Backpropagation Algorithm for Classification problem

James Hoang
School of Information System
vdthoang.2016@phdis.smu.edu.sg

ABSTRACT

Neural Networks (NN) have been known as an important data mining technique used in classification, clustering, etc. A neural network is defined by a set of input, hidden and output layer. Each layer contains several nodes depending on user requirements. There are connection between nodes in input, hidden and output layer. Those connection represent weights between nodes. In this paper, we describe Backpropagation (BP) algorithm which is one of the most popular algorithm in NN. Moreover, we also discuss an advantage and disadvantage of this algorithm. Finally, we show that BP is quite often used in practice.

Keywords

Neural network; backpropagation; classification.

1. INTRODUCTION

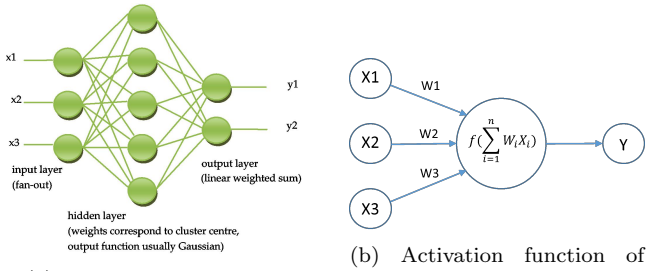
In machine learning, classification is the problem of identifying to which of a set of categories a new observation belongs, based on the basis of a training set of data containing observations whose category membership is known. An example would be assigning a given email into "spam" or "non-spam" classes or assigning a diagnosis to a given patient as described by observed characteristics of the patient (gender, blood pressure, presence or absence of certain symptoms, etc.).

NN is one of data mining technique used in classification problem. In this paper, we explain NN architecture, activation function and briefly describe BP algorithm.

2. NEURAL NETWORKS

2.1 Architecture

NN are constructed from input, hidden and output layer. Each node from input layer is connected Figure 1a shows architecture of an simple NN. Each node from input layer is connected to a node from hidden layer, and every node from hidden layer is connected to a node in output layer. A number of nodes for each layer vary depeding on user requirements. However, Larose [2] show that choosing a number of nodes for each layer may be challenging task due to overfitting problem and computation time.



(a) Simple neural network

(b) Activation function of neural node

Figure 1: An example of neural network

2.2 Activation Function

Activation function maybe linear, sigmoid or threshold function depeding on our problem. Normally we used sigmoid function for hidden layer [2]. Figure 1b shows the activation function used in network model. The output nodes (x_i) are multipled with connection weights (w_i) and put throught sigmoid function (f). Hence, the ouput can only be a number between 0 and 1 [1].

3. BACKPROPAGATION ALGORITHM

The BP algorithm is decomposed in the three main steps:

- Feed-forward computation: It includes two processes. First is getting the values of hidden layer nodes, and second is using these values to compute value of output nodes.
- Backpropagation to the output layer: This step calculates the error between true output node and predicted output node. Error values are used to update connection weight from output layer to hidden layer.
- Backpropagation to hidden layer: Now we need to update connection weights from hidden layer to output layer.

These processes will be repeated many times until the error is satisfy.

© 2016 ACM. ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

4. APPLICATION, PROS AND CONS

NN have been used more and more recently, and therefore BP algorithm also have been used a lot. In particular, financial sector has been known for using BP in classifying credit rating and market forecasts [4]. Marketing is another application which BP has been used for customer classification (i.e., identifying new markets for certain products, etc.) [3].

An advantage of BP algorithm is that it can adapt to new scenarios, moreover it is fault tolerant and may deal with noisy data. However, the BP is very time-consuming since it requires that all connection weights need to be updated for each iteration. Moreover, the BN also requires a very large sample dataset to accurately train model.

5. CONCLUSION

In this paper, we brieftly describe NN architecture and then show how BP algorithm works. Intuitively, the most important characteristics of BP is its ability to construct model from training data. The model is used to classify a new data. In some fields like pattern recognition, the results of using BP are significant improvement compared to tradition algorithms (i.e., logistic regression, random forest, etc.).

6. REFERENCES

[1] K. Gurney. *An Introduction to Neural Networks*. Taylor & Francis, Inc., Bristol, PA, USA, 1997.

[2] D. T. Larose. *Discovering Knowledge in Data: An Introduction to Data Mining*. Wiley-Interscience, 2004.

[3] H. L. Poh. *A Neural Network Approach for Marketing Strategies Research and Decision Support*. PhD thesis, Stanford, CA, USA, 1991. UMI Order No.GAX91-22576.

[4] R. R. Trippi and E. Turban, editors. *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance*. McGraw-Hill, Inc., New York, NY, USA, 1992.