Design and Implementation of Enterprise Management Intelligent Decision System Based on Data Analysis

Wenting Xue[1]

147398403@qq.com

[1]Wuhai Vocational and Technical College, Wuhai, Inner Mongolia, China

Abstract: In recent years, the severe situation at home and abroad has led to the increasingly complex trend of China's market development and the increasingly fierce competition among enterprises. If an enterprise wants to stand in such a market environment stably for a long time, it must formulate relevant coping strategies according to the changing laws and demands of the market. The design and development of enterprise management intelligent decision-making system based on data analysis in this paper is to help managers in different departments of enterprises collect relevant data and analyze the changing trend and development law of the most useful information mining market, and make scientific decisions according to the data, so as to ensure the accuracy of decision-making. The Hadoop cluster is used to collect and analyze the most favorable information, extract effective data, and then combine the time series algorithm in data mining technology to predict the risks and costs of enterprises according to the collected data. Finally, Java technology is used to visualize the data for users to view. The development of this system helps enterprise decision makers to analyze the required data and improve their management and decision-making ability.

Keywords: intelligent decision; data analysis; Hadoop; Java

1 Introduction

At present, the severe international and domestic situation has not only had a significant impact on the development of China's market and economy, but also brought enormous pressure and challenges to various enterprises. In the fierce market game, the enterprises in the same industry who can know the development trend and change trends of the market and industry in time and capture the needs of customers can stand out from the competition and quickly stand on the market. If an enterprise wants to survive and develop, it is far from enough to rely solely on employees' efforts, but also to enhance its core competitiveness. The core competitiveness can be divided into many parts according to its form, of which the decision-making competitiveness is the key [1]. This competitiveness comes from the ability of enterprise managers to identify market opportunities, find out market rules and respond to market changes. An enterprise that is not competitive in decision-making is like a walking corpse, with a superficial appearance but no soul. At present, the decision-making ability of some enterprise leaders is poor, because on the one hand, they can't collect valuable information, and on the other hand, they lack the ability

ICIDC 2022, June 17-19, Qingdao, People's Republic of China

DOI 10.4108/eai.17-6-2022.2322710

to diagnose problems and analyze causes, which makes them unable to make correct decisions. In the process of enterprise development, it is not only necessary to collect the data of internal management and external market as a whole, but also to fully guarantee the comprehensiveness and accuracy of the data, so as to provide strong data support for later data processing and enterprise management decisions. Therefore, the role of data analysis is particularly important. Data analysis needs to go through many links, such as data collection, cleaning, analysis and extraction [7]. The role of data analysis can be roughly divided into the following two aspects: First, it can intuitively reflect the internal objective situation of the company and the external market trend, which is more complete and systematic than the form of data reports, and can be quickly understood, absorbed and utilized by people. Secondly, data analysis provides high-precision, comprehensive and powerful data support for company managers to make various planning decisions. Data analysis depends on big data technology and data mining technology. After collecting data, big data tools integrate and analyze the data, remove the false and seek truth from the collected information, and extract the most valuable and accurate information, so that the value of data can be brought into full play. Data mining technology preprocesses the collected data sources through big data tools, and then calls the pre-written algorithm to perform related mining tasks. When the data mining is finished, the corresponding results will be generated. The flexible application of big data and data mining technology to enterprise decision-making not only broadens the channels for enterprises to obtain data sources, but also helps enterprises improve their decision-making ability by integrating and classifying the collected data to complete data analysis [5]. The author thinks that the intelligent decision-making system of enterprise management based on data analysis uses Hadoop big data tool to capture the information needed by the enterprise, collate, calculate and analyze the data, and combine the trend extrapolation method in the time series algorithm model of data mining technology to predict the direction of new products, funds, market environment, etc. to be decided according to the analyzed data. Finally, these data and charts are transmitted to Echarts through Java language to present a visual interface for the reference of managers of various departments of the enterprise. On the premise of saving cost and improving efficiency, the purpose of this platform development is to help managers in different departments of enterprises to find and diagnose the problems existing in the company, explore the development and change rules of the market and the personalized needs of users, etc., and on this basis, formulate accurate and scientific enterprise strategies to increase the core competitiveness of enterprises in the same industry market.

2 Technical overview

2.1 Hadoop

Nowadays, in the open network environment, every platform will produce a large amount of network data every day. Faced with such a huge amount of data, one computer and one application program can't handle it, so Apache Foundation has developed Hadoop big data processing cluster. Hadoop cluster can work on several computer hosts at the same time, and distribute the data collection, storage, calculation and analysis to different tool components to complete. At the same time, the high-speed operation greatly improves the processing efficiency of massive data. This paper mainly uses the following Hadoop tools: flume, kafka, scrpy, HDFS, hive, Mapreduce, sqoop. Among them, flume, kafka and scrpy are used together to capture data

such as web logs and pictures. HDFS, as the underlying data support of Hadoop cluster, is used to store the collected massive data. Mapreduce is one of the core components of Hadoop, which is divided into two processes: Map and Reduce. It can be used to realize distributed computing of data. Hive can be regarded as the transmission medium between HDFS and Mapreduce. Through hive, the data in HDFS is taken out and the corresponding Mapreduce program is started for distributed calculation. Sqoop is the "middleman" between HDFS storage and MySQL database, and it is responsible for the mutual transfer of data between the two data storage systems [8].

2.2 Time series algorithm

Time series algorithm is a group of random numerical variables generated by sorting according to equally spaced time periods. This algorithm observes a potential process according to a given sampling rate and obtains the results. In real life, time series algorithm is used to predict data in many fields. Time series algorithm includes a variety of algorithm models. In this paper, the trend extrapolation method is commonly used. Its principle is to explore the law of things' development and change according to the past state and present situation, so as to predict the future development trend of things. The trend extrapolation method is used to predict, and the prediction with this algorithm has to go through five stages: the first stage, select the parameters to be predicted. The second stage is to collect the necessary data. In the third stage, the curve is fitted with data. The fourth stage is to extrapolate the trend, make a forecast explanation and give the forecast result. In the fifth stage, the manager studies the feasibility of application in decision-making according to the forecast explanation and forecast result. The trend extrapolation method contains many commonly used function models, as shown in Figure 1, which is the code of JAVA language combined with Logistic function to realize iterative calculation of optimal parameters. The jsp writing controller is responsible for file upload and task submission, and accesses Mapreduce Mapper API provided to users to request Mapreduce calculation program to call this code for data mining prediction. The forecasting work of data mining in this paper is mainly to extrapolate the forecasting trend based on time series algorithm and ARMA model. Figure 2 shows the forecasting workflow after the time series algorithm enters the Mapreduce calculation program. The original data (the sales volume of products in the previous quarter or the first half of the year) is imported, and the stationarity of the time series is detected and identified according to the preset function model, parameter values and seasonal variation rules. After the detection, the corresponding ARMA model is identified and established according to the rules of identifying time series models. After the modeling is completed, the parameters are tested to estimate whether it has statistical significance, and then the qualified serial model is applied to forecast. Finally, according to the results of the serial model forecast, the corresponding sales trend chart of the products in the next quarter is returned.

```
public static double sumMatrixRowMutiWeight(DenseMatrix64F src,int inx,double [] weights) {
        double rs =0;
        for(int i=0;i<src.numCols;i++) {
            rs += src.get(inx, i)*weights[i];
        }
        return rs;
    }
    public static double[] adjustWeights(DenseMatrix64F src,int inx,double [] weights,double
alpha,double error) {
            double [] rs = new double[weights.length];
        for(int i=0;i<weights.length;i++) {
            rs[i] = alpha*error*src.get(inx, i)+weights[i];
        }
        return rs;
    }
    public static double[] stocGradAscentl(DenseMatrix64F datas,double[] classLabels,int
numIter) {
        Random r=new Random();
        double[] rs = new double[datas.numCols];
        for(int i=0;i<rs.length;i++)
            rs[i]=1;
        for(int i=0;i<numIter;i++) {
            for(int j=0;j<datas.numRows;j++) {
                double alpha=4/(1.0+j+i)+0.01;
                int randIndex = r.nextInt(datas.numRows-j);
                double h=sigmoid(sumMatrixRowMutiWeight(datas,randIndex,rs));
                double error=classLabels[randIndex]-h;
                rs = adjustWeights(datas,randIndex,rs,alpha,error);
```

Figure 1: Java code

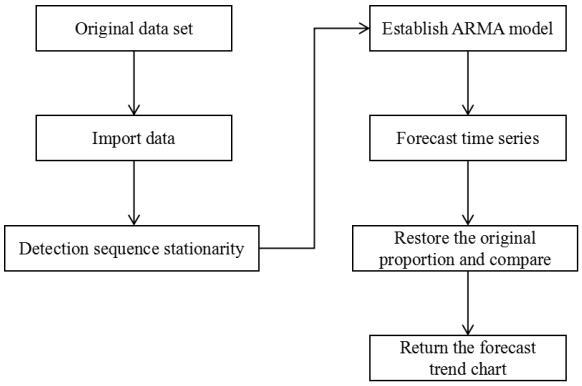

Figure 2: The algorithm implementation process

2.3 Java

Java is an object-oriented computer programming language developed based on C++ and its advantages include easy learning, distributed, robust, safe and portable, multi-threading, etc. The most important feature of Java language is object-oriented. The so-called object-oriented (OO) is a message-driven programming method that regards objects as the center. It supports three ways: encapsulation, polymorphism and inheritance. Java also has these three characteristics. Java programming environment is generally divided into two parts: JRE and JDK. The former is the running environment and the latter is the development environment. In JDK, you can write Java programs that you want to realize, while the written Java programs run in JRE. When writing and running a Java application, it is necessary to consider the Java programming language, the format of its class files, the virtual machine environment and Java API. Java is widely used in many fields, such as server program, website system, embedded field, scientific application, big data technology, etc. Hadoop used in this paper is developed with Java language.

2.4 SSM framework

SSM framework is a J2EE enterprise-level framework that integrates Spring, Spring MVC and MyBatis, and is often used in the development and construction of large-scale enterprise-level application systems. SSM framework integrates the core ideas of the three frameworks, provides developers with an easy-to-learn and simplified development process [3], and saves Java code writing time. SSM framework design is based on MVC architecture, which divides the whole program development into four layers, namely: view layer, data persistence layer, business logic layer and control layer (presentation layer). Among them, the view layer is used to show the pages written by JSP. The persistence layer encapsulates some operations used to interact with the database. First, define its interface, and then design the class to be implemented by this interface in the configuration file of Spring. The business layer is responsible for the design and implementation of business logic, designing interfaces and concrete implemented classes, and calling the interfaces defined in advance by the persistence layer. The presentation layer calls the interface designed by the business layer to control the business process, as shown in Figure 3, which is the interaction process of the four-tier architecture.

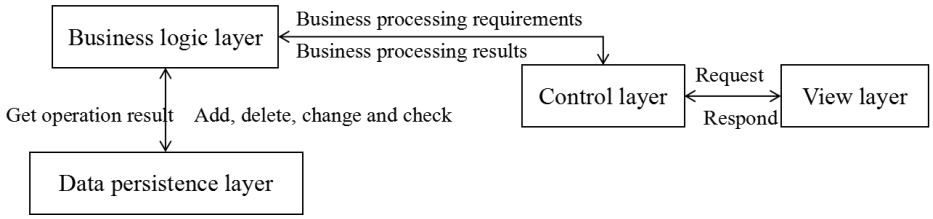

Figure 3: Four-tier architecture

2.5 Development environment

The technology used in this system and the development environment of cluster tools should be configured and installed in Linux operating system.

Hadoop cluster installation and deployment under Linux operating system need to prepare several Linux virtual machines, and the running memory of each machine should be allocated as much as possible to ensure the normal operation of Hadoop cluster. As Hadoop cluster depends on Java language, each virtual machine needs to configure Java environment first, and then install and deploy Hadoop. Enter Java official website, select the version of jre-8u331-linux-x64.tar.gz for Linux, download and unzip it, configure the environment variables for it, restart the system and check the installation, then modify the host name and IP mapping, and set SSH as secret-free login. Go to official website to download Hadoop installation package, select Hadoop-version 2.8.0, unzip and configure Hadoop environment variables after installation: go to CD/usr/local/Hadoop--2.8.0-CDH5.7.0/etc/Hadoop folder, and change ${JAVA_HOME} to $ {Java _ home} in the line of export Java _ home. Load the installation packages of required components into Hadoop and configure relevant environment variables: configure hdfs after formatting namenode, execute the command "hdfs namenode -format" on spark1, and enter the command "start-dfs.sh" to start HDFS. Save hive's installation package to the /usr/local directory of spark1 and unzip it. Rename hive directory: mv apache-hive-0.13.1-bin hive. Copy the kafka installation package to the /usr/local directory of spark1, unzip the kafka installation package, and then change the directory name: mv kafka_2.9.2-0.8.1 kafka is the kafka configuration ID. Upload the installation package of flume to the /usr/local folder of sparkproject1, unzip and rename the directory, and configure the environment variables for it. Install eclipse, add Map/Reduce functional area, configure maven dependency, modify Hadoop home/etc/Hadoop/HDFS-site.xml file, add port 50070 and update startup firewall. Import the installation package of sqoop into the /opt folder and unzip it. Change the name of the folder to sqoop and configure the environment variables for it. Enter the command "pip3 install scrapy" to install the scrap framework and establish the soft connection of scrap. Finally, restart Hadoop. As shown in Figure 4, copy the deployed hadoo files to other sub-machines to realize synchronous operation.

```
scp -r /bigData/hadoop- 2. 8.0 root@s lave1:/bi gData/
scp -r /bigData/hadoop- 2. 8.0 root@s lave2:/bigData/
```

Figure 4: Copy command

3 Requirement analysis

3.1 System requirements analysis

In the complex and changeable market environment, the fierce competition among the same industries has brought great pressure to the business development of enterprises. If enterprises want to occupy a favorable position in the market competition, they must have the ability to manage decisions scientifically and reasonably, and they need to collect relevant data before making decisions. In the traditional management decision-making mode of enterprises, enterprise managers will collect relevant information and data for reference. The statistics of these data mostly come from the internal financial situation and business management status of enterprises. These data have the disadvantages of large error and single data source channel [9]. Therefore, how to extract valuable and accurate data information from massive data, analyze and predict it to provide guidance for enterprise management decision-making has become the

key demand of this kind of system development. The system developed in this paper can collect a wide range of accurate data sources through the enterprise decision-making system that integrates big data tools and data mining technologies. The integration of these technologies changes the relationship between data from causality to correlation. This system will help the managers of different departments of the enterprise to collect information and data at all levels, such as the internal business situation and external market environment, and diagnose various problems within the company. Then, the system will sort out and analyze these data to build relevant data models, and further explore the correlation between data based on the data models and management decision models of different departments of the enterprise. Using the Logistic curve (growth curve) function in trend extrapolation algorithm to predict the future probability or change trend of an event according to its history and current process will play a positive role in promoting the decision-making ability and level of managers in various departments of enterprises.

3.2 Global design

The design of enterprise management intelligent decision-making system is to use Hadoop big data cluster to complete data analysis, among which flume, kafka, scrpy and sqoop are several data grabbing tools to get the information needed by the enterprise, put the collected data into HDFS distributed memory for storage, then take out the data in HDFS by hive tool and start the corresponding Mapreduce program for distributed calculation and analysis, and extract the useful data for the enterprise. The JAVA server calls jsp controller to initiate a request to Mapreduce calculation program through the request statement, requesting it to call the trend extrapolation algorithm [2] written in advance in Java language, and predict the direction of the project, capital, market environment, etc. to be decided according to the predicted value set by the algorithm. The calculated data is transferred to MySQL database through sqoop tool, and the data is called by sending a request to the database through Java language, and a visual interface is presented for the reference of managers of various departments in enterprises. Figure 5 shows the process of data visualization of enterprise management intelligent decision system: Java client initiates a request for information, calls data to MySQL database, MySQL responds and transmits the result back to the business layer, encapsulates the data of Service collection and transmits it to JSP, and uses Jquery as a medium to transmit the data to Echarts to present a visual image [6].

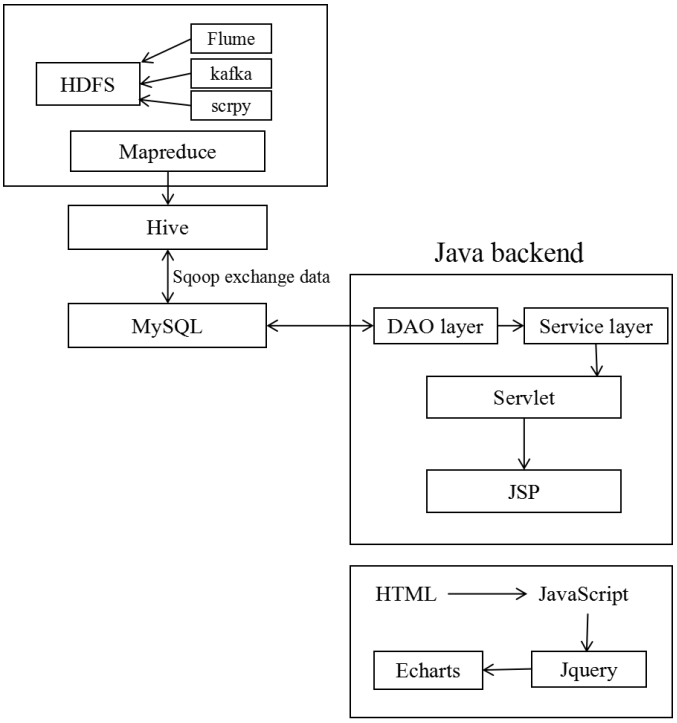

Figure 5: Interaction flow chart

4 Function realization

According to the requirements of enterprise managers in decision-making, the system design is divided into three modules as shown in the figure: problem diagnosis, collection and analysis, and prediction, as shown in Figure 6.

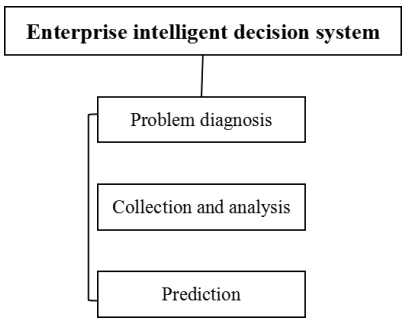

Figure 6: Function module

Problem diagnosis: The focus of problem diagnosis is the main contradictions faced by enterprises at various stages of development and the fundamental obstacles that restrict the development of enterprises. Different managers click on the diagnostic function interface, such

as business strategy management, product research and development management, production management, marketing management, financial management, personnel management and other departments, to view different internal information of the enterprise. The personnel department of the company will check the attendance of employees and the positions of various departments to analyze the current human resources problems in the company. Department managers can check the financial statements, the operation of funds, etc. to analyze the problems of the company's revenue and expenditure and fund scheduling [4]. The sales management department analyzes and diagnoses the problems existing in the product sales of enterprises according to the sales volume and total sales amount of products. Diagnose departments according to different problems to determine specific decision-making objectives.

Collection and analysis: With the specific decision-making direction, different managers can click on this functional system to control Hadoop cluster tools to collect the data and information needed for enterprise decision-making on various websites and platforms as comprehensively as possible. For example, the manager of the sales department can collect information about competitors in the same industry, purchasing power information of customers, etc. Managers of R&D departments can obtain information such as the characteristics of similar products and the demand of current consumers for products. Department managers can collect information such as post salary and welfare benefits of different enterprises. The information collected by the system will be cleaned and analyzed, and the most direct and effective data will be extracted to provide different aspects and sufficient information guarantee for decision makers in different departments to make corresponding decisions.

Prediction: One of the indispensable links in the decision-making of an enterprise is to make a basic prediction of the development trend of the enterprise and the change of the market based on complete and reliable information data. Managers of different departments can use the forecasting function of the platform to forecast the internal operation of enterprises and the changing trends of external markets in different directions. For example, the sales management department can view the sales rate forecast and product profit rate forecast of related products. Financial management can view information such as production cost forecast, investment forecast and market risk forecast. Managers of the strategy department can check the changes of supply demand, relevant policies and price mechanism. According to different forecast information, each department can formulate corresponding measures in advance to avoid all kinds of unnecessary risks.

5 Conclusions

The survival and development of enterprises can not be separated from the scientific decision-making of managers. The design and implementation of enterprise intelligent management decision-making system has effectively changed and solved the limitations and problems existing in traditional enterprise decision-making. By collecting and integrating the internal and external data of the enterprise, and analyzing and forecasting the integrated effective data [10], the system provides reliable data support for enterprise managers to make scientific decisions. The development of this platform has played an auxiliary role for enterprises to realize standardized management, real-time data, comprehensive analysis, accurate prediction and scientific decision-making, improved the efficiency, decision-making ability and management

level of enterprise managers, and provided a powerful impetus for the sustainable development of enterprises.

References

[1] Deng Yingying. (2022) Research on the Influence of Big Data on Enterprise Management Decision-making. Management Vertical and Horizontal.03.

[2] Jiao Xiangyu, Huang Kanghui, Lu Zheng. (2020) Hadoop+JavaWeb Big Data Analysis Visualization System. Management & Technology of SME.02.

[3] Liu Hai, Wang Zhuangzhuang, Qiao Zhaoyuan et al. (2020) Design and Development of Campus Help Platform Based on SSM Framework. Digital Users.10.

[4] Lou Dehan, Yang Jianghai, Deng Haisheng. (2021) Research on Enterprise Financial Management System Based on Big Data Hadoop. Software Development.07.

[5] Shi Fangxia, Gao Yi. (2021) Hadoop Big Data Technology Application Analysis.Modern Electronic Technology.05.

[6] Wang Dan. (2015) Research on Time Series Data Processing Method Based on Hadoop Platform. University of Science and Technology of China.05.

[7] Wang Weifeng, Yang Lin. (2017) Design and Implementation of Big Data Analysis System for Postal Delivery Based on Hadoop. Journal of University of Chinese Academy of Sciences.05.

[8] Yang Haimin, Zhang Zhisong, Bai Wei. (2019) Overview of Time Series Prediction Methods. Computer Science.01.

[9] Zhang Guodong. (2014) Design and Implementation of Telecom Big Data Analysis Platform Based on Hadoop Technology. Shanghai Jiao Tong University.05.

[10] Zheng Lanlan. (2016) Design and Implementation of Construction Enterprise Data Analysis and Decision Support Platform. Beijing University Of Technology.12.