ATLANTIS PRESS

Advances in Economics, Business and Management Research, volume 80

1st International Conference on Business, Economics, Management Science (BEMS 2019)

Application of Data Mining Algorithm in Tax Source Management

Shenghui Yan[1], Wenbao Zhao[1], Lei Yu[1], Bin Zhang[2], Chi Zhang[1], Yuexin Pang[3]

[1]Hebei University of Economics and Business, Hebei, China

[2]Beijing University of Technology, Beijing, China

[3]Capital University of Economics and Business, Beijing, China

Abstract. The advent of the era of big data, data from all walks of life in the big background, the related technology and data can be effective screening and application of data mining in different types of tax source management provides effective reference, and to face and solve the problem of tax management provides decision-making reference for specific applications. This paper analyzes and discusses the development of tax source management and the application of data mining algorithm.

Keywords: Tax source management; Data mining; Strategic analysis.

1. Development Status of Tax Source Management

As an important source of national finance, it is an important way to effectively regulate national income and regulate economic operation. In order to further promote the development of national economic construction, it is necessary to carry out effective distribution of national income and improve fiscal and taxation systems. Not only to ensure the stability of the national tax, but also to carry out and implement the relevant laws and regulations. Therefore, it is not necessary to establish an efficient and accurate tax system [1].

1.1 Content and Meaning of Tax Source Management

Tax source management is to organize, regulate, make decisions and supervise the whole process of tax collection in accordance with relevant national laws and regulations on tax collection and administration, so as to ensure the stability of national tax distribution activities. Therefore, tax administration is a very important part in the process of national economic management. Tax source management in general can be divided into the following several parts: the content of the management main body of the sources for the state, the object of tax source management for the social from all walks of life in the tax to participate in the activities of all relevant national allocation process, not only that, management can also be sources for the allocation of tax related activities organization, supervision and decision making and coordination, etc., and according to the tax as specified by the legal basis and objective distribution law of tax, security tax source management function reasonable and stable.

2. The Role of Tax Source Management

Tax source management and its management function are inseparable. The function of taxation plays a role in regulating economy, revenue, management and supervision. Therefore, the perfection of tax source management system determines whether its function can be effectively played. First of all, taxation can provide a stable fiscal revenue for the state to organize and raise construction funds. Secondly, taxation is an important way and means for the macro-adjustment of the country's overall economy. Especially under the dual influences of economic globalization and reform and opening up, the rational and scientific distribution of resource market can be realized by utilizing tax source management and tax intelligence. Whether tax revenue is reasonable even affects the changes of income and distribution among various industries, regions and regions in society. Therefore, the effective play of tax functions is an important guarantee for tax source management.In the process of

socialist economy high speed development in our country, there are still a variety of economic system and the status of the diversified production means to coexist, as a result, the tax authorities and relevant personnel to the taxpayers and society economic activity in the process of the tax, the need to play to the tax sources management in the process of the role of supervision and management. We will seriously deal with taxpayers' violation of tax laws and regulations, even tax evasion and evasion, and further strengthen economic management and accounting [2].

3. Application of Data Mining Algorithm

3.1 Core Technology of Data Mining

The main source of data mining is database, which has the following important characteristics: first, the data content in the database is huge[3]. Therefore, in the process of data processing, its scale needs to reach a certain amount of data to ensure the rationality and accuracy of its data. Second, the data stored in the database larger dimension and high complexity of the data is more, and with the accumulation of time, the data will show several different types, including time series data will appear in the change of time more link relation data, law, heterogeneous database, multimedia data, spatial data and space-time describe data types. However, the traditional data mining method is too simple to effectively conduct in-depth data mining according to its different characteristics. Therefore, the following technologies and methods should be used for data mining operations [4].First of all, the data relationship model in the database can be used to conduct structured query language, association rules and other related algorithms, extensibility technology and other effective mining and computing. Secondly, hierarchical clustering and approximate information retrieval can also be carried out according to the similarity of data. Finally, the application of computer algorithm is carried out through relevant theoretical methods such as data structure and algorithm design and analysis, and assisted by regression analysis, neural network and other related algorithms. With the progress and improvement of technology, the application of data mining technology is also deepening and innovating along with the emergence of new theoretical ideas and research methods, such as scientific computing visualization and evolutionary computing.

3.2 Specific Functions of Data Mining

Data mining bureau technology can solve many practical problems in practical work, and its specific functions mainly include the following aspects. When people describe a person or a thing, they tend to describe its representative characteristics, which is conducive to distinguish it from other objects at the same time. Therefore, the targeted recognition of features can be extracted in a standardized way by summarizing and comparing. At the same time, data mining technology can extract regular or frequent information from massive unanalyzed data, explore the regularity of phenomena or things that appear in randomness, and provide favorable theoretical analysis basis for decision-making. Clustering analysis and the analysis of outliers is also one of the functions of data mining shows, in the case of classified information is not clear, usually with the method of cluster analysis for data classification, through the clustering analysis and relevant way, can be in the same or similar object minimizes the differences, and then improve the effect of the data mining technology [5]. However, liqun store analysis is often applied and the data information analyzed lacks regular data point rules. Therefore, accident monitoring with small probability and event situation can be processed and analyzed.

3.3 Basic Introduction of Layer Clustering Algorithm

Hierarchical clustering algorithm is to decompose the given data in the hierarchy until the relevant conditions are met. Generally speaking, hierarchical clustering algorithm is to form the object data into a clustering tree and analyze it from top to bottom or from bottom to top according to the different levels. The hierarchical algorithm is also divided into two methods: splitting method and condensing method. The method of splitting is to gather data objects in a cluster at the initial stage, and gradually and meticulously decompose them into better clusters, until each final data object can be set up as an

independent cluster, or finally meet certain requirements. In short, the algorithm is more widespread across applications.

3.4 Basic Introduction to K-Means Clustering Algorithm

means clustering algorithm is also a widely used algorithm. In general, (1) gives n data samples, and let I randomly select K initial clustering centers.

$$Z(I), j = 1, 2, , K$$

(2) if the above conditions are met, the algorithm is entered and the distance from the sample clustering is entered.

$$D[xi , zi (I)], i = 1,2,.....N, \quad j = 1,2.....K ;$$

$$D[x_i, Z_j(I)] = \min\{D[x_i, Z_j(I)], i = 1, 2, \cdots, n\} \tag{1}$$

(3) xi ∈ wk $Z_j(2) = \frac{1}{n}\sum_{i=1}^{n_j} x_i \quad j = 1,2,\cdots,K$, and the value of the error squared sum criterion function Jc, $J_c = \sum_{j=1}^{K}\sum_{k=1}^{n_j}\left\|x_k - Zj\,(2)\right\|^2$.

(4) judgment: if $\left|J_c(I+1) - J_c(I)\right| < \varepsilon$ That's the end of the algorithm.

This algorithm is also known as iterative clustering algorithm. In general, when the distance between two data objects is small, the similarity between the two data objects is large. Therefore, the data objects that are close to each other are formed into corresponding clusters, and independent and compact clusters are taken as the final analysis target. Because this algorithm is simple, it is also one of the classical algorithms. In the process of continuous development and innovation of k-means clustering algorithm, researchers of relevant disciplines in different fields have conducted targeted analysis and research on this algorithm, and extended multiple variants, all of which have made effective contributions to social development [6].

4. Specific Application of Clustering Analysis

In the process of tax source management, specific financial data can be reduced in dimension through cluster analysis, and effective cluster analysis can be carried out according to the results of different dimensions. Relevant research results show that in financial data, the data with reduced dimensions are more reasonable in the fitting process. Although the methods of dimensionality reduction are different, the ultimate purpose is to conduct highly targeted analysis, screening and elimination according to the value of independent variables, and conduct a detailed study on the principal components of the main data variables. Therefore, in the application and processing of tax source management, the data mining algorithm of its clustering analysis provides more guarantees for the stability of the tax industry.

5. Conclusion

As a practical tool, in-depth data mining through the financial data of various enterprises not only provides an effective way for tax enterprise supervision, but also simplifies the indicators of tax source management. Future tax source management provides a new way of thinking. Therefore, in

the future development, it is very important and necessary to constantly optimize and improve the tax database.

References

[1]. Xin han, Application of data mining algorithm in tax source management under the background of big data [D].

[2]. Shuang liu . Application of data mining technology in tax work [D].

[3]. Jing xu , Ying sun . Application of data mining algorithm in financial management software [J]. Science and technology economics guide, 2017(14).

[4]. Dan li . Research on anomaly data mining algorithm and its application in taxation.

[5]. Zheng chui-yong, Xu Li, dubious svala, et al., an improved clustering algorithm and its application in the tax analysis [J]. Journal of PLA university of science and technology (natural science edition), 2010, 11 (2) : 202-206.

[6]. Jia-ningzhang. Application of association analysis data mining algorithm in online tax declaration system [J]. Computer knowledge and technology, 2007, 1(2):299-300.