



Application of Data Mining in Public Management Classification

Fuyan Hong and Louyan Sun^(✉)

Harbin University of Commerce, Harbin 150028, China
sly13555099111@163.com

Abstract. Data mining technology has become a very powerful tool to improve organizational competitiveness, organizational performance and service level. Many fields of public management, such as e-government, government performance management and public crisis management, are also in urgent need of introducing data mining technology. Its significance lies in the fact that the introduction of data mining technology in public management is the need of government decision-making, can provide technical support for government information sharing, and is an important guarantee for government informatization. Starting from the definition and characteristics of data mining, this paper demonstrates the feasibility and importance of the application of data mining in the classification of public management through the method of literature, and provides a new idea and solution for government departments to make public decisions.

Keywords: data mining · Public management · app · application

1 Introduction

In recent years, with the increasing integration of electronic information in China's economy, science and technology, society and other aspects, a large number of data sets have been generated, which have a profound impact on the government's operation, the development of the national economy and the people's livelihood security. Under this background, the government's primary concern is to encourage and support various departments to mine and use data resources, adjust and improve various management systems, management ideas and methods, and use big data technology to mine and sort out various information, so as to provide reference and basis for the government to make public decisions. The application of big data technology can improve the openness of information and work efficiency, and the transparency of government management and work is gradually enhanced. In addition, after the reform of government management and information work, the government behavior mode will be innovated [1]. At the initial stage of database development, the application of office automation system (OA), business transformation (TPS / EDPS), information management (MIS), decision support system (DSS) and expert system (ES) in public administrations have supported the savings of cost management, the improvement of management efficiency and the supply

of goods good business for the public for the public. However, since the 1990s, the development and realization of information technology and the development of technology, Public management departments and public managers are caught in the ocean of data. They can't find knowledge and rules from the rich data, so they lament that "data is rich but knowledge is poor". So today, public managers call the databases and data warehouses that they built by themselves and once regarded as "treasures" as "datatombs". How to mine knowledge and discover rules from abundant data has become one of the core issues of public management today, which has attracted the general attention of public management departments, public managers and public management researchers [2].

The data mining technology, which has made great progress in the past 10 years, is an information technology that automatically converts data into knowledge by solving how to extract (mine/discover) unknown but useful information and knowledge from a large number of incomplete, noisy, fuzzy and random data. However, the application of this knowledge discovery technology mainly focuses on the field of business management. For these reasons, this paper first introduces the structure and functions of data mining system, and then illustrates how to apply data mining technology in public management field with examples. Finally, the application of information technology in public administration is discussed briefly and hopefully. Document Management (DM) is a new technology developed in recent years. Its topics include machine learning, knowledge discovery in artificial intelligence, data warehousing, online analytics, multi-databases, and more. It is considered one of the most effective ways to solve the problems of "rich information" and "poor knowledge" in the current information industry. Data is just a trivial process of extracting information and knowledge hidden in data sets that are unknown to people before, but potentially useful, and share it accordingly as a model that can be understood in the end. Information can not only learn the existing knowledge from a lot of incomplete information, without noise, vague and random, but also discover new unknown knowledge. The knowledge gained is not only easy to understand, but also easy to keep and use. This knowledge can be used in many areas such as information management, question answering, decision support, process management, etc. Information systems are simply human-computer interaction processes that are connected and repeated through many steps. The complete data mining process includes problem description, data selection, data preparation, data conversion, running data mining algorithm and analysis results. Assessment, etc. [3]. As shown in Fig. 1:

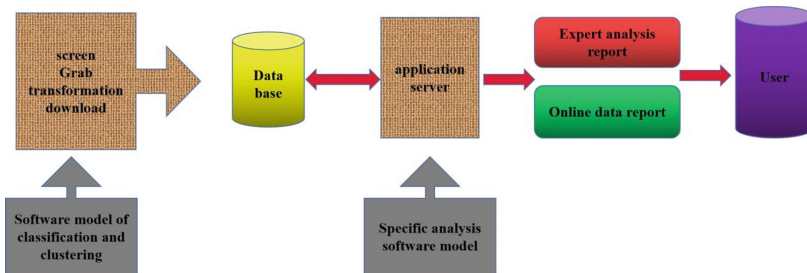


Fig. 1. Data mining process diagram

2 Objects and Functions of Data Mining

The purpose of the data can be all kinds of data centers, especially including relational data, data centers, data centers, product-centers database, database, multimedia data, heterogeneous data, heritage The functions of data mining such as WWW web site data, time database and time log data refer to the types of models that should be found in the processing of data mining. Generally speaking, it can be divided into description only and prediction only. Specifically, it includes: ① Concept/class description is the most basic form of descriptive mining. It describes the relevant data sets of a given task in a concise and summarized form, and provides interesting general features of the mined data. It consists of characterization and differentiation. Conceptual data can be visualized by users in various forms. Compared with OLAP (On-Line Analytical Processing), concept description can more automatically handle complex attributes of data types and their clustering [4]. ② association analysis is the mining of association rules between data. Association rules show the conditions that attribute values frequently appear together in a given database, and it is widely used in shopping basket analysis and transaction data analysis. ③ clusteranalysis is the grouping of data objects into multiple classes or clusters (the difference from classification is that the data to be divided is unknown): loyalty in the same cluster. Objects have high similarity, and there is a big gap between objects in different clusters. Through clustering, dense and sparse areas can be identified, which is convenient to find the global distribution pattern. ④ outlieranalysis is to dig the general model of the data that often or occasionally appear in the database and are significantly different from other data. It can be widely used in fraud monitoring, abnormal information detection and other fields. ⑤ dataclassification is a model/function that distinguishes data classes or concepts on the basis of description, So as to predict the object classes with unknown class markers. The basic technologies of data classification include decision tree induction, Bayesian classification, neural network, genetic algorithm, K-nearest neighbor analysis, rough set method, fuzzy set method and case-based reasoning, etc. ⑥ Prediction is to mine some data values (not class markers) that are desired but are vacant or unknown, It can also identify the methods for forecasting the distribution trend based on available data, such as linear regression, polynomial regression, logarithmic regression and Poisson regression. ⑦ evolution analysis, which describes and models the laws or trends of related data (including characterization, differentiation, association, classification and clustering) whose behaviors change with time. The typical evolution analysis is periodic pattern analysis [5].

3 Application of Data Mining in Public Management Classification

3.1 Application of Data Classification

Classification of data plays an important role in the field of public management. The fuzzy logic is effectively quoted, the fuzzy boundary is reasonably defined, the precise definition is not based on the set, and a truth table is selected in a reasonable range to show a specific value, thus promoting the convenience and reliability of high-level abstract processing.

3.2 Application of Time Series Analysis

The typical time series analysis is periodic analysis. On the basis of mining the periodic patterns, we find the repeated patterns and mine their inherent laws. In the field of public management, time series analysis can mine the whole cycle and part of the cycle, and also clarify the cycle or cycle association rules, including the tidal law and the effect of every day of the year on the seasonal cycle, etc., so as to provide services for the national public management needs, facilitate the adjustment of the national conference plan, and promote the smooth development of various national affairs activities [6].

3.3 Application of Cluster Analysis

The in the field of public management is mainly reflected in the regional ethnic autonomy and the formulation of public policies. At present, in the process of social operation and development, social problems have certain complexity, and the formulation of public policies should be targeted, and the types and regional characteristics of policy objects should be fully considered to promote the effective solution of social problems. Generally speaking, cluster analysis plays an important role in the design of highway and railway transportation network, urban planning, urban river system planning and bus route design in the public domain [7].

3.4 Application of Correlation Analysis

Association analysis is a technique of mining association rules, which can accurately judge the relationship between two factors in the system. The effective application of correlation analysis can judge whether high salary is effective or not, and help to effectively solve the problem of park or urban green square site selection, and improve the public environment and the comprehensive benefits of the city.

3.5 Application of Concept Class Description

In the field of public management, conceptual description is good at finding the effectiveness of different poverty alleviation methods, including micro-loans, charity relief transfer relief and other poverty alleviation methods. Therefore, when applying data mining technology in the field of public management, we should select the appropriate poverty alleviation methods according to the characteristics of different vulnerable groups, so as to promote the effective realization of poverty alleviation goals. At the same time, in the field of public administration, if you want to know the performance of civil servants with different academic backgrounds in government departments, you can reasonably use the concept class description on the basis of investigation and statistics, connect the databases of all relevant attributes in an orderly manner, and convert the tasks into initial target class working relationship and initial comparison relationship query. After relevant analysis of various data, irrelevant data results will be deleted, leaving relevant data items for further analysis. Then, we can show the comparative description of the result categories in the form of visualization or rules, such as charts, so as to draw relevant conclusions, and carry out targeted management for civil servants with different academic backgrounds, so as to comprehensively improve the effectiveness of public domain management [8].

4 Innovation Trend of Data Mining in Public Management Classification

The era of big data provides technical support for the government to realize information management. Applying big data technology to the management of various government departments has greatly improved the management efficiency of various departments. The openness and openness of information under big data also enhance the enthusiasm of the public to participate in political decision-making, which makes social governance diversified. In the past, the bureaucratic system in government management organizations led to slow transmission and processing efficiency of government information and high information error rate, which could no longer meet the actual needs of the management work of various departments. In the context of big data, cooperation and collaboration between government departments can be strengthened by relying on the use of big technology. The management model has become more collaborative and efficient. The government should use all the benefits and functions of big data technology to understand the nature of the overall management function. Realize the social coverage of the resource sharing system, and promote the substantial improvement of government management efficiency [9].

As can be seen from Fig. 2, through the flat decision-making mode of data system, a dynamic circular system of data collection-decision analysis-decision-policy influence can be formed more quickly and effectively. The emergence and development of big data has created conditions for the diversified governance of society. In the era of big data characterized by openness and diversity of data, the government can get all kinds of data from the public more easily with the help of the Internet and mobile networks, which can serve as an important basis for government decision-making. At the same time, the application of big data technology also provides conditions for the communication between social organizations, individuals and governments [10]. By using big data information and technology, the government can quickly grasp the implementation of

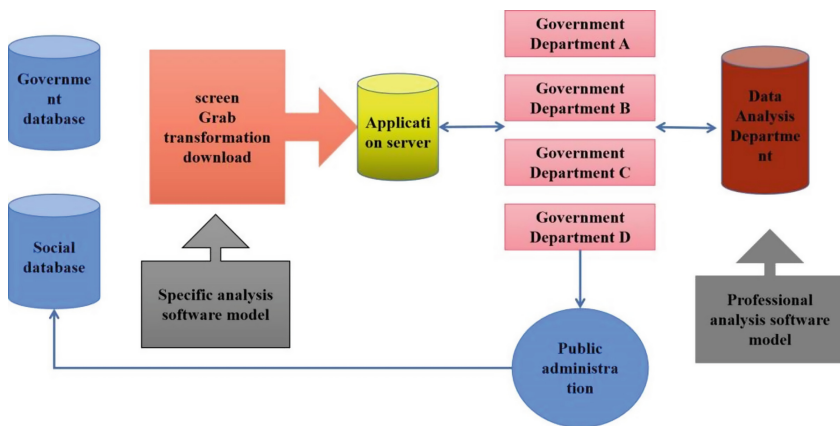


Fig. 2. Government decision-making “flattening” model

relevant policies, evaluate the implementation results of relevant policies more quickly and conveniently, and objectively and effectively evaluate the impact of policies.

5 Conclusion

In a word, data mining is a modern information technology, which is good at mining valuable knowledge and information from unobtrusive data, meeting the needs of social groups, and providing quality services for social operation. It plays an important role in biomedical, business administration, financial telecommunications, public management and other fields. Effective use of data mining technology will help improve the level of information in the government, improve the efficiency of the government, and promote sustainable development in society.

At present, there are still some constraints in the application of data mining technology in the field of public management classification (such as the limitation of information technology for managers, the limitation of interface design and visualization of data mining software), and data mining will also bring problems such as information security and privacy invasion, but compared with the benefits it brings, we believe that the prospect of applying information technology in public administration will be very good.

References

1. Tangut, B. . (2021). Cloud technology management in public and big data mining. (1).
2. Jiang, Y. , & Atif, Y. . (2021). A selective ensemble model for cognitive cybersecurity analysis. *Journal of Network and Computer Applications*, 193(1), 103210.
3. Wang, J. , Li, J. , Gao, Z. , Han, Z. , & Wang, X. . (2021). Resource management and pricing for cloud computing based mobile blockchain with pooling. *IEEE Transactions on Cloud Computing*, PP(99), 1–1.
4. Wu, D. , & Guo, M. . (2021). Application of data mining in traditional benchmark evaluation model for buildings energy consumption. *Scientific programming (Pt.8)*, 2021.
5. Moretto, V. , Elia, G. , Schirinzi, S. , Vizzi, R. , & Ghiani, G. . (2022). A knowledge visualization approach to identify and discovery inner areas: a pilot application in the province of lecce. *Management decision* (60–4).
6. Kunis, S. , Hnsch, S. , Schmidt, C. , Wong, F. , & Weidtkamp-Peters, S. . (2021). Omero.mde in a use case for microscopy metadata harmonization: facilitating fair principles in practical application with metadata annotation tools.
7. Mathan, K. , Kumar, P. M. , Panchatcharam, P. , Manogaran, G. , & Varadharajan, R. . (2022). A novel gini index decision tree data mining method with neural network classifiers for prediction of heart disease (retraction of vol 22, pg 225, 2018). *Design automation for embedded systems*(2), 26.
8. Martin, A. , Capilla, R. M. , & Anquela, A. B. . (2022). Big data architecture and data mining analysis for market segment applications of differential global navigation satellite system (gnss) services: case study of the analysis of the demand for navigation and agriculture. *Journal of navigation* (2), 75.
9. Liu, J. , Shi, G. , Zhou, J. , & Yao, Q. . (2021). Prediction of college students' psychological crisis based on data mining. *Mobile Information Systems*, 2021(23), 1-7.
10. Li, H. , Huang, W. , Zha, Z. , & Yang, J. . (2021). Application and platform design of geospatial big data. Copernicus GmbH.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

