Prediction of Yuan to IDR Exchange Rate using General Regression Neural Network

1st Evi Febrion Rahayuningtyas
Informatics
University of Muhammadiyah Malang
Malang, Indonesia
febrionrevi@webmail.umm.ac.id

2nd Galih Wasis Wicaksono
Informatics
University of Muhammadiyah Malang
Malang, Indonesia
galih.w.w@umm.ac.id

3rd Didih Rizki Chandranegara
Informatics
University of Muhammadiyah Malang
Malang, Indonesia
didihrizki@umm.ac.id

***Abstract*— The exchange rate is the value or price of a currency in front of other currencies divided into selling rates and buying rates. The differences and alteration of exchange rates are caused by interest rates, inflation, and many other factors. The General Regression Neural Network method is applied to build a prediction system for the Yuan to IDR exchange rate, using the input to determine the output. The dataset is taken from the Bank Indonesia website with 191 records after pre-processing. Based on the resulting test, we found that the MSE score is 106.13, the RMSE score is 10.30, and the MAE score is 8.73. The model can find and recognize training data patterns to provide excellent data output with the results given.**

Keywords—prediction, exchange rate, general regression neural network

I. INTRODUCTION

In buying and selling transactions, people use money as a tool in transactions to replace bartering that was previously used by traditional communities[1][2][3]. Based on the type, money is divided into foreign money and non-foreign money[3]. The exchange rate is value or price of that currency in front of other currencies[4]. In the real world, the exchange rate is divided into the selling rate and the buying rate. The selling rate is often used in import transactions. The buying rate is used in commerce for export activities. The exchange rate is an essential indicator in its application to an open economic system[5][6]. The exchange rate of each country is always different. Several factors trigger the difference. Among them are interest rates and inflation. Interest rates and inflation are changed by the increase in the price of money in a country. It is also affected by fluctuating interest rates[7][8][9].

China is the largest importing country in the world. China also the biggest importers for Indonesia. This fact is evidenced by the existence of a net import value in December 2019 of USD 4.1 billion. However, in January 2020, the net import value will be USD 4 billion[10]. Unpredictable changes can lead to unanticipated inflation and currency exchange rates. Thus, the proposed idea is to create a prediction system for the exchange rate of the Yuan to Rupiah with several contributing factors, namely inflation and reference interest rates.

Previous research was conducted by Ciptya Rahma Almira using the Adaptive Neuro-Fuzzy Inference System to produce RMSE 10.76 on the daily prediction system and 47.51 on the monthly prediction system[11]. Other research was also conducted by Safira Yasmin Amalutfia using the Fuzzy Time Series Markov Chain, resulting in an error of 0.42% and 0.41%[12]. Nova Kusumawati also conducted research. However, using the USD dataset using the Linear Regression method produces 75% accuracy[13].

General Regression Neural Network is a regression prediction method that determines how the Artificial Neural Network works by using the training data's output as a guide in finding patterns in testing data to produce the appropriate output. Using the Yuan to Rupiah dataset, the research proposal is implemented in the General Regression Neural Network method to explain its problems. Building a prediction system that aims for this study using the General Regression Neural Network can be a study with better accuracy than previous studies.

II. RESEARCH METHODS

This research was conducted in several stages. Each of these stages is described in detail in Figure 1.

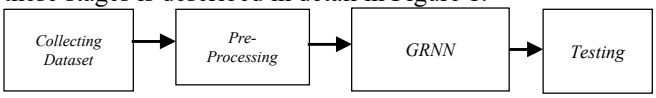

Fig 1. Research Stages

A. Collecting Dataset

The dataset is collected from the official website of Bank Indonesia (bi.go.id). Consists of 3 separate files that go through the data integration stage to produce complete unified data. The entire dataset has three dependent variables and an independent variable. Several samples in each dataset are described in Table I, Table II, and Table III.

The sample dataset in Table I is The Inflation Dataset Sample which amounts to 44 records. The sample dataset in Table II is Sample Reference Interest Rate Dataset which amounts to 45 records. The sample dataset in Table III is the Sample Dataset of Selling Rate and Buying Rate, which amounts to 878 records.

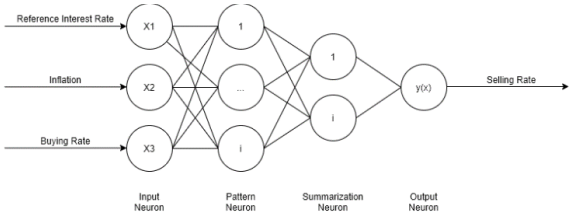

Fig 2. General Structure of the General Regression Neural Network Used in the Research

TABLE I. INFLATION DATASET SAMPLE

No.	Period	Inflation
1	April, 2017	4.17 %
2	May, 2017	4.33 %
….	….	….
….	….	….
43	October, 2020	1.44 %
44	November, 2020	1.59 %

TABLE II. SAMPLE REFERENCE INTEREST RATE DATASET

No.	Date	BI-7Day-RR
1	April 20th, 2017	4.75 %
2	May 18th, 2017	4.75 %
….	….	….
….	….	….
44	October 13th, 2020	4.00 %
45	November 19th, 2020	3.75 %

TABLE III. SAMPLE DATASET OF SELLING RATE AND BUYING RATE

No.	Value	Selling Rate	Buying Rate	Date
1	1,00	Rp1,949.11	Rp1.929,02	April 3rd, 2017
2	1,00	Rp1,949.18	Rp1.929,09	April 4th, 2017
….	….	….	….	….
….	….	….	….	….
877	1,00	Rp2.169,41	Rp2.147,69	November 19th, 2020
878	1,00	Rp2.174,18	Rp2.152,48	November 20th, 2020

In regression, the main requirement in implementing it is to know the correlation between the variables used in the study. Thus, Pearson Correlation is used to determine the correlation of variables. With Pearson correlation, the dataset's characteristics can be known, making it easier to understand the relationship between the variables used. Karl Pearson discovered Pearson Correlation in 1990. Karl defined the finding with the following formula:

$$r = \frac{\sum x_i y_i}{\sqrt{\sum {x_i}^2}\sqrt{\sum {y_i}^2}} \quad (1)$$

or,

$$r = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{\sqrt{n\sum {X_i}^2 - (\sum X_i)^2}\sqrt{n\sum {Y_i}^2 - (\sum Y_i)^2}} \quad (2)$$

with information,

n : number of records

$\sum x$: total number of variables x

$\sum y$: total number of variables x

$\sum x^2$: the square of the number of variables x

$\sum y^2$: the square of the number of variables y

$\sum xy$: the product of the total number of variables x and variable y

x_i : $x_i - \bar{x}$

y_i : $y_i - \bar{y}$

B. Pre-Processing

Separate data are put together with the same period, namely the weekly period. The statistical "mean" calculation is given to fill the blank records and change each row's daily period in the dataset into weekly periods, especially in the exchange rate dataset. Dropping unnecessary columns such as "Week to", "Date". The last process of pre-processing is the use of normalization using the MinMax Scaler.

C. Modeling using General Regression Neural Network

General Regression Neural Network was first introduced to the public by Specht in 1991. The way this model works is to mimic the Artificial Neural Network that humans have. GRNN is a predictive model that uses a non-linear transformation with the output calculated from the input values [14][15][16][17]. GRNN has a general structure that shows Figure 2.

In the structure in Figure 2, there are four neurons or layers, which are described in detail as follows[18]:

- Input Neuron

  Retrieval of input information occurs at this layer[19]. The Input information distributes to the next layer. There is no data processing at this layer.

- Pattern Neuron

  This layer is the first hidden layer where a data processing stage is obtained from the previous layer's inputs distribution. At this layer, you get output θ which will be processed in the next layer.

- Summarization Neuron

  This layer has neurons called numerator neurons and denominator neurons. A summation calculation is carried out, which involves weighting. The analysis result has output θ on the previous layer. The results from this layer are used to calculate the output for the next layer.

- Output Neuron

  The last layer of GRNN calculates the results of the summarization layer and produces the output $y(x)$[16][20][19].

D. Testing

The performance of GRNN tested by several test parameters such us in the form of MSE, RMSE, and MAE, which are described in detail as follows:

- Mean Square Error (MSE)

  MSE was obtained from the error squared and divided by the average. Thus, the bigger the error in the model, the greater the MSE that will be generated. The MSE formula is written as follows[21]:

  $$MSE = \frac{\sum(Y' - Y)^2}{n} \quad (1)$$

with information:

Y' : predicted
Y : actual
n : number of data

- Root Mean Square Error (RMSE)

The RMSE is the MSE result given the square root. Similar to MSE, the smaller the error, the lower the resulting RMSE value. The RMSE formula is written as follows[22]:

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

with information:

Y' : predicted
Y : actual
n : number of data

- Mean Absolute Error (MAE)

The same is the case with MSE. The MAE value is obtained from the squared error. However, the difference between the two lies in the presence and absence of absolute values. Any errors in the MAE will be made an absolute value. The MAE formula is written as follows[22]:

$$MAE = \frac{\sum|Y'-Y|^2}{n} \quad (3)$$

with information:

Y' : predicted
Y : actual
n : number of data

III. Results and Discussion

This chapter will discuss the dataset results, pre-processing, modeling, and testing, which are the research stages. The results of the study are described as follows:

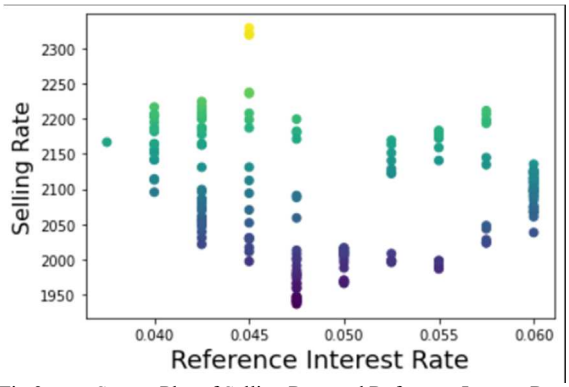

Fig 3. Scatter Plot of Selling Rate and Reference Interest Rate

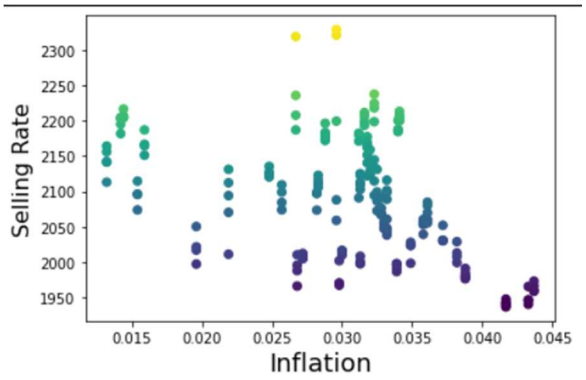

Fig 4. Scatter Plot of Selling Rate and Inflation

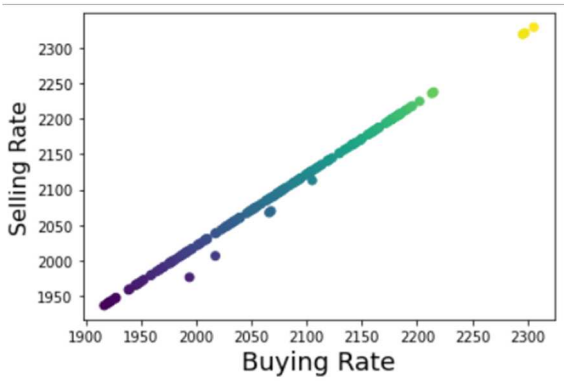

Fig 5. Scatter Plot of Selling Rate and Buying Rate

A. Dataset

The data used in this study use actual data. The data is taken from the official website of Bank Indonesia. The dataset is public which can be accessed by everyone. The owner of the dataset is Bank Indonesia. Data is presented in real-time from April 2017 to November 2020. The total dataset consists of 967 data and divides into three excel documents. The inflation rate in Indonesia is also recorded in one of the three excel documents. Each row in the dataset is presented with a monthly period. To produce a dataset of 44 records.

The second file is historical data on reference interest rates. This data is downloaded with a data range from April 20, 2017, to November 19, 2020. Just like the inflation dataset, the reference interest rate dataset is also presented in a monthly period. The benchmark interest rate is generally expressed in monthly terms. However, there is a reference interest rate given two times a month, so this download produces a dataset of 45 records. The third and final file of the entire dataset to download is historical data on the yuan exchange rate against the rupiah. There are two types of codes in the yuan exchange rate dataset. We use the CNH coded dataset because this code is used for transactions outside Mainland China, such as China and others. This dataset has a daily period. Exchange rate changes do occur every day. However, the dataset presented is incomplete, so it needs to go through a complicated pre-processing stage. The CNH coded yuan exchange rate dataset that was successfully downloaded in this study was 878 records.

The three files produce the attributes used in the research process. However, the three downloaded files still must go through the pre-processing stage to make it easier to process, especially at the GRNN modeling stage.

Data plotting was used, so the relationship between variable x and y and the data's characteristics could be found. This method can determine each data's features and the relationship between variables. There are three primary colors used when plotting data, namely purple, green and yellow. If the point on the plot is close to yellow, the higher the selling rate will be. This plot applies to purple. The closer to purple, the lower the selling rate. Data plotting is presented in Figure 2, Figure 3, and Figure 5.

The relationship between the selling rate and the reference interest rate shown in Figure 3 illustrates that the distribution is evenly distributed over several values. When the benchmark interest rate was at 4.5%, the selling rate shot above IDR 2,300. This value is the highest peak of the selling rate in the data range period that has been taken. The lowest selling rate is at less than IDR 1,950, which is the reference interest rate in the range of 4.75%. Unlike the case with the reference interest rate. The highest reference interest rate is at the 6% point, where the selling rate is in the range of IDR 2,000 to IDR 2,150. Meanwhile, the lowest reference interest rate is less than 4% at a point when the selling rate is in the range of IDR 2,150 to IDR 2,200.

It can be seen in Figure 4 that the relationship between the selling rate and inflation is different from the relationship between the selling rate and the reference interest rate. The spread of the inflation rate is at 3% to 3.5%, with the selling rate at IDR 2,000 to IDR 2,250. Seen in the distribution of the scatter plot points. The highest selling rate in the downloaded dataset is at IDR 2,300, with an inflation rate of 3%, but the lowest selling rate is between 4% and 4.5%. Meanwhile, the highest inflation is at 4% to 4.5%, with a selling rate of less than IDR 1,950. On the other hand, the lowest inflation is at less than 1.5%, with a selling rate of IDR 2,100 to IDR 2,200. It can be concluded that the lowest selling rate is the moment when the inflation rate is at its highest peak.

The last data plot is in Figure 5. This plotting explains the relationship between the selling rate and the buying rate. The relationship is positive at most of the points. The points are generated on the scatter plot to form a linear graph line. So, it can be concluded that the higher the buying rate, the higher the selling rate. However, some points are outside the linear graph line, namely at the buying rate of IDR 1,950 to IDR 2,150 and IDR 1,950 to IDR 2,150.

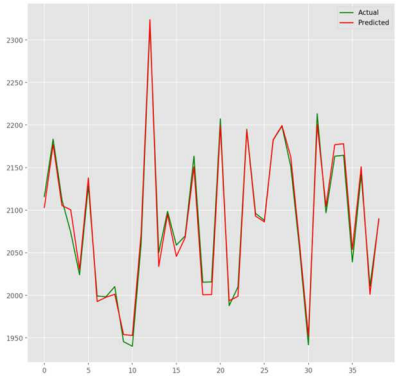

Fig 6. Line Plot of Prediction Results

TABLE IV. DATASET AFTER GOING THROUGH PREPROCESSING

Week to	Date	BI-7Day-RR	Inflation	Selling Rate	Buying Rate
1	02-Apr-17	4,75%	4,17%	Rp 1.947,39	Rp 1.927,51
2	09-Apr-17	4,75%	4,17%	Rp 1.936,35	Rp 1.916,91
….	….	….	….	….	….
….	….	….	….	….	….
190	15-Nov-20	4,00%	1,59%	Rp 2.164,44	Rp 2.142,77
191	22-Nov-20	3,75%	1,59%	Rp 2.166,68	Rp 2.144,98

TABLE V. CORRELATION WITH PEARSON CORRELATION

Correlation	BI-7Day-RR	Inflation	Selling Rate	Buying Rate
BI-7Day-RR	1	0,228	-0,130	-0,129
Inflation	0,228	1	-0,437	-0,436
Selling Rate	-0,130	-0,437	1	-0,436
Buying Rate	-0,129	-0,436	0,999	1

In Table V, we can find the relationship between the variables used in the study. Based on the correlation number index, a very weak correlation occurs in the selling rate with the reference interest rate and the buying rate with the reference interest rate with a value of -0.130. Weak correlation occurs in the attribute of the benchmark interest rate with inflation. The correlation value is 0.228. There is a relation between the selling rate and inflation, and the buying rate with inflation. The two values are 0.437 and 0.436. However, it cannot be denied that there is a perfect correlation value obtained from the attributes of this study, namely the correlation value between the selling rate and the buying rate of 1. A negative value in the correlation indicates that there is a relationship between the two variables.

On the other hand, positive values have a positive relationship. So, in this case, it can be said that the higher the buying rate is, the higher the selling rate. This condition happened because the selling rate and the buying rate have a positive relationship.

B. Pre-processing

Because the dataset downloaded and used in this study still has a lot of noise, a data cleansing stage is needed, which is the pre-processing stage. Data cleansing is done to remove all noise in the dataset, especially in the yuan exchange rate dataset against the rupiah. In this dataset, many records are empty. The "mean" calculation is given, which aims to provide a value for a blank record. So that the document will be complete, and the data can be said to be data cleaning.

According to the dataset details, it was explained that the result has a different period from the dataset required. To be used in this study and the dataset also comes from additional files. It is necessary to equalize each dataset period and unify the dataset into one file so that it is easier to process it. This stage starts from the equalization of the periods for each file in the dataset. The period used in this study is the weekly period. So, especially in the inflation attribute, four records have the same value. The reference interest rate attribute also

experiences the same thing. However, because the reference interest rate is set on a specific date even though it only occurs once a month, the determination of each record's value in the reference interest rate attribute follows the date of conclusion of the reference interest rate given. In the exchange rate attribute, which initially has a daily period in each record, a statistical "mean" calculation is provided to determine a fixed value for each week.

After each attribute in the dataset has the same period, each feature goes through the data integration stage. The dataset used can be a complete unit and processed into a data frame before being used in modeling. The resulting dataset is 191 records which are described in Table IV.

The next step in pre-processing is splitting the dataset. This step is crucial in the modeling process. The process divided the dataset into two-part consist of data train and test. The data test size approximately 20% of the total data or equivalent with 39 records. The last stage of pre-processing is the data normalization stage. At this stage, the MinMax Scaler's normalization is used to aim that the original data has a balanced nature because the balance between the data affects the pattern-finding process in training and affects the score when testing.

C. *Modeling using General Regression Neural Network*

The modeling in this study uses the General Regression Neural Network, which is illustrated in Figure 6.

There are three variables x, namely the reference interest rate, inflation, and the buying rate as in Figure 6 above, with variable y is the selling rate. The first stage in modeling starts from the training process. The data used are data that have gone through the splitting dataset in the pre-processing process, producing 152 training data or 80% of the total data and 39 testing data or 20% of the entire data. Normalization is carried out using the MinMax Scaler so that the data is expected and balanced and is also given a standard deviation of 0.1. The prediction results obtained are illustrated in the sample below

The prediction results using GRNN in this study are visualized with a line plot with two colors; red is the predicted data or the prediction result. The green color is actual data or actual data.

From Table VI, the modeling with GRNN in this study went very well. The model can find patterns in the training data to provide results like the actual data when testing. The model can recognize the direction of the data pattern.

D. *Testing*

The performance of the GRNN has been tested with three parameters, MSE, RMSE, and MAE, with different standard deviation measures. The results are shown in Table VII. Experiments were carried out aimed at comparing the standard deviation measures used in this study. It can be seen in the works in Table VI that the standard deviation of 0.1 results in an accuracy test that has the highest score beating other standard deviation measures. The study uses a standard deviation of 0.1 to get the best accuracy test results.

IV. CONCLUSION

The prediction using the General Regression Neural Network with the non-linear regression model works very well. Using training data of 80% to look for patterns in the data and tested with testing data of 20% shows that GRNN can produce the accuracy of the test using several test parameters with an excellent score.

TABLE VI. SAMPLE PREDICTION RESULTS USING GENERAL REGRESSION NEURAL NETWORK

No.	Actual	Predicted	Prediction Error	Percentage of Prediction Error
1.	2115,914	2103,10	-12,814	-0,61 %
2.	2183,290	2177,01	-6,280	-0,29 %
3.	2112,214	2105,42	-6,794	-0,32 %
4.	2073,820	2100,34	26,520	1,28 %
5.	2024,056	2030,50	6,444	0,32 %
6.	2128,650	2137,82	9,170	0,43 %
7.	1999,190	1992,67	-6,520	-0,33 %
8.	1998,232	1997,60	-0,632	-0,03 %
9.	2010,174	2001,23	-8,944	-0,45 %
10.	1945,480	1953,88	8,400	0,43 %

TABLE VII. TESTING RESULT

No.	Testing Parameters	Result		
		Standard Deviation		
		0,1	0,2	0,3
1.	MSE	106,13	243,52	812,61
2.	RMSE	10,30	15,61	28,51
3.	MAE	8,73	13,29	23,47

REFERENCES

[1] Avi Nela Vitrina, “Tinjauan Hukum Islam Terhadap Sistem Barter,” Institur Agama Islam Negeri Salatiga, 2017.

[2] S. Endriani, “Konsep Uang: Ekonomi Islam VS Ekonomi Konvensional,” Anterior J., 2015, doi: 10.33084/anterior.v15i1.201.

[3] Bank Indonesia, "Kamus Bank Indonesia," Bank Indonesia, 2020. [Online]. Available: https://www.bi.go.id/id/Kamus.aspx?id=U. [Accessed: 23-Nov-2020].

[4] U. Mardiyati and A. Rosalina, “Analisis Pengaruh Nilai Tukar, Tingkat Suku Bunga dan Inflansi Terhadap Indeks Harga Saham,” J. Ris. Manaj. Sains Indones., 2013.

[5] K. L. Center, “Seri Makro Ekonomi-Perekonomian Terbuka.” [Online]. Available: https://klc.kemenkeu.go.id/seri-makro-ekonomi-perekonomian-terbuka/. [Accessed: 23-Nov-2020].

[6] T. M. Langi, V. Masinambow, and H. Siwu, “ANALISIS PENGARUH SUKU BUNGA BI, JUMLAH UANG BEREDAR, DAN TINGKAT KURS TERHADAP TINGKAT INFLASI DI INDONESIA,” J. Berk. Ilm. Efisiensi, 2014.

[7] KBBI, “KBBI - Kamus Besar Bahasa Indonesia,” Kamus Besar Bhs. Indones., 2007.

[8] S. K. Junaidi, S. Sulasmiyati, and F. Nurlaily, “Pengaruh pendapatan nasional, inflasi dan nilai tukar Yuan terhadap impor Indonesia dari China periode 2010-2017,” Adm. Bisnis, 2018.

[9] S. Sukirno, Teori Pengantar Ekonomi Makro. 2011.

[10] B. P. Statistik, “Eksport Import,” Badan Pusat Statistik, 2020. [Online]. Available: https://www.bps.go.id/exim/. [Accessed: 24-Nov-2020].

[11] Almira Ciptya Rahma, “Prediksi Nilai Tukar Rupiah Terhadap Yuan Menggunakan Adaptive Neuro Fuzzy Inference System,” Institut Teknologi Sepuluh November, 2018.

[12] S. Y. Amalutfia, “Analisis peramalan nilai tukar Rupiah terhadap mata uang Dollar dan Yuan dengan menggunakan metode Fuzzy Time Series Markov Chain,” UIN Sunan Ampel Surabaya, 2020.

[13] N. Kusumawati, F. Marisa, and I. D. Wijaya, “PREDIKSI KURS RUPIAH TERHADAP DOLLAR AMERIKA DENGAN MENGGUNAKAN METODE REGRESI LINEAR,” J I M P - J. Inform. Merdeka Pasuruan, 2017, doi: 10.37438/jimp.v2i3.79.

[14] D. . Specht, "A General Regression Neural Network," IEEE Trans. Neural Networks 2, pp. 568–576, 1991.

[15] R. E. Caraka, H. Yasin, and A. Prahutama, “Pemodelan General Regression Neural Network (GRNN) Pada Data Return Indeks Harga Saham Euro 50,” Gaussian, 2015.

[16] R. E. Caraka, “Pemodelan General Regression Neural Network (GRNN) Dengan Peubah Data Input Return Untuk Peramalan Indeks Hangseng,” in Trusted Digital Indentitiy and Intelligent System, 2014.

[17] R. E. Caraka and H. Yasin, “Prediksi Produksi Gas Bumi Dengan General Regression Neural Network (GRNN),” in National Seminar Statistics IV “ Implementation Statistics in the Exploration Energy,” 2014.

[18] M. Theodosiou, "Disaggregation & aggregation of time series components: A hybrid forecasting approach using generalized regression neural networks and the theta method," Neurocomputing, 2011, doi: 10.1016/j.neucom.2010.10.013.

[19] B. Warsito, A. Rusgiyono, and M. A. Amirillah, “PEMODELAN GENERAL REGRESSION NEURAL NETWORK UNTUK PREDIKSI TINGKAT PENCEMARAN UDARA KOTA SEMARANG,” MEDIA Stat., 2012, doi: 10.14710/medstat.1.1.43-51.

[20] B. Warsito, “PERBANDINGAN MODEL FEED FORWARD NEURAL NETWORK DAN GENERALIZED REGRESSION NEURAL NETWORK PADA DATA NILAI TUKAR YEN TERHADAP DOLAR AS,” in Prosiding SPMIPA, 2006, pp. 127–131.

[21] H. C. So, Y. T. Chan, K. C. Ho, Y. Chen, and T. S. Expansion, "Simple Formulas for Bias and Mean Square Error Computation," no. June, pp. 162–165, 2013.

[22] T. Chai, R. R. Draxler, and C. Prediction, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," no. 2005, pp. 247–1250, 2014, doi: 10.5194/gmd-7-1247-2014.