Analysis of the Application of Data Mining Technology in Economic Statistics

Yao Chen

Yunnan Technology and Business University, Kunming, Yunnan, 651701, China

Keywords: Data Mining Technology; Economic Statistics; Application

Abstract: From the current situation in China, the social economy is in the process of rapid development. Constantly generating data and information in the economic sector can help us make more sound economic decisions and create momentum for future economic development. This has become the most important development method at this stage. Data mining technology can find hidden events with high value in big data, and make more scientific economic decisions based on various technologies such as artificial intelligence and statistics. This article has comprehensively elaborated and analyzed the relevant content. Firstly, the basic situation of China's data mining technology and current economic statistics survey data are introduced, and then the application of data mining technology in economic statistics survey is analyzed. It is hoped that through the elaboration and analysis of the relevant content of this article, the effectiveness of relevant research can be further improved and more power can be created for the development of China.

1. Introduction

There is a large amount of data in the economic statistics database. Most data is used by mathematical statistics [1]. This approach results in a single form and shallow depth. The only work done is to save and organize the existing materials. In the economic statistics process of data mining technology, it is difficult to accurately analyze these data if it relies only on traditional economic statistics methods [2]. Therefore, data mining technology should be more widely used in economic statistics. Data mining technology can make the data obtained by economic statistics more extensive and in-depth. Data analysis can also be more accurate. Thereby improving the quality of the data. The information obtained is also more accurate and extensive.

2. Data Mining Technology Overview

Data mining technology is only a deeper mining of data information [3]. The technology is mainly to simplify the database, it has many different changes, and analyzes the simplified data information to find out the data that can be used. Use information through continuous analysis and organization. The effect is further enhanced [4]. China's current economic statistics survey data is very rich, and many of the data are not complete, leading to the use of some commonly used statistical methods to achieve better results. Data mining technology can effectively solve this problem. Through software and program algorithms, comprehensive analysis of incomplete data, and finally develop a more reasonable statistical data table, so that relevant users can directly extract and use the data. In this process, it is usually divided into three steps. The first is to prepare the corresponding data, then the data mining work, and finally the analysis of the results.

3. Characteristics of Mining Technology

Data mining technology can process large amounts of raw data to make it more valuable [5]. Data mining technology has the following features.

1) Data mining technology is easy to face a large number of economic statistics.Data mining technology can continuously expand its own database. This satisfies the search needs. You can also continuously improve the accuracy of your search data during the search process. It allows users to find the data they need faster.

  DOI: 10.25236/scmc.2019.018

2) Data mining technology can perform a Lenovo search on keywords input by users.

This is a more convenient search method for users who are unsure of the search content. And the information found by the content input by the user can be arranged in an orderly manner.

3) Data mining technology can find the association between data and data in a large amount of information in the database. Then contact other aspects of the data to predict future economic development.

4) Data mining technology can handle very large data. Data mining technology can be easily handled in the face of large amounts of data [6].

4. Advantages of Applying Data Mining Technology to Economic Statistics

4.1. Strong comprehensive ability

With the development of China's economy, it is inseparable from data statistics, and the complementary effects of the two, data statistics play a more important role. Relevant departments need to refer to economic data statistics in the process of implementing major decision-making, but the management forms between departments are very different, resulting in large differences in data types, forms and requirements [7]. To some extent, work cannot be done in a normal and orderly manner. To this end, in economic statistics, data mining technology can be used to freely convert data forms to meet the needs of different departments, thereby promoting rapid economic development.

4.2. Actual statistical effects

In the process of statistical economic data, data mining technology plays a greater role in it, real-time statistical analysis of economic data in the database, and effective processing on this basis to improve the effectiveness of data applications [8]. And science. In addition, data mining technology can effectively explore valuable information, thereby playing the role of economic statistical analysis, maximizing the efficiency of data information management, ensuring the authenticity and reliability of economic statistical information, and thus effectively implementing Economic statistics activities.

4.3. Wide range of uses

Data mining technology can be effectively applied to different management departments, to maximize the effective integration of data information, to meet statistical needs, to make the data parameters of economic activities accurate and true, so as to effectively improve the smooth implementation of economic activities, maximize data mining the role of technology [9].

5. Application of Data Mining Technology in Economic Statistics

5.1. Preprocessing of economic data

The pre-processing of economic data is the most basic way to deal with statistical activities of economic data [10]. The reason for the pre-treatment of mining technology economic statistics is that the mining technology itself is limited by various economic conditions and it is impossible to replace the collection function of the economic system. It is just an intelligent analysis technology based on providing basic information. There are many things to deal with, such as inaccuracies in processing data, unreality in processing data, and errors in processing information. The most important thing is to process data information with a large gap. The so-called data cleansing is the processing and analysis of problematic databases. There are three main methods for cleaning data: estimation methods, averaging methods, and smoothing methods. Among them, the mean method is mainly used for data without numerical value and noise. It is a new form of application of fuzzy concepts in modern analytical techniques. Known data in the database is used to make up for unknown data to fill in the gaps. Ensure that the mining system and the underlying data analysis organizations are functioning properly to obtain more accurate statistical analysis data, As shown in

Fig. 1:

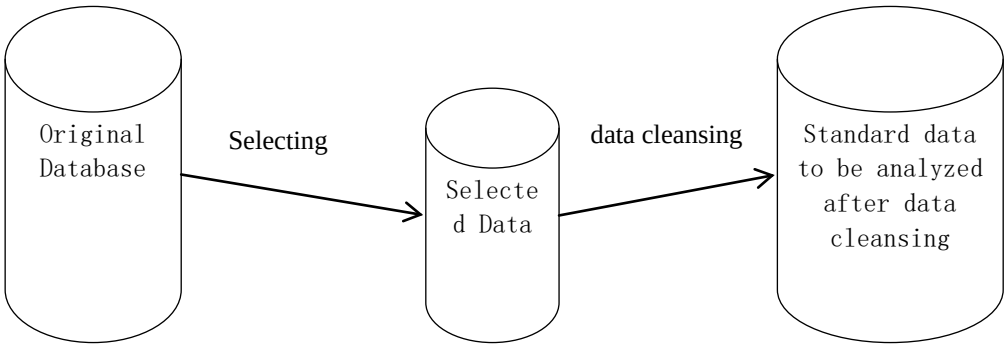

Fig.1. Data preprocessing[11]

5.2. Integrated structure aspects

After applying the data mining technology, even if the data in the same area is the same, if the statistical subject is different from the standard, the corresponding data integration system must be adopted to ensure the optimization of the data integration effect and the accuracy of the data. Statistical data enhancements. This is also data mining. The goal that technology has been pursuing.

On the one hand, the integration structure should be deeply integrated. Due to the large amount of data and relatively complex types of economic activities, it is necessary to systematically and in-depthly process the integration of economic data information and the representation of data. Among them, in the use of data mining technology, in order to compare and analyze std-id and std-no, it is necessary to compare the metadata module with Ranger to ensure that the entity recognition efficiency is guaranteed and meets the quality standards.

On the other hand, the problem of human redundancy should be analyzed in depth, because data mining technology is essentially a process of processing data in depth. In this case, in order to ensure the integrity of economic statistics, it is necessary to ensure that the technical model is in the simplest state, and focus on the real-time concentration and simplified operation of the positive correlation data, comprehensively consider the redundancy attributes, and refine the problem data. .

5.3. Building a decision tree

Due to the current situation, decision trees respond very quickly to situations and are therefore widely used in data mining techniques. In general, there are two steps to building a complete and accurate decision tree: The first step is to simplify and validate the decision tree method to create valid data that can be used for data output analysis. The second step is to make full use of the decision tree of the build and effectively analyze and classify the data from the basis of the decision, and then slowly extend it to all parts of the decision tree [12]. The data will not run until certain conditions are met. In order to terminate the segmentation in a particular operation, we must first implement two specific conditions: First, when two segments are in progress, if the data on the segmentation point is accurately segmented, the decision will be successfully completed; secondly, it is impossible to continue classification, but can divide the data. However, the above situation will bring some unnecessary trouble to the decision, so in order to reduce the negative impact on the decision data itself through some unreasonable methods, it is necessary to establish a decision tree. The constructed decision tree is pruned according to the specific usage to make the decision tree more perfect.

5.4. Genetic algorithm aspects

The so-called genetic algorithm combines the natural and genetic mechanisms of the organism and randomly extracts the algorithm. In practical applications, comprehensive consideration of social issues, effective collection of information on designated populations, and final results based on implicit information integration and analysis. Because genetic algorithms have some implicitness, they can be effectively combined with other models to collect hidden data. Subsequently, the

existing mining data information is analyzed in depth and applied in practice. It should be noted that economic issues are a matter of development and change, so internal linkages are very complex. Through the genetic algorithm as an important reference, the genetic algorithm can be extended downward at the source, the data can be obtained efficiently, and the overall analysis of the data information can be guaranteed. Economic issues are more direct and specific, ensuring that relevant staff members deal with problems more intuitively and demonstrate hidden problems. In this way, statistical work can be guaranteed to be simpler and simpler.

6. Conclusion

As far as China's current development work is concerned, economic statistics surveys have improved. With the continuous application of data mining technology, statistical data analysis has achieved good results and achieved high quality. At this stage, data mining has reached a relatively stable state and is being studied. However, in the actual application process, there are still many problems. It is necessary to study it through more effective methods, improve its actual effect, make China's data mining technology get more progress, and create more for China's economic statistics. power.

References

[1] Harvey R, Mcbean E. A Data Mining Tool for Planning Sanitary Sewer Condition Inspection. Conflict Resolution in Water Resources and Environmental Management. Springer International Publishing, 2015.

[2] Abdullah A, Albidewi I. Using data mining for analysis of meat production as source foods and the challenges in 2030: comparison of six South Asian countries. Journal of Applied Animal Research, 2016 44 (1) 180-188.

[3] Edwards M, Rashid A, Rayson P. A systematic survey of online data mining technology intended for law enforcement. Acm Computing Surveys, 2015 48 (1) 1-54.

[4] Lee E B, Kim J, Lee S G. Predicting customer churn in mobile industry using data mining technology. Industrial Management & Data Systems, 2017 117 (1) 90-109.

[5] Yan L, Hu P, Li C, Lei F. The performance prediction of ground source heat pump system based on monitoring data and data mining technology. Energy and Buildings, 2016 127 1085-1095.

[6] Liu X, Huang Z, Tong B. Review on the data mining technology and the applications on financial analysis area International Conference on Communication & Electronics Systems. IEEE, 2017.

[7] Zhang K, Li G, Gong J Z, Zhang M. Research on Assembly Data Mining Technology of Complex Mechanical System. Procedia CIRP, 2016 44 97-101.

[8] Yau-Ren S, Ching-Hsing T, Yung-Hsiang H, Yu-Ting K. The Application of Data Mining Technology to Build a Forecasting Model for Classification of Road Traffic Accidents. Mathematical Problems in Engineering, 2015 2015 1-8.

[9] Haijian Z. Analysis and Application of Consumer Features with Cloud Computing and Data Mining Technology International Conference on Intelligent Computation Technology & Automation. IEEE, 2015.

[10] Jin M, Wang Y, Zeng Y. Application of Data Mining Technology in Financial Risk Analysis. Wireless Personal Communications, 2018.

[11] Chunguang Y. Data Mining Technology in the Implementation of the Library Service System in Colleges and Universities International Conference on Intelligent Systems Design & Engineering Applications. IEEE Computer Society, 2015.

[12] Usmani Z A, Khan S, Kazi M, et al. Zaimus: A department automation system using data mining and web technology 2017 International Conference on Nascent Technologies in Engineering . IEEE, 2017.