

Retraction

Retracted: Design and Application of Intelligent Management Platform Based on Big Data

Security and Communication Networks

Received 26 December 2023; Accepted 26 December 2023; Published 29 December 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] Y. Su, M. Dai, and H. Zhou, "Design and Application of Intelligent Management Platform Based on Big Data," *Security and Communication Networks*, vol. 2022, Article ID 1790678, 14 pages, 2022.

Research Article

Design and Application of Intelligent Management Platform Based on Big Data

Yunuo Su,¹ Minhui Dai ,^{2,3} and Haoyu Zhou¹

¹College of Public Administration and Law, Hunan Agricultural University, Changsha 410125, China

²Teaching and Research Section of Clinical Nursing, Xiangya Hospital of Central South University, Changsha 410008, China

³Department of Ophthalmology, Xiangya Hospital of Central South University, Changsha, China

Correspondence should be addressed to Minhui Dai; daiminhui@csu.edu.cn

Received 17 June 2022; Revised 15 July 2022; Accepted 10 August 2022; Published 11 October 2022

Academic Editor: Hangjun Che

Copyright © 2022 Yunuo Su et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Big data technology has greatly promoted the construction of intelligent administrative management and improved the decision-making ability continuously. Data mining also lays a solid foundation for the construction of administrative management platform and reflects the potential value of data. In this study, an intelligent management platform based on big data is designed and implemented. First, the problems of the intellectualization of administrative management are discussed, and the big data platform and functional framework of administrative management are introduced. Second, in order to apply data mining to administrative management, the object of data mining in administrative management is defined, and a data mining system is designed. Finally, the application of machine learning methods such as cluster analysis in administrative management is analyzed in detail. The research results show that the application of intelligent management platform based on big data can promote the construction of intelligent administration and lay a good foundation for the development of more perfect intelligent administrative management.

1. Introduction

With the rise of Internet and big data technology, human society is gradually developing into an interconnected and highly integrated society [1]. Big data makes the development of social, economic, and political life full of new vitality and has become a powerful engine to promote the intellectualization of administrative management [2]. Big data can provide strong technical support for the intellectualization of administrative management. Due to the fragmentation of traditional administration and other reasons, it is impossible to comprehensively and effectively grasp social problems such as social risks [3, 4]. Therefore, applying big data platform and its information processing system to administrative management and realizing multiagent collaborative governance is one of the ways to realize low-cost, fast, and effective management [5]. In intelligent administration, information technology promotes a more reasonable organizational structure of departments through self-

iterative upgrading. It will make the governance behavior more efficient and optimized [6]. Accordingly, with the help of the information technology platform, relevant departments have realized the efficient combination of work organization and work process. It also provides transparent and high-quality services for the whole society, so as to achieve integrated governance [7]. Big data can easily and quickly collect and analyze a wide range of data information and apply it to administrative management. Thus, the administrative department can accurately grasp the goal of social public opinion and realize social risk prediction and real-time optimization of events. In addition, due to the rapid and accurate prediction of risks, the ability of the government or enterprises to control risks is greatly enhanced. This helps to formulate the corresponding treatment plan at the first time, complete the decision more efficiently according to the latest information and development trends, and improve the timeliness, pertinence, and foresight of its response [8]. How to fully explore and realize the value of big

data in administrative management has gradually become a topic of common concern to scholars. It is very necessary to build a reasonable and scientific intelligent management platform based on big data.

In general, domestic and foreign countries have done a lot of research on the definition of the concept, business processes, and mechanism guarantee of big data administrative management platform [9]. However, there is still much room for improvement if we put the development of the management platform into the context of social development and government innovation. It is mainly reflected in the following aspects. First of all, the research on the development of big data administrative platform and the promotion of overall government construction is relatively weak [10]. At present, the construction of the platform is more aimed at the optimization of its own internal business processes, while the research on how to break the shackles of the traditional management system and departmentalism for the platform and realize the integration and optimization of governance resources of various departments is relatively scarce and vague. Second, as a product of the application of big data technology in the field of public management, the government management big data platform is lack of sensitivity and responsiveness to the arrival of the big data era [11]. The powerful data collection, collation, and analysis capabilities of big data will, to a considerable extent, break through the ministry pattern and information island phenomenon and directly promote the transformation of government decision-making mode and business process reengineering. Therefore, the research and application of related aspects are still in the ascendant [12].

The specific contributions of this study are as follows:

- (1) This paper analyzes the problems existing in the intellectualization of administrative management and constructs the big data platform of administrative management.
- (2) Process the data of big data platform, build the functional framework of the platform, and establish the administrative management system structure.
- (3) Applying data mining technology to administrative management, this paper constructs the administrative data warehouse system, explains the objects of data mining, and focuses on the data preprocessing module and data mining module.
- (4) Cluster analysis is applied to administrative management, and the application of clustering algorithm in text analysis is analyzed.

The rest of the paper is organized as follows. Section 2 discusses related studies, and Section 3 introduces the construction of administrative big data platform. Section 4 describes application of data mining in administration chiefly. Application of cluster analysis in administration is discussed in Section 5, and Section 6 serves as a conclusion, providing the summary and future directions of research.

2. Related Work

With the practical application of big data and the development of social diversity, the exploration and research on the intellectualization of administrative management have attracted extensive attention of scholars. Especially in developed countries, through the introduction of big data and artificial intelligence technology, the government and academia have effectively improved the efficiency of administrative management in dealing with complex social issues such as public safety, social security, medical, and health care. Some experts and scholars believe that big data will shape the political phenomenon in the new information age and transform the behavior patterns and relationships of political actors such as government, citizens, and political parties. Koren et al. [13] applied the classical *K*-means algorithm to both numeric and categorical attributes in big data platforms. They first presented an algorithm that handled the problem of mixed data. Then they used big data platforms to implement the algorithm. In terms of administration, the wide application of big data in the field of administration makes it more convenient, intelligent, and efficient. Zhao et al. [14] suggested that big data has great potential in improving policy description and strengthening policy prediction ability. Pintye et al. [15] studied the relevant policy framework of big data and proposed that the whole social governance, policies, and aspects closely related to big data should be considered as a complete system, involving user privacy, data accuracy, data collection methods, and social equity.

Zhang [16] employed the future development of intelligent administrative management, believed that the current development of modern social governance is driven by the integration of social Internet and big data, and emphasized how to establish a modern social governance working mechanism by using big data technology to effectively promote and promote administrative management. Seop and Lee [17] pointed out that at present, the application and embedding of big data in promoting administrative management has built a certain material and technical foundation for the government information system and formed an integrated development mechanism combining government public service information disclosure and policy public service supply. Zhou et al. [18] suggested that appropriate governance technology is an effective medium for the modernization of governance system and the modernization of governance ability. Big data can optimize the ecological environment of governance process and expand the elastic space of system design. It is a good opportunity to induce institutional innovation and governance transformation [19]. Big data technology has broad application prospects in the intellectualization of administrative management. Dong [20] pointed out that the sharing and interworking of data is the key in the intellectualization of administrative management and put forward the innovative development strategy of system and mechanism, so as to promote the interworking and cooperation of information and data. On the basis of in-depth study of the historical background,

characteristics, and relevant theoretical data sources of “Internet + Administration,” Wang et al. [21] focused on the deep-seated problems faced in the current construction process and put forward solutions. The application of big data is conducive to promoting the transformation of administrative management mode and improving public satisfaction, providing a theoretical basis for scientific decision-making, and promoting the improvement of government management efficiency and the transparency of supervision [22, 23]. However, big data may also cause some problems in administrative management, such as information security. Therefore, we should objectively, rationally, and comprehensively understand the two sides of big data. In the process of administrative management, we should not only carefully deal with the impact of big data and actively explore effective measures to avoid big data risks but also make full use of the opportunities brought by big data in the process of management to improve the level of administrative management [24, 25].

3. Construction of Administrative Big Data Platform

The unified big data platform can effectively solve the problems of information island and information fragmentation and promote the unified integration, sharing, and utilization of departmental data and information. The establishment of administrative big data platform can effectively ensure the orderly flow and sharing of data among departments, realize the connection between information resources and management platform, reduce human interference to data information, and solve the important problem of data separation [26]. Therefore, it is particularly important to build an administrative big data platform. The construction of administrative big data platform should obey the following principles [27, 28]: (1) The unification and standardization of data to ensure the quality of data; (2) to achieve comprehensive integration, the platform should cover all kinds of data information of each department and integrate multisource heterogeneous data across levels, regions, systems, departments, and businesses; (3) the system has objective feasibility and conforms to the current social environment, technical conditions, and economic foundation and has strong operability.

3.1. Administration Platform. The realization of intelligent administrative management depends on the big data platform with excellent design and perfect functions. The two ends of the platform are connected with data business and data application, respectively. The general big data platform is composed of data layer, application layer, and cloud platform. The cloud platform includes Software Defined Data Center (SDDC) and intelligent computer room. The data engine is mainly composed of data acquisition layer, data storage layer, data management layer, data subject layer, and data service layer. It is mainly responsible for data collection, data storage, the establishment of data warehouse and data packaging, operation, and maintenance. The

platform has analysis and display functions, such as the analysis of big data by terminal and application layer, which is realized by application layer and application layer. In addition, data engine management can also be divided into operation and maintenance management, data quality management, metadata management, and security management to complete core services such as data storage, processing, and protection. Users can obtain, store, and analyze data on the platform through client software, so as to apply it to specific businesses. The overall structure of the big data platform is shown in Figure 1.

3.2. Technology Architecture. Big data involves the process of data collection, management, storage, analysis, and visualization. According to the processing stage of big data, big data technology can be divided into data collection, and structured, semi-structured, and unstructured data can be collected through crawler or other system interfaces; data cleaning, processing repeated data and noise data, standardizing data, storing data, storing data in data warehouse, distributed database, etc. Data analysis: analyze the collected and processed data, such as basic analysis, multidimensional analysis, or use data mining technology for analysis; data application and visualization: the platform can be used for marketing, report making, operation, and index application by analyzing the data [29, 30]. Based on this, this paper proposes a process based big data technology architecture, as shown in Figure 2.

3.3. Administrative Big Data Integration Technology. Data integration is mainly the ETL work of data, which transforms heterogeneous data into isomorphic data under data warehouse through data extraction, data transformation, and data loading [31–35]. The main function of data extraction is to collect data from the data source and provide it to the subsequent data warehouse environment. After the data are successfully extracted, it can be transformed and loaded into the data warehouse. The extraction-transformation-loading (ETL) architecture is shown in Figure 3.

For the collection of government data, it is necessary to first clarify the data to be collected and then use big data-related technologies to collect and store the data scattered in departments at all levels. It includes actual data acquisition, batch data acquisition, and data filling, as described below:

- (1) Real time data acquisition: for the business generated in real time such as personnel travel and requiring high timeliness, the real-time data acquisition mechanism is selected and then the changed data resources are collected and published through Kafka component. Finally, the front-end database stores the data in the form of subscription to complete the real-time data acquisition.
- (2) Batch data acquisition: for the business data generated by the daily business of each department with low requirements for data timeliness, select the data batch acquisition mechanism, use ETL tools such as kettle to configure the data collection task regularly,

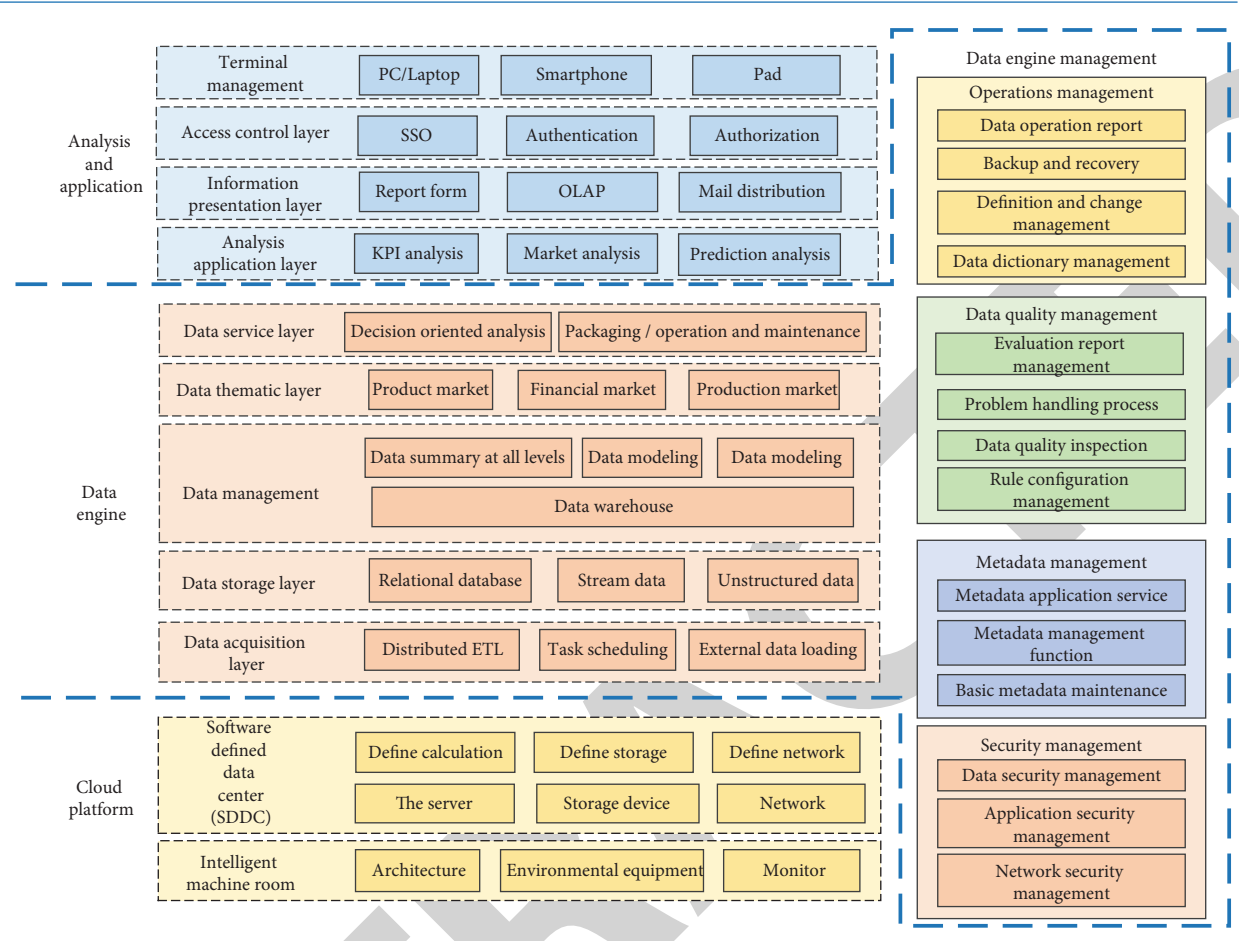


FIGURE 1: Big data platform structure.

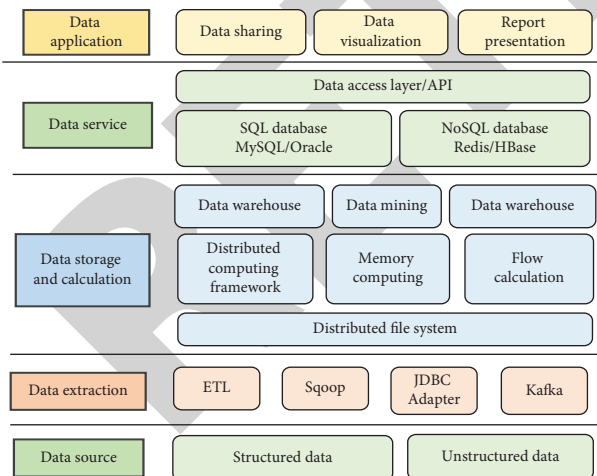


FIGURE 2: Big data technology architecture.

and periodically push the business data to the front database.

- (3) Data reporting: for departments with low development of data intelligence and no independent business system, select the data reporting mechanism to directly report the data stored in the form of

electronic documents and paper documents to the front-end database through the data reporting system.

By integrating administrative data resources, we can accurately record the data resource structure of each department and business system and better realize the integration and sharing of administrative data resources. In addition, the data resource directory is the centralized embodiment of data resource integration, and its specific structure is as follows:

- (1) Department directory. According to the department to which the data belongs, the data of departments and departments at all levels are divided. Different departments connect different front-end databases. In order to facilitate the maintenance and management of data, the department data directory is directly oriented to governments or enterprise departments at all levels.
- (2) Business system directory. According to the business system to which the data belongs, divide the data and classify the different business data in the directory, which helps to improve the business collaboration efficiency between different departments and improve the convenience of data use.

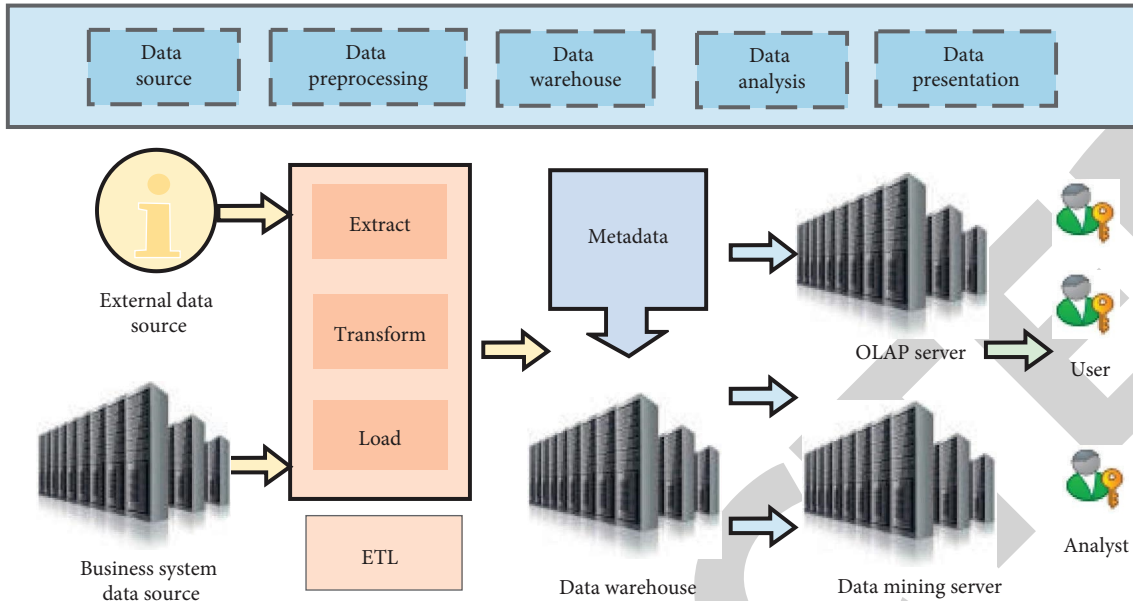


FIGURE 3: ETL architecture.

- (3) Thematic data directory. According to the thematic indicators corresponding to different data, the fused data are removed from the departments and business systems for directory division. The division of the thematic data directory is conducive to the in-depth integration of administrative data.

3.4. *Functional Framework of Administrative Big Data Platform.* The application function framework of administrative big data platform mainly includes infrastructure, data standards and specifications, directory management platform, data sharing demand management subsystem, shared exchange platform, government big data resource pool, data governance subsystem, data service platform, big data analysis and application platform, visual implementation, and application support platform. Relying on the above functional framework, the platform can provide data governance services, data sharing services, etc. As the main body of data resources of the administrative big data platform, the government big data resource pool provides unified and standardized storage support and update management for the data resources of the platform. The government information resource sharing and exchange platform implements data exchange within the response time of business needs to data definition, so as to ensure the data timeliness of the resource pool of the administrative big data platform. The data management subsystem is used to comprehensively improve the data quality and ensure the availability of data in the resource pool. The data service platform encapsulates the data service interface based on the timely government data in the government big data resource pool and provides data services for various business management systems and application support platforms. The data service itself belongs to one of the three ways of sharing and exchanging government information resources. By

using the government information resource sharing and exchange platform, the data will be decrypted and shared to the big data analysis and application platform on the premise of ensuring data security. Subsequently, these data can be used for various thematic and thematic data analysis and support various business applications, auxiliary decision-making, and macro analysis. The functional architecture of the administrative big data platform is shown in Figure 4.

3.5. *Administrative Management System Structure.* The functional modules of the administrative management architecture mainly include several modules as follows: data collection, data aggregation, data storage, data governance, data directory, data standard, data service, and data visualization.

The data acquisition system is the unified data entry of the whole big data platform, which realizes the management of heterogeneous data sources, the configuration and management of data acquisition services and so on.

The data aggregation system realizes the data collection and aggregation under the complete transaction from the source data to the target through data collection and outputs the calculation results to the target of data service.

The data directory system is the organization and description of all the data of the platform. The data entering the platform first need to be registered and catalogued in the directory system. In addition, combined with the scientific simulation model, it can realize the three-dimensional presentation of the description object and relationship model, realize the whole domain digitization, and at the same time, it can be goal oriented and provide various detection, evaluation, and reports based on the template.

With the increasing number of data resources and the change of business form, it is necessary to establish a complete data governance system from the perspective of

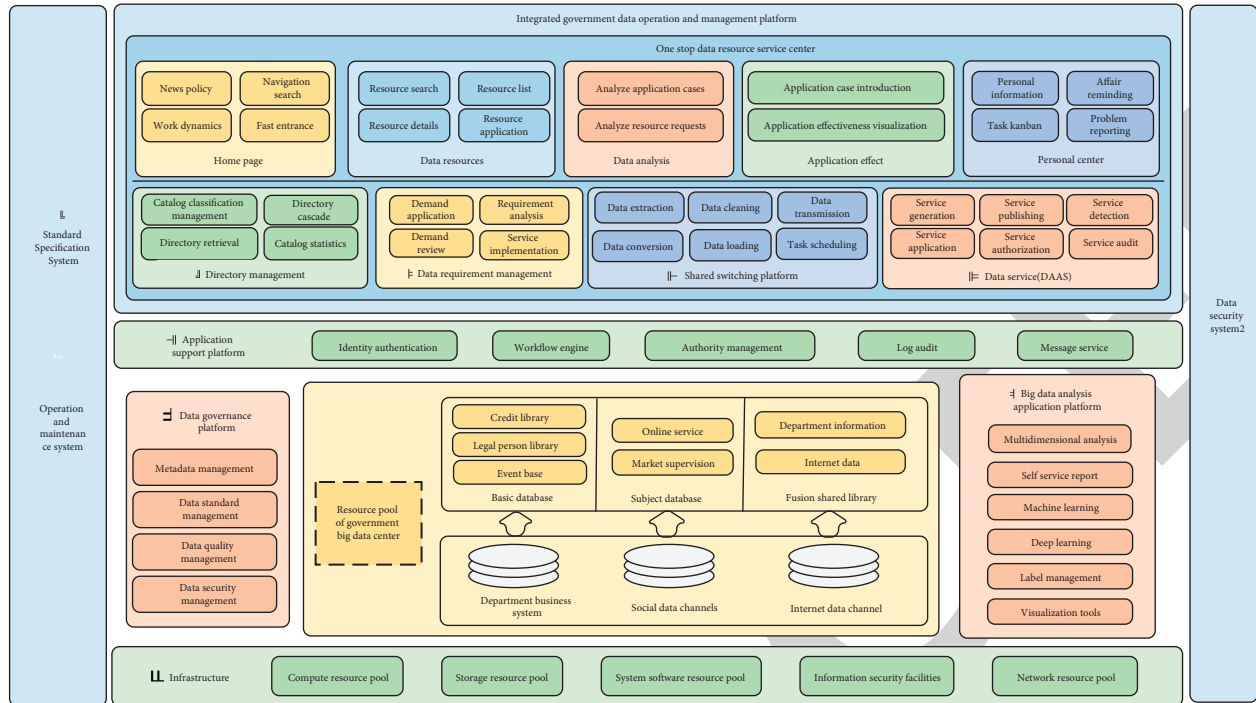


FIGURE 4: Functional framework of administrative big data platform.

data quality improvement and data governance, ensure the quality of data content, improve the platform's ability of data management and governance, and effectively mine data value. The data storage system builds an accurate, complete, consistent, and logically unified data storage system around the process of planning, construction, warehousing, storage, update, and management of information resources, focusing on the construction and use of databases, and realizing the unified storage planning management and external services of the platform.

The data service system is to provide data services for all kinds of data and applications of the data center internally and externally through different service methods, such as industry departments or the public, and open it to the society and industrial chain. Through data services, it can drive the overall development of the upstream and downstream of the data industrial chain, promote information consumption, and drive the development of big data industry.

The data visualization system realizes the visual display of data and the release of applications and other related functions. Through data modeling and visualization tools, it shows the application of each system and monitors the business collaboration, support relationship, and operation status between each system, as shown in Figure 5.

The long-term and stable operation of the data center is inseparable from efficient and scientific management mechanism and management measures. From the perspective of management, it is necessary to realize all-round integrated management control, so as to improve the management level and ability of the big data platform. At present, the business of data management has problems such as unknown resources, uncontrollable data, and lack of

supervision of data. It is necessary to establish a data control system, fully realize the control intervention and scheduling of each component system by the platform, improve the control level and data control efficiency of the administrative big data platform, and realize the long-term and stable operation of the data center.

4. Application of Data Mining in Administration

4.1. Data Mining and Knowledge Discovery. Data mining (DM) is a process of extracting hidden, unknown but potentially useful information and knowledge from a large number of incomplete, noisy, fuzzy, and random data. Knowledge discovery (KDD) is the process of transforming information into knowledge. Data mining is an important part of knowledge discovery. It can extract useful information and knowledge from a large number of incomplete and fuzzy data by using some specific algorithms. KDD flow chart is shown in Figure 6.

The process of knowledge discovery is to connect multiple stages and conduct human-computer interaction for many times. The specific process is as follows:

- (1) Identify the areas to learn: knowledge and objectives;
- (2) Establishment of target data set: focus on one data set or subsets of multiple data sets;
- (3) Data cleaning and preprocessing: use specific algorithms to remove noisy, irrelevant data and blank data space;
- (4) Data conversion: convert or merge data to form a description form suitable for data processing;



FIGURE 5: System interface of data visualization.

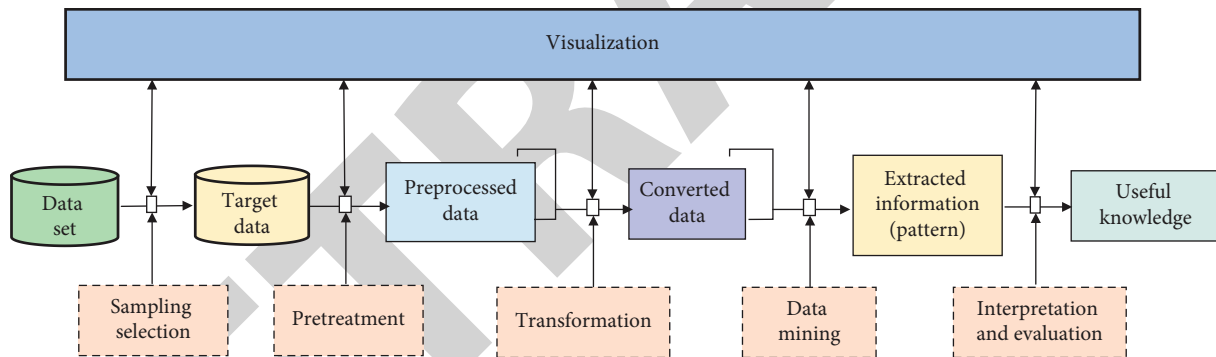


FIGURE 6: Flow chart of knowledge discovery.

- (5) Select data mining algorithms according to features such as summary, clustering, classification, and regression;
- (6) Implementing data mining: mining out potentially useful information;
- (7) Using interpretation mode to extract the information people need and then transform it into knowledge;
- (8) Conduct knowledge evaluation and study and demonstrate the extracted knowledge in practice.

- (2) Carry out administrative data mining on the extracted data. The technology of data mining breaks through the traditional data query and finds the value of data at a deeper level.

The data warehouse architecture in administrative management is shown in Figure 7.

The detailed data processed by ETL are stored in the data warehouse and classified according to the subject. Operational data store (ODS) is an optional part of data warehouse structure. ODS has some characteristics of data warehouse and OLTP (on line transaction processing) online transaction processing system. It is “subject oriented, integrated, current, or close to current and changing” data.

4.2. *Government Data Warehouse.* Data mining is completed in data warehouse. For data mining, database is a data source with rich data. Generally, intelligent decision-making is implemented under the following steps:

- (1) Multidimensional analysis of administrative data warehouse;

4.3. *Objects of Administrative Management Mining.* Data mining is widely used in administrative system. Different businesses have different data structures, such as hierarchical

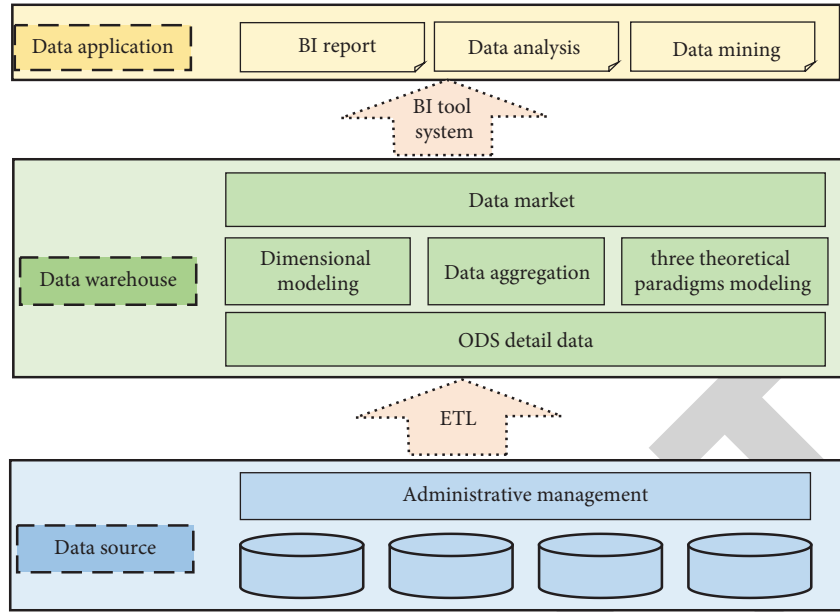


FIGURE 7: Data warehouse architecture in administrative management.

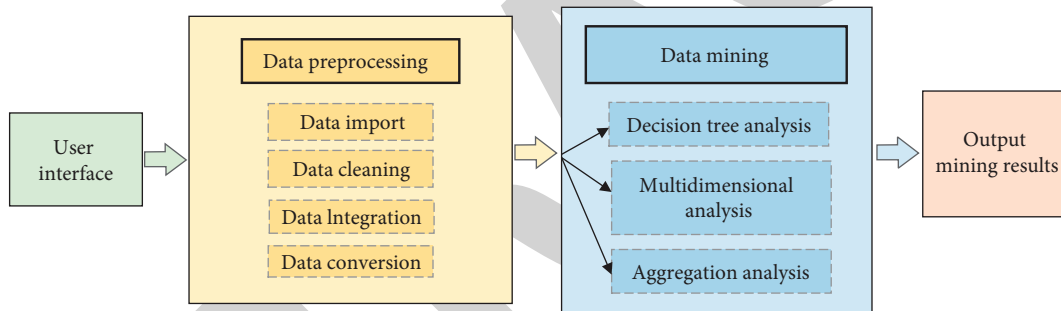


FIGURE 8: Overall framework of data mining.

data, network data, relational data, and object-oriented data. The mining object of administrative management should be determined according to the knowledge content and form required to process data in the administrative business process. Its processing objects mainly face the following aspects:

- (1) Relational database. Relational database is the most complete and abundant database system in administrative management, from which a large amount of relevant knowledge can be mined. Therefore, it is the main data form of administrative data mining.
- (2) Transaction database. Transaction databases usually consist of files, and records represent transactions. Transactions usually contain only the transaction flag and the list of items that make up the transaction. Therefore, transaction database is the largest part of data mining.
- (3) Data warehouse. Data warehouse is the main object of data mining. It is a collection of multiple data elements collected and sorted in the same topic. Thus, data mining from data warehouse can reduce a lot of data preparation time and workload.

4.4. Design of Data Mining System. The overall design of data mining is mainly divided into two parts: data preprocessing module and data mining module. First, preprocess the data input into the system, including data cleaning, data integration, data conversion, etc., and then analyze and process the data through data mining algorithms such as decision tree analysis, multidimensional analysis, and aggregation analysis. The overall framework of data mining is shown in Figure 8.

4.4.1. Data Preprocessing Module. Data preprocessing is the basis of data mining. In this module, first collect various data according to the needs and then find out the relationship between data through data mining. Data preprocessing mainly includes data cleaning, data integration, and data conversion. The processed high-quality data are a good foundation for data mining.

Data cleaning: data cleaning mainly solves incomplete, wrong, and repeated data by filling in missing values, smoothing noise data, identifying or deleting outliers, so as to make the data consistent.

Data integration: data integration is the logical or physical integration of data from different sources, formats, and characteristics. It can provide comprehensive data sharing, including periodic integration, real-time integrations, batch integration, and change data capture integration.

Data conversion: data conversion operations convert or merge data to obtain a form suitable for data processing. Data conversion includes smoothing processing, total processing, data generalization processing, formatting, and attribute construction.

4.4.2. Data Mining Module. Data mining refers to the process of revealing hidden, previously unknown and potentially valuable information from a large number of data in the database. Through data mining, governments, or enterprises can automatically analyze data and conduct inductive reasoning to mine potential patterns. Thus, it can help them adjust management strategies, reduce risks, and make correct decisions. The specific steps are as follows:

- (1) Clarify the purpose of excavation. Before starting to mine knowledge, understand the characteristics of data, use these professional background knowledge, clarify the problems to be solved, and determine the objectives to be mined, so as to select the data to be mined from the massive administrative information data.
- (2) Data preparation. Data preparation must focus on the variables defined in a certain target stage to prepare data, remove redundant, invalid and irrelevant data in the data, and provide high-quality data.
- (3) Data conversion. Convert the obtained data into a format acceptable to data mining software and a form of computer data storage.
- (4) Build the model. According to the actual situation of the data, select the appropriate mining technology and establish the corresponding model. Modeling is an iterative process. We need to carefully examine different models to determine which model is the most suitable.
- (5) Result analysis. Confirm and analyze the mining results and evaluate their value.

There is a large amount of data in the administrative database, which contains a wealth of knowledge. However, it is difficult to find the value of the data and extract useful information from a large amount of data for predicting the development trend or making decisions. Therefore, data mining technology is used to extract these relationships from a large amount of data. It is of great significance to apply it to administration.

5. Application of Cluster Analysis in Administration

The traditional classified catalogue system of administrative information resources is no longer suitable for the

management and application needs of government big data. Information mining, dynamic resource analysis, and personalized directory generation of cluster analysis can better meet this development demand.

5.1. Comparison of Clustering Analysis Algorithms. Cluster analysis is a process of dividing data objects into different classes or clusters according to their characteristic attributes according to the principle of “similarity compatibility,” so that data objects in the same cluster have greater similarity and objects in different clusters have greater dissimilarity. The process of cluster analysis is to find out some statistical values that can measure the similarity between objects or variables according to multiple observation indexes of data objects, take them as the classification basis, aggregate some objects with high similarity of features into one class, and aggregate other objects with high similarity of features into another class until all objects are aggregated to form a classification system.

Cluster analysis is a method of collecting objects through data modeling. The traditional statistical cluster analysis methods include systematic clustering or hierarchical clustering, decomposition, addition, dynamic clustering, ordered sample clustering, overlapping clustering, and fuzzy clustering. The following is the performance of different clustering algorithms on the toy data set, as shown in Figure 9. The last data set shows a special case of null, and the data are isomorphic. According to the intuition in the figure, the clustering algorithm with better effect is DBSCAN and average linkage algorithm, which can better cluster 2D data, but this intuitive method is not suitable for high-dimensional data. At the same time, the application of clustering algorithm also needs to be selected according to specific conditions.

There is plenty of unstructured text information in administrative management. Cluster analysis divides a large number of detected core or subject words into several groups by tracking the information of each information source and then counts their feature series for calculation and processing. Several main clustering methods are partition method, hierarchy method, density method, grid method, and model-based method.

5.2. The Value of Cluster Analysis in Administration. In the field of administrative management, cluster analysis can be used for big data analysis and auxiliary decision making. At present, the amount of data shows an exponential growth trend with the development of social information and business. If there is no good method to analyze the data, it will cause a phenomenon of “data explosion and information scarcity.” In view of this situation, scholars have proposed many clustering algorithms to solve the problem of big data feature collection. Cluster analysis can be used for hot spot clustering of all contents collected by the resource engine, or hot spots in a field, a vertical system, a comprehensive department, etc. In many people’s livelihood decisions, hot events and hot issues are usually direct driving factors.

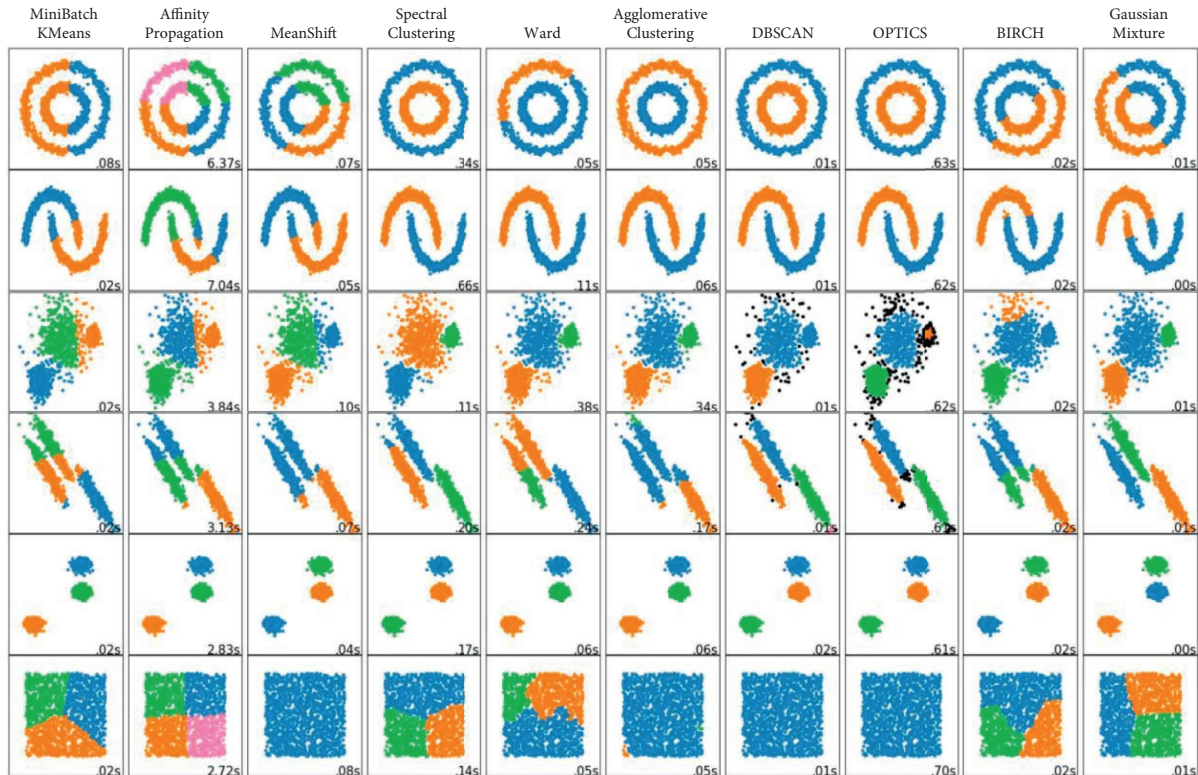


FIGURE 9: Visual display of different clustering algorithms on toy data sets.

From the perspective of time series, the initiation, initiation, development, growth, and extinction of many social events, industrial and economic phenomena have a complete life cycle, which not only conforms to the general cycle law but also has its unique characteristics. The occurrence, evolution, and extinction of some important events will be reflected in various public media and will enter the administrative big data platform through the resource engine. Through information clustering, decision makers can judge the development trend of hot information according to the trend of hot information and take appropriate measures at the appropriate time. At the same time, decision-makers often have to review and refer to the handling of similar events in history and the correctness and suitability of response measures, so as to summarize experience and lessons and make the current decision making more scientific, timely, and reasonable. For example, the course and measures taken by various countries to deal with major epidemics such as SARS, mad cow disease, highly pathogenic avian influenza, and COVID-19 can be used for reference. Therefore, when making decisions, it is often necessary to cluster the information flow of similar events in history, find out its evolution context, and evaluate various coping strategies, measures, and performance at that time, so as to reduce the current decision-making risk. Therefore, finding and analyzing the causes of hot spots and tracking their development are of great significance for scientific decision making, formulating policies, and taking measures.

In the field of administrative management, the advantage of clustering analysis is that it can simply and intuitively

obtain useful information from massive data and transform it into knowledge resources. Then, the knowledge resources of decision analysis are established to help decision makers have an overall insight into the overall situation and form an overall view and depth view. Then, it automatically identifies and tracks the beginning, development, and trend of various events, as well as the evolution process of various themes or the integration with other factors. Finally, help decision-makers establish links between seemingly isolated events and eliminate the knowledge gap. It can find and count some key signal words and subject words in time and respond to the changes of various macro resource gathering.

5.3. Application of Clustering Algorithm in Text Analysis.

There are a large number of types of rich data in administrative management, such as text data, and multimedia data, different algorithms should be selected for different types of data. Because there are a lot of unstructured text information in administrative management, clustering text information can find hot events and hot issues, judge their development trend according to the trend, and take corresponding measures. Therefore, this paper gives an example of the application of clustering algorithm in text information.

5.3.1. Text Similarity Calculation. The similarity measurement of text is generally divided into two steps: text vectorization representation and similarity calculation [36]. Text vectorization representation is to transform the text

into points in high-dimensional vector space and use high-dimensional vectors to represent the text; similarity calculation is based on the representation form of high-dimensional vectors to calculate the distance between vectors.

The text space vector model is as follows:

- (1) Document: generally refers to general text or fragments in text. Generally refers to an article, represented by D .
- (2) Item: the content features of a text are often represented by the basic language units, such as words and phrases. These basic language units are collectively referred to as the item of the text, which is represented as $D(t_1, t_2, \dots, t_n)$, where t_k is the item and $1 \leq k \leq n$.
- (3) Item weight: for text $D(t_1, t_2, \dots, t_n)$ with n items, t_k is often given a certain weight. w_k represents their importance in the text, which is recorded as $D = D(t_1 w_1, t_2 w_2, \dots, t_n w_n)$ and abbreviated as $D = D(w_1, w_2, \dots, w_n)$. At this time, the weight of item t_k is w_k , $1 \leq k \leq n$, and $D = D(w_1, w_2, \dots, w_n)$ is the vector representation of text D .
- (4) Similarity: the content correlation between two texts D_1 and D_2 is usually measured by similarity $\text{Sim}(D_1, D_2)$, and the inner product between common vectors represents the text similarity:

$$\text{Sim}(D_1, D_2) = \sum_{k=1}^n W_{1k} W_{2k}. \quad (1)$$

Also, it can be expressed by cosine value of included angle:

$$\text{Sim}(D_1, D_2) = \cos \theta = \frac{\sum_{k=1}^n W_{1k} W_{2k}}{\sqrt{\sum_{k=1}^n W_{1k}^2} \sqrt{\sum_{k=1}^n W_{2k}^2}}. \quad (2)$$

The more representative the items selected in the text are, the higher the language level is, the richer the information they contain. Because vocabulary is the most basic representation item of the text, its occurrence frequency is high in the text and presents a certain statistical law, so words or phrases are selected as feature items, and the more important items account for more weight. For specific government affairs, dealing with similar texts often requires high interpretability of clustering results.

Therefore, the method based on word bag should be selected for text vector representation, and the TF-IDF weight calculation can be defined as $W_{ik} = tf_{ik} * idf_k$. The common calculation method is given by

$$W_{ik} = tf_{ik} \cdot \log\left(\frac{N}{N_k} + 0.01\right), \quad (3)$$

where tf_{ik} represents the number of occurrences of item t_k in text D_i and idf_k is an indicator of the frequency of t_k in a document set according to document statistics. Besides, N represents the number of texts in all training sets, and N_k represents the frequency of t_k in training texts.

Considering the influence of text length on the weight, it should also be normalized to between $[0, 1]$. Thus, the TF-IDF weight can be obtained by

$$W_{ik} = \frac{tf_{ik} \cdot \log(N/N_k + 0.01)}{\sqrt{\sum_{k=1}^n (tf_{ik})^2} \cdot \log^2(N/N_k + 0.01)}. \quad (4)$$

5.3.2. K-Means Clustering Algorithm. K -means clustering is a typical clustering algorithm based on partition. It calculates the distance between the sample points and the cluster centroid and the sample points close to the cluster centroid are divided into the same cluster. K -means measures the similarity between samples by the distance between them. The farther the distance between two samples, the lower the similarity, otherwise the higher the similarity.

In the K -means algorithm, k clustering centroids $\mu_1, \mu_2, \dots, \mu_k \in R^{(n)}$ are randomly selected for each $x^{(i)} \in R^{(n)}$ of the training sample $\{x^{(1)}, \dots, x^{(m)}\}$. Then, for each sample, calculate the class it should belong to: $c^{(i)} = \arg \min_i \|x^{(i)} - \mu_j\|^2$. For each class j , recalculate the centroid $\mu_j = \sum_{i=1}^m 1\{c^{(i)} = j\} x^{(i)} / \sum_{i=1}^m 1\{c^{(i)} = j\}$ of the class and repeat the above process until convergence. When $k = 3$, the K -means clustering effect is shown in Figure 10.

5.3.3. Canopy Clustering Algorithm. However, K -means algorithm has the problem that the initial clustering center point is sensitive, so it often uses the mixed form of canopy+ K -means algorithm for model construction: first use canopy algorithm for rough clustering to get clustering center points and then use K -means algorithm to take k clustering center points as the initial center points for fine clustering [37].

Canopy algorithm has fast execution speed, does not need a given value, and has many application scenarios. It can alleviate the sensitivity of K -means algorithm to the initial clustering center point. Therefore, it is often selected as the acceleration scheme of K -means algorithm. The application flow of canopy algorithm in text clustering is shown in Figure 11. After receiving a new text, the system first converts the text into TF-IDF vector in high-dimensional space. Next, we need to search the data close to the vector from the historical data, use canopy algorithm for rough clustering, and screen out a few candidate data from the massive historical data. Then calculate the cosine similarity of TF-IDF vector with the new text, respectively. If the similarity of candidate data is greater than the threshold, the new text will be classified into the category to which the data belongs. Otherwise, a new category will be created, and the new text will be used as the initial member of the new category.

Canopy algorithm is a coarse clustering algorithm with fast execution speed but low accuracy. The execution steps of the algorithm are as follows:

- (1) Given sample list $L = x_1, x_2, \dots, x_m$ and a priori value T_1, T_2 ($T_1 > T_2$);

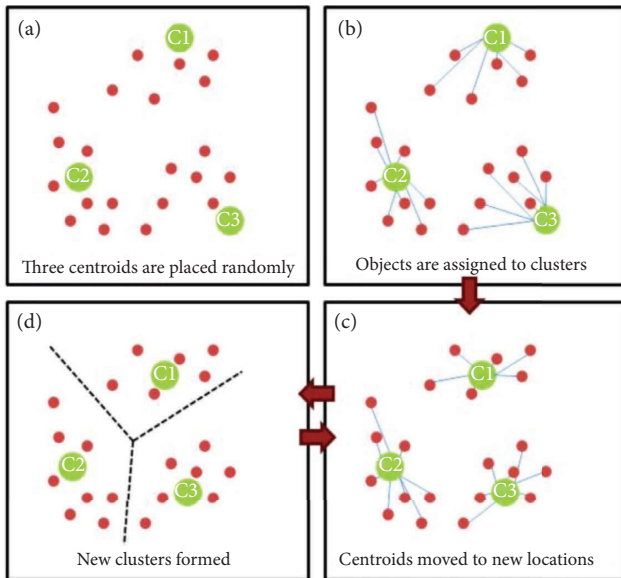


FIGURE 10: *K*-means clustering effect with $k = 3$.

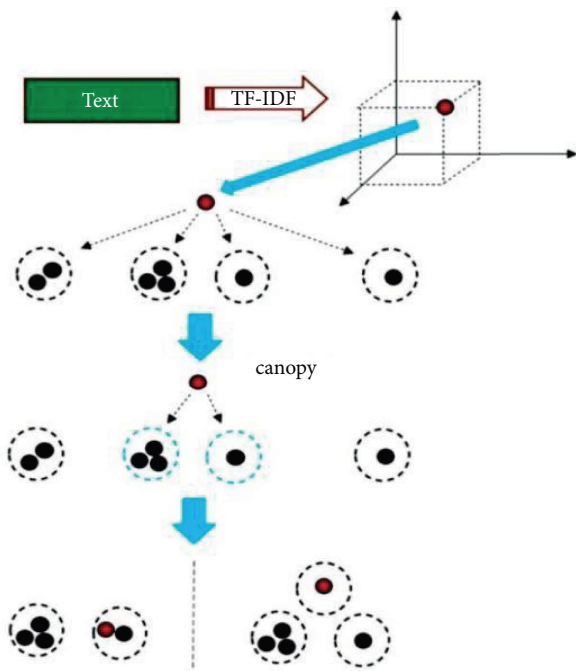


FIGURE 11: Application flow of canopy algorithm in text clustering.

- (2) Obtain a node P from the list L , calculate the distance from P to all cluster centers and select the minimum distance $D(P, a_j)$;
- (3) If the distance D is less than T_1 , it means that the node belongs to the cluster and is added to the cluster list;
- (4) If the distance D is less than T_2 , update the center point of the cluster to the center point of all samples of the cluster and delete P from the list L ;
- (5) If the distance D is greater than T_1 , node P forms a new cluster and deletes P from list L ;

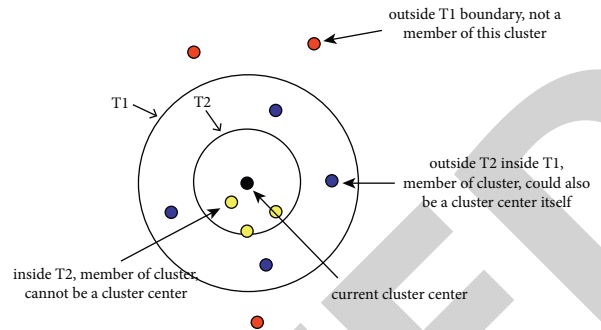


FIGURE 12: Diagram of clustering process of canopy algorithm.

- (6) Execute the loop operation until the element data in list L are no longer changed or the number of elements is 0.

The clustering process of canopy algorithm is shown in Figure 12:

Administrative management involves a large number of complex problems in different forms, such as science and technology and production. Cluster analysis has broad application space for this. But at the same time, because the problem types of government design are complex, cluster analysis should have good scalability to ensure the clustering effect; the field of administration needs to deal with various types of data and should also have the ability to deal with different types of data; minimize the domain knowledge that determines the input parameters and reduce the burden and time cost of users; and it has the ability to process noise data and reduce the sensitivity of the algorithm to noise data. At the same time, users often hope that the cluster analysis results are interpretable, understandable, and useable. Therefore, the visual display of cluster analysis results in intuitive, dynamic, multifactor, plane, and three-dimensional forms that will achieve better results.

6. Conclusion and Future Work

In this study, an intelligent management platform based on big data is designed and implemented. The platform can effectively solve the problem of information's island and fragmentation, promote data sharing, realize the connection between information resources and management platform, and reduce the interference of human factors on data information. We analyzed the application of data mining technology in administration, analyzed the objects of data mining, and designed a data mining system to analyze and process data. In the interim, it mainly includes data pre-processing module and data mining module. Effectively integrate structured, semi-structured, and unstructured data, analyze, and tap its potential value, so as to effectively improve administrative efficiency and competitiveness. Finally, this paper also introduces the application of cluster analysis in administrative management and expounds the value of cluster analysis. Because government data contain a lot of text information, this paper applies clustering algorithm to the text analysis of big data platform. In terms of demand analysis and big data platform construction, some

details need to be further studied. In the future work, we will further improve and optimize the system, carry out clustering analysis for specific businesses, and continue to explore other efficient clustering algorithms.

Data Availability

The simulation experiment data used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

All authors participated in the design, interpretation of the studies, and analysis of the data and review of the manuscript. Yunuo Su designed the study. Minhui Dai wrote the original draft. Haoyu Zhou implemented the system. Minhui Dai supervised the study.

Acknowledgments

This work was supported in part by the National Science Foundation of China (Grant no. 61972147).

References

- [1] M. Qi and J. Wang, "Using the internet of things E-government platform to optimize the administrative management mode," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 2224957, 11 pages, 2021.
- [2] Y. Xu, "Design system construction of intelligent enterprise management system based on big data technology," in *Proceedings of the 5th International Conference on Smart Grid and Electrical Automation (ICSGEA)*, pp. 363–366, Zhangjiajie, China, June 2020.
- [3] X. He and P. Liu, "Empirical study on the construction of A unified information system with big data analysis A case of social work organizations and civil affairs departments," in *Proceedings of the 20th IEEE International Conference on Computational Science and Engineering (CSE)/15th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC)*, pp. 228–233, Guangzhou, China, July 2017.
- [4] M. Li, C. Wang, L. Yan, and S. Wei, "Research on the application of medical big data," in *Proceedings of the 14th International Conference on Computer Science and Education (ICCSE)*, pp. 478–482, Toronto, ON, Canada, August 2019.
- [5] W. Chen, "Analysis and application: sports education information management under the background of big data," in *Proceedings of the 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pp. 543–546, Beihai, China, January 2021.
- [6] L. Zhao, "Business intelligence implementation in the framework of enhanced learning application," in *Proceedings of the 2nd International Conference on Advanced Electronic Materials, Computers and Materials Engineering (AEMCME)*, pp. 563–571, Changsha, China, April 2019.
- [7] C. S. Wang, S. L. Lin, T. H. Chou, and B. Y. Li, "An integrated data analytics process to optimize data governance of non-profit organization," *Computers in Human Behavior*, vol. 101, pp. 495–505, 2019.
- [8] L. H. Zhou, G. Du, R. Wang et al., "A tensor framework for geosensor data forecasting of significant societal events," *Pattern Recognition*, vol. 88, pp. 27–37, 2019.
- [9] M. K. Y. Hong, A. R. Skandarajah, and I. P. Hayes, "Administrative data: what surgeons should know about big data," *Anz Journal of Surgery*, vol. 87, no. 9, pp. 650–651, 2017.
- [10] G. H. Park and D. H. Son, "The determinants for the usage of big data in administrative service: perspectives of the quality control of data," *The Journal of Internet Electronic Commerce Research*, vol. 17, pp. 331–339, 2017.
- [11] H. Li and J. Zhai, "Constructing investment open data of Chinese listed companies based on read-write linked data," *International Journal of Electronic Government Research*, vol. 13, no. 4, pp. 15–33, 2017.
- [12] Z. Li, S. Pei, and G. Feng, "The model design of medical data life cycle based on big data platform," *Journal of Physics: Conference Series*, vol. 1865, no. 4, pp. 042088–042093, 2021.
- [13] O. Koren, C. A. Hallin, N. Perel, and D. Bendet, "Decision-making enhancement in a big data environment: application of the K-means algorithm to mixed data," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 4, pp. 293–302, 2019.
- [14] D. Zhao, B. Ning, C. Yang, and I. C. Soc, "Application research on application performance management system in big data of power grid," in *Proceedings of the 40th IEEE International Conference on Distributed Computing Systems (ICDCS)*, pp. 1358–1363, Singapore, July 2020.
- [15] I. Pintye, E. Kail, P. Kacsuk, and R. Lovas, "Big data and machine learning framework for clouds and its usage for text classification," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 19, pp. 2311–2318, 2021.
- [16] D. Zhang, "Analyze the current direction of administrative management reform from the big data perspective," in *Proceedings of the 2nd International Conference on Forthcoming Networks and Sustainability in the IoT Era (FoNeS-IoT) Electr Network*, pp. 281–287, Nicosia, Cyprus, January 2022.
- [17] P. J. Seop and S. M. Lee, "Big data analysis of busan civil affairs using the LDA topic modeling technique," *Informatization Policy*, vol. 27, no. 2, pp. 66–83, 2020.
- [18] G. Zhou, K. Chen, and J. Tu, "Research on big data open intelligent platform of guizhou province E-government service," in *Proceedings of the 3rd International Conference on Computer Science and Application Engineering (CSAE)*, pp. 333–339, Sanya, China, October 2019.
- [19] Y. Liu, "Research on the construction of school education management decision system based on Data Mining Framework," in *Proceedings of the 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, pp. 2215–2218, Harbin, China, December 2020.
- [20] Y. Dong, "On the evaluation of propagation force of new media for government affairs based on the theory of information acceptance technology," *13th International Conference on Multimedia and Ubiquitous Engineering (MUE)/14th International Conference On Future Information Technology (Future Tech)*, vol. 590, pp. 140–146, Springer, Singapore, 2020.
- [21] Y. Wang, H. Liu, and Q. Liu, "Application research of web log mining in the E-commerce," in *Proceedings of the 32nd Chinese Control And Decision Conference (CCDC)*, pp. 349–352, Hefei China, August 2020.

- [22] X. Yifei, "Research on religious management based on big data," *Cluster Computing*, vol. 22, no. S2, pp. 3531–3542, 2019.
- [23] O. Koren, M. Binyaminov, and N. Perel, "The impact of distributed data in big data platforms on organizations," *Future Technologies Conference (FTC)*, vol. 881, pp. 1024–1036, 2019.
- [24] X. Li and S. He, "Research and analysis of student portrait based on campus Big data," in *Proceedings of the IEEE 6th International Conference on Big Data Analytics (ICBDA)*, pp. 23–27, Xiamen, China, March 2021.
- [25] Y. Lin, "Innovative social Governance mode of information sharing platform in the era of Big data," in *Proceedings of the 3rd IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 1712–1715, Chongqing, China, October 2018.
- [26] Y. Fuguang, "Research on campus network cloud storage open platform based on cloud computing and big data technology," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 2, pp. 1215–1223, 2020.
- [27] S. J. Moon and N. J. Lee, "A review of big data analytics and healthcare," *Journal of the Korean Applied Science and Technology*, vol. 37, no. 1, pp. 76–82, 2020.
- [28] D. Jiang, G. Li, C. Tan, L. Huang, Y. Sun, and J. Kong, "Semantic segmentation for multiscale target based on object recognition using the improved Faster-RCNN model," *Future Generation Computer Systems*, vol. 123, pp. 94–104, 2021.
- [29] C. Cai, J. Wang, F. Zhang, X. Liu, P. Zhang, and Y. G. Zhou, "A multichannel wireless UAV charging system with compact receivers for improving transmission stability and capacity," *IEEE Systems Journal*, vol. 16, no. 1, pp. 997–1008, 2022.
- [30] R. Jiang, Y. Xin, Z. Chen, and Y. Zhang, "A medical big data access control model based on fuzzy trust prediction and regression analysis," *Applied Soft Computing*, vol. 117, pp. 108423–108530, 2022.
- [31] L. Dong, M. N. Satpute, W. Wu, and D.-Z. Du, "Two-phase multidocument summarization through content-attention-based subtopic detection," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 6, pp. 1379–1392, 2021.
- [32] Y. Sun, J. Hu, G. Li et al., "Gear reducer optimal design based on computer multimedia simulation," *The Journal of Supercomputing*, vol. 76, no. 6, pp. 4132–4148, 2020.
- [33] W. Wei, B. Zhou, D. Polap, and M. Wozniak, "A regional adaptive variational PDE model for computed tomography image reconstruction," *Pattern Recognition*, vol. 92, pp. 64–81, 2019.
- [34] Y. Yu, M. Rashidi, B. Samali, A. M. Yousefi, and W. Wang, "Multi-image-feature-based hierarchical concrete crack identification framework using optimized SVM multi-classifiers and D-S fusion algorithm for bridge structures," *Remote Sensing*, vol. 13, no. 2, p. 240, 2021.
- [35] W. Shu, K. Cai, and N. N. Xiong, "Research on strong agile response task scheduling optimization enhancement with optimal resource usage in green cloud computing," *Future Generation Computer Systems*, vol. 124, pp. 12–20, 2021.
- [36] W. Cao, Z. Zhang, C. Liu et al., "Unsupervised discriminative feature learning via finding a clustering-friendly embedding space," *Pattern Recognition*, vol. 129, Article ID 108768, 2022.
- [37] M. F. Leung, J. Wang, Z. B. Deng, and Q. W. Jin, "Cardinality-constrained portfolio selection based on collaborative neurodynamic optimization," *Neural Networks*, vol. 145, pp. 68–79, 2022.