Multithreading with separate data to improve the performance of Backpropagation method

Mulia Dhamma¹, Muhammad Zarlis², Erna Budhiarti Nababan³

¹, ², ³ Department of Computer Science, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia, 20155

Email: muliadhamma@gmail.com¹, m.zarlis@yahoo.com², ernastp@yahoo.com³

Abstract. Backpropagation is one method of artificial neural network that can make a prediction for a new data with learning by supervised of the past data. The learning process of backpropagation method will become slow if we give too much data for backpropagation method to learn the data. Multithreading with a separate data inside of each thread are being used in order to improve the performance of backpropagation method. Base on the research for 39 data and also 5 times experiment with separate data into 2 thread, the result showed that the average epoch become 6490 when using 2 thread and 453049 epoch when using only 1 thread. The most lowest epoch for 2 thread is 1295 and 1 thread is 356116. The process of improvement is caused by the minimum error from 2 thread that has been compared to take the weight and bias value. This process will be repeat as long as the backpropagation do learning.

1. Introduction

Artificial Neural Network(ANN) is an algorithm that can make prediction which is able to learn from non linear data and also learn data base on time series. ANN does prediction in two ways that is learning by supervised and learning without supervised. Backpropagation method is one of the method that implement the learning by supervised. This method has two phase learning cycle that are propagate forward and propagate backward. This method will propagate forward from input layer to output layer to compare the actual output with the desire output. The error value in output layer will be propagated back from output layer to get a new weight and bias value. The learning process will stop if the error value is lower than the expected error value. The performance of backpropagation method depend on the weight value, bias value, learning rate, expected error value and also data quantity. Learning with many data will reduce the performance of this method. The error that has been produced in each epoch will affect the performance. the smaller the error produced for each epoch then the learning speed of this method will become faster. Adjusting the right weight value will also improve the performance of the performance of backpropagation method. Khairani said that a backpropagation method can be speed up with many technique such as parallel training where the learning process done in parallel. Using single thread for backpropagation to learn many data will consume too much time. Researcher will use multithreading in order to improve the performance of backpropagation method.

2. Literature Review

2.1. Thread

Thread is the smallest unit in a process that can be scheduled by operating system. Multithreading is a type of model execution that can make multiple threads run at a time. The thread runs independently but they share a process that is origin from the main source. Multithreading can also be called threading.
The concept of multithreading consists of two that are concurrency and parallelism. Concurrency is a multithreading process that runs only on one processor where the processor can switch the execution on another thread in a very fast time and can be divided base on the number of cores that exist.

Parallelism is a multithreading process that can run more than one processor / multiprocessor simultaneously which can be called as parallel processing. Any thread on a different processor can run a concurrency process. Normally, parallel processing makes the program run faster because the performance of the processor is not split into each thread. Practically it would be quite difficult to divide the program on different cpu which is also run one data set without any connection between each parallel. Most computers have only one cpu, but there are also computers with multiple cpu, there is also the possibility to do the parallel processing by connecting computers on a computer network but it requires a pretty good software which is called distributed processing software.

3. Research Methodology
Backpropagation method will consume so much time in learning many data so in order to improve the performance of this method the researcher will do multithreading with separating data into each thread. The total data that have been taken are 39 data from www.finance.yahoo.com for Adaro Energy Company(ADRO).
Table 1. Adaro Energy (ADRO) Data

| Number | Date       | X1  | X2  | X3  | T   |
|--------|------------|-----|-----|-----|-----|
| 1      | 2016-12-30 | 1660| 1705| 1660| 1695|
| 2      | 2016-12-29 | 1700| 1705| 1680| 1690|
| 3      | 2016-12-28 | 1695| 1700| 1675| 1690|
| 4      | 2016-12-27 | 1655| 1685| 1655| 1675|
| 5      | 2016-12-23 | 1680| 1705| 1655| 1655|
| 6      | 2016-11-15 | 1595| 1650| 1465| 1465|
| 7      | 2016-11-14 | 1650| 1680| 1605| 1605|
| 8      | 2016-11-11 | 1590| 1695| 1550| 1660|
| 9      | 2016-11-10 | 1610| 1685| 1610| 1650|
| 10     | 2016-11-09 | 1635| 1645| 1475| 1570|
| 11     | 2016-11-08 | 1700| 1720| 1610| 1630|
| 12     | 2016-11-07 | 1685| 1710| 1675| 1685|
| 13     | 2016-10-26 | 1545| 1595| 1545| 1555|
| 14     | 2016-10-25 | 1540| 1540| 1520| 1525|
| 15     | 2016-10-24 | 1500| 1540| 1495| 1525|
| 16     | 2016-10-21 | 1510| 1545| 1450| 1500|
| 17     | 2016-10-20 | 1530| 1540| 1505| 1510|
| 18     | 2016-10-19 | 1485| 1545| 1485| 1530|
| 19     | 2016-10-18 | 1480| 1520| 1440| 1440|
| 20     | 2016-10-17 | 1435| 1480| 1435| 1460|
| 21     | 2016-10-14 | 1405| 1425| 1390| 1425|
| 22     | 2016-10-13 | 1415| 1430| 1405| 1405|
| 23     | 2016-09-06 | 1230| 1245| 1215| 1225|
| 24     | 2016-09-05 | 1180| 1220| 1180| 1220|
| 25     | 2016-09-02 | 1175| 1190| 1160| 1170|
| 26     | 2016-09-01 | 1150| 1190| 1135| 1185|
| 27     | 2016-08-31 | 1145| 1190| 1140| 1150|
| 28     | 2016-07-11 | 885 | 940 | 885 | 925 |
| 29     | 2016-07-01 | 860 | 895 | 860 | 885 |
| 30     | 2016-06-30 | 870 | 875 | 850 | 850 |
| 31     | 2016-06-29 | 865 | 885 | 845 | 855 |
| 32     | 2016-06-28 | 835 | 865 | 830 | 855 |
| 33     | 2016-06-27 | 830 | 840 | 820 | 835 |
| 34     | 2016-06-24 | 850 | 855 | 805 | 840 |
| 35     | 2016-06-23 | 760 | 825 | 760 | 815 |
| 36     | 2016-06-02 | 725 | 770 | 720 | 755 |
| 37     | 2016-06-01 | 710 | 725 | 705 | 720 |
| 38     | 2016-05-31 | 700 | 710 | 690 | 710 |
| 39     | 2016-05-30 | 700 | 710 | 695 | 695 |
| 40     | 2016-01-04 | 515 | 515 | 494 | 494 |
3.1. Research Plan
   The researcher will compare the backpropagation with 2 thread learning process and the
   backpropagation with the normal learning process.
   The value that will be used in backpropagation are :
   • Weight and bias value = random value between (-0.5) – (0.5).
   • Neuron input = 3.
   • Hidden neuron = 4.
   • Output neuron = 1.
   • Error = 0.001.
   • Max epoch = 100000.
   • Learning rate = 1.

3.2. Learning Process
   The data that has been randomized to be inserted into each thread will be normalized using the min-
   max normalization method. Chamidah said that the min-max method can increase the speed of
   convergence by scaling its value between 0 and 1.

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

   \(x'\) = normalization result, \(x\) = original value, \(a\) = lowest value, \(b\) = highest value.
   The parameter will consist of :
   • X1 = open price.
   • X2 = highest price.
   • X3 = lowest price.
   • Y = target.

   Each thread contains data in unorder sequence.
   Learning process with normal backpropagation will run only on 1 thread. Learning process with 2
   thread will start the learning of 2 thread in every epoch simultaneously. Every epoch in a thread can
   possibly stop the learning process if the error is lower than the expected. If not than the learning will
   continue to the next epoch again. The error in 2 learning thread will be compared in each epoch. The
   thread with the most lowest error will be selected. The selected thread will replace all its current
   weight and bias value to other thread except itself. This learning process of the thread will continue
   until the error is lower than the expected. If the learning process had done then the new data will be
   tested.

3.3. Normalization process
   The normalization process using mix-max normalization will return a value between 0 – 1. The lowest
   value = 437 and the highest value = 1850.

| No | Date       | X1         | X2         | X3         | T            |
|----|------------|------------|------------|------------|--------------|
| 1  | 2016-12-30 | 0.8903     | 0.8762     | 0.8762     | 0.8868       |
| 2  | 2016-12-29 | 0.8620     | 0.8832     | 0.8620     | 0.8762       |
| 3  | 2016-12-28 | 0.8620     | 0.8832     | 0.8620     | 0.8762       |
3.4. Process of Separating Data
The data that has been normalized will be shuffled and divided equally into each thread.

Table 3. Separating Data Into 3 Thread For 39 Data

|   | P1     | P2     | P3     |
|---|--------|--------|--------|
| 13 data |        | 13 data | 13 data |

4. Result And Discussion
The research that has been done for stock predicting is used multithreading. The total data is 39 data which is took from www.finance.yahoo.com. The total testing are 5 times experiment with separate data into 2 thread.
The result of normal learning backpropagation for the smallest epoch are in the second testing = 356116 epoch and the biggest in the first testing = 587263. The average epoch is 453049.8
The result of 2 thread learning backpropagation for the smallest epoch are in the forth testing = 1295 epoch and the biggest in the second testing = 13600. The average epoch is 6490.

Tabel 4. Total Epoch for Normal Learning backpropagation & 2 Thread Learning

| Model                  | Testing | Total Epoch |
|------------------------|---------|-------------|
| 1 Thread / Normal Learning | 1       | 587263      |
|                        | 2       | 356116      |
|                        | 3       | 364589      |
|                        | 4       | 572351      |
|                        | 5       | 384930      |
| 2 Thread               | 1       | 6650        |
|                        | 2       | 13600       |
|                        | 3       | 4970        |
|                        | 4       | 1295        |
|                        | 5       | 5935        |

5. Conclusion
Learning process in backpropagation using 2 thread with separate data can decrease the number of epoch. Data will be normalized before the learning process begin with min – max normalization which change the value between 0 and 1, then continue to separate the date into each thread. The result showed that the process of improvement is caused by the minimum error from 2 thread that has been compared to take the weight and bias value in each epoch for the most lowest error value and replace it on another thread.
6. Reference

[1] Gonzales J F 2012 *Java 7 Concurrency Cookbook* (Birmingham: Packt Publishing)

[2] Silberschatz A, Galvin P B & Gagne G 2013 *Operating System Concepts* (New Jersey: John Wiley & Sons)

[3] Khairani M 2014 *Improvisasi Backpropagation Menggunakan Penerapan Adaptive Learning Rate dan Parallel Training* (Medan: University of North Sumatera)

[4] Ghosh A W, Chakraborty M 2016 Hybrid optimized backpropagation learning algorithm for multi-layer perceptron. *International Journal of Computer Applications* 57: 0975 - 8887

[5] Wagh S M & Pawar 2017 GPU parallelization of backpropagation neural network. *International Journal of Engineering science and innovative technology* 6: 2319 - 5967

[6] Chamidah N, Wiharto & Salamah 2015 Pengaruh Normalisasi data pada jaringan syaraf tiruan backpropagasi gradient descent adaptive gain(bpgdag). *Jurnal Itsmart* 1: 2301 - 7201

[7] Suresh S, Omkar S N & Mani V 2005 Paralel implementation of backpropagation algorithm in network of workstations. *IEEE*