

The Research and Analysis of Data Mining Technology

Mulin Miao

Hebei University of Engineering, Hebei, China

Abstract: With the rapid development of information technology, data analysis, machine learning, artificial intelligence and other technologies arises at the historic moment, followed by the exponential growth of all kinds of information, in the current era of big data, the human brain has been unable to deal with huge amounts of daily data, which leads to the phenomenon of rich data but lack of information, thus prompted the emergence of data mining technology. Computer data mining technology has greatly improved people's quality of life, met people's material and spiritual needs, and promoted the progress and development of the society.

Keywords: Information Technology, Data Mining, Applied Technology.

1. Computer Data Mining Technology Development Technology

1.1. Traditional methods

The traditional methods developed by computer data mining technology include sampling analysis method, multiple statistical analysis method and statistical prediction method, which reduces data mining workload through sampling analysis, and accurate data analysis results through regression analysis and sequence analysis.

1.2. Commonly used algorithms

Common algorithms developed in computer data mining technology include regression algorithms, classification algorithms, cluster analysis, neural network methods, association rules and Web data mining. These algorithms are efficient, superior and comprehensive, and can use various application directions of current computer data mining technology. Regression algorithm mainly analyzes the rules or similarity between data; the classification algorithm mainly classifies the differences between data attributes; cluster analysis mainly classifies the similarities between data attributes; neural network method processes data through centralized testing; association rules processes data through the connection between data; and Web data mining collects various data through the web page.

1.3. Axis-type data mining method

Computer data mining technology development, axis data mining method refers to the axis summary data mining information input, information processing, information research, information output process, the information input in the axis data mining method as a starting point, information output in the axis data mining method as an end, from the beginning to the end can finish data mining.[1]The advantages of axis data mining method are strong timeliness, simple technical principle and convenient operation, but this technology does not classify the mined data. Therefore, the long data mining time is the disadvantage of this method. In general, the application of axis data mining method in the development is relatively common.

1.4. Circular data mining method

Computer data mining technology development, circular

data mining method refers to the information input, information processing, information research, information output each data mining process and real-time operation, will each data mining process into different data mining closed loop, the final information output process can get relatively accurate information data.[2]The advantages of the circular data mining method are strong persistence, strong target and high data mining efficiency, but the circular data mining process of this method is more complex than the axis data mining process. Therefore, the disadvantage of this method is the high requirements for development technicians and the low practical application feasibility.

1.5. Decision tree data mining method

In the development of computer data mining technology, the decision tree data mining method refers to the mining of huge data according to the decision tree statistical classification method, which includes the application of other computational statistical methods, otherwise the scientific and accurate classification results cannot be obtained.[3]Decision tree data mining method is the advantage of can handle huge data, and can directly use statistical methods for mining and classification of data, data mining of target is strong, but the technology of computer technology dependence, the computer technology is still developing, the future the technology is still great progress and development space.

1.6. Visualization technology

Computer data mining technology has greatly improved the application effect of visualization technology. Visual technology can fully show the data features through more intuitive ways such as charts, bringing people a stronger visual feeling. And computer data mining technology itself cannot show the data characteristics, only combining the technology and visualization technology, can improve the efficiency of visualization technology and computer data mining technology, and fully show the advantages of both, so the visualization technology in the use of computer data mining technology development is more common.

1.7. Online data processing

Computer data mining technology can greatly improve the efficiency of data information processing, when the data mining process encounter volume huge data efficiency is

often unsatisfactory, you can through data online processing fully use computer data processing capacity, in order to achieve the purpose of improving the efficiency of the whole data mining technology. Under normal circumstances, this method is only used when the data volume of computer data mining technology is large. The use of this method has high requirements on the synergy of computers, so the data online processing will not have a negative impact on the data processing effect.

1.8. Decision trees and Neural network

Computer data mining technology development can be decision tree and neural network algorithm, decision tree is a kind of efficient statistical method of data integration and classification, neural network can through input layer, hidden layer and output layer get more scientific and accurate data, decision tree and neural network computer data mining technology can analyze the mapping relationship between data attributes and data value, with decision tree classification of a single output. At present, this technology is relatively common in the development of computer data mining technology, and the researchers can predict and analyze the data through the decision tree and the neural network.

2. Data Mining Applications

Data mining is widely used in all walks of life. At the beginning, data mining was mainly applied and commercial. With the development of data mining technology, the application of data mining in other fields began to develop rapidly. In terms of shopping mall sales, we analyze users' historical consumption records, analyze consumer habits and characteristics to make appropriate recommendations and improve sales performance. In the business field of banks, banks and credit can decide whether to lend to individuals, predict fraud and other behaviors. In the transportation field, data mining can help in planning pathways. In the field of scientific research, a large number of experimental data can be analyzed and find out the relevant laws and knowledge. In the manufacturing field, various data of parts are analyzed to find the key factors of improving yield, and optimize production. In the medical field, diseases can be predicted and analyzed. In social networks, users can be grouped, recommend topics and groups interested in, and improve better services for users.

In short, data mining has played an indispensable role in our daily life. In the future, data mining will be integrated into every aspect of our life and play a crucial role for our better life experience.

3. Development and Trend of Data Mining

The field of data mining has made great progress, but in the long run, there are still many directions to be improved, and there are still many problems to be solved urgently. The development trend of data mining is mainly in the following aspects:

Data mining language standardization: There are already many mature, commercial or free data mining systems available, and the collaboration between different systems can promote the application of these systems in the society and in enterprises. This requires the language standardization of these systems. **Complex data:** Usually there are many properties in each table in the database, but not all attributes

are suitable for solving a given data mining problem, some properties even reverse the result, so you need to remove some attributes, how to decide to remove those attributes is also a problem. At the same time, traditional data mining algorithms are mostly for text or digital attributes, while images, maps, video, audio and even spatial and temporal data. These data structures are complex. In these data mining, traditional data mining methods have poor effect, so it is necessary to put forward better mining algorithms for various fields. There is also the data flow, the data in the database is static, while the data flow is real-time and orderly, fast arrival, requiring real-time analysis. These complex data require proposing more refined solutions.

Data processing and unification: Because the problems of data mining are often imprecise, the data mining results often require the intervention of the field experts, who usually ask the questions and help predict the results. Too much or too little data used in data mining will lead to incorrect results. Some data may be wrong or incomplete. These data need to be preprocessed before mining execution. There is still no unified statement on whether the incomplete data is to be completed or abandoned. There are also some data that may have different meanings on the surface, but their essence is the same. How do you distinguish these data and merge them together? At the same time, there are usually some anomalous values in the data set, which are not well adapted to the model, and even affect the results of the model training. How to judge and remove these anomaly points is a very necessary problem to be solved in some large databases. All of these problems occur in the data preparation stage, and the problems are often encountered, and good data sets are crucial to data mining, so better and more systematic solutions are needed.

Privacy protection: It's not a technical issue, but a social one. Data mining at the individual level usually requires various kinds of data related to this person to mine some potential information of the person, but due to the technical or management deficiencies and misuse or abuse of data mining may lead to user sensitive information leakage. Therefore, the protection and processing of mining data privacy is a very important concern.

Distributed data mining: With the development of information technology, a large amount of data is generated every day. A single computing resource is no longer able to process these rapidly growing massive data, and distributed technology provides a lot of good support for these big data. The combination of distributed technology and data mining technology is the direction of future data mining and also an important development trend.[4]

Results interpretation and visualization: Visualization is an indispensable part of data mining. Sometimes the results of data mining may require field experts to explain and intervene, and ordinary users cannot understand its deep meaning, requiring visualization technology to allow experts to directly intervene in the mining process and adjust the mining results. And enable users to understand the data results.

4. Conclusion

Computer data mining technology in the field of science and technology production can be based on neural network and decision tree algorithm extraction parameter features, and then based on the corresponding database, finally to produce the precision detection products after the data transmission to the input layer of the neural network algorithm, after comparison with the hidden layer data parameters, finally can

output in the output layer accuracy parameters meet the requirements of the conclusion. In addition, the regression algorithm, classification algorithm, clustering analysis and other computers can also be used through the computer data mining technology to analyze the reasons why the precision parameter data does not meet the requirements, so that the algorithm can be used to replace the manual decision.

To sum up, the importance of the computer data mining technology is becoming more and more high with the development of the society, and the various industries and fields are becoming more and more dependent on the computer data mining technology. Therefore, each enterprise must strengthen the research and application of computer data mining technology, and try to use the technology to improve the enterprise economic benefits and market competitiveness, so that the enterprise decision-making and management level

is enhanced, and then promote the sustainable development of enterprises.

References

- [1] Sun Zili. The Application of Python Language in Web Data Mining-Sun Zili [J]: 1-2
- [2] Shi Youzhao. Research and Application of Data Mining Technology-Shi Youzhao [J]: 1-2
- [3] Meng Weicheng. The Development and Application Analysis of Computer Data Mining
- [4] Technology-Meng Weicheng [J]: 1-2 Application of [4]Python in Big Data Mining and Analysis-Fang Ji [J]