Design and Construction of Enterprise Financial Data Sharing Service Center Based on Big Data Technology

Yuxin Wang

wangyx070322@163.com

Dongbei University of Finance and Economics, Dalian City, Liaoning Province, China

Abstract: With the rapid rise of digital economy, strengthening the construction and optimization of enterprise financial data sharing service center with the new generation of digital information technology is an important issue in the new period of high-quality development of enterprises. Therefore, this paper takes the financial data sharing service center as the research object, gives full play to the practical characteristics of big data technology, network information technology and computer application technology, and builds an enterprise financial data sharing service platform with Hadoop cluster as the data analysis and processing server and Javaweb as the interactive application system. The whole platform adopts B/S architecture design, and completes the development and deployment of various functional modules and API interfaces according to MVC pattern. On the basis of traditional accounting work, the platform will be further integrated into high-value financial work such as financial management analysis and cost review, which will make the platform's functional system more perfect and powerful, especially in the aspect of in-depth mining of financial data value. The system will give full play to its functional advantages and support K-means, CART decision tree and other data models to complete the identification and investigation of enterprise financial risks. The construction of the platform is not only conducive to promoting the realization of enterprise financial integration, but also makes outstanding contributions to improving the effectiveness of enterprise internal management.

Keywords: Big Data, Financial Data Sharing Service Center, Data Mining, Javaweb, Computer Application.

1 INTRODUCTION

Financial Shared Service Center (FSSC), as a brand-new enterprise financial operation and management mode, can highly meet the management requirements of multi-format or multi-branch group enterprises. The essence of the shared financial service center is financial centralization, that is, the accounting work of each branch is integrated, and the centralized operation is carried out in the service center, and all kinds of accounts and reports are automatically processed according to unified standards and norms, thus strengthening the financial control of the group headquarters over the branches. At the same time, it can also improve the efficiency of financial management, greatly reduce the financial system and labor costs, strengthen the sharing of data and information resources within enterprises, and promote the information transformation and upgrading of enterprise internal management.

BDEDM 2023, January 06-08, Changsha, People's Republic of China

DOI 10.4108/eai.6-1-2023.2330332

However, the rapid popularization of digital technology and the rapid rise of digital economy have brought unprecedented influence to the economy and society. The new generation of digital technology represented by big data technology is permeating all industries and fields, making data a brand-new factor of production, and constantly promoting the emergence of new industries, new formats and new models, providing an important driving force for the transformation and upgrading of enterprises in the new era. [1] As a natural data center within an enterprise, the financial department can directly grasp all kinds of data and information resources, which is the key to the innovation and integration of big data technology. Similarly, the financial sharing service center should combine the characteristics and development trend of big data technology, promote the financial sharing service center to gradually evolve into a financial data sharing service center, and realize the update and improvement of the functional system of the financial sharing service center. It provides a set of feasible solutions for the actual effect, security and value problems in the process of data storage, call and mining. In view of this, this paper holds that in the new era, the enterprise financial sharing service center should adhere to innovation drive, give full play to the application advantages of big data technology, network information technology and computer application technology, and build an enterprise financial data sharing service platform with Hadoop cluster as data analysis and processing server and Javaweb as interactive application system. The whole platform is presented as a Web application, which can widely support enterprise managers, financial personnel and other business personnel to log in and use, remould various work processes, and overcome the shortcomings of inconvenient user operation and cumbersome business initiation process. It organically combines the financial system with the business system, strengthens the management of data interfaces among modules, promotes the circulation and interaction of various data, and realizes the sharing and utilization of resources, thus promoting the value creation of the integration of industry and finance under the financial sharing mechanism of enterprises, and making a beneficial attempt to improve the effectiveness of internal management of enterprises. [7]

2 INTRODUCTION OF KEY TECHNOLOGIES

2.1 Big Data Technology

With the rapid development and wide application of network information technology, Big data gradually emerges and rises, which is a new stage of information technology extension and development. Big data is a concept of continuous and dynamic change, which generally refers to a huge amount of data set that cannot be acquired, managed and processed by traditional information technology and software and hardware tools in a tolerable time. [10] The characteristics of big data, such as huge scale, various data types, fast processing speed, low value density and complex content, determine the application direction and development trend of big data, and also put forward new requirements for data processing technology.

Big data technology, namely big data processing application technology, is a collection of technologies covering the whole life cycle of big data, and it is also a new mode different from traditional data processing. Big data technology includes collection, cleaning, storage, analysis and mining, visual display and application, which provides a complete processing paradigm for

the acquisition and reflection of the value of big data, and constantly expands the application scenarios of big data.

2.2 Hadoop Technical Framework

Hadoop is an open source technical framework implemented by Java language. Hadoop can run on large-scale clusters, can be used as the core of the underlying storage and analysis, and provide users with reliable, scalable and distributed computing big data services. [5] The technical core of Hadoop is distributed storage and distributed computing, which are implemented by HDFS and MapReduce respectively.

2.2.1 Kafka

Kafka is in the transport layer of big data technology stack heap, which is essentially a distributed information queue. Its core application is to realize the asynchronous transmission and transmission of data and reduce the coupling between data sources and data storage units. [3] Kafka has excellent data throughput, and can accept many clients or data sources at the same time. It is widely used in large-scale real-time data calculation and log collection and analysis application scenarios. Kafka architecture is mainly composed of producers, consumers and brokers.

2.2.2 Sqoop

Sqoop is a data migration component under Hadoop framework, and its target object is the conventional relational database, which aims at assisting the rapid and direct data transmission and transfer between RDBMS and Hadoop. Sqoop can support the bidirectional flow of structured data, that is, data can be imported into HDFS under Hadoop from Mysql, Oracle, Postgresql and other databases, or data can be extracted from Hadoop framework and exported to the corresponding database. The underlying program of Sqoop is extracted, transformed and loaded by MapReduce, which can ensure that Sqoop has high concurrency efficiency and high fault tolerance in the running process, and greatly improve the performance of Hadoop in calling data resources. [9]

2.3 Javaweb

JavaWeb is the sum total of technologies that use Java technology to solve related Web and Internet fields, and it is the core technology that uses Java language to realize dynamic Web application development. The key work of JavaWeb application development lies in the design and development of background Web Server, that is, the construction and deployment of each functional module, business logic control and database application model. The development of JavaWeb has gone through many stages in different periods, from the earliest Servlet technology to the subsequent JSP technology, and then to the extensive application of a large number of development frameworks. The development of Web applications is further hierarchical, structured and modular, so as to make the development of Web applications more efficient and concise. [6]

2.4 Data Mining

Data mining (DM), as a kind of computer science and technology, aims at the complex process of extracting and mining the hidden and valuable patterns or laws from a large number of

incomplete and noisy actual data. The birth of data mining technology meets people's demand for the analysis and processing of massive information data in the current information explosion era. It is a technology that integrates and comprehensively applies database technology, information science, statistics, machine learning, data timeliness and other technologies. [4] Compared with traditional data analysis, data mining technology can handle more complex data objects, including structured data, semi-structured data and heterogeneous data. The main process is shown in Figure 1, in which database, data warehouse, Internet and other data sources represent the collection of raw data, while the data preprocessing stage includes three basic steps: data cleaning, data integration, data selection and transformation.

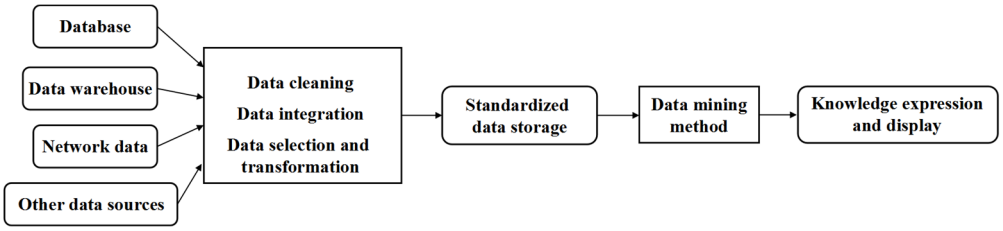

Figure 1: Data mining process

The construction of data mining model is the core of the whole data mining work. The construction of data mining model corresponds to data analysis method, which not only determines the application direction of data mining results, but also determines the construction of data mining model. Common data analysis methods include classification analysis, prediction analysis, cluster analysis, valuation analysis and correlation analysis, as shown in Table 1. [8]

Table 1: Type of common data mining algorithms

Classification	Data analysis method	Data mining model
Guided data mining	Classification analysis	Decision tree, random forest, neural network
	Prediction analysis	Regression tree and rough set method
	Valuation analysis	SVM, bayesian method
No guided data mining	Correlation analysis	Pearson correlation coefficient and aprior algorithm
	Cluster analysis	K-means clustering algorithm, lineage clustering

2.5 Development Process

According to the application requirements of the above related application technologies, complete the configuration and deployment of the development environment of the enterprise financial data sharing service platform. The development content of the system is divided into two parts. One is to build Hadoop cluster to complete the collection and storage management of financial data. Secondly, under the Java development environment, the development of Web Server is completed with Spring framework, and the business system, financial system and data analysis system of the financial sharing service center are built to form a standard Web application.

First of all, Hadoop cluster architecture needs the support of hardware and software. The underlying operating system is Linux, CentOS 6.7(x86_64) is the version, and jdk-8u291-linux-x64 is the JDK version. According to the application requirements of the system, Hadoop cluster will be set up into seven nodes, named Master1, Master2, Slave1, Slave2, Slave3, Slave4 and Slave5 respectively. Master is the master node and Slave is the slave node. Hadoop version is 2.7.7, which is installed in each node, and components such as Yarn, HDFS, Zookeeper, HBase and Kafka are also deployed in each node. In addition, the system will use Sqoop component to import financial data, business data and other data into HDFS system under Hadoop framework to realize distributed storage and form original data. With the help of Java, the sample information table is constructed, and the data preprocessing operation is completed and imported into MySQL to obtain sample data.

Secondly, for the development of each functional module of the shared service platform, the workflow engine will be introduced to complete the process service control and performance expansion of each function. The system selects Activiti open source workflow engine, and gives the core class ProcessEngine process engine and some Service classes in Activiti to Spring container for management, so as to realize the call and processing of workflow by Web Server. In addition, for the data mining analysis function, the system will build different data algorithm models and package them into classes that can be called directly, as shown in Figure 2, which is the key code to realize the K-means clustering algorithm model.

```
public class Kmean
{     public List<Point> points;
      FileWriter out = null;
      DecimalFormat dFormat = new DecimalFormat("00.00");
      public KMeansCluster kMeansCluster;
      public int numCluster = 5;
      public int numIterator = 200;
      public int numPoints = 50;
      public static final String FILEPATH="f:/kmeans/res.txt";
      public static void main(String[] args)
    {     Kmean kmeans = new Kmean(100, 5, 200);
          kmeans.init();
          kmeans.runKmeans();
          kmeans.printRes();
          kmeans.saveResToFile(FILEPATH);      }
```

Figure 2: Model building code of the K-means clustering algorithm

Finally, for the development of Web application server, the basic development environment is Java, JDK version requires 1.6 or above, Java development environment is MyEclipse 2014, Web server is Tomcat 8.0, and database server is MySQL 5.5. And the project object model (Maven) is used to manage the project structure. Maven chooses Apache-Maven-3.2.1 version. In the process of building the overall development environment, the installation of JDK and the configuration of environment variables are completed first to build the foundation of Java application development. Secondly, the installation of MyEclipse and the installation of Tomcat, the Web server, and the configuration of Tomcat is completed in the Preference option under MyEclipse. Then, based on Spring architecture, the integration and encapsulation of the whole system is completed. Through the introduction of the above key technical theories, we have determined the overall environment of system development, the configuration of related

software and tools, and the technical feasibility of the overall project of enterprise financial data sharing service platform.

3 FUNCTIONAL IMPLEMENTATION

3.1 Business System

Under this function module, the platform will plan and integrate all businesses such as procurement, asset management, customer relationship management, project management, logistics and warehousing, sales and operation according to the standardized process formulated by the enterprise, and design corresponding forms according to different business properties, each of which contains several key fields. When the user initiates the business process, the user can input, upload or scan the corresponding content according to the form. The system automatically starts the workflow engine according to the standardized business process to accept the business data and forward it to the approval verification node. After the approval verification is correct, the business data and the approval result will be returned to the business system for display. At the same time, the approval and verification node converts the business form into financial information voucher, and transmits it to the financial system. After review, the receipt and payment of funds are completed.

3.2 Financial System

Under this functional module, the platform will be highly compatible with the settings of various functional modules of the enterprise financial management software, mainly providing functional interfaces such as voucher management, account table inquiry, summary accounting and report filling. After financial personnel log in to the system, they can intuitively see the initiation and execution of various business processes, and the corresponding vouchers and forms will be naturalized into different subsystems according to their categories, which will become the data source of the financial data sharing service platform, thus realizing data sharing and providing the foundation for subsequent data value mining and application.

3.3 Data Analysis System

Under this function module, system users can synthesize financial data, calculate financial indicators and evaluate risk control for related projects, data information and departments. For example, after selecting different departments, projects, time and other basic information in the system in turn, users can calculate the analysis indicators from four aspects: business ability, debt paying ability, profitability and equity analysis according to the analysis and calculation rules. Table 2 shows some index information, in which KMO is the measured value, which provides the necessary weight proportion for the subsequent calculation of scores. [2] After calculation, the enterprise's abilities are graded accordingly, and this is the key feature, and the K-means clustering algorithm model is selected to complete the data classification, thus generating the classification standard, which is compared according to the standard. If the distribution of a certain ability is quite different from that of previous years, the existence of financial risks can be determined. As shown in Table 3, the analysis table of enterprise financial risk clustering results shows that the greater the positive index value, the greater the financial risk, and the greater the reverse index value, the smaller the financial risk.

Table 2: Calculation formulas of some indexes

Category	Specific items	Analysis index	Computational formula	KMO factor
Operation capacity	Assets	Proportion of current assets	Current assets / Total assets	0.188
	Cash flow	Cash turnover rate	Main business income / Cash balance	0.291
Debt paying ability	Creditor's rights	Quick ratio	(current assets-liabilities)/Current liabilities	0.271
Profitability	Profit contributio n	Net interest rate on sales	Net profit / Total profit	0.213
Equity analysis	Owner's equity	Total asset growth rate	Total assets year-on-year	0.403

Table 3: Analysis Table of Enterprise Financial Risk Clustering Results

Category	Analysis index	Cluster 1	Cluster 2	Cluster 3	Evaluate
Operating ability	Current assets ratio	3.38	2.11	1.56	Reverse index
	Cash turnover ratio	11.84	4.66	7.85	Reverse index
Solvency	Quick ratio	0.15	0.33	0.17	Reverse index
Profitabilit y	Net profit rate of sales	0.21	0.32	0.22	Reverse index
Equity analysis	Total asset growth rate	0.66	0.61	0.26	Positive index

4 CONCLUSIONS

With the purpose of promoting the transformation of the operation mode of enterprise shared financial service center, this paper constructs a Web-based enterprise financial data sharing service platform based on the functional characteristics of big data technology, network information technology and computer application technology. The function of the platform combines the business system and financial system of an enterprise, completes the sharing and interaction of all kinds of data, and provides a brand-new data analysis system. It is convenient for users to improve the processing efficiency of various businesses with concise, convenient and efficient operation, increase the value embodiment and rational application of financial data, strengthen the internal financial management ability of group enterprises, and promote the digitalization and intelligent construction of internal management and control of enterprises.

REFERENCES

[1] Chen Xiaohong, Li Yangyang, et al. (2022) Theoretical System and Research Prospect of Digital Economy[J]. Management World.02.

[2] Li Chengfeng, Huang Kangmei, et al. (2021) Application of Data Mining Technology in Financial Risk Early Warning[J].National Circulation Economy.01.

[3] Liu Bang, Yu HuapingLiu Bang, Yu Huaping. (2019) Research on High Performance of Kafka Distributed Message Queue[J]. Computer Knowledge and Technology.11.

[4] Meng Weicheng. (2022) Development and Application Analysis of Computer Data Mining Technology[J]. Electronics World.01.

[5] Shi Fangxia, Gao Yi. (2021) Hadoop Big Data Technology Application Analysis[J]. Modern Electronics Technique.09.

[6] Shi Feng. (2021) JavaWeb Development and Application Based on MVC Pattern[J]. Electronic Technique.05.

[7] Tan Chunhua. (2021) Construction of Shared Financial Center under the Background of Digital Transformation[J].Tax Paying.12.

[8] Wang Lili. (2021) Application of Data Mining Technology in the Background of Big Data[J]. Computer & Network.10.

[9] Yang Bin. (2017) Application of Sqoop Data Collection and Warehousing System[J].Practical Electronics.11.

[10] Zhao Peng, Zhu Yilan. (2022) Overview and Development Prospect of Big DataTechnology[J]. Astronautical Systems Engineering Technology.01.