# Implementation Extreme Learning Machine for Rainfall Forecasting

Laksmita Puspaningrum[1], Ayundyah Kesumawati[2]
14611059@students.uii.ac.id[1], ayundyah.k@uii.ac.id[2]

Statistic Deparment, Universitas Islam Indonesia, Yogyakarta 55584, Indonesia[1,2]

**Abstract.**Indonesia is one of country that have a large number of rainfall days in a year. So, make any sense that forecasting is important to get any strategies to overcome the problem of erratic rainfall. There are a lot of method that can conduct forecast rainfall, the new one is Extreme Machine Learning. Extreme Learning Machine (ELM) is a new learning method of artificial neural networks. ELM is an easy-to use and effective learning algorithm of single-hidden layer feed-forward neural networks (SLFNs). Therefore, ELM has the advantages of fast learning speed and good generalization performance. This research conducted by using rainfall data in Sleman city to get forecasting in one year. It found that the ELM method has a smaller value compared to the GARCH method for all six rainfall station stations in Sleman region.

**Keywords:** Rainfall, Forecasting, Extreme Learning Machine,Neural Network, GARCH.

## 1 Introduction

This Rainfall is the amount of rain water that falls in an area inside certain time. Rainfall in Indonesia is variable and complex, besides that factors that affect rainfall in Indonesia are temperature factors cloud tops, latitude, altitude, water source distance, temperature difference land and land area. Forecasting rainfall is one important requirement for planning and management of many applications, including, agriculture, flood and traffic [1].

The highest amount of rainfall in Indonesia generally occurs in September – March, while the lowest amount of rainfall occurs in April – August. But Badan Meteorologi dan Geofisika (BMKG) Yogyakarta recently announced that there is a condition in most regions of Indonesia had a severe heavy rainfall in second week November 2017 and were also the wettest in 2016 in same month. BMKG predicted that the peak of rainy season or heavy rainfall occurs in December 2017 – February 2018. It is necessary to be aware the impact of heavy rainfall that will be any disaster such as floods, landslides, storm, and the other. references.

There are several indications caused of heavy rainfall and the uncertainty of climate/weather changes in Indonesia, one of is tropical cyclone seeds in the South Indian Ocean of East Java. With the uncertainty of changes in rainfall in Indonesia, serious attention, supervision and handling is needed to reduce the risk of the impact of the disaster caused by heavy rainfall. Based on picture 1 it can be seen that the highest rainfall is in November 2016 for Sleman Region with an average annual rainfall of 351,33 mm.
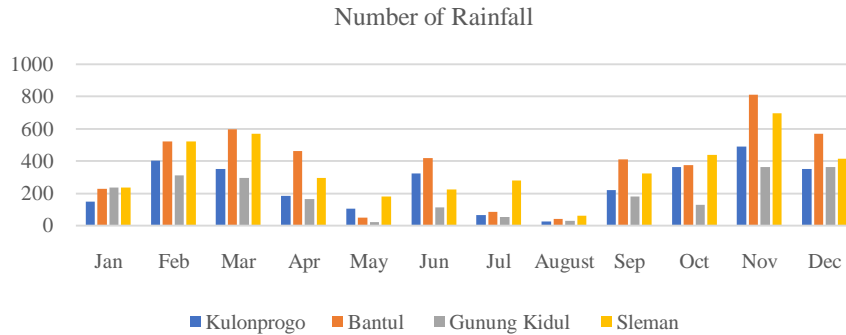
**Fig. 1.**Rainfall in Yogyakarta District on 2016

Forecasting is a prediction of some future event or events. There are many methods in forecasting that can used to predict the rainfall and make an early warning to handle with extreme weather. One of them is Extreme Learning Machine. Extreme Learning Machine (ELM) is a new learning method from Artificial Neural Network (ANN) [2]. Extreme Learning Machine is an ANN feedforward by using the Single Hidden Layer Feedforward Neurons Networks (SLFNs). This method used to overcome the weakness of the network artificial nerves especially in learning speed. The advantages of the Extreme Learning Machine is suitable for almost all data with nonlinear function and continuous function part by part which is like a disconnected, differential part non function differential. And it suitable for complex activation functions.

Based on the problem above, this research will conduct to determine forecasting weekly rainfall in Sleman Regency using Extreme Learning Machine and comparing with time series method to get the best result.

## 2 Method

### 2.1 Extreme Learning Machine

Extreme Learning Machines (ELMs) is the name for a collection of neural network models, which employ randomization of the hidden layer weights and a fast training algorithm [3].The ELMs function represented by the following equation

$$f_L(x) = \overset{L}{\underset{i=1}{\overset{\circ}{a}}} b_i h_i(x) = h(x)b \tag{1}$$

Where $b = \begin{bmatrix} b_1,...,b_L \end{bmatrix}^T$ is represents the output weight vector between the L nodes hidden layer to the $m \, 3 \, 1$ output nodes,
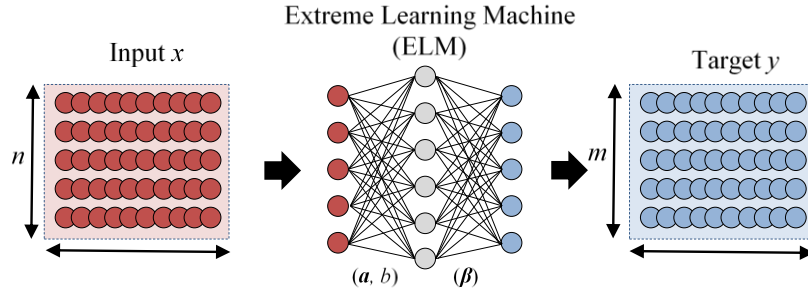
**Fig. 2.** Extreme Learning Machine Algorithm

## 2.2 Generalized Auto-Regressive Conditionally Heteroscedastic (GARCH)

In parallel to the idea of allowing for time-varying variances in a sequence of normal distribution underlying the basic ARCH model, it is possible to allow the scale parameter in a sequence of Stable Paretian distributions to change over time [4]. The GARCH (p, q) model is defined by [5]:

$$y_k = \sigma_k \varepsilon_k \qquad (2)$$

$$S_k^2 = W + \sum_{i=1}^{p} a_i y_{k-i}^2 + \sum_{j=1}^{q} b_j S_{k-j}^2 \qquad (3)$$

Where $W > 0, a_i \geq 0, b_j \geq 0$ and the innovation sequence $\{e_i\}_{i=-\infty}^{\yen}$ is independent and identically distributed with $E(e_0) = 0$ and $E(e_0^2) = 1$. ARCH and GARCH model have been developed to a wide range of time series model, its succes and have been developed in many studies in finance [6].

## 2.3. Flowchart
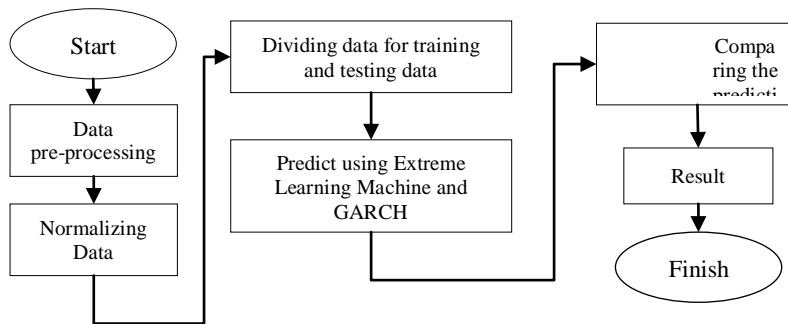The flowchart of this research as follow:



**Fig. 3.** Flowchart of the research

Based on Figure 3, this research conducted by the amount raifall in Yogyakarta. start with data preprocessing, normalizing data, dividing data for training and testing, predict using ELM and GARCH, comparing the result and conclusion.

## 3  Result and Discussion

This section contained Descriptive Analysis the time series of Rainfall of 6 district in Sleman Regency, result of GARCH Model and Extreme Learning Machine and result of comparison between ELM and GARCH.

### 3.1    Descriptive Analysis

The descriptive of time series data conducted by figure 4 as follow:
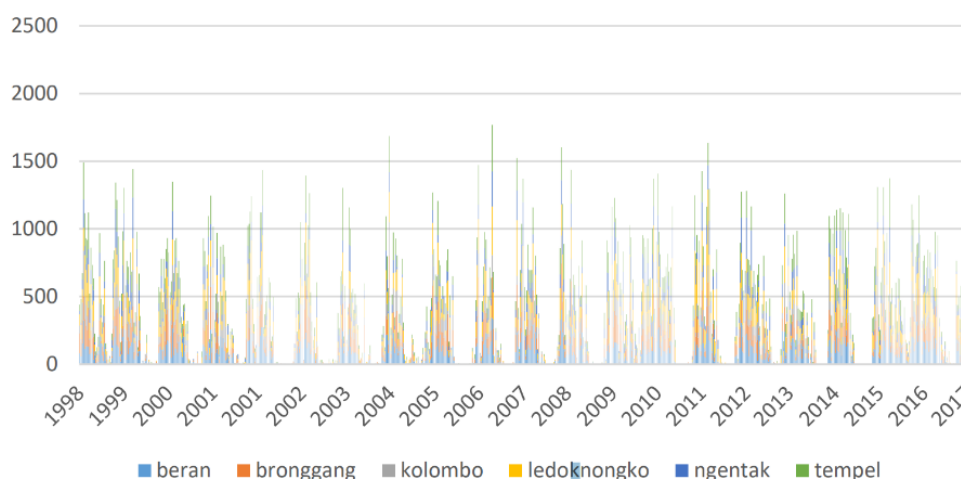


**Fig. 4.**Descriptive of Rainfall in Sleman District

It can be seen from the figure that the x axis is the year from 1998 to 2017, and the y axis is the amount of weekly rainfall in millimeters (mm). Based on Figure 5.1, it can be seen that weekly rainfall in Sleman Regency from 1998 to 2017 which has rainfall including heat, namely in 1998, 2010, 2016 and 2017. This is because there is a phenomenon of climate change in Indonesia, namely the La Niña, La Weak to moderate Nina contributes to higher rainfall than usual throughout Indonesia. The continued impact of La Niña, normal rainfall is expected to occur above normal, increasing the risk of flooding, landslides and associated damage. Considering that Indonesia has experienced an unusually high number of floods and landslides since mid-2016. In addition, at the end of 2017 Cempaka typhoon brought heavy rain and extreme winds to the coast of southern Java which also occurred in the Special Region of Yogyakarta. (BMKG, 2018)

### 3.2    Extreme Learning Machine Result

First step is dividing data became training and testing data. As it can be seen that there are 6 districts so the result as follow:

Table 1 Training and Testing Data

| Rainfall Station | Training | % Training | Testing | % Testing |
|---|---|---|---|---|
| Beran | 683 | 94.86% | 37 | 5.14% |
| Bronggang | 684 | 95% | 36 | 5% |
| Kolombo | 691 | 95.97% | 29 | 4.03% |
| Ledoknongko | 683 | 94.86% | 37 | 5.14% |
| Ngentak | 684 | 95% | 36 | 5% |
| Tempel | 684 | 95% | 36 | 5% |

First, after dividing the dataset became training and testing here there are the network architecture of Extreme Leaning Machine for testing data:
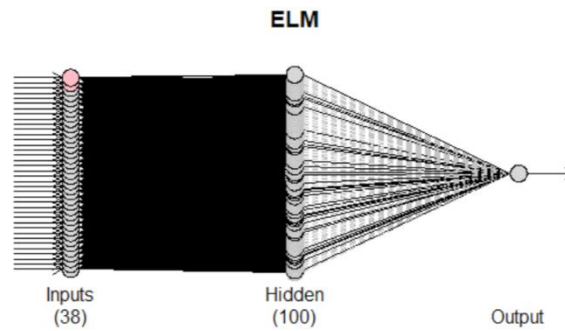


Fig. 5.Network Architecture of ELM

based on figure 5 it can be seen that there are 38 inputs and 100 hidden layer to predict one item in this case is one period. Here is the result of forecasting for 1 period.
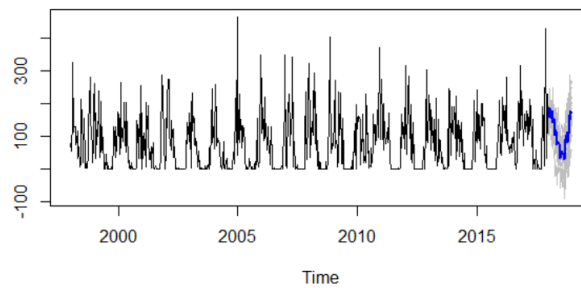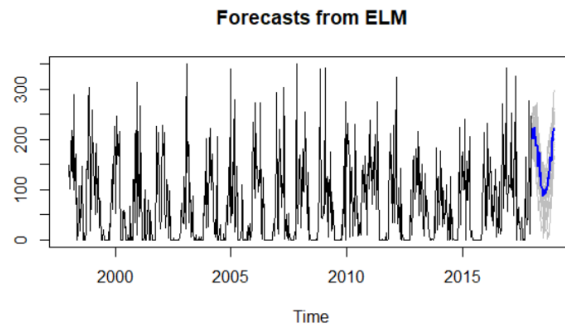


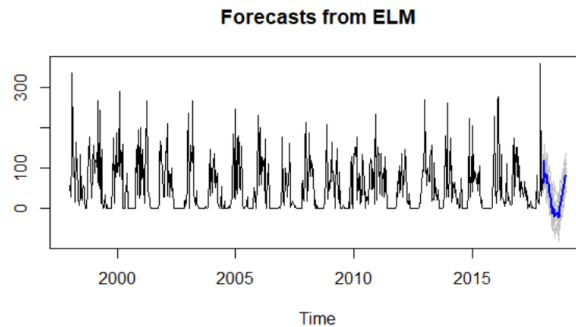Fig. 6.Beran Station Forecasting Plot

Factual data is shown by a black chart, while the blue chart shows the data forecasting rainfall weekly Beran station in a year. Rainfall forecasting results has an average of 110.44. The lowest rainfall results occur in the second week in a month at 29.44 mm, while the highest rainfall occurs in January 3rd week with an estimated amount of rainfall per week of 187.21 mm.
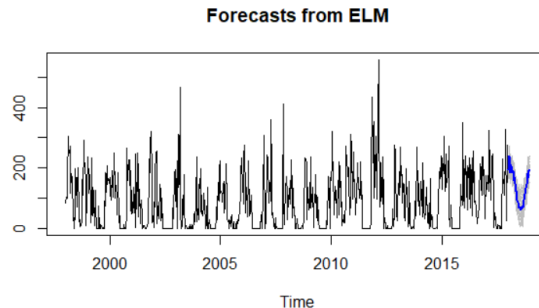


**Fig. 7.** Bronggang Station Forecasting Plot

Rainfall forecasting results weekly has an average of 164.78. The lowest rainfall results occur in the second weekly month at 87.99 mm, while the highest rainfall occurs in January 3rd weekly with an estimated amount of rainfall per dasarian of 224.36 mm.
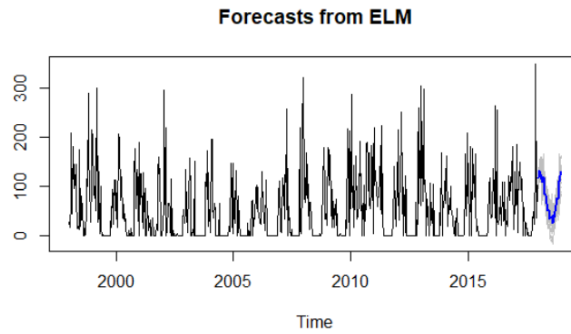


**Fig. 8.** Kolombo Station Forecasting Plot

Rainfall forecasting results of Kolombo station weekly has an average of 31.97. The lowest rainfall results occur in the second weekly month at 23.21 mm, while the highest rainfall occurs in January 3rd weekly with an estimated amount of rainfall per dasarian of 116.46 mm.
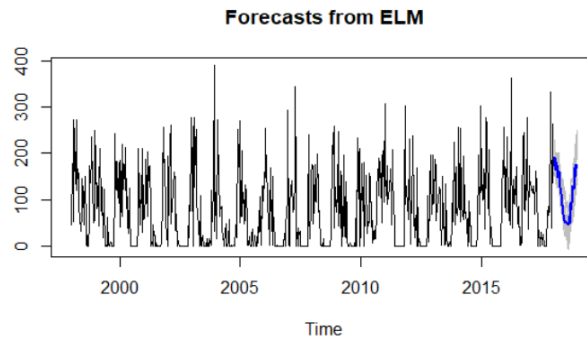
Rainfall forecasting results of Ledongnongko station weekly has an average of 139.23. The lowest rainfall results occur in the second weekly month at 61.27 mm, while the highest rainfall occurs in January 3rd weekly with an estimated amount of rainfall per weekly of 240.71 mm.



**Fig. 10.** Ngentak Station Forecasting Plot

Rainfall forecasting results of Ngentak station weekly has an average of 79.34. The lowest rainfall results occur in the second weekly month at 25.5 mm, while the highest rainfall occurs in January 3rd weekly with an estimated amount of rainfall per weekly of 132.91 mm



**Fig. 11.** Tempel Station Forecasting Plot

Rainfall forecasting results of Tempel station weekly has an average of 115.65 mm. The lowest rainfall results occur in the second weekly month at 45.73 mm, while the highest rainfall occurs in January 3rd weekly with an estimated amount of rainfall per weekly of 190.51 mm

Table 2 Comparison GARCH and ELM

| Station | Method | MSE | RMSE | MAE |
|---------|--------|------|------|-----|
| Beran | GARCH | 5889.03 | 76.74 | 51.14 |
| | ELM | 4813.79 | 69.38 | 46.24 |
| Bronggang | GARCH | 4332.01 | 65.82 | 45.04 |
| | ELM | 4623.78 | 68 | 47.22 |
| Kolombo | GARCH | 3316.98 | 57.59 | 36.59 |

| | | | | |
|---|---|---|---|---|
| | ELM | 2691.63 | 51.88 | 30.88 |
| Ledoknongko | GARCH | 7512.95 | 86.68 | 57.42 |
| | ELM | 6015.52 | 77.56 | 48.3 |
| Ngentak | GARCH | 4238.81 | 65.11 | 42.24 |
| | ELM | 3242.44 | 56.94 | 34.08 |
| Tempel | GARCH | 5421.89 | 73.63 | 50.98 |
| | ELM | 4372.15 | 66.12 | 43.47 |

From the results of table 2. It can be seen that the results of accuracy with the ELM method are smaller than using the model from the SARIMA / GARCH method. This means that the ELM method is suitable for analyzing rainfall data in Sleman Regency.

## 4      Conclusion

From results of the rainfall forecasting comparison with the original data that the ELM method has a smaller error value compared to the error value with the SARIMA / GARCH method in 5 stations, namely the Beran, Colombo, Ledoknongko, Ngentak and Tempel rainfall stations. However forecasting results with the ELM method cannot be used to forecast weekly rainfall data because there are illogical forecasting results, because there are forecasting values that are negative and there are several rainfall values that have a considerable difference between the forecasting value and the original data.

## 5      References

[1]M.S. Sreekanth, R. Rajesh and J. Satheeshkumar.: Extreme Learning Machine for the Classification of Rainfall and Thunderstorm. Journal of Applied Sciences, 15: 153-156. (2015).

[2] Albadr, M. A. A., & Tiun, S.: Extreme Learning Machine A Review. International Journal of Applied Engineering Research Vol 12. (2017)

[3]Huang, G. B., Zhu, Q. Y., & Siew, C. K. : Extreme Learning Machine Theory and Application. Neurocomputing 70, 489-501. doi:10.1016/j.neucom.2005.12.126. (2006).

[4] Engle, R.: Autoregressive Conditional Heterosckedasticity with Estimates of the Variance of United Kingdom Inflation. journal of 111 Econometric Society, vol. 50(No. 4), 987-1007. Retrieved from http://www.jstor.org/stable/1912773 (2006).

[5] Bollerslev.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307-327. (1986)

[6] Engle, R.: The Use of ARCH/GARC Models in Applied Econometrics. Journal of Economic Perspectives Volume 15 Number 4. (2001).