Research on Classifications of Mobile Networking Device Users Based on BP Neural Networks

Xiaoling Xia[1, a, *], **Ruyu Han**[1, b]

[1]Schoole of Computer Science and Technology, DongHua University, Shanhai 200000, China;

[2]Schoole of Computer Science and Technology, DongHua University, Shanhai 200000, China;

[a]sherlysha@dhu.edu.cn, [b]2161714@mail.dhu.edu.cn

Keywords: BP neural networks, model, user classification.

Abstract: In order to make full use of users' networking data reported by mobile networking devices and offer a reference for companies to develop targeted and personalized marketing strategies, the author trains related attributes extracted from data and neutral networks which can make classifications towards users according to the model principles of BP neural networks, trying to make predictions towards potential consumption ability of users or develop personalized marketing strategies according to the classification results. The results show that the BP neural networks can be well applied to problems of this classification, and the results of user classifications also have certain values for reference and practice.

1. Introduction

With the advent of the information age, the network has become an indispensable part in people's life. In order to provide users with more convenient, fast network experience, a variety of networking devices emerge in endlessly, and thus a large number of network data come into being. Through network sniffer, we can extract useful information from network data. The information is extremely useful and can be used in fields such as network security, big data analysis, or data classifications.

BP (Back Propagation) neural networks are one kind of artificial neural networks, as well as the most widely used one. They are mainly used for data compression, data mining, function approximation, pattern recognition and classification[1]. Based on the above background, this paper will apply BP neural networks to user classifications of mobile networking devices. By extracting the information fields that can really reflect the attributes of users as input of neural networks and by training, corresponding association rules are generated and can be used for simple user classifications.

2. Principles of BP neural networks

BP neural networks are neural networks combining multilayer feedforward networks with reverse error pass learning (gradient descent). As shown in Fig. 1, they are composed of input layers, hidden layers and output layers, in which, input layers and the output layers are unilaminar, and hidden layers can be unilaminar or multilaminate. The nodes located on the same network layer exist independently and do not interfere with each other. The nodes located on contiguous network layers are connected in pairs and correspond with their respective connection weights. The networks consist of two phases: forward pass of input vectors and reverse feedback of errors[2].

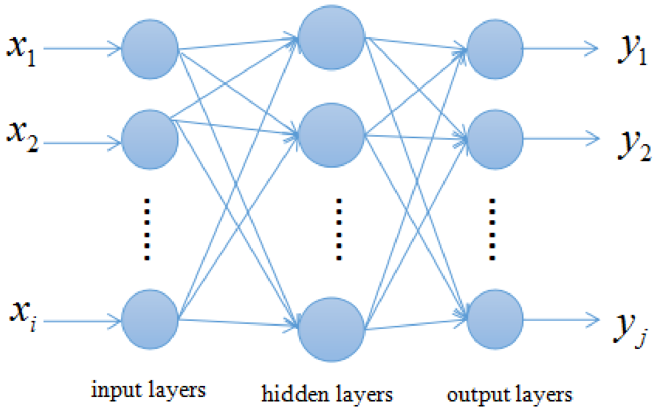


Fig. 1 Schematic diagram of neural networks

The main process of BP neural network:

1) Initialize

An existing network model has been established. Determine the training set and test set in the sample data. Giving the initial parameters, including the weights, thresholds, learning operators, and allowable error.

2) Calculate outputs of hidden layers and networks

BP networks are actually complex multi-layer sensors, and the output of hidden layers is the independent variable of $z = f(\sum_{i=1}^{n} w_i x_i + b)$. As for the calculation function of each hidden layer, that is, a mapping. The independent variables of this layer are dependent variables of the upper layer. The difference is that the Sigmoid excitation function is generally nonlinear, which can solve more complex problems compared with the linear function of sensors.

3) Calculate errors

The error between actual output and expected output is E. If the error is within the allowable range, receive the learning results. Otherwise, enter the second phase.

4) Phase of error reverse feedback

Feedback layer by layer, conduct BP weight and threshold correction, and aim to adjust the model to an ideal state.

5) For the next sample, enter the cycle from 2) to 4), until the sample data in all training data sets is completed. This moment, the error is within the allowable range or reaches the specified times of learning.

3. BP neural networks applied to user classifications and their implementation procedures

3.1 User classifications

In order to provide the user with referable user data, after pretreating online real-time network data of users reported by mobile networking devices, it considers to extract information fields which can measure using time, rates and degrees of dependence in terms of users' use of devices, in order to construct BP neural networks to classify users. Through market research and analysis, users are divided into three categories: vip users, growth users and ordinary users.

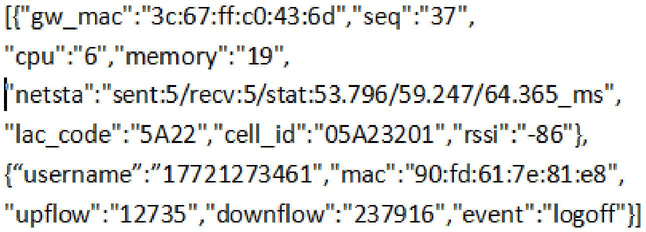
3.2 Sample processing

Networking devices will send a request of http at fixed time and on demand. The request contains the device state and using information. So the original data are json data reported to backstage in device operation, stored in files of log type.

Data processing of samples is divided into the following steps:

1) Extract the required field data

Fig. 2 Shows the raw data, and extract useful data information fields: username, user online time (logontime), user offline time (logofftime), online uplink traffic (upflow) and online downlink traffic (downflow). In addition, query the times of users who buy Internet traffic the same month from the database.



[{"gw_mac":"3c:67:ff:c0:43:6d","seq":"37",
"cpu":"6","memory":"19",
"netsta":"sent:5/recv:5/stat:53.796/59.247/64.365_ms",
"lac_code":"5A22","cell_id":"05A23201","rssi":"-86"},
{“username”:”17721273461","mac":"90:fd:61:7e:81:e8",
"upflow":"12735","downflow":"237916","event":"logoff"}]

Fig. 2 Raw data

2) Calculate input value

Users’ online duration per day on average (avgontime):

$$avgontime = \frac{\sum_{i=1}^{count} ontime_i}{count} = \frac{\sum_{i=1}^{count} (\log offtime_i - \log ontime_i)}{count}$$

(1)

Users’ usage amount of Internet traffic per day on average (avguseflow):

$$avguseflow = \frac{\sum_{j=1}^{count} useflow_j}{count} = \frac{\sum_{j=1}^{count} (upflow_j - downflow_j)}{count}$$

(2)

Where variable count is the number of users online times the same day.

3) Data normalization and quantification

In order to reduce the influences of larger numbers directly input into networks, which makes the network difficult to converge, unify the evaluation criterion, normalize data: $\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$, it should let $\bar{x}_i \in [0,1]$. Quantify and recode attributes of users’ categories towards discrete users as expected output.Table 1 shows the data after a series of pretreatment, including the input of the BP neural network and the expected output.

Table 1 Style of the data

username	avgontime (hour)	avguseflow (MB)	buycounts (times)	categories of users	quantization of the categories
177****3461	2.7	190	1	ordinary users	（0，0，1）
182****0867	2.5	350	1	growth users	（0，1，0）
136****5747	3.6	230	1	growth users	（0，1，0）
183****1526	4.3	280	2	vip users	（1，0，0）
180****8270	2.6	180	1	ordinary users	（0，0，1）
...	...	...	...	...	...

3.3 Build networks and conduct training

The construction of neural networks is the key step in experiments. As for three attributes and three categories, this paper constructs a network with three input nodes, three output nodes and a hidden layer. Fig. 3 shows the structure. However, nodes of hidden layers will affect the convergence speed of the network. Therefore, appropriate values need to be chosen. On this issue, there is such a proposal: $num = N * m/(m+n)$ or $num = 1.5*(mn)^{1/2}$, where, m is the number of input nodes, n is the number of output nodes, and N is the number of samples. In this procedure, eight hidden layer neurons are used.

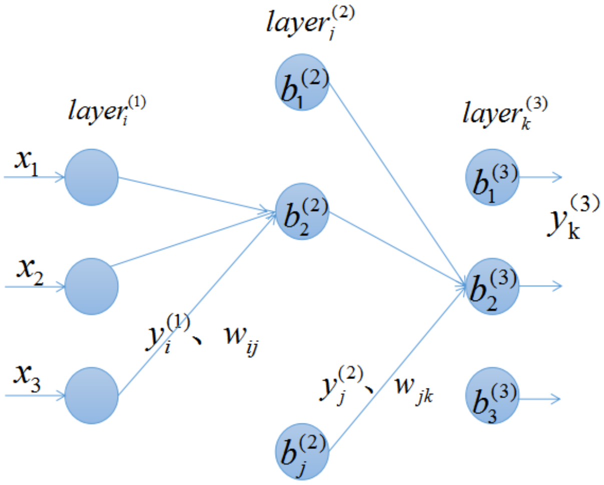


Fig. 3 Network structure diagram

The asymmetric Sigmoid function $f(x_i) = \frac{1}{1+e^{-x_i}}$ (where i = 1, 2, ..., m) is used as the excitation function. Before training, weights and threshold values are initialized into random values in the range of [-1,1][3].

After determining the learning operator λ and the allowable error value ε, send the NO.1 sample data to the network input layers, start training, calculate the input, output and error of each layer, and make comparison between calculated error and allowable error to judge whether it needs to be adjusted. After adjustment, start training the NO.2 sample, and cycle the same procedure til a certain suspensive condition is met with.

The output of $layer_i^{(1)}$, $layer_j^{(2)}$ and $layer_k^{(3)}$ is:

$$y_i^{(1)} = g(layer_i^{(1)}) = g(x_i) = x_i$$

(3)

$$y_j^{(2)} = f(layer_j^{(2)} + b_j) = f(\sum_{i=1,j}^{m} w_{ij}^{(1)} y_i^{(1)} + b_j)$$

(4)

$$y_k^{(3)} = f(layer_k^{(3)} + b_k) = f(\sum_{j=1,k}^{8} w_{jk}^{(2)} y_j^{(2)} + b_k)$$

(5)

Use gradient descent to change towards negative gradient direction, correct weights and thresholds and calculate error:

$$E(t) = \frac{1}{2}\sum_k (y - y_k)^2$$

(6)

For the output layer, quantity of einmal error correction is that $\Delta w_{jk}(t) = -\lambda \frac{\partial E_k(t)}{\partial w_{jk}(t)}$, and the learning operator $\lambda \in (0,1)$ will affect convergence speed. From the above formula, the obtained correction quantity of connection weight is that $\Delta w_{jk}(t) = \lambda \delta_k y_j$. In the above formula:

$$\delta_k^{(3)} = (y - y_k^{(3)}) y_k^{(3)} (1 - y_k^{(3)})$$

(7)

where y is the expected output, and y_k is the actual output. so the modified amount of threshold is:

$$w_{jk}(t+1) = w_{jk}(t) + \Delta w_{jk}(t)$$

(8)

And for input layer and hidden layer, parameters of substitution is:

$$\delta_j^{(2)} = y_i^{(1)} y_j^{(2)} (1 - y_j^{(2)}) \sum_k \delta_k^{(3)} w_{jk}^{(2)}$$

(9)

The correction quantity of the threshold is that $\Delta b = \lambda \delta$, in which between input layer and hidden layer, between hidden layer and output layer, δ takes different types of formulas for calculation. Reverse transfer correction and adjust networks until the end of the training.

4. Classification results and analysis

Generally, 80% of the sample data are used for training, while 20% for testing. Use the Visual C ++ tool to write experimental codes and read the processed sample data for training. This training uses 130 sets of sample data, taking 26 sets as test data. Define the input layer node of parameter: IN_NUM = 3, the output layer node: OUT_NUM = 2, the hidden layer node: Hidden_NUM = 8, the learning operator: learnRate = 0.9, the allowable error: accuracy = 0.001, and the maximum number of training: maxLoopCout = 10000.

Fig. 4 shows that training to 864 times will make the system converge to the allowable error value and the training time is about 0.4 seconds. From the data we can see that at the beginning, the error is larger and then makes rapid decline.With the increase in the number of training, rate of decline becomes gradually steady and slow.

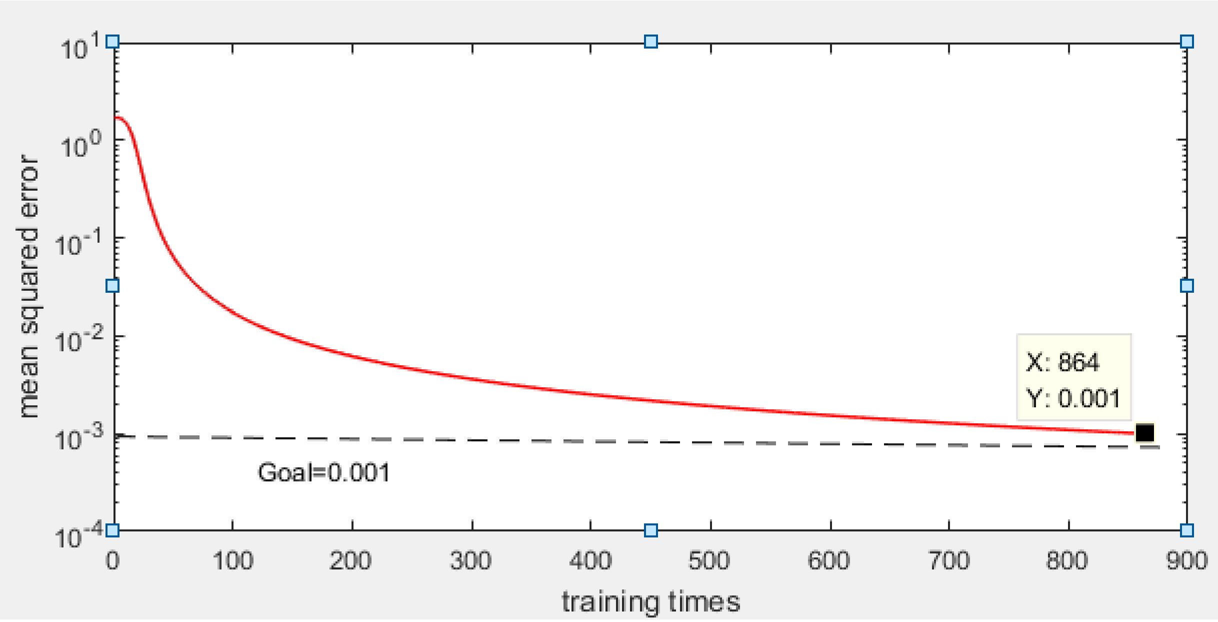


Fig. 4 Variation diagram of training errors

Networks after training can construct association rules about input and output. Using this network model to predict and classify towards users of mobile networking devices (car wifi and portable wifi) can get the category of each user and the proportion of users of each type. Fig. 5 shows that vip users take up small parts and more than half are ordinary users, besides growth users also occupies small parts. Thus, from the perspective of commercial marketing analysis, we should consider whether product qualities or data charges and other aspects need to be improved. From the view of developers, they can develop personalized recommendations or customization functions when users buy data packages so that users can make full use of resources in devices so as to increase their preferences towards products, and change them from ordinary users into growth users or even vip users.

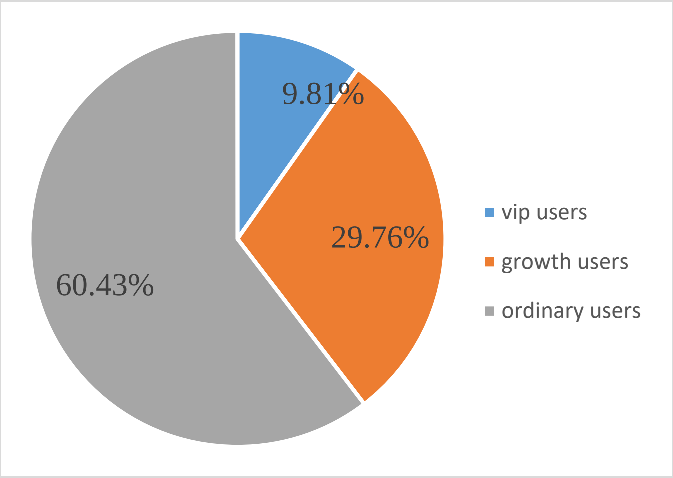


Fig. 5 Sector diagram of user categories in one month

5. Conclusion

This paper introduces the work and learning process of BP neural networks, as well as their current application fields, and carries on the classification experiment of mobile networking device users based on BP neural networks. This experiment shows that the constructed neural networks possess convergence of fast speed and great attributes after training, which can well solve the classification problem of this kind and be applied to predict the user categories. BP neural networks based on the classification results have certain practical values, and the network itself can also improve attributes in the process of constant improvement and optimization and be popularized for use. In the future work, we can also develop the personalized recommendation function towards users based on user classification.

References

[1] Huang Xuzhao, Liang Yueming & Li tonglin. Data Analysis Methods of Airborne Geophysical Exploration Based on BP Neutral Networks. Geophysical and Geochemical Exploration, 2010，34 (4): 486.

[2] Liu Jia. Study on classifications of Heat Rate of Articles from Social Network Sites Based on Artificial Neutral Networks. Journal of Tonghua Normal University (Natural Science), 2015，36 (6): 56.

[3] Yin Jiangning, Xiao Keyan, Li Nan & Zou Wei. Application of BP Networks in Data Classifications of Geochemical Exploration. Geological Bulletin of China, 2010，29 (10): 1565.