An Improved BP Neural Network in Internet of Things Data Classification Application Research Feng WANG

Feng Wang
School of Computer and Information Engineering
Fu Yang Teachers College, Fu Yang, 236037
An Hui, China
wffync@fync.edu.cn

ABSTRACT: *For the problem that traditional BP neural network exists slow convergence, easy to fall into local minimum value which will restrict the application of BP neural network in the Internet of Things(IOT), this paper proposes an improved BP algorithm which integrate into variable learning rate. In order to verify the effectiveness of the improved algorithm, the paper makes use of neural networks toolkit and GUI interface to classify collected data experiment based on MATLAB environment. Experimental result show that the algorithm has better data classification performance.*

Keywords: BP Neural Network; Internet of Things(IOT); Data Classification

Received: 13 March 2016, Revised 24 April 2016, Accepted 2 May 2016

1. Intoduction

BP (Back Propagation) network was proposed by a team of scientists led by Rumelhart and McCelland in 1986, whose principle is a multilayer feedback network by error backward propagation algorithm, and it is one of the most widely used neural network model. BP network can learn and reserve a lot of input-output model mapping relationship without mathematical equations. Its learning rule is to use the fast descent method and adjust continuously the network weights and thresholds through reverse delivery, so that the network square error reaches the minimum. BP neural network topology includes the input layer, hidden layer and output layer.

With the rapid development of wireless communication technology, IOT (The Internet of things) become the current hot topic, its concept was proposed in 1999[1]. That is, through radio frequency identification (RFID), infrared sensors, global positioning systems, laser scanners and other information sensing device, according to the agreed protocol, it is connected anything to the Internet to exchange information and communication in order to achieve intelligent identification, track, monitor and manage a network. In short, IOT is " the connection between the object and the object network."

A huge number of objects in the physical world, various forms, constant motion,? distributed in various locations and being susceptible to external influences, which lead to obtain data with massive, heterogeneous, high-dimensional, redundancy, time series and spatial location associated dispersion characteristics. IOT offers a variety of information services to users based on

knowledge discovery from massive data, classification technology is suitable for massive data networking whose feature is a prerequisite to achieve things information services[2].

2. Improved BP neural network

2.1 The basic principle of BP neural network

BP neural network classification is based on the learning error reverse delivery algorithm. The basic principle is each sample of the training sample set, input from feedback neural network input layer, by the weight of the connection between the layers to calculate the output of each neuron, to reach the output layer of the network layer by layer delivery. Finally, we compare the output results with the actual data, the error is passed back to the each layers, and then modify the weights to approach target results. In the field of data classification, BP neural network classification method has strong generalization ability, simple learning processing etc. It has good effect especially in multidimensional nonlinear data processing[3-4].

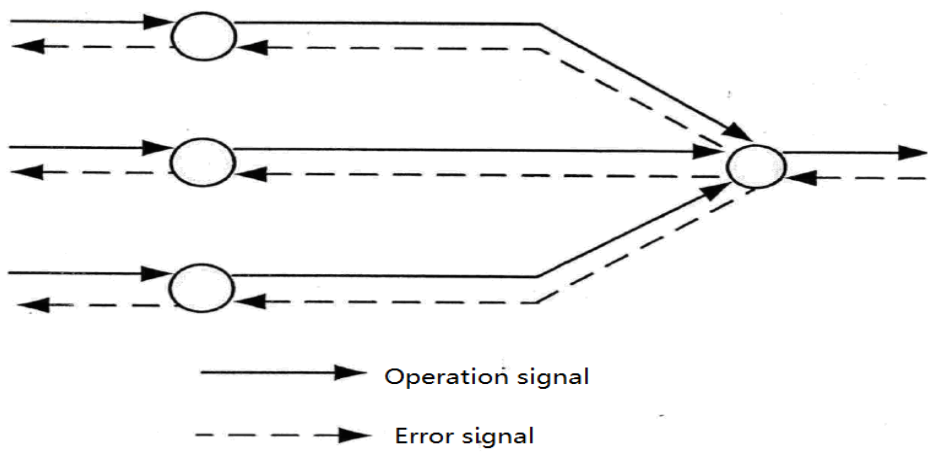

Figure 2-1. Propagation of the signals of BP algorithm

2.2 Traditional BP neural network defects

Although BP network has a solid theoretical foundation and rigorous derivation process, but it still has some flaws, as follows:

(1) slow convergence

BP is actually a standard gradient descent algorithm, but it requires the optimization objective function which is very complex. Since the error surface there will be a relatively flat area, when the adjusted weight values falls into this area, the change of error gradient will be very small which make the convergence rate stagnated, the entire training process almost halt.

(2) over-fitting problems

Usually, if the training effect is not good, its generalization ability certainly is not ideal. For BP network, each weight adjustment is learning again, if learning ability is not strong enough, of course, the final effect is certainly not ideal. But there is a limit to this trend, that is, when training exceeds this limit, the generalization capability of the network doesn't become stronger with increasing training times . Theoretically this phenomenon is called "over-fitting." Neural networks does not reflect the inherent laws due to learning too much sample details.

(3) difficult to determine the network structure

The BP network topology neurons selection is currently lack of a theoretical guidance, especially the hidden layers number and the number of neurons, its selection usually is based on experience. But different topologies have a great impact on the performance of BP neural network.

(4) easy to fall into local minima

On the training learning process of BP neural network, each layer weights always gradually is changed from a starting point to

reach the minimum of errors. However, for a very complex network, taking into account the weights, thresholds and the number of training we can conclude that the error function is a multi-dimensional curved space, there may be many local minima in the bottom of the surface. The standard BP algorithm always approach the point from one direction and has no ability to jump out, so the error value of training is no longer mutative even if the training continues as usual[3].

2.3 Improved BP neural network algorithm

Many experts and scholars had done a lot of work for improving the performance of BP algorithm proposed some improved methods. For example, a small learning rate method[4], local learning rate adaptively based on symbols transform [5], the learning rate automatically adjusted [6] and so on. For these shortcomings of traditional BP algorithm, this paper proposes an improved method that we use a variable learning rate method to solve the problem of BP network training so that the entire training network will not fall into local minima.

In practical applications, The initial value of the network learning rate has a great impact on convergence and effectiveness of BP algorithms. Its optimal is often associated with the specific problems, so there is no determination of a learning rate for any problems. In fact, even in a particular issue, we can not find a suitable learning rate from beginning to end. Because BP algorithm is very sensitive to the learning rate of change in performance, if the learning rate is set too small, the convergence rate will be very slow. If the learning rate is too large, although the convergence rate can be accelerated, but it may lead to weights adjustments oscillation phenomenon. So we should modified the learning rate of the network according to the specific circumstances of training. as follows:

$$w_{ji}(n+1) = w_{ji}(n) + \theta(n)\eta(n) \tag{1}$$

where $\theta(n) = -\dfrac{\partial E}{\partial w_{ji}} = \delta_j x_i$ it is the negative gradient

$$\eta(n) = 2^{\lambda}\eta(n-1)^{\lambda} = \mathrm{sgn}\left(\theta(n)\theta(n-1)\right) \tag{2}$$

When a new weights is got by the current learningate and the error function, we do not rush to throw away the original error, until the new error generation, then comparing the new with the old error. If the difference is greater than the set value , then we need to reduce the learning rate, the specific adjustment by the formula (1) and (2). When the difference is too small, we again use this formula to increase the learning rate. Thus, using this method we can maintain the learning rate to reach a steady value range to maintain the normal speed of convergence.

In the course of using this algorithm, if two consecutive iterative gradient direction is same, it indicates that the energy decline is too slow, then you can consider increasing the learning rate, and improve the training convergence speed. When the energy drops too fast, there will be two continuous iterations gradient direction converse, then the learning rate should be reduced in order to prevent shock.

3. Application of improved BP network in IOT data classification

3.1 IOT data classification necessity-the uncertain massive data

IOT data sources is very broad, because there are a lot of random factors in data storage, transmission, gathering and production process. In many cases, because of inaccurate perception means, outdated data sources, network delay and sampling errors and other reasons, so that uncertainties data is ubiquitous. as follows:

(1) data collection generated uncertainty. Including data generation, transmission, and collection process, since the sensing device or the defect of data transmission technology, generate a large amount of uncertainty data. In real life, many measuring instruments exist more or less measurement error. For example, in the data collection processing the sensor has a lot of uncertain information, temperature, humidity, pressure and other external factors will affect the data collection. In addition to the data collection, data transmission in some cases also decrease the quality of the raw data. For example, when the wireless WIFI data transmission, data accuracy will be affected transmission delay and network bandwidth.

(2) the data integration process generates uncertainty . In reality a lot of different data sources with different format, such as intelligent home network systems, for refrigerator node information, the output signal should have temperature, power, and

usage and so on multi-dimensional information. It is difficult to integrate them into smart home system.

(3)Privacy protection. In many areas, people will deliberately add some uncertainty information in the original data. For example, in the military field with a wireless sensor network, when the wireless sensor nodes are distributed in the target area, they definitely need to encrypt communications. at this time, If the region also has enemy sensor nodes, then the sensor networks base station must be able to distinguish received data which belongs to own.

In summary, there is a lot of uncertainty in IOT system, so it is urgent to classify and deal with these uncertain data.

3.2 Data Classification

Data classification is a process with many types of existing data sets for analysis, and the same type of data features are extracted and divided new and unknown types of data according to the extracted data feature. So far, the classification technology development has made great achievements, in general, we can put them into three categories: one is the use of statistical principles, such as KNN, support vector machine, regression model, maximum entropy model Bayesian networks and other methods; the second is to use the principle of connection, such as artificial neural network algorithm proposed in this paper; Another is based on the principle of classifying certain rules, such as rough set theory, association rules and decision trees and so on.

Core content of data classification is the classification model discovery and structure, generally we obtain data classification rules on the basis of algorithm, by analyzing a large number of known data types and data sets. we commonly call the training set used to learn known types of data set. In the data classification process, because we need to constantly analyze the training set, so data classification belongs to a form of supervised learning from the perspective of machine learning.

3.3 Data classification application of IOT

In order to improve data processing efficiency, it is necessary to classify these original and high-dimensional uncertain data. After data classification, data fusion ,statistics can be better operated so that it can improve the performance of the overall network. In addition, the Internet of Things data obtained from the sensor network are often associated with the results each other.

Data classification directly affect the overall performance of the network processor. Therefore, we believe to increase data classification processing module in the system in order to reduce the amount of data processing time. Co-processor, by reducing the complexity of the task, increase clearly defining and computing intensive tasks to improve the network processor (NP) performance [7], as shown in Figure 3-1. We join the co-processor settings in the traditional network processor which not only realizes data classification, but also integrate other intelligent data processing technology.

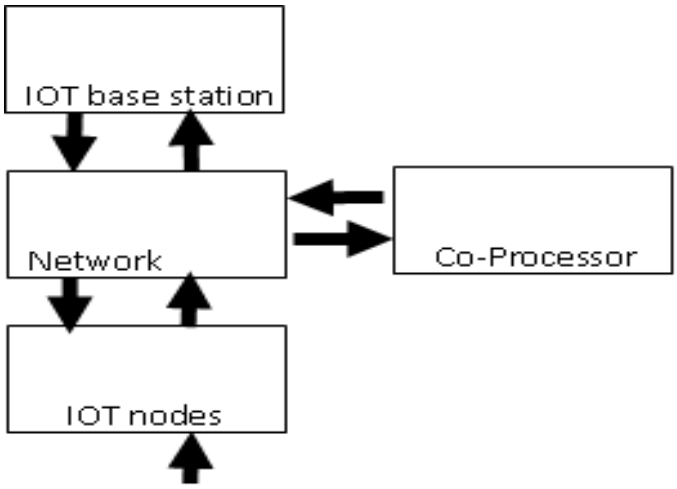

Figure 3-1. Packets Classification Model

4. Experiments and results analysis

4.1 Experimental environment and data sources

Currently, there are already a lot more mature neural network development package. Among them, the neural network toolbox is most widely applied in MATLAB environment, this experiment use GUI (Graphical User Interface) functions provided by MATLAB. The experimental data use IOT sensor simulation system by inputting a series of temperature, soil moisture, carbon dioxide concentration, light intensity and so on, and then use the BP neural network training.

4.2 Data Normalization

Through the analysis of sample data, since each unit of data collection is not consistent, in order to facilitate the realization of MATLAB BP network training, so the data needs to be normalized [-1,1]. Data normalization is actually a simplified calculation method. In theory the dimensionless expression transform a dimension expression, thus avoiding the input parameters of the physical meaning generated inequalities.

Normalization has many ways, the paper adopts a linear function conversion, the expression is as follows:

$$y = (x - MinValue) / MaxValue - MinValue$$

Where x is the data value before conversion, y is the value of the converted data, and MaxValue and MinValue are respectively the maximum and minimum taken from the samples property.

4.3 Network Training Results and Improved Algorithm Comparison

After data normalization completes, we can train a neural network. First, it is necessary to select the BP training neural network structure. Based on the structural analysis of the data set, the experiment selects 10 junctions as the input layer, 6 junctions as the output layer. Different hidden layer nodes will have an impact on the experimental results, the paper first sets 20, 2 times the input layer. In later experiments, This article will compare to the classification error and the networks convergence speed in a different hidden layer nodes number in order to determine the optimal network structure.

	Classification Type	training set number	test set number	The number of correct classification	Accuracy rate
experimental data sets	First Classification result	1000	1000	970	97%
	Cross-validation result	1000	1000	990	99%

Table 4-1 The comparison table between learning rate and number of training

	Classification Type	training set number	test set number	The number of correct classification	Accuracy rate
experimental data sets	First Classification result	1000	1000	970	97%
	Cross-validation result	1000	1000	990	99%

Table 4-2. The comparison table between learning rate and number of training

Before you start training, the experimental system also requires the user to input the target network accuracy and maximum

training times, when training BP neural network reaches one of these two conditions, the training will be terminated. Network accuracy of this experiment is set to 0.01, the maximum number of training 10,000 times.

according to the convergence results, we can see that when the train iteration times reaches the maximum 10000, the current BP network learning rate is 0.0132, it training time is about 2.3s, while the 1000 group classification test set consumes time in less than 1s. It shows the improved BP neural network in a way to enhance the network convergence rate.

5. Conclusion

Data classification is to alleviate the pressure because of data collection, transmission and data processing huge data in IOT. As an important topic in data mining, data classification aims to generate a classification function or model, it is able to map data in a specific category of a certain type. In the specific network training experiments, we found that the improved algorithm in a certain extent improves the convergence speed of the network. Final data classification results show that BP neural network is clear that the network cannot reach the target accuracy of 0.01. When training times is between 500 and 5000, network training accuracy has been significantly improved. By the training results, when the train reaches close to 10000 times iterations, the BP neural network precision has reached the target accuracy of 0.01. For a more comprehensive response to BP network training, this paper uses traditional BP algorithm and improved BP algorithm contrast. The result is shown in Table 4-1 (next page):

From the final result, we can conclude that the proposed algorithm improves the convergence performance of BP neural network to some extent. In addition, the two training results of BP neural network indicate that the convergence speed is quick in the beginning, with the increasing of training times, the convergence rate significantly becomes slower. When training more than 50,000 times, the amount of change in both the accuracy of network algorithms are very small, the improved algorithm in this interval the convergence rate does not show a clear advantage. Thus, when the BP neural network error is reduced to a certain extent, the learning rate must be required to maintain at a lower value, if increasing learning rate at this time it is likely to occur shock phenomenon.

Finally, the statistics in Table 4-2 (next page)show that the final classification accuracy rate is over 98%.

Acknowledgements

This work is supported by National Natural Science Foundation of China under Grant No. 61401101, The National Bureau of statistics project under Grant No. 2012LY009, and Natural Science Foundation of Anhui Province under Grant No. 1408085QF122, KJ2013Z261. In addition, the authors would particularly like to thank the anonymous reviewers for helpful suggestions.

References

[1] Atzori, Luigi., Iera, Antonio., Morabito, Giacomo (2010). The Internet of Things: A survey. Computer Networks, 54. 2787-2805.

[2] WANG, Mei-ling., WANG, Nian-ping. (2009). BP Improvement and Application of neural network algorithm, *Computer Engineering and Applications*,45 (35) 47-48

[3] Maciej, A., Mazurowski, Piotr A., Habas, Jacek., Zurada, M (2008). Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. *Neural Networks*, 21 (2-3) March-April 2008, 427-436.

[4] Veloso, A. Jr., W.M., Zaki, M.J. (2006). Lazy Associative Classification, *In*: Proceedings of the Sixth International Conference on Data Mining. 645-654.

[5] Knobbe, A., Siebes, A., Blockeelet, H. (2001). Multi-Relational Data Mining, using UML for ILP. *In*: Proceedings of the 5th European Conference on Principles and Practice in Knowledge Discovery in Databases, Freiburg, Germany, 1-12.

[6] Blockeel, H., Jensen, D., hHrame, S (2006). Introduction to the Special Issue on Multi-relational Data Mining and Statistical Relational Learning. Machine Learning, 62 (2) 3-5.

[7] Rumelhart, D.E., Hinton, G.E., Williams, R.J.(1996). Learning internal representations by error propagation. Parallel Distributed processing: Explorations in the Microstructure, 48 (1) 71-79.