Analysis of Artificial Neural Network Backpropagation Using Conjugate Gradient Fletcher Reeves In The Predicting Process

Anjar Wanto¹, Muhammad Zarlis², Sawaluddin³, Dedy Hartama⁴, Jaya Tata Hardinata⁵, Herlan F Silaban⁶

¹¹¹ STIKOM Tunas Bangsa, Jl. Sudirman Pematangsiantar, North Sumatera, 21127
¹²² University of North Sumatera, Faculty of Computer Science and Information Technology, University of Sumatera, Medan, Indonesia, 20155
³³³ Faculty of MIPA, University of North Sumatera, Jl. Dr. Mansur 9 Medan Indonesia
⁶⁶⁶ AMIK Parbina Nusantara, Jl. Sutomo Pematangsiantar, North Sumatera 21136

Email: anjarwanto@amiktunasbangsa.ac.id, m.zarlis@yahoo.com, sawal@usu.ac.id, dedyhartama@amiktunasbangsa.ac.id, jayatatahardinata@gmail.com, herlan76ban@gmail.com

Abstract. Backpropagation is a good artificial neural network algorithm used to predict, one of which is to predict the rate of Consumer Price Index (CPI) based on the foodstuff sector. While conjugate gradient fletcher reeves is a suitable optimization method when juxtaposed with backpropagation method, because this method can shorten iteration without reducing the quality of training and testing result. Consumer Price Index (CPI) data that will be predicted to come from the Central Statistics Agency (BPS) Pematangsiantar. The results of this study will be expected to contribute to the government in making policies to improve economic growth. In this study, the data obtained will be processed by conducting training and testing with artificial neural network backpropagation by using parameter learning rate 0.01 and target error minimum that is 0.001-0.09. The training network is built with binary and bipolar sigmoid activation functions. After the results with backpropagation are obtained, it will then be optimized using the conjugate gradient fletcher reeves method by conducting the same training and testing based on 5 predefined network architectures. The result, the method used can increase the speed and accuracy result.

1. Introduction

Artificial Intelligence is one area that is quite reliable in solving problems such as prediction (forecasting). Artificial Neural Network (ANN) is one of the artificial representations of the human brain that always tries to simulate the learning process in the human brain [1][2]. ANN approach can imitate any complex and non-linear relationship through non-linear units (neurons) and has been widely used in the forecasting area [3]. Prediction (forecasting) is basically a presumption about the occurrence of an event or event in the future. Prediction (forecasting) is very helpful in planning and decisionmaking activities of a policy [4]. One of the sub areas of Artificial Intelligence that can be relied upon in doing a prediction is the artificial neural network backpropagation. To produce a good Backpropagation, the selected parameters must be precise. Therefore, an algorithm that can help accelerate Backpropagation training, one of the reliable algorithms is Conjugate Gradient Fletcher Reeves. This algorithm is expected to improve the performance of the system, because the algorithm is
able to optimize so that it can minimize a function, where the search is based on the direction of conjugation orthogonal value [5]. Due to its orthogonal searching nature, so this algorithm can quickly reach convergence on the solution sought. These methods will be used to predict. The data to be predicted is the data of Consumer Price Index of Pematangsiantar that come from the Central Bureau of Statistics. The Consumer Price Index is one of the common economic indicators used to measure the rate of price change (inflation / deflation) at the consumer level, especially in urban areas. In Indonesia, the inflation rate is measured from the percentage change of the Consumer Price Index and announced to the public at the beginning of each month (first business day) [6]. The expenditure group from the Consumer Price Index can be seen in Table 1 below:

| No | Group                                             |
|----|---------------------------------------------------|
| 1  | Foodstuff                                         |
| 2  | Food So, Drink, Cigarette And Tobacco             |
| 3  | Housing, Water, Electricity, Gas and Fuel         |
| 4  | Clothing                                          |
| 5  | Health                                            |
| 6  | Education, Recreation and Sports                  |
| 7  | Transport, Communication and Financial Services    |

Table 1. Consumer Price Index Expenditure Groups

Based on table 1, each group consists of several subgroups. The subgroup data from each sector of Consumer Price Index can be seen in table 2 below:

| No | Group                          | Subgroups                                                                 |
|----|--------------------------------|--------------------------------------------------------------------------|
| 1  | Foodstuff                      | Grains (Bulbs), Meat, Fresh Fish, Fish Preserved, Eggs and Milk, Vegetables, Nuts, Fruits, Seasonings, Fats and Oils, Other Foodstuffs |
| 2  | Food So, Drink, Cigarette And Tobacco | Food, Non-Alcoholic Beverages, Tobacco and Alcoholic Beverages             |
| 3  | Housing, Water, Electricity, Gas and Fuel       | Housing, Fuel (Illumination) and Water Supply, Household Supplies, Household Implementation |
| 4  | Clothing                         | Men’s Clothing, Women’s Clothing, Kids clothing, Personal Goods and other Clothing |
| 5  | Health                           | Medical Services, Pharmaceuticals, Physical Services, Physical and Health Care |
| 6  | Education, Recreation and Sports  | Educational Services, Courses / Training, Educational Equipment / Equipment, Sports Recreation |
| 7  | Transport, Communication and Financial Services | Transport, Communication and Shipping, Transport Support Facilities, Financial Services |

Table 2. Subgroups of the Consumer Price Index Sector

Table 3. Previously Research

| No | Author        | Title                                           | Results | Deficiency          |
|----|---------------|-------------------------------------------------|---------|---------------------|
| 1  | Ramanda, K [7]| Prediction of Premature Birth Using Neural Network and PSO | 95%     | Learning rate is too big ie 0.4 |
2. Material and Methods

2.1. Data Used

The data used in this paper is the Consumer Price Index (CPI) data based on the Foodstuffs of Pematangsiantar from 2014 to 2016, from January to December. Consumer Price Index Data (CPI) is derived from the Central Bureau of Statistics Pematangsiantar.

Table 4. Input Data Used Before Normalization

| Year | Month | Consumer Price Index 2014-2016 |
|------|-------|-------------------------------|
|      | Jan   | Feb   | Mar   | Apr   | May   | Jun   | Jul   | Aug   | Sep   | Oct   | Nov   | Dec   |
| 2014 | 116.22| 116.03| 117.54| 113.89| 120.12| 119.79| 120.05| 119.22| 119.98| 123.53| 126.17| 127.07|
| 2015 | 125.95| 119.60| 118.37| 117.91| 122.16| 127.04| 125.53| 124.15| 122.17| 122.85| 123.72| 128.40|
| 2016 | 130.65| 128.53| 130.70| 128.30| 130.83| 131.66| 134.01| 135.67| 138.00| 141.85| 144.06| 148.06|

Based on table 3. It can be explained that, the Consumer Price Index (CPI) dataset based on Foodstuff Sector on 2014-2015 is used as training data, while dataset on 2015-2016 is used as testing data. The data presented are January to December data each year.

2.2. Stage Data Processing

At this writing, it was Created a pattern recognition system and predicted the Consumer Price Index (CPI). This process has two stages where the first stage is to do pattern recognition by finding the best architecture of the Artificial Neural Network model created. The process of training and testing the data to get the best model is obtained from the Consumer Price Index (CPI) data at the Central Bureau of Statistics Pematangsiantar as much as 24 Patterns counted from 2014 to 2016, with the criteria Pattern 1 to Pattern 12 is training data (training) , While Pattern 13 through Pattern 24 is the Test data. The second stage is to make predictions with the best architectural patterns obtained in the first stage.

2.3. Data Normalization

Normalization of the data is done so that the output of the network appropriate to the activation function used. The activation function that the author uses in this writing is the sigmoid activation function. Sigmoid function is asymptotic function (never reach 0 or 1), then data transformation is done at smaller interval [0,1; 0.9], shown by the equation (following) [9].

\[
x' = \frac{0.8(x - a)}{b - a} + 0.1 \quad \ldots \ldots \ldots (1)
\]

Based on table 4, Then we will get the result of transformation as follows:
Table 5. Training Data After Normalization

| Data | Input | Target |
|------|-------|-------|
|      | Jan   | Feb   | Mar   | Apr   | May   | Jun   | Jul   | Aug   | Sep   | Oct   | Nov   | Dec   |
| 1    | 0.2285| 0.2180| 0.3012| 0.1000| 0.4435| 0.4253| 0.4396| 0.3939| 0.4358| 0.6315| 0.7771| 0.8267| 0.7649|
| 2    | 0.2180| 0.3012| 0.1000| 0.4435| 0.4253| 0.4396| 0.3939| 0.4358| 0.6315| 0.7771| 0.8267| 0.7649| 0.4148|
| 3    | 0.3012| 0.1000| 0.4435| 0.4253| 0.4396| 0.3939| 0.4358| 0.6315| 0.7771| 0.8267| 0.7649| 0.4148| 0.3470|
| 4    | 0.1000| 0.4435| 0.4253| 0.4396| 0.3939| 0.4358| 0.6315| 0.7771| 0.8267| 0.7649| 0.4148| 0.3470| 0.3216|
| 5    | 0.4435| 0.4253| 0.4396| 0.3939| 0.4358| 0.6315| 0.7771| 0.8267| 0.7649| 0.4148| 0.3470| 0.3216| 0.5560|
| 6    | 0.4253| 0.4396| 0.3939| 0.4358| 0.6315| 0.7771| 0.8267| 0.7649| 0.4148| 0.3470| 0.3216| 0.5560| 0.8250|
| 7    | 0.3939| 0.4358| 0.3939| 0.4358| 0.6315| 0.7771| 0.8267| 0.7649| 0.4148| 0.3470| 0.3216| 0.5560| 0.7418|
| 8    | 0.3939| 0.4358| 0.3939| 0.4358| 0.6315| 0.7771| 0.8267| 0.7649| 0.4148| 0.3470| 0.3216| 0.5560| 0.6657|
| 9    | 0.4358| 0.6315| 0.7771| 0.8267| 0.7649| 0.4148| 0.3470| 0.3216| 0.5560| 0.8250| 0.7418| 0.6657| 0.5565|
| 10   | 0.6315| 0.7771| 0.8267| 0.7649| 0.4148| 0.3470| 0.3216| 0.5560| 0.8250| 0.7418| 0.6657| 0.5565| 0.5940|
| 11   | 0.7771| 0.8267| 0.7649| 0.4148| 0.3470| 0.3216| 0.5560| 0.8250| 0.7418| 0.6657| 0.5565| 0.5940| 0.6420|
| 12   | 0.8267| 0.7649| 0.4148| 0.3470| 0.3216| 0.5560| 0.8250| 0.7418| 0.6657| 0.5565| 0.5940| 0.6420| 0.9000|

Training Data on 2014-2015 is done using rotary rotation, it means each dataset has the same rights to achieve the target. The data value in data 1 is taken from the CPI of the foodstuff sector in 2014. While the Target value is taken from the CPI of the foodstuff sector in January 2015. The data values in data 2 are taken from the CPI of the foodstuff sector in 2014 from February to December and the month dataset January 2015. Target value in data 2 was taken from the CPI of the foodstuff sector in February 2015. Data values in data 3 were taken from the CPI of the food sector in 2014 March-December and the dataset from January to February 2015. Target value on data 3 is taken from the CPI food sector in March of 2015. So on until all values are completed in play. The maximum value (b) of the dataset is 128.40. While the minimum value (a) is 113.89.

By using the normalization formula as follows:

\[ u = \frac{x - \text{min}(x)}{\text{max}(x) - \text{min}(x)} \]

Then will get result Normalization data 1 for January 0.22885. So on for all data, normalized by using the same formula in equation (1).

Table 6. Testing Data After Normalization

| Data | Input | Target |
|------|-------|-------|
|      | Jan   | Feb   | Mar   | Apr   | May   | Jun   | Jul   | Aug   | Sep   | Oct   | Nov   | Dec   |
| 1    | 0.3460| 0.1517| 0.1141| 0.1000| 0.2300| 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298|
| 2    | 0.1517| 0.1141| 0.1000| 0.2300| 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249|
| 3    | 0.1141| 0.1000| 0.2300| 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249| 0.4913|
| 4    | 0.1000| 0.2300| 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249| 0.4913| 0.4179|
| 5    | 0.2300| 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249| 0.4913| 0.4179| 0.4953|
| 6    | 0.2300| 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249| 0.4913| 0.4179| 0.4953|
| 7    | 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249| 0.4913| 0.4179| 0.4953| 0.5207|
| 8    | 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249| 0.4913| 0.4179| 0.4953| 0.5207|
| 9    | 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249| 0.4913| 0.4179| 0.4953| 0.5207|
| 10   | 0.3793| 0.3331| 0.2909| 0.3030| 0.2511| 0.2777| 0.4209| 0.4298| 0.4249| 0.4913| 0.4179| 0.4953| 0.5207|

3. Analysis And Results
3.1. Analysis
Previously, the data to be tested should be divided into two (2) sections, where the first part is for training data and the second part is for the test data. Standard Backpropagation method uses gradient decrease algorithm (descent gradient). Variations on the standard model are done by replacing the algorithm with another algorithm. The parameters used in this training and testing: Architecture = 1 hidden layer, Input neurons = 12, Activation Function = Sigmoid, Initialization weights = Random, Target Error Minimum= 0.001 – 0.09, Maximum Epoch = 10000, Learning Rate = 0.01.

3.2. Results
This research uses 5 architecture. Among others are 12-6-1, 12-15-1, 12-24-1, 12-33-1, and 12-34-1. From the 5 architectures, this is the best architecture is 12-15-1. Whether it is using standard backpropagation or merging between backpropagation with conjugate gradient fletcher reeves. The architecture is better because of the high accuracy and stable results, the lesser Epoch and the smaller MSE (Mean Square Error).

![Training and testing with standard backpropagation](image-url)
Figure 2. Training and testing with backpropagation with Fletcher Reeves

Table 7. Architectural Results Backpropagation and Backpropagation + CGFR

| Method                            | Architecture | Epoch  | MSE (Mean Square Error) | Accuracy |
|-----------------------------------|--------------|--------|-------------------------|----------|
| Backpropagation                   | 12-6-1       | 5308   | 0.0228736153            | 50%      |
|                                   | 12-15-1      | 821    | 0.0142803691            | 75%      |
|                                   | 12-24-1      | 4999   | 0.0205544065            | 58%      |
|                                   | 12-33-1      | 961    | 0.0366150738            | 25%      |
|                                   | 12-34-1      | 1491   | 0.0175775766            | 25%      |
| Backpropagation + Conjugate       | 12-6-1       | 6      | 0.0185594035            | 58%      |
| Gradient Fletcher Reeves          | 12-15-1      | 2      | 0.0090116088            | 67%      |
|                                   | 12-24-1      | 15     | 0.0261824790            | 50%      |
|                                   | 12-33-1      | 16     | 0.0162138928            | 42%      |
|                                   | 12-34-1      | 149    | 0.0356825278            | 33%      |

Figure 3. Graph of Iteration and Accuracy between standard backpropagation with BP + CGFR

The results of standard backpropagation methods have been optimized using Conjugate Gradient Fletcher Reeves. Optimization is done by using the Traincgf function. Traincgf is a network function that updates the weights and biases according to the backpropagation gradient convergence with the Fletcher-Reeves update. Based on the results of tests conducted with 5 experiments on backpropagation with different architectures, there was an average of 2716 iterations. While with the same architecture using Backpropagation with Conjugate Gradient Fletcher Reeves there was an increase of iteration with an average of 38. For accuracy of training conducted with the same 5 experiments, there was increased accuracy of training in the first (1st) experiment with the network architecture 12-6-1, the fourth (4th) experiment with 12-33-1 network architecture and fifth (5th) experiment with 12-34-1 network architecture. The decline in the level of accuracy of training in second (2nd) experiment with network architecture 12-15-1 and third (3rd) experiment with the network architecture 12-24-1. For accuracy testing conducted with the same 5 experiments, increased accuracy testing in first (1st) experiment with network architecture 12-6-1, fourth (4th) experiment with network architecture 12-33-1 and fourth (5th) experiment with network architecture 12-34-1. The decline in the level of accuracy testing in second (2nd) experiment with the network architecture 12-15-1 and third (3rd) experiment with the network architecture 12-24-1.
4. Conclusion
From the above discussion can be drawn conclusion with the same architectural parameters on the backpropagation that the backpropagation network architecture affect the ups and downs of the rate of learning and the level of accuracy of training and testing.

References
[1] A. Wanto, A. P. Windarto, D. Hartama, and I. Parlina, “Use of Binary Sigmoid Function And Linear Identity In Artificial Neural Networks For Forecasting Population Density,” International Journal Of Information System & Technology (IJISTECH), vol. 1, no. 1, pp. 43–54, 2017.
[2] S. F. Weng, J. Reps, J. Kai, J. M. Garibaldi, and N. Qureshi, “Can Machine-learning improve cardiovascular risk prediction using routine clinical data?,” PLoS ONE, vol. 12, no. 4, pp. 1–15, 2017.
[3] D. Huang and Z. Wu, “Forecasting outpatient visits using empirical mode decomposition coupled with backpropagation artificial neural networks optimized by particle swarm optimization,” PLoS ONE, vol. 12, no. 2, pp. 1–18, 2017.
[4] S. P. Siregar and A. Wanto, “Analysis Accuracy of Artificial Neural Network Using Backpropagation Algorithm In Predicting Process ( Forecasting ),” International Journal Of Information System & Technology (IJISTECH), vol. 1, no. 1, pp. 34–42, 2017.
[5] X. Li, W. Zhang, and X. Dong, “A class of modified FR conjugate gradient method and applications to non-negative matrix factorization,” Computers & Mathematics with Applications, vol. 73, no. 2, pp. 270–276, 2017.
[6] M. S. Frits Fahridws Damanik, SST and S. Magdalena Sinaga, “Analysis of Consumer Price Index (CPI) Pematangsiantar City,” Economic Census, pp. 1–63, 2015.
[7] K. Ramanda, “Performance Improvement of Neural Network Based Particle Swarm Optimization for Predicting Premature Birth Case Study of Cipto Mangunkusumo General Hospital Jakarta,” in National Seminar on Innovation and Trend (SNIT) 2015, 2015, pp. 203–208.
[8] Amrin, “Analysis of Consumer Price Index (CPI) Pematangsiantar City Comparative Analysis of Neural Network Backpropagation And Multiple Linear Regression For Forecasting Inflation Rate,” Journal of Computer Engineering, vol. II, no. 2, pp. 1–6, 2016.
[9] Sumijan, A. P. Windarto, A. Muhammad, and Budiharjo, “Implementation of Neural Networks in Predicting the Understanding Level of Students Subject,” International Journal of Software Engineering and Its Applications, vol. 10, no. 10, pp. 189–204, 2016.