of 3. The time series dataset was converted into supervised learning, it aimed to be recognized and studied by the neural netowk model.

To do the learning process, the optimal parameters of the neural network architecture were determined first. Determination of each parameter, such as learning rate, momentum, activation function and the number of hidden layers was determined based on the results of the experiment.

Optimal parameter values were used for the training process using the RNN model, and the obtained results were the value of the error rate or MSE compared to the results with different neural network architecture methods. The cross validation process was carried out at the end, to determine the validation of the training results for each fold. This research applied a fold amount of 5. The final result was denormalized, so that the actual value is obtained..

A. Preprocessing data

Preprocessing data on the tourist visit prediction model, there are three main processes including such as ;

- Normalization the data with min-max normalization,
- Change the tourist visit dataset into a univariate time series dataset with windows size = 3
- Convert the time-series dataset to a supervised learning dataset
- Determination of optimal parameters in the liberal network architecture.

The process begins with normalizing the dataset, out of 120 datasets valued at x_{max} = 378,729, x_{min} = 87,400, D = 1, C = 0. By using the normalization equation, the calculation can be explained as follows;

$$x' = \frac{(189.809 - 87.400)}{(378.729 \ - 87.400)}(1-0)+0 \qquad (4)$$

Normalization results are outlined in Table 1, the highest value of normalization results is 1, while the lowest is 0.

Table- I: Result of normalization data

Date	*Data*	*Normalisasi*	*De-Normalisasi*
2009-01	189.806	0,351513	189.806
2009-02	87.400	0	87.400
2009-03	122.688	0,121128	122.688
…	…	…	..
2016-12	378.729	1	378.729
…	…	…	…
2018-12	268.165	0,620484	268.165

The denormalization process is carried out to return the value, returning to the original range. The results of denormalization are shown in Table I, the denormalization process can be carried out in the following way.

$$x_1 = \frac{0.35(378.729 - 87.400)}{(1 \ - 0)} + 87.400 = 189.806 \qquad (4)$$

he normalized data was then converted into a sliding windows form by using the number of windows size of 3. Changing the data form to x_1, x_2, x_3 is the input value, while x_4 becomes the output value y_{out}. Overall results of sliding windows can be described in Table I.

The time series dataset shown in Table II, was subsequently changed into the form of supervised learning in order to be trained with the neural network model. In the supervised learning model the value of x_1, x_2, x_3 becomes the input value and the value of x_4 becomes the target value of y_{out}. The hidden layer value, in the form of the number and number of neuron units, is determined by several experiments.

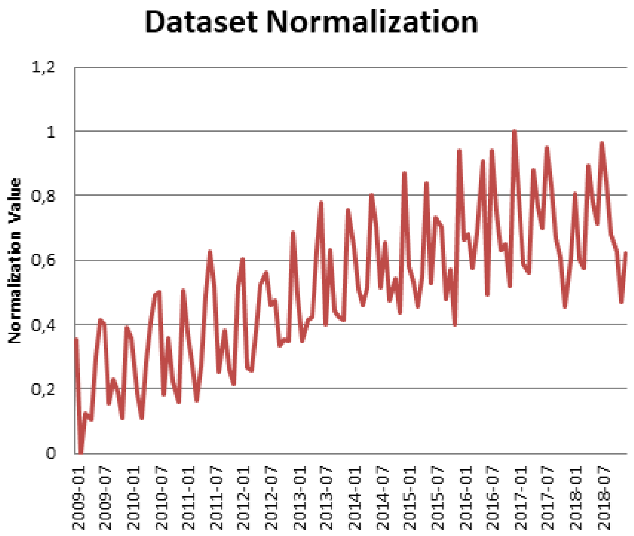

Fig. 7.Grafik of normalisasi dataset

Table- II: Time series dataset

Date	x_i	x_2	x_3	*y*
2009-01	0,3515	0	0,1211	0,1025
2009-02	0	0,1211	0,1025	0,2972
2009-03	0,1211	0,1025	0,2972	0,4155
…	…	…	..	
2016-12	1	0,7963	378.729	0,5625
…	…	…	…	
2018-09	0,6831	0,6244	0,4713	0,6204

The next process after the dataset in the form of supervised learning was to determine the optimal neural network architecture. The determination of the optimal parameters in the neural network architecture, obtained from experimental results. The experimental results showed the optimal parameter values shown in Table III. The learning rate value used is 0.001, the number of hidden layers used in architecture, the number of neuron units used is 15, and the momentum value used is 0.001.

Table- III: Neural network parameters for data training

Parameter	*Value*	Information
Unit Layer	*10*	Number unit layer
Learning Rate	*0,01*	Learning rate NN
Input Layer	*1*	Number of input layer
Hidden Layer	*3*	Number of hidden layer
Output Layer	*1*	Number of output layer
Momentum	*0,001*	Value of momentum
Transfer Function	*Sigmoid Binner*	
Weight	*Random (0,1)*	

The activation function used is sigmoid binner, and the determination of the weight value is a random number with a range of values from 0 to 1 Table III was used next for training on neural network models. The RNN method used in subsequent studies compared the results with several neural network architectures, such as Jordan RNN, Fully RNN, Feed froward NN.

Retrieval Number: A1833109119/2019©BEIESP
DOI: 10.35940/ijeat.A1833.109119
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

B. Comparison Model Neural Network

The comparison result of RNN optimization model for the prediction process with several other models were shown in Table III. Values that were used as reference points were the error rate (MSE) and training time generated from each model. Prediction results showed the results of the training show Elman RNN of 0.0703 with training time 00:17:36, while the testing value was 0.025317. The Elman RNN model had the smallest training time compared to other models.

Table- IV: Comparison of performance using various prediction methods

Algoritm	Training (MSE)	Testing (MSE)	Training Time
Elman RNN	0,0703	0,025317	00:17:36
Jordan RNN	0,0639	0,029027	00:19:02
Fully RNN	0,0639	0,029012	00:22:02
Feedforward	0,0707	0,025171	00:18:16

Prediction results using Elman RNN had a better result compared to neural network models with different architecture in terms of training time. The highest training value of 0.0707 was produced by the feed forward NN model while the lowest training value was generated by the Elman RNN model, but with a very small difference of 0,0004. The time needed to carry out the training process refers to the number of epochs of 10^5.

C. Validation data training and testing with K-fold Cross validation

Cross-validation is used to assess / validate the accuracy of a model that is built based on a specific dataset. Modeling usually aims to make predictions and classifications of new data that have never appeared in the dataset. The data used in the model development process is called training data, while the data used to invariably model is called test data. In this study, it used K value of 5, this is based on several experiments that have been done before.

Table- IV: Comparison of performance using various prediction methods

Algoritm	Training (MSE)	Testing (MSE)	Training Time
K-Fold 1	0,04827	0,04554	783 s
K-Fold 2	0,06300	0,02371	788 s
K-Fold 3	0,05174	0,02992	768 s
K-Fold 4	0,04581	0,05137	849 s
K-Fold 5	0,04646	0,03895	908 s
Rata-rata	0,05105	0,0379	819 s

Each dataset had 21 data records for the training process. Prediction results were foreach dataset (fold) was then taken an average value from the above experiment. The average value of training was 0.05105, the value of testing was 0.0379, and the average training time was 819 seconds. The value obtained has a large enough difference, if compared without using the k-fold validations model, this is because the amount of data used is different so that the average taken is the result of training after being validated and averaged.

IV. CONCLUSION & FUTURE WORK

Tourism is an activity that is directly connected and involves the community, thus it brings various impacts on the local community. Tourist visits to Tanah Lot tourism objects have fluctuating data patterns, so we need a prediction system that can recognize the prediction patterns of tourist visits. The application of the Elman RNN model was able to recognize the pattern of tourist visit data, and the average training result was 0.05105 and testing was 0.0379, with an average training time of 819 s. The test results also showed the Elman RNN model had a better training time than the Jordan model RNN, fully RNN and feedforward NN.

The prediction model of tourist visits to Tanah Lot attraction by using the Elman RNN model can be added by combining the model (Hybrid Model) with other models such as wavelets, clustering models and classification models. In addition, the RNN model can also be optimized by using several models such as Genetic Algorithms, particle swarm optimization, and evolutionary computation. In future research, the process of prediction of tourism visit can also do some optimization on RNN architectural parameters. This research can also use deep learning models such as GRU and LSTM as a reference for comparing RNN optimization model.

REFERENCES

1. Dispardabali, "Pertumbuhan Kunjungan Wisatawan Nusantara Tahun 2015," 2015. [Online]. Available: http://www.disparda .baliprov.go.id/id/Statistik3.
2. S. R. Sanam, "WISATA TANAH LOT," *Politek. Int. Bali*, 2018.
3. P. Sugiartawan, R. Pulungan, and A. K. Sari, "Prediction by a Hybrid of Wavelet Transform and Long-Short-Term-Memory Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 2, 2017, pp.326-332.
4. M. A. Hazazi and A. Sihabuddin, "Extended Kalman Filter In Recurrent Neural Network : USDIDR Forecasting Case Study," vol. 13, no. 3, 2019, pp. 293–300.
5. I. B. N. Pascima and S. Hartati, "Sistem Prediksi Harga Nilai Tukar Mata Uang Menggunakan Elman Recurrent Neural Network dengan Algoritma Genetika sebagai Metode Pembelajaran," *BIMIPA*, vol. 25, no. 3, 2018, pp. 275–287.
6. R. Arizon, "Prediksi curah hujan menggunakan recurrent neural network rainfall prediction using recurrent neural network," 2011.
7. P. Sugiartawan, S. Hartati, and A. Musdholifah, "Prediction of Multivariat Time Series Data Using Fully Recurrent Neural Network.",2018
8. P. Sugiartawan, S. Hartati, and A. Musdholifah, "Tourist Visits Prediction With Fully Recurrent Neural Network", 2018 pp. 156–165.
9. P. Sugiartawan and I. G. S. Arta, "Peramalan Tingkat Kunjungan Wisatawan dengan Metode Average Based Fuzzy Time Series dan Markov Chain Model di Sriphala Resort & Hotel," 2015.
10. P. Sugiartawan, R. Pulungan, and A. K. Sari, "Prediction by a Hybrid of Wavelet Transform and Long-Short-Term-Memory Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 2, 2017, pp. 326–332.
11. N. Davey, S. Hunt, and R. Frank, "Time Series Prediction and Neural Networks," ,1999, pp. 1–6.
12. B. Warsito, Subanar, and Abdurakhman, "Pemodelan Time Series Dengan Maximal Overlap Discrete Wavelet Transform," *Pros. Semin. Nas. Stat.*, no. 1998, 2013,pp. 605–613.
13. L. F. Ortega, "A Neuro-wavelet Method for the Forecasting of Financial Time Series," *Wavelet*, vol. I, 2012.
14. I. Khandelwal, U. Satija, and R. Adhikari, "Forecasting seasonal time series with Functional Link Artificial Neural Network," *2nd Int. Conf. Signal Process. Integr. Networks, SPIN 2015*, 2015 pp. 725–729.
15. Z. Zhang, "Research on Prediction Method of API Based on The Enhanced Moving Average Method," *2012 Int. Conf. Syst. Informatics*, 2012, pp. 2388–2392.
16. R. Adhikari, R. K. Agrawal, and L. Kant, "PSO based Neural Networks vs. traditional statistical models for seasonal time series forecasting," *2013 3rd IEEE Int. Adv. Comput. Conf.*, 2013 pp. 719–725.

17. A. Q. Munir, "Forecast Model Time Series Data to Dengue Fever Phenomena with Statistical Approach Model to Endemic Area Determination,", 2013, pp. 1–6.
18. Z. Kasiran, Z. Ibrahim, M. Syahir, and M. Ribuan, "Mobile Phone Customers Churn Prediction using Elman And Jordan Recurrent Neural Network," *7th Int. Conf. Comput. Converg. Technol.*, 2012, pp. 673–678.
19. P. Sugiartawan, S. Hartati, and A. Musdholifah, "Prediction of Multivariate Time Series Data Using Fully Recurrent Neural Network."
20. H. Wang, "Forecasting the Quality of Water Diversion in the Water Source of Lake Taihu,", 2011 pp. 984–988.
21. W. Setiawan, "Prediksi Harga Saham Menggunakan Jaringan Syaraf Tiruan Multilayer Feedforward Network Dengan Algoritma Backpropagation," in *Konfrensi Nasional Sistem & Informatika 2008*, 2008, pp. 108–113.
22. H. Wang, Y. Gao, Z. Xu, and W. Xu, "An recurrent neural network application to forecasting the quality of water diversion in the water source of Lake Taihu," *2011 Int. Conf. Remote Sensing, Environ. Transp. Eng.*, vol. 1, 2011, pp. 984–988.
23. C. L. Chen and Y. P. Bau, "Establishing a multi-criteria evaluation structure for tourist beaches in Taiwan: A foundation for sustainable beach tourism," *Ocean Coast. Manag.*, vol. 121, 2016 pp. 88–96.
24. J. Hakim and S. Hartati, "Optimasi Prediksi Kehadiran Pegawai Untuk Intensif Kehadiran Menggunakan Jaringan Saraf Tiruan-Backpropagation," *SEMNASTEKNOMEDIA ONLINE*, vol. 1, 2013, pp. 1–22.
25. Z. Fangwen, L. Xiaowei, Y. Zehong, and S. Yixu, "Financial time series prediction based on Echo State Network," *Nat. Comput. (ICNC), 2010 Sixth Int. Conf.*, vol. 8, 2010, pp. 3983–3987.
26. J. Wang and Q. Zhu, "Short-term Electricity Load Forecast Performance Comparison Based on Four Neural Network Models," *27th Chinese Control Decis. Conf. (2015 CCDC)*,2015, pp. 2928–2932.
27. G. Dominguez, M. Guevara, M. Mendoza, and J. Zamora, "A Wavelet-Based Method for Time Series Forecasting," *2012 31st Int. Conf. Chil. Comput. Sci. Soc.*, 2012, pp. 91–94.
28. R. Chandra, "Competition and collaboration in cooperative coevolution of elman recurrent neural networks for time-series prediction," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 26, no. 12, 2015, pp. 3123–3136.
29. X. Wang, C. Zhang, and S. Zhang, "Modified Elman Neural Network and Its Application To Network Traffic Prediction,", 2015 pp. 2–6.
30. S. K. Panda and P. K. Jana, "Efficient task scheduling algorithms for heterogeneous multi-cloud environment," *J. Supercomput.*, vol. 71, 2015, pp. 1505–1533.
31. S. A. Abdulkarim, "Time Series Prediction With Simple Recurrent Neural Networks," vol. 9, no. 1, 2016, pp. 19–24.
32. S. G. K. Patro and K. K. Sahu, "Normalization: A Preprocessing Stage," *arXiv Prepr. arXiv1503.06462.*, 2015, p. 4.
33. M. Vafaeipour, O. Rahbari, M. A. Rosen, F. Fazelpour, and P. Ansarirad, "Application of sliding window technique for prediction of wind velocity time series," *Int. J. Energy Environ. Eng.*, vol. 5, no. 2–3, 2014, pp. 1–7.
34. V. V Kondratenko and Y. a Kuperin, "Using Recurrent Neural Networks To Forecasting of Forex," *Time*, p. 23, 2003.
35. Z. Liu, "Chaotic time series prediction based on MPMR," *J. Comput. Inf. Syst.*, vol. 1, no. 4, 2005, pp. 733–739.
36. S. L. Badjate and S. V. Dudul, "Prediction of Duffing Chaotic time series using focused time lagged recurrent neural network model," *TENCON 2008 - 2008 IEEE Reg. 10 Conf.*, 2008.
37. S. Borra S and A. D. Ciaccio, "Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty method," *Comput. Stat. Data Anal.*, 2010, pp. 468–475.
38. L. V. Fausett, "Fundamentals of Neural Networks: Architectures, Algorithms, and Applications," , 1994, p. 461.
39. L. R. Medsker, L. R. Medsker, L. C. Jain, and L. C. Jain, "Recurrent neural networks," *New York*, 2001, pp. 1–18.
40. R. Begg, J. Kamruzzaman, and R. Sarker, *Neural Networks in Healthcare : Potential and Challenges*. 2006.
41. F. Gorunescu, "Data mining: Concepts, models and techniques," *Intell. Syst. Ref. Libr.*, vol. 12, 2011, pp. 467–534.
42. P. May, E. Zhou, and C. W. Lee, "Learning in fully recurrent neural networks by approaching tangent planes to constraint surfaces," *Neural Networks*, vol. 34, 2012, pp. 72–79.
43. I. A. Gheyas and L. S. Smith, "A novel neural network ensemble architecture for time series forecasting," *Neurocomputing*, vol. 74, no. 18, 2011, pp. 3855–3864.
44. Z. C. Lipton, J. Berkowitz, and C. Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning,", 2015, pp. 1–35.
45. L. R. Medsker and L. C. Jain, "Recurrent neural networks: design and applications," , 2000, p. 414.
46. L. B. Neto, P. H. G. Coelho, M. L. F. Velloso, J. C. C. B. De Scares Mello, and L. A. Meza, "Monthly flow estimation using elman neural networks," *ICEIS 2004 - Proc. Sixth Int. Conf. Enterp. Inf. Syst.*, no. January, 2004, pp. 153–158.
47. T. Kolarik and G. Rudorfer, "Time series forecasting using neural networks," *ACM SIGAPL APL Quote Quad*, vol. 25, 1994, no. 1, pp. 86–94.
48. R. Frank, N. Davey, and S. Hunt, "Time series prediction and neural networks," *J. Intell. Robot. Syst.*, vol. 31, no. 1–3, 2001 pp. 91–103.
49. A. A. Ferreira, T. B. Ludermir, and R. R. B. De Aquino, "Comparing recurrent networks for time-series forecasting," *Proc. Int. Jt. Conf. Neural Networks*, 2012, pp. 10–15.
50. N. Susanti, "Penerapan Model Neural Network Backpropagation untuk Prediksi Harga Ayam," *Semin. Nas. Teknol. Ind. dan Inform.*, 2014, pp. 325–332.
51. B. Sutijo, "PEMILIHAN HUBUNGAN INPUT-NODE PADA JARINGAN SARAF (Input-Nodes Link Selection on Radial Basis Funtion Neural Network),", 2006, pp. 55–61.

AUTHORS PROFILE

Putu Sugiartawan, M.Cs is a PhD candidate at Department of Computer Science and Electronics Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia. He did his master in Department of Computer Science and Electronics Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia. His research and publication interest include Artificial Intteligence (Neural Network) and Decision Support System. He is a mamber of the association Computer Science, Electronics and Instrumentation Support Society, National and Asosiasi Perguruan Tinggi Informatika dan Komputer (APTIKOM). He also teaches in STMIK STIKOM Indonesia as lecturers, courses taught include discrete mathematics, artificial intelligence, statistics, and business intelligence. Previous related research is Prediction of Multivariate Time Series Data Using Fully Recurrent Neural Network, and Prediction by a Hybrid of Wavelet Transform and Long-Short-Term-Memory Neural Network.

Prof. Sri hartati, M.Sc., Phd is a senior lecturer and Professor in Department of Computer Science and Electronics Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia. She did her master and PhD in Faculty of Computer Science, Univ of New Brunswick, Canada. Her areas of interest are in the Artificial Intelligence (Knowledge Based Syst., Reasoning Syst., Expert Syst, Fuzzy Syst., Pattern Recog, Decision Support System, Medical Computing and Soft Computation (ANN, Fuzzy Logic, Genetic Algorithm)cluster - Emergency Nursing) and Computer Reasoning. She is chairman of the association Computer Science, Electronics and Instrumentation Support Society, National. Her research and publication interests include Artificial Intelligence and Computer Reasoning. Previous related research is Prediction of Multivariate Time Series Data Using Fully Recurrent Neural Network