Application of Neural Network in Computer Big Data Mining

Zhang Guoming

Nanchang Institute of Science & Technology, Nanchang 330108, China

Keywords: dynamic; big data mining algorithm; interdisciplinary; neural network; anti-noise; big data

Abstract: At present, in the process of mining large data, it is very practical to analyze, train and summarize large data so as to find out the important information in the neural network. The combination of neural network and big data mining is also a hot research topic at present. In this paper, the concepts, current situation and development trend of big data mining and neural network are expounded. At the same time, the large data mining algorithm and its algorithm process in the neural network environment are also studied. It enables people to understand the big data mining algorithm under the neural network more clearly. It also provides reference for the improvement of big data mining algorithm in the future.

1. Introduction

With the continuous development of computers and the Internet, the acquisition of relevant information becomes increasingly simple. However, at present, because the large data capacity stored in large databases is increasing rapidly and important information is hidden in large databases, it is difficult to use the previous methods for information retrieval and query. So an intelligent big data mining technology has become the technology needed in this era. At the same time, people are interested in various fields of the development of neural networks. Neural network is a complex network system. It studies biological neural network, then simplifies, summarizes and refines its neurons and neural network. It uses non-linear mapping, parallel processing and structured input and output of information related to the expression of the neural network itself. At first, the application of mining algorithm in neural network is not very optimistic, because of its poor interpretability and long training time. However, with the continuous development of network extraction and other algorithms, the mining algorithm of neural network is becoming more and more popular.

2. Neural Network

2.1 Introduction of Neural Network

Neural network, alias connection model, is an algorithmic mathematical model, which is mainly used to process distributed parallel information. As can be seen from the word "nerve" in the neural network, it is a simulation of the way neurons think. Neural network has three characteristics: parallel processing, neural network structure and non-linear mapping. Its main method of processing information is to change the connection of internal nodes to each other. At the same time, it has the advantages of high anti-noise performance, low error rate and so on. At present, the research of neural network and other interdisciplinary fields is very hot.

2.2 Current Situation and Development Trend of Neural Networks

Since the 1980s, neural networks have attracted the attention of academia all over the world. At the same time, the research of neural networks has begun in China. After decades of development, the development and application of neural networks in China have achieved many successes. Neural network is a kind of complex network system. It has the advantages of high anti-noise performance and low error rate that other systems do not have. With the development of science and technology, these advantages are sought after by people, making its research more and more hot. Generally, the

research of neural networks is divided into eight categories: loosely coupled model, tightly coupled model, transformation model, synthesis model, chaotic model, fuzzy set theory, genetic algorithm and hybrid neural network.

3. Big Data Mining

3.1 Introduction of Big Data Mining

Since the 1990s, big data mining, an extended discipline developed from the field of big databases, has begun to develop rapidly. Large data mining is a process of extracting important information hidden in a large amount of large data. It is mainly aimed at large databases, in which large data are extracted, analyzed and processed to find important information from the extracted information. Big data mining, which finds rules in a large number of large data, discovers important information that has been neglected, and assists decision-making. Big Data Mining is also a subject that can cross with other disciplines. It improves people from simple query operation to mining important information in large data. It is a new technology based on large database, neural network and statistics. Similarly, it has attracted many researchers to study it, making it one of the hot technologies recently. Figure 1 shows a typical large data mining system.
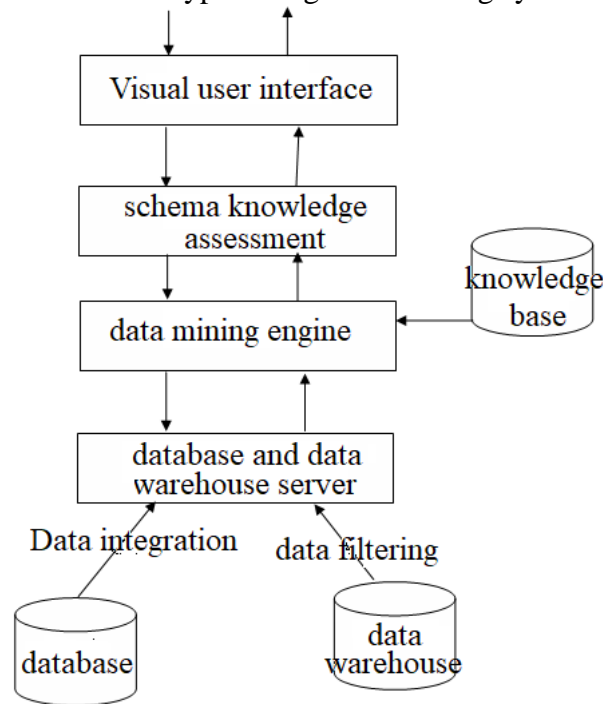

Fig.1 A Typical Large Data Mining System

3.2 Evolution of Big Data Mining Technology

Big data mining technology has been regarded as a new technology with the development of big database technology and information technology. Since the big database was first put forward in 1989, after more than 20 years of development of big database technology, big data mining technology has also made a lot of significant achievements. Table 1 details the evolution of large data mining research.

Table 1 The Evolution of Data Mining Research

phylogenetic scale	Support technology	Product manufacturer	Product characteristics
Big Data Collection(1960s)	Computers, tapes and disks	BM,CDC	Providing historic and static large data information
Big Data Access(1980s)	Big relational databases,Structured Query Language, ODBC	Oracle,Sybase,Informix, BM,Microsoft	Providing historical and dynamic big data information at the record level
Large Data Warehouse and Decision Support (1990s)	On-line analytical processing, Multidimensional large database Large Data Warehouse	Pilot,Com share ,Arbor,Cognos, Micro strategy	Providing retrospective and dynamic large data information at all levels
Big Data Mining	Advanced Computing,Multiprocessor computers, massive databases	Pilot,Lockheed,BM,SGI, Other startups	Providing predictive information

3.3 Current Situation and Development Trend of Big Data Mining

In the world, the technology of big data mining combined with other fields has been widely used, such as SKICAT system, AS system and so on. The paper on big data mining has become one of the hot papers published in international academic journals. At home, because it started a little later than abroad, but the research on big data mining is as hot as abroad, such as AMINER big data mining toolkit, CASDM model and so on. In terms of papers, papers on big data mining published by major domestic research institutions are also ubiquitous. Because big data mining involves many other disciplines, the hot research directions at present are spatial big data mining, multimedia big data mining, time series big data mining, Web big data mining, uncertain big data mining and so on.

At present, big data mining is facing the problems of mining methods and user interaction, performance and diversity of large database types. These problems need to be solved urgently in the development of big data mining.

4. Large Data Mining Algorithms of Neural Networks

There are ten general algorithms in big data mining, they are C4.5, K-means, SVM, Apriori, EM, pagerank, Adaboost, KNN, genetic algorithm and neural network. In this paper, only the large data mining algorithm in the neural network environment is studied. The process of mining is divided into the following steps: the first step is the preparation of large data; the second step is the mining of large data and the elaboration of mining results. Its process is shown in Figure 2.
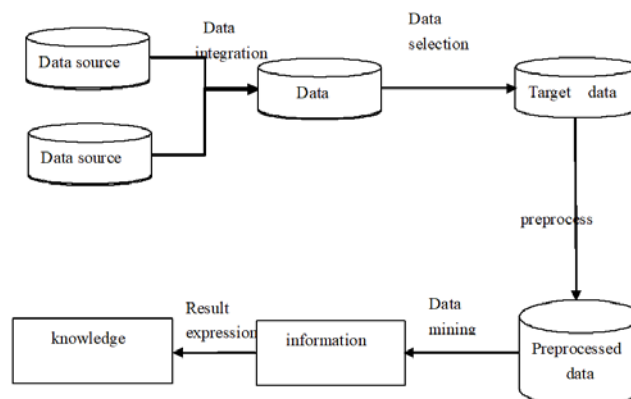

Fig.2 Data Mining Process

4.1 Preparation of Big Data

From the previous big data mining process, we can see that big data preparation is the first step of big data mining. Usually, big data comes from big databases. In the process of reading from these large databases, errors often occur, leading to the emergence of large data errors. When preparing large data, we should preprocess large data. Using statistics and other methods to repair and correct large data, and then the correct large data should be converted, and its programming can be accepted by the algorithm format. Finally, the converted large data sets are managed.

4.2 Big Data Mining Algorithms

The big data mining of neural network is generally attribute rules, and its big data mining process is shown in Figure 3.
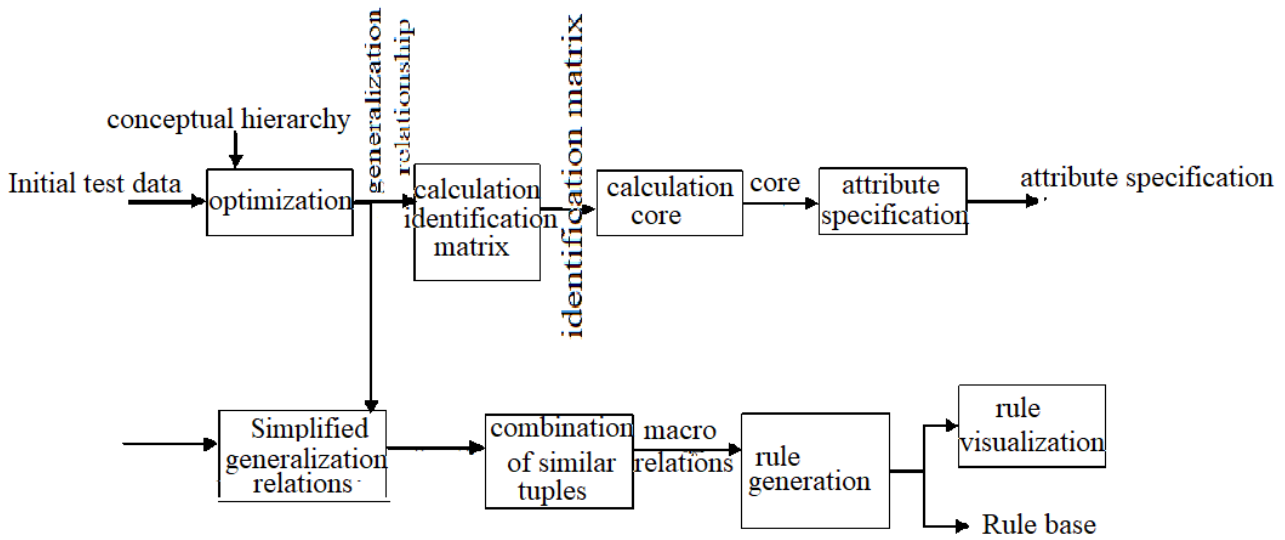

Fig.3 Data Mining Processes

Choosing the number of large training data is a major difficulty of the algorithm. The selection of large training data is related to the accuracy of the neural network. Root mean square error (RMSE) is usually used to evaluate learning performance. Its formula is as follows:

$$e = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(d_{ij} - y_{ij)}}{m*n}}$$

Among them, the number of training samples is m, and the number of neural network output is n.

According to the relationship between training data and error, the error formula is obtained as follows:

$$e = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(d_{ij} - y_{ij})}{Xm*n}}$$

Then the function is substituted as follows:

$$y = x/(1 - \frac{1}{\sqrt{x}})$$

Where x is the coefficient and Y is the cost index. Table 2 shows the relationship between them.

Table 2 Cost Indicators and Their Derivatives and Coefficients

x	1.25	1.5	1.75	2	2.25	2.5	2.75	3	3.25
y	11.84	8.17	7.17	6.83	6.75	6.80	6.92	7.10	7.30
y'	-30	-6.7	-2.24	-0.71	0	0.38	0.61	0.75	0.85

According to the table of the cost function mentioned above, the basis for selecting training large data sets is known. For large data processing, 2.25 is the best factor. Then the algorithm is described, and the flow chart of the algorithm is shown in Figure 4.
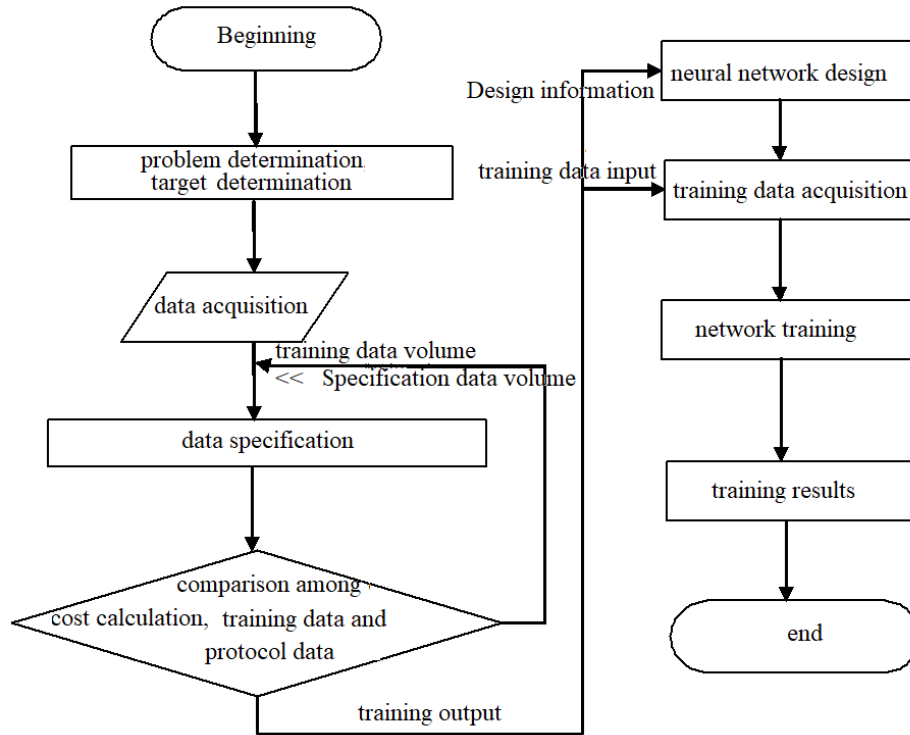

Fig.4 The Flow Chart of Algorithms

In summary, we can understand from the flow chart a process step of large data mining algorithm based on neural network when dealing with large data. At the same time, it clearly shows the whole process of extracting large data from large data mining algorithm.

5. Conclusion

With the development of interdisciplinary research, the medical terminology of neural network has been applied more and more widely in many interdisciplinary fields. Because of its good anti-noise performance and high acceptance of complex large data, the application of neural network in large data mining is becoming more and more popular. At present, people pay great attention to the process of large data mining based on neural network. This paper makes a detailed study on adaptive neural network, and also makes a corresponding study on the key technologies and methods of large data mining based on neural network. Through these studies, the process of large data mining of dynamic neural network is described. At the same time, we hope to provide some reference for the following research on large data mining algorithm in neural network.

Acknowledgement

References

[1] Santosh Thakur, Ramesh Dharavath. Artificial neural network based prediction of malaria abundances using big data: A knowledge capturing approach [J].Clinical Epidemiology and Global Health, Volume 7, Issue 1, March 2019, Pages 121-126.

[2] Alcides Fonseca, Bruno Cabral.Prototyping a GPGPU Neural Network for Deep-Learning Big Data Analysis [J].Big Data Research, Volume 8, July 2017, Pages 50-56.

[3] Qifa Xu, Xingxuan Zhuo, Cuixia Jiang, Yezheng Liu.An artificial neural network for mixed frequency data[J].Expert Systems with Applications, Volume 118, 15 March 2019, Pages 127-139.

[4] Kangyi Wang.Network data management model based on Naïve Bayes classifier and deep neural networks in heterogeneous wireless networks [J].Computers & Electrical Engineering, Volume 75, May 2019, Pages 135-145.

[5] Javeria Amin, Muhammad Sharif, Mussarat Yasmin, Steven Lawrence Fernandes. Big data analysis for brain tumor detection: Deep convolutional neural networks [J].Future Generation Computer Systems, Volume 87, October 2018, Pages 290-297.

[6] Qingchen Zhang, Laurence T. Yang, Zhikui Chen, Peng Li.A survey on deep learning for big data[J].Information Fusion, Volume 42, July 2018, Pages 146-157.

[7] Abir Jaafar Hussain, Panos Liatsis, Mohammed Khalaf, Hissam Tawfik, Haya Al-Asker.A Dynamic Neural Network Architecture with Immunology Inspired Optimization for Weather Data Forecasting[J].Big Data Research, Volume 14, December 2018, Pages 81-92.

[8] Mohammed S. Hadi, Ahmed Q. Lawey, Taisir E. H. El-Gorashi, Jaafar M. H. Elmirghani.Big data analytics for wireless and wired network design: A survey[J].Computer Networks, Volume 132, 26 February 2018, Pages 180-199.

[9] Zebin Yang, Dennis K. J. Lin, Aijun Zhang. Interval-valued data prediction via regularized artificial neural network [J]. Neurocomputing, Volume 331, 28 February 2019, Pages 336-345.

[10] Cristina Zuheros, Siham Tabik, Ana Valdivia, Eugenio Martínez-Cámara, Francisco Herrera.Deep recurrent neural network for geographical entities disambiguation on social media data[J].Knowledge-Based Systems, Volume 173, 1 June 2019, Pages 117-127.

[11] Xin Liu, Yanju Zhou, Zongrun Wang. Recognition and extraction of named entities in online medical diagnosis data based on a deep neural network[J].Journal of Visual Communication and Image Representation, Volume 60, April 2019, Pages 1-15.

[12] J. H. Pacheco-Sánchez, R. Alejo, H. Cruz-Reyes, F. Álvarez-Ramírez. Neural networks to fit potential energy curves from asphaltene-asphaltene interaction data[J].Fuel, Volume 236, 15 January 2019, Pages 1117-1127.