

Research on university library information level measurement based on genetic algorithms and neural networks

Sheng You

Library of Zhongzhou University, Zhengzhou 4500044 China

Corresponding author's e-mail: YouShengzhzhu@163.com

Received 01 October 2013, www.cmnt.lv

Abstract

Setups a comprehensive indexes system for the evaluation of academic library digitalizing level, presents a new evaluation model for academic library digitalizing level by combining genetic algorithm and BP neural network. Experimental results demonstrate the developed model is valid and practical. The model can give a scientific theoretic guidance for the evaluation of academic library digitalizing level. On the basis of the establishment of the University library digitalizing level evaluation index system based on the proposed neural network based on genetic algorithm Mechanism, and thus give a new university library information level evaluation methods. Measured results show that the evaluation model Effectiveness and practicality of the University Library Information Evaluation provides a more scientific theory.

Keywords: university library, information technology, genetic algorithm, neural network

1 Introduction

Back in the late 1980s, "electronic library", "virtual map Library" attention in developed countries, and the implementation of the class library Related projects and programs, and achieved great success. 1993 by the US States National Science Foundation (NSF), US Department of Defense Advanced Research Projects Agency (DARPA), the National Aeronautics and Space Administration (NASA) jointly launched Digital Library Initiative (Digital Library Initiative), proposed the concept of digital libraries. Since then opened a worldwide digital Strong research and practice Library. In traditional libraries abroad Transition to a modern library digital library, the library electrical Son, virtual, digital, information technology has become crucial. With the map Library information technology, the digital transition process, how to evaluate the letter Library. The level of interest, what kind of evaluation and evaluation methods for books Museum reasonably evaluate the level of information to become an urgent problem to ask Title. University Library to the level of information objectively and accurately A comprehensive evaluation, the evaluation results truly reflect the university library information. The actual level of development, so as to promote university library information by evaluating Technology development [1], is considered the use of neural networks for university libraries letter Assess the level of interest, but the method of slow convergence, and there Bureau Section minimum. Aiming at the University Library's special evaluation of the level of information Point, the genetic algorithm [2] introduced the level of information to evaluate the University Library, That the university library information technology to provide a scientific basis for decision-making.

A neural network for handwriting recognition is defined by a set of input neurons which may be activated by the pixels of an input image. After being weighted and transformed by a function (determined by the network's designer), the activations of these neurons are then passed

on to other neurons. This process is repeated until finally, an output neuron is activated. This determines which character was read.

Like other machine learning methods - systems that learn from data - neural networks have been used to solve a wide variety of tasks that are hard to solve using ordinary rule-based programming, including computer vision and speech recognition. Computational devices have been created in CMOS, for both biophysical simulation and neuromorphic computing. More recent efforts show promise for creating Nano-devices [3] for very large scale principal components analyses and convolution. If successful, these efforts could usher in a new era of neural computing [4] that is a step beyond digital computing, because it depends on learning rather than programming and because it is fundamentally analog rather than digital even though the first instantiations may in fact be with CMOS digital devices.

Between 2009 and 2012, the recurrent neural networks and deep feed forward neural networks developed in the research group of Jürgen Schmidhuber at the Swiss AI Lab IDSIA have won eight international competitions in pattern recognition and machine learning [5]. For example, multi-dimensional long short term memory (LSTM) [6-7] won three competitions in connected handwriting recognition at the 2009 International Conference on Document Analysis and Recognition (ICDAR), without any prior knowledge about the three different languages to be learned.

2 Establish indicators of library information system

The purpose of the University Library is to assess the level of information to give high School library information technology to provide a reliable basis for decision-making, in order to effectively level university library information comprehensive evaluation, we must first build Established a scientific and comprehensive evaluation system. In this paper, according to the index body Department of purpose-

built, scientific, objective, systematic, comparable, real the basic principles of usability, simplicity, etc. [8], drawing on existing enterprise information And university information technology research [9], University Library combines information [8]. On the basis of their own characteristics, evaluation of university library constructed water information level indicator system, which is mainly from the strategic position of the index system, information technology infrastructure, information resources, information technology, the use of information technology people only, information security, information technology aspects of the effectiveness of a total of seven of the library. The level of information to evaluate the content of a specific assessment has been shown in Table 1.

TABLE 1 Indicators of library information system

No.	Second level indicator	Indicator composition
1	Information organization	Special organization in charge
2	Importance	Related meetings and knowledge
3	regulation	Effective and timeless
4	Density of computer network	Total length/square
5	Coverage	Hardware/number of readers
6	Density of Internet	Total length/square
7	Auxiliary devices	Printer, hub and so on
8	Computers per reader	Computers for reader/reader
9	Computers per stuff	Computers for stuff/stuff
10	Seats per reader	Seats for reader/reader
11	Seats per stuff	Seats for stuff/stuff
12	Database	Total amount/number of readers
13	Database records	
14	Electronic publications	
15	Consumption fee	
16	Consumption amount	
17	Magazine	
18	Newspapers	
19	Books	
20	Paper books	
21	Electronic books	
22	Digital ratio	Related items/total items
23	New publications	
24	Consult ratio	
25	Borrowing ratio	
26	Buying ratio	
27	Self-construction ration	
28	technicians	
29	Specialist	
30	Technical employ	
31	Special employ	
32	Specialist among leaders	
33	regulation	Effective, implementation
34	input	fee used on security
35	Security of digital source	Break down time/365
36	Security management	importance
37	Scientific and technological achievements	Per reader
38	Patent	Per reader
39	Employment rate	College students
40	The effectiveness of teaching	Scores of students
41	Contribution on management	Support and help
42	Contribution on service quality	improvement

Genetic algorithm is the late 1960s and early 1970s, by John Holland and his colleagues at the University of Michigan American student research A method of forming the study. In this method, mechanism simulation of biological evolution Model to construct artificial system, has been widely used in recent years [8]. From Traditional BP neural network algorithm slow convergence and easy to fall into Local minima shortcomings, this paper based on genetic algorithm BP Network mechanisms to improve the convergence speed of the network, and then using the improved BP Neural networks for evaluation of the level of the university library information.

In this paper, three-layer BP network, shown in Figure 1. BP network three Layer nodes are represented, m input nodes x_i (each two indicators in Table 1 Value), in a hidden layer node h_j , 1 output node y (University Library Assess the level of information). Network power input nodes and hidden layer nodes Value w_{ij} , network nodes and the output power hidden layer nodes is t_j .

Each index can score from reviewer's subjective scoring method after obtaining. The data to be using Equation (1) is normalized

$$\bar{x}_i = \frac{x_i - b_i}{a_i - b_i} \tag{1}$$

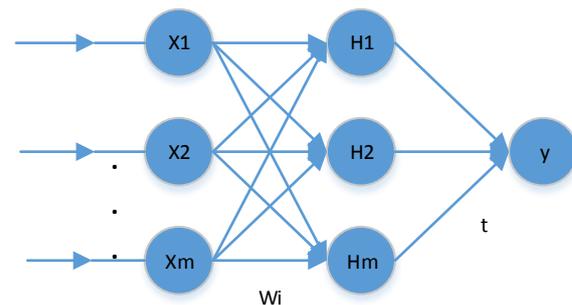


FIGURE 1 Three BP neural network

Where x_i, \dots, \bar{x}_i Table 1, respectively, the i -th index and the actual value Standard value; a_i, b_i are the maximum, minimum, the i -th index. Known evaluation indexes m, n hidden layer nodes depending Problems and experimental data to determine, you can also experience the value of the Equation (2) the decision

$$n = \log_2 m \tag{2}$$

Hidden node output is calculated as follows:

$$h_j = f \left(\sum_{i=1}^m w_{ij} x_i - \theta_j \right), \tag{3}$$

where θ_j is the threshold value for hidden node.

The output of the output node is calculated as follows:

$$f \left(\sum_{i=1}^m w_{ij} x_i - \theta_j \right) = f \left(f \left(\theta_j \right) \right), \tag{4}$$

where in θ is an output node threshold.

Equation (3) and Equation (4) in the transfer function is generally expressed as (0,1) interval of S-type function:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{5}$$

Using genetic algorithm to train the neural network can be used in binary code, code String value from the hidden layer to the input layer connection weights w_{ij} , hidden layer and output layer connection Weights t_j , hidden layer threshold θ_j , the output layer threshold θ . So connected together to form into a long string (the string corresponding to each position of a group of network weights and threshold Value), constitutes an individual. It can generate an initial population of N individuals.

Guiding the evolutionary process toward the region of space may contain the best individual conduct. Suitable Select the function should have a great impact on the training results. This paper fitness Function uses a calculation error of the neural network (see Equation (6)), and the error. Always cutting along the neural network output error fewer search direction, the convergence target is the minimum output error.

$$E = \frac{1}{N} \sum_{s=1}^N \sqrt{\left[\left(\overset{s}{y} \right)^2 - \left(\overset{\sim s}{y} \right)^2 \right]} \tag{6}$$

where $\overset{s}{y}$, $\overset{\sim s}{y}$, respectively, for the desired output and the actual output, N for training.

Training samples, then the fitness function is:

$$f(X_s) = \frac{1}{E} \tag{7}$$

which, X_s represents one individual.

The basic genetic algorithm by selection, crossover and mutation operators constitute three.

Select operation is based assessment in groups of individual fitness value Conduct random roulette method adopted in this algorithm on the basis of estimates Selection. Each chromosome number of offspring produced. According to individual fitness the order value is converted to the corresponding selection probability, a large fitness value corresponding High probability of selection, small fitness value corresponding to a low probability of selection [2], each chromosome Selection probability p_s . The Equiton is:

$$p_s = c(1 - c)^{s-1} \tag{8}$$

which c is the probability of an individual to select the first sorting 1.

In the selection wheel when using binary search method to effectively reduce A small number of comparisons to ensure that the corresponding roulette found (N in O (log N) for population size, the number of training samples).

When using consistent cross crossover method (ie, word by setting shield The entire body of inherited genes determine the two parent individuals in which the individual The corresponding gene, wherein the shield is a randomly

generated word and individual encoding result 0/1 sequence identical configurations).

When variation in a population of randomly selected individuals, mutation operator Loci child adoption, the basic steps are as follows: start with groups of individuals compiled Code string randomly selected one or more loci, and then follow the variation The probability of each generation after the completion of genetic manipulation, for a new generation of individual school Xi error and evaluate the fitness value, to identify the best individual, the optimal Individual two-step gradient descent local search, and put out a search of the Directly into the body of the new generation of the population, if the learning error achieve accuracy or up To specify the genetic algebra calculation is terminated. So after more than genetic luck Count, you can get BP neural network error minimum set of complete initial Weights and thresholds.

Section 2.2 of genetic networks by BP's has initialized iterations to get the right value and threshold for normal network training decoded offspring into BP neural network. This method using genetic algorithms to search after the near global optimum, then using BP algorithm local search Such BP algorithm can quickly and accurately converge near the solution point, final Get the problem global minimum. In this way, both to avoid the BP neural network into local minimum value and BP neural network can quickly close Convergence, greatly improving the accuracy and speed of the BP neural network.

3 Model testing

This work in MatLab 7.0 environment, the use of GA toolbox can be achieved. Simulation experiments take crossover probability $p^*=0.79$, mutation probability p' value of 0.16, the initial population size N is 100, hereditary algebra 120, an accuracy of 0.01, the experiment BP network hidden layer nodes n take Value 7,42 input nodes (Table 1 two indicators) and an output node Points. In this case each of weights obtained with a threshold value, wherein the hidden layer and the output node comprising Network weights between nodes, as shown in Table 2.

TABLE 2 The genetic algorithm to optimize the weighting factor

Index	Weights
Strategic position	0.128
Information infrastructure	0.229
Information Resources	0.127
IT use	0.125
Information technology professionals	0.136
Information Security	0.143
The effectiveness of information technology	0.112

Optimized by genetic algorithm initial weights and threshold values into BP Neural network, BP network setting is set to 500 times the number of cycles, training Error of 0.001. 10 experts to comment on the level of the Library Information Divide the network as a test sample, sample root mean square error of 0.05, was Library level of information and experts predicted values to score more consistent. Table 3 shows the test to validate the model.

TABLE 3 The model test results

Number of samples	Score	Predictive value	Error	Cycles
1	0.81	0.81003	0.0070	8
2	0.91	0.9098	0.0191	9
3	0.64	0.6401	0.0113	14
4	0.87	0.8707	0.0349	17
5	0.84	0.8392	0.0367	21
6	0.79	0.8001	0.1267	30
7	0.85	0.8493	0.0345	9
8	0.77	0.7687	0.0447	16
9	0.73	0.731	0.0382	22
10	0.88	0.8803	0.0230	43

Table 3 in 10 test samples for RMSE 0.0376. Achieve the accuracy requirements, indicating that the model is valid, can be better University library reflects the level of information to the indicators and evaluation Mapping relationship between performance, is a more reasonable and feasible Library Digitalizing level evaluation model and compare the simple use of neural networks Evaluation model has a faster convergence rate.

In order to illustrate the actual effect after using genetic algorithms and neural networks, we choose different loading nodes to do the test. The comparison before and after using genetic algorithms and neural networks in library information disposal can be seen from Figure 2 and Figure 3. The result shows that in the same experimental time, after using the genetic algorithms and neural networks it can achieve better performance in calculating time and calculating complex than before using it.

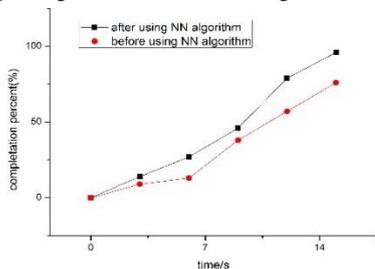


FIGURE 2 The comparison before and after using genetic algorithms and neural networks in calculating time

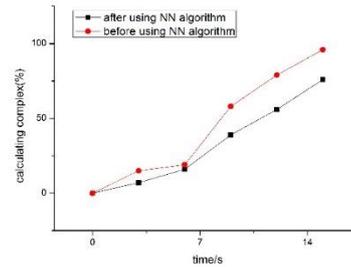


FIGURE 3 The comparison before and after using fractal algorithm in calculating complex

5 Conclusions

Traditional BP neural network algorithm slow convergence and easy Shortcomings in local minima, we use genetic algorithms to obtain participate Optimum weights and thresholds of BP network training, in order to improve the convergence network Speed, and then use the improved BP neural network Letter university library The level of interest for evaluation. Measured results show that the evaluation method better Reflect the nonlinear characteristics of fast convergence is an effective University Library digitalizing level evaluation method, able to believe university library Information development strategy to provide theoretical guidance science.

Acknowledgments

This work is supported by soft-scientific project of Henan Province (No. 102400440083).

References

- [1] Chen J 2006 University Library Information Level Measurement of neural networks *Now On behalf of the intelligence* 8 124-6
- [2] LIM-q 2002 Waiting GA basic theory and application *Beijing: Science Press* 62-7
- [3] Azurite, fees army, Yu Li hua university digitalizing level fuzzy comprehensive evaluation 2004 *Meter Computer Engineering and Applications* 40(28) 202-4
- [4] Chen Aijuan, Xiaoyan Ren, Huang Duo 2006 Our evaluation of the level of higher education information Summary *Information Science* 24(9) 1321-4
- [5] Zhao Q-c, Zhao G-j, Wang Ju 2004 Ying university education level of information and comprehensive measure *Cooperative Research Intelligence magazine* 6 101-204
- [6] Tai J-z, Meleis W, Zhang J-m 2013 Adaptive Resource Allocation for Cloud Computing Environments under Bursty Workloads *Northeastern University, Boston. USA* 978-87
- [7] Tirado J M, Higuero D, Isaila F 2011 Predictive Data Grouping and Placement Ivor Cloud-based Elastic Server Infrastructures *11th IEEE/ACM International Symposium on Cluster, Cloud and grid Computing. IEEE* 281-94.
- [8] U Labs S. I 2013 Private cloud computing with open Nebula *Software system: private cloud computing with open nebula* 21-30
- [9] Uranovetter M 1973 The strength of weak ties *American Journal of Sociology* 78(6) 1360-80

Author



You Sheng, born in 1977, Liuyang, Hunan, China.

Current position, grades: Librarian, Postgraduate.
Scientific interests: Library Information Construction.
Publications: 2 papers