Application of Data Mining Prediction of Electricity Deviation Flow Using Metode Backpropogation at PLN Binjai Area

A M H Pardede¹, Febi Wahyu Andrayani¹, Nuredin², Dahlan Abdullah²*, Cut Ita Erliana³, Raditya Danar Dana⁴, Arif Rinaldi Dikananda⁴, Firmanda Fasya⁴, Nining Rahaningsih⁵, Meri Andriani⁶, Purwo Subekti⁷, Erlin Elisa⁸, Sohibun⁹, Muhammad Noor Hasan Siregar¹⁰, Ponidi¹¹, Edy Winarno¹² and I Ketut Sudarsana¹³

¹College of Information and Computer Management Kaputama, Medan, Indonesia
²Department of Informatics, Universitas Malikussaleh, Aceh, Indonesia
³Department of Industrial Engineering, Universitas Malikussaleh, Aceh, Indonesia
⁴Department Of Information Technology, STMIK IKMI Cirebon, Indonesia
⁵Department Of Accounting Information Management, Universitas Malikussaleh, Aceh, Indonesia
⁶Department of Industrial Engineering, Universitas Samudra, Langsa, Indonesia
⁷Department of Mechanical Engineering, Universitas Pasir Pengaraian, Riau, Indonesia
⁸Department of Information System, University of Putera Batam, Batam, Indonesia
⁹Department of Industrial Engineering, Universitas Pasir Pengaraian, Riau, Indonesia
¹⁰Department of Management, Universitas Graha Nusantara, Indonesia
¹¹Department of Mechanical Engineering, Universitas Muhammadiyah Surabaya, Surabaya, Indonesia
¹²Faculty of Information Technology, Universitas Stikubank, Semarang, Indonesia
¹³Department of Religious Education, Institut Hindu Dharma Negeri Denpasar, Indonesia

* dahlan@unimal.ac.id

Abstract. PLN is the only state-owned electricity company that provides and regulates electricity to the public. However, in the provision of electricity to every community, there are many frauds or irregularities in the utilization of electricity by irresponsible parties was found by PLN. Electricity deviation caused by irresponsible parties, like the theft of electricity can have a tremendous impact due to loss of electrical energy. Electricity Deviation, using Electricity Deviation data for 3 years, from 2012 to 2014. Sales data processed by the network, in the form of amount of Kwh deviation per month. The data needed in the process of analysis of this Neural Network is pure database results by taking the input variable in the form of Amount of Kwh Deviation and Amount of Kwh Deviation for the next Month for the following year as output. Electricity for 3 years is processed into Matlab using GUI facilities, to produce information on the predicted results of the Kwh Deviation in November 2014 to March, May, July, August, November, December 2015 experiencing has increased in Kwh, while in April, June, September, and October 2015 decreased.

1. Introduction

PLN is the only state-owned electricity company that provides and regulates electricity to the public. However, in providing electricity to every community there are many frauds or irregularities in the use
of electricity by irresponsible parties that occur in electricity users found by PLN, one of them is the theft of electricity. 

Electricity deviation carried out by irresponsible parties, the theft of electricity can have a tremendous impact because it causes loss of electrical energy. The impact on PLN from the electricity theft was in the form of material losses that could reach hundreds of millions or even billions of Rupiah due to the loss of some electricity to the user. While the impact on the community/customer is can reduce the capacity of the power that should be distributed to home, frequent outages or black-out lights and the risk of fire due to a short circuit of electrical current. 

Data Mining is a process that employs one or more computer learning techniques to analyze and extract knowledge automatically (Hermawati, 2013)[1]. According to Larose (2006) data mining can be divided into several groups based on the tasks performed, such as: description, estimation, prediction, classification, cluster and association. In making predictions, one method that can be used is backpropagation [2]. According to Puspitaningrum (2006) Backpropagation is a gradient reduction method to minimize the output error squared. There are three stages that must be carried out in network training, that is the phase of forward propagation, backpropagation, and weight and bias changes. The network architecture consists of input layer, hidden layer, and output layer [3].

With the prediction system that will be accomplish to predict the utilize of electricity, there will be an estimate of electricity flow cheating so that it can be seen increasing or decreasing and it can be countered by the PLN in making decisions in operational work part to handle the consumption of electricity.

2. Related Works

By using the Backpropagation Artificial Neural Network can expected to provide another alternative in estimating and predicting the rate of Electricity Usage Deviation in the Binjai area PLN. In addition, the application of Artificial Neural Networks is also used for forecasting the number of production requests (Febriana et al, 2013)[4] and is also used in weather prediction research using Artificial Neural Networks. In a simply, the Artificial Neural Network is a non-linear statistical data modeling tool that can be utilized to model complex relationships between input and output to find patterns in data[5].

3. Research Methodology

In designing artificial neural network systems using backpropagation algorithms, the authors want to predict the utilization of electricity deviations. The data needed in this neural network analysis process is the result of pure data by taking input variables in the form of electricity deviation data for 3 years from 2012 to 2014 as input and the total result is output. The data is taken randomly from the data of electricity deviation by using data during 3 years from 2012 to 2014, the data deviation of electricity will be divided into two, the first data used to train achieve convergent data. The second data will be used as test data, which is the data has never been trained to produce a prediction of the amount of electricity deviation in a month in the next year. Each training data uses 20 patterns and has 20 training data targets. While the test data uses 12 patterns and has 12 test data targets. Artificial Neural Network Architecture using the Backpropagation method consists of:

1. Input Layer (X_i) consist of 2 neuron,
2. Hidden Layer (Z_i) consist of 4 neuron,
3. Output Layer (Y_i) consist of 1 neuron.
4. Learning Rate (α) = 0.2
5. Target Error = 0.01
Information:

X<sub>i</sub>: Input Layer 2 neuron (Number of variable in January and February, and so on)
Z<sub>j</sub>: Hidden Layer 4 neuron (Number of Kwh)
Y<sub>k</sub>: Output Layer 1 neuron (Number of Much or Less Prediction Result)
V<sub>ij</sub>: Weight of Hidden Layer
W<sub>ij</sub>: Weight of Output Layer
V<sub>0j</sub>: Bias of Hidden Layer
W<sub>0j</sub>: Bias of Output Layer
i, j, k: 1, 2, 3, ..., n
n: Number of neuron in a layer
1: Bias constant = V01=0.1, V02=0.2, V03=0.4, V04=-0.1

The early weight connects neurons in the input layer and hidden layer (V11, V1-n, V21, V2-n) and the bias weights of V01 and V0n are chosen randomly. Likewise the first weights that connect neurons in the hidden layer and the output layer (W11, W12, W13, ..., W41) and the W0 bias weights are also randomly selected.

The following is a training calculation using the Backpropagation method.

Initialization is set:

1. Learning Rate (α) = 0.2
2. Target Error = 0.01
3. Maximum Epoch = 10000
4. Target (T) = (0.282973)

4. Result and Discussion

Data was trained by the affect of the network algorithm model used. Number of hidden layers, number of Kwh deviations. To produce the fastest iteration, the values of the number of hidden layers are changed. The results of the training were carried out in 2 phases, that is training of data that was trained and testing of new data that had never been trained consisting of 20 data, like the number of Kwh deviations in 2012 and 2013. After that, the network will be tested with 14 new data, that is the...
number of Deviations Kwh in 2014 after the data that had been trained was identified. This function is to test a large number of ANN to recognize new data.

The data was taken from the Electricity Deviation for 3 years, during in 2012 to 2014 which has trained on the network as much as 2 x 20 data patterns and 1 x 20 as target data. And the training data is 2 x 14 and the 1 x 14 target are used to test the accuracy of the system in recognizing other data inputs.

Table 1. Determination of Input and Output Variables for Amount of Kwh Deviation

| Input Variable | Output Variable |
|----------------|-----------------|
| X1.1 X1.2      | Y1              |
| X1.2 X1.3      | Y1              |
| X1.3 X1.4      | Y1              |
| X1.4 X1.5      | Y1              |
| X1.5 X1.6      | Y1              |
| X1.6 X1.7      | Y1              |
| X1.7 X1.8      | Y1              |
| X1.8 X1.9      | Y1              |
| X1.9 X1.10     | Y1              |
| X1.10 X1.11    | Y1              |
| X1.11 X1.12    | Y1              |

All existing variable data are separated into two parts, particularly input and output. Input data serves as a training and testing process. While the output data as the target of achieving the process. The data used is derived from data on the number of electricity deviations at PT. PLN Binjai Area, during 2012 to 2014 as shown in the table below.

The table below is the data that inputted into the matlab software. The data divided by 1000000 so the data can be read in Matlab.

Table 2. Data of Electricity Deviation in 2012-2014

| Month     | Year 2012 | Year 2013 | Year 2014 |
|-----------|-----------|-----------|-----------|
| January   | 359724    | 95641     | 109761    |
| February  | 377704    | 168565    | 33461     |
| March     | 282973    | 434289    | 86128     |
| April     | 459273    | 411992    | 160707    |
| May       | 298351    | 323261    | 98576     |
| June      | 428721    | 94165     | 178832    |
| July      | 344421    | 20512     | 215673    |
| August    | 191531    | 68095     | 111149    |
| September | 241934    | 108089    | 736851    |
| October   | 447630    | 122372    | 730910    |
| November  | 343818    | 89393     | 298464    |
| December  | 365109    | 82108     | 197037    |
The use of the Artificial Neural Network Algorithm development program using Matlab (Matrix Laboratories) is carried out to determine the Kwh Deviation Prediction. Tests are accomplish to determine the extent of the system's performance in processing data so it is able to produce the coveted information. Apart from the results of the trials that have been fulfilled, it will be analyzed in case the design can fill the desires to be achieved.

The results of the training were carried out in 2 stages: training data that was trained and testing on new data that had never been trained. The training data consists of 20 data, such as the number of Kwh deviations. After that the data tested, it is consists of 14 new data. After the data has been tested, the Neural Network recognizes how much the new data. The Following Prediction of Electricity Use Flow Deviation.

| Hidden layer activation function | Logsig |
|---------------------------------|--------|
| Output activation function      | Purelin|
| Epoch                           | 20000  |

Where the logsig activation function is used to calculate the weight of artificial neural networks using the backpropogation method, and the purelin activation function is used to get the output that approaches its input.

Table 3. Results Data of Electricity Flow Deviation in 2012-2014

| Month     | Year 2012 | Year 2013 | Year 2014 |
|-----------|-----------|-----------|-----------|
| January   | 0.359724  | 0.095641  | 0.109761  |
| February  | 0.377704  | 0.168565  | 0.033461  |
| March     | 0.282973  | 0.434289  | 0.086128  |
| April     | 0.459273  | 0.411992  | 0.160707  |
| May       | 0.298351  | 0.323261  | 0.098576  |
| June      | 0.428721  | 0.094165  | 0.178832  |
| July      | 0.344421  | 0.020512  | 0.215673  |
| August    | 0.191531  | 0.068095  | 0.111149  |
| September | 0.241934  | 0.108089  | 0.736851  |
| October   | 0.447630  | 0.122372  | 0.730910  |
| November  | 0.343818  | 0.089393  | 0.298464  |
| December  | 0.365109  | 0.082108  | 0.197037  |

Table 4. Results of Kwh Deviation Prediction

| Month     | Old Data (Year 2014) | Prediction Result (Year 2015) |
|-----------|----------------------|-------------------------------|
| January   | 89393                | 89197                         |
| February  | 82108                | 113678                        |
| March     | 109761               | 110852                        |
| April     | 33461                | 92165                         |
| May       | 86128                | 114438                        |
| June      | 164707               | 97462                         |
| July      | 98576                | 130891                        |
| August    | 178832               | 131344                        |
| September | 215673               | 158950                        |
| October   | 111149               | 150109                        |
| November  | 736851               | 214358                        |
| December  | 197037               | 196697                        |
Table 4 In the month column shows that will be the test target and that will be predicted, that is the Amount of Kwh Year 2014 Data and what will be predicted is the Number of Kwh Deviations in the following month in 2015. In the 2014 test target column shows that the Old data starts from January to December 2014 is an Electricity Deviation data. In the predicted column of 2015, the prediction results of Kwh deviation can be known starting from January to December 2015. The results of the prediction show that from February to May, July, October, 2015 has experienced an increase in Kwh deviation, while in January, June, August, September, November and December 2015 has decreased.

5. Conclusion
From the results of research conducted on Electricity Deviation Flow Prediction, the authors can draw the following conclusions:
1. Analyzing the Deviations of Electricity Flow Data to Predict the Number of Deviations in the Next Month in case, there is an increase or decrease, to facilitate anticipation of electricity deviations.
2. The prediction results show that the prediction results of Kwh deviation can be known from January to December 2015. The results of the prediction show that from February to May, July, October, 2015 has experienced an increase in Kwh deviation, while in January, June, August, September, November and December 2015 has decreased.
3. With the existence of Electricity Deviation Flow Predictions, it is expected to provide very useful information for PLN Binjai Area in making decisions. So that it can better anticipate the Electricity Theft.

References
[1] F. A. Hermawati, “Pengolahan citra digital,” Yogyakarta Penerbit ANDI, 2013.
[2] D. T. Larose, Data mining methods & models. John Wiley & Sons, 2006.
[3] D. Puspitaningrum, “Pengantar Jaringan Syaraf Tiruan,” 2006.
[4] M. Febriana, F. Arina, and R. Ekawati, “Production Demand Forecasting Using Backpropagation Artificial Neural Network (ANN) Method,” J. Tek. Ind. Untirta, vol. 1, no. 2, 2013.
[5] S. A. F. Haryanto, E. Ernawati, and D. Puspitaningrum, “Implementation of Backpropagation Neural Network to Predict Weather (Case Study: Bengkulu City),” Rekursif J. Inform., vol. 3, no. 2, 2015.