Apply Data Mining Technique to E-Business Platform

Ge Guo¹ and Zhenzheng Chen²,*

¹ School of Management, Wuhan Donghu University, Wuhan, China
² School of Management, Wuhan Donghu University, Wuhan, China

*Corresponding author e-mail: 270336487@qq.com

Abstract. Network technology continues to advance, cloud computing technology develops the internet provides us with a variety of data services, and the development of online shopping has gradually matured, E-commerce platform has become the most mainstream shopping place now, rich network data provides a lot of useful information for e-commerce, but the current data mining technology of e-commerce platforms is not mature enough, unable to meet the growing demand for online shopping.

This article is to solve the problem of how to better apply cloud computing technology and data mining technology to e-commerce platforms, analyzes the current status of data processing in e-commerce platforms, combining the research results of domestic and foreign scholars, combining data mining technology with specific applications in e-commerce platforms, after research analysis, finally summarized the advantages and disadvantages of applying data mining technology to e-commerce platforms, and how to solve the disadvantages, guidance on the application of data mining in e-commerce.

Keywords: Data mining technique, E-business platform, Cloud computing

1. Introduction
In recent years, internet construction is improving day by day, especially with cloud computing, Cloud storage the development of space technology, the data that can be collected on the network covers almost the entire world, It's an era of big data."¹ In business, the rapid development of online shopping, has gradually become the mainstream domestic shopping method. E-commerce platform uses information network technology as a means, business activities centered on commodity exchange, on various e-commerce platforms, with the development of network information, users can post information on the Internet, the network also records information about their activities on the Internet. For this data, there are many data analysis and processing methods in the Internet age, data processing can extract and derive data that is valuable and meaningful to certain people from a large number of potentially disorganized and incomprehensible data, solve these people's problems.² Currently, we can find the use of data mining technology in many fields such as finance, healthcare, marketing, retail,
manufacturing, justice, engineering and science.

Views of domestic scholars:(1) In 2018, a scholar of Nanjing University of Aeronautics and Astronautics named JiangLi, research on applying data mining technology to advertising recommendation system, she points out that text mining is extremely important for analysis, computational analysis of the collected merchant information by using the improved K-Means clustering algorithm, finally, give overall design and module design of advertising recommendation system in e-commerce platform, provided guidance for applying data mining to e-commerce platforms. (2) In 2019, a scholar of ChangJiang university named ZhangLei point out that e-commerce has produced many different models, such as b2b,c2c,b2c etc. However, due to technical conditions and data constraints, e-commerce platforms cannot make full use of the data and information collected by themselves. The advent of big data and data mining technologies has solved this problem, it can more accurately mine the business opportunities hidden behind the data. Views of foreign scholars: (1) Abroad, Ashraf Darwish research on the combination of cloud computing and IoT, they apply these two technologies to the medical field, he mentioned the idea of Cloud to healthy, talking about combining cloud computing, cloud technology and health care to improve health protection, it also tells the future prospects and directions of this concept, this provides us with ideas for combining cloud technology, cloud computing and other fields such as e-commerce platforms, is the future development direction of e-commerce platforms. (2) Foreign scholar Vishruti Kaknad and Hitarth Shah, research on the application of game theory in network security and cloud computing, by establishing a model, the validity of the combination of the two is explained more accurately and clearly, it is one of the important research directions for the development of cloud computing and data mining technology in the future.

There also have some problems in it. First, unable to mine deep, hidden data information provided by platform historical data, second, based on data processing, when e-commerce platforms make product recommendations for different customers, the problem of recommending products does not meet customer needs or customers who are not interested in recommending products. [3]

2. Solution— Applying data mining technology to e-commerce platforms

2.1 Introduction
Data mining is the process of searching for information hidden in a large amount of data through algorithms, usually associated with computer science, and achieve the goal through many methods such as statistics, information retrieval, online processing analysis, machine intelligent learning, expert systems, and pattern recognition.

Data mining technology uses advanced computer technology to apply data mining to actual. Considering the data itself, it specifically includes 8 steps of information collection, data integration, data specification, data cleaning, data transformation, data mining process, pattern evaluation, and knowledge representation, and it's a continuous cycle, every link matters, if you make any mistake you have to remake it. [4]

In addition, with the development of cloud computing technology, we could use cloud computing to store a large amount of data, it have the functions that allow us to ultrahigh speed(10 trillion operations per second) calculation, and also have strong compatibility which means it can be used on most hardware, and can effectively predict market trends, make our data mining deeper, more accurate data calculation, and data analysis is more targeted.

2.2 Introduction to common data mining software
In reality, the more commonly used data mining software are SAS/EM and IBM's SPSS.

2.2.1 SAS/EM. SAS EM (SAS/enterprise miner) it processes data mining requests from the SAS / enterprise miner client and the results of the processing analysis are transmitted back to the client through SAS / connect,SAS EM use the Graphical interface, Menu driven and Drag-and-drop
operating, very friendly and convenient for software users, and it can handle huge and complex data, has a very powerful information mining function.

The basic operation steps of the software are as follows:

1) Use "Data Acquisition Tool" in the system, set the data variables to be used in data mining, variable divided into interval variable and class variable, the software itself also helps users review the rationality of numerical settings, ensure initial quality; (2) Use “Data sampling tools” to sample data, software has a variety of sampling methods, for example, isometric sampling, random sampling, classified sampling and stratified sampling, let users understand the purpose of data mining; (3) Use “Data screening tools” to eliminate unwanted data from the sample, make sample data more suitable for data mining goals; (4) Use “Data Variable Conversion Tool” to transform some data, make it fit better with future models; (5) After the collection and screening of these data is completed, enter them into the database which is made for data mining; (6) After all of these, we can formally entered into the "data mining process" now, this process use the extensive mathematical method in the SAS/EM, and the rich data mining and flexible algorithms it provides, build the most suitable model for mining, analyzing, and matching target data and related information in the database, during operation you also need to use the SAS / EM Ensemble node in the software sometimes to integrate multiple models, make the generated model more stable and efficient.

![Figure 1. SAS/EM Operation flowchart](image)

2.2.2 Cloud computing . Cloud computing it was developed on the basis of the increasing establishment, use and delivery of Internet-related services, provides dynamically scalable, often virtualized resources via the Internet. Cloud means it has virtually no boundaries in virtual storage space, can store large amounts of data, cloud computing is based on such a huge database, can reach 10 trillion accurate calculations per second, can be used for prediction market development trends, nuclear explosion, weather etc. It also has strong compatibility features, can switch the resources needed by the enterprise to the required applications.

This shows, have complete, diverse and detailed process data mining software, after having the
blessing of cloud computing technology, can mine more hidden data information, can also improve the speed and accuracy of software data calculations, it also makes the analysis of data more accurate.[5]

2.3 The specific application of data mining in e-commerce platform[6]
In order to make data mining technology fully applicable to e-commerce platforms, We should make the application of data mining technology in e-commerce more detailed and targeted, the following researches two aspects of data mining applied to e-commerce sales and recommendation systems.

2.3.1. Apply to Sales[7]. When data mining technology is applied to e-commerce, the most important thing is to collect and analyze data on merchants, consumers, goods, transaction activities, etc. The platform analyzes these sales data, can better analyze the shopping habits of each customer or group when making decisions, increase customer satisfaction with platform recommendations, increased satisfaction will bring more customer resources to the platform, in order to achieve the purpose of increasing profit income and market influence.[8]

(1) Statistical sales. They can record sales of their stores, store it in cloud space, can prevent loss, use of statistical tools and cloud computing technology in data mining systems, make accurate and fast analysis of changes in sales volume, so you can see the trend of sales volume over time, submit processed information to decision makers, allow him to design a time-targeted sales plan based on this data.

(2) Targeted analysis of goods. Using data mining technology, count the total sales of each product, monthly and daily sales, compare individual products within the system, calculate the sales of goods in different aspects, data analysis for decision makers to make more reasonable production plans, reduce unnecessary costs, at the same time, you can also investigate the degree of support and credit of the product, use data mining to analyze the support and confidence of each product, study the links between different products, generate a "1 + 1> 2" product combination form and let the decision maker formulate a marketing strategy for combination marketing.

(3) Analyze consumer purchase concentration. This indicator can clearly show the customer's contribution sales and their loyalty to the store, high concentration of purchases, explain that the customer's contribution to the store is also high in sales and loyalty, this data enables decision makers to take well-maintained countermeasures against highly concentrated customers, become a Loyal Consumer.

2.3.2 apply to product recommendation system[8]. For today's e-commerce platforms, if the customer did not search for a keyword in the search, then this merchant will not appear, in this way, merchants also lost some passers-by customers, E-commerce platforms have lost part of their trading activity, so, using the recommendation system on the e-commerce platform can bring more transaction value to the e-commerce platform, increase profit income.[9]

From the user experience perspective right now, too large and numerous products make it difficult for customers to choose, the products recommended by the platform are still inappropriate, failure to meet customer needs, then the application of data mining technology is particularly important, mining hidden information, make classification statistics more accurate, leverage a large number of efficient mathematical methods, calculate the best recommendation, solve the problem of unreasonable recommendation on the platform.

College students are now the major players in online shopping, below are the usual shopping statistics for 100 male college students and female college student, and also collected the consumption situation of the e-commerce platform for college students for one year, calculate the average and draw a line chart, as follows:
3. Impact of the application of data mining technology on consumers and platforms

3.1 Impact on consumers

(1) Benefit: Online shopping is currently the way most people shop, so the development of e-commerce platforms, consumers are big beneficiaries, the gradual increase in product categories allows consumers to solve most of their shopping needs on e-commerce platforms. Make life more convenient and faster, product recommendation using data mining technology is more accurate, consumer search on the platform is more convenient, you can also find more products that suit your interests, the platform can also recommend combined products, make shopping more cost-effective, unnecessary spending left for consumers.

(2) Disadvantage: Consumers need to provide a lot of private information to the platform, including personal preferences, phone numbers, ID numbers and bank card numbers, etc. Data mining can use
this information to provide consumers with thoughtful services, on the other hand, it may also leak information, or perform a vicious information analysis, make consumers vulnerable to privacy violations and stolen deposits.

3.2 Impact on e-commerce platforms

(1) Benefit: In e-commerce platforms, data mining makes the information link between buyer and seller closer, after an in-depth analysis of the information on both sides, provide a trading platform that meets the requirements of both parties, let consumers and businesses feel the convenience of platform operation, to attract more users, also keeps regular users, finally bring huge profits to e-commerce companies. It also promotes the development of related technologies in the era of big data.

(2) Disadvantage: If these efficient data analysis techniques cannot be properly controlled, will cause platform information confusion, criminals steal information, causes social panic and even worse adverse results.

4. Prevention of the shortcomings after the application of data mining technology

The possible adverse effects of this technology have been discussed above, it can be roughly summarized that cloud storage and data mining systems store a lot of private information, if there is a problem with the privacy of this information, will provide criminals with information that causes social problems.

So, the protection of information in the cloud space is a very important part of the development of the future big data era, it provides the most critical basic protection for a large number of data analysis applications based on cloud space.

In addition, any technology will have more or less problems in the process of using it, then when we apply data mining technology and cloud computing technology to e-commerce platforms, need to have a complete system maintenance program, and regular inspection and repair of equipment, to prevent the bad phenomenon of platform paralysis, even if there is a major accident that the platform is paralyzed, there must also be an emergency plan for this, technical staff in this area need to be on call anytime.[10]

So, only to improve the reliability of the platform to ensure the normal use rights of users.

Acknowledgement

This work was supported by the grants from Hubei Provincial Collaborative Innovation Centre of Agricultural E-Commerce (Wuhan Donghu university research[2019] No.17 Document).

References

[1] Nanjun Ye, Analysis of the application of data mining technology in e-commerce, information recording materials.6(2019)137.

[2] Jing Yu, Research on Data Mining and Application in Big Data Era, Science and Technology & Innovation.24(2018)160.

[3] Fen Liu, Research on the Status and Development of Data Mining in China, Shandong Industrial Technology.17(2018)123.

[4] Haiyan Tang, Application of data mining technology in software engineering, Electronic Technology & Software Engineering.(2019) 141.

[5] Tuqian Zhang, Key database technology in cloud computing technology, Electronic Technology & Software Engineering. 22(2019)142-143.

[6] Ming Dong, Wang Kun, Application and Research of Data Mining Technology in E-commerce, E-commerce.23(2018)34.

[7] Feng Wang, Research on Application of Data Mining Technology in Supermarket Merchandise Marketing, Digital Communication World.11(2019)198.

[8] Yumei Zhang, Design of personalized recommendation system based on web data mining, Information Technology and Informatization.6(2019)12-15.
[9] Shanshan Wang, Research on Application of Intelligent Recommendation System in Personalized Data Mining, The Journal of Shandong Agriculture and Engineering University. 36(2019)28-29.

[10] Liu Kang, Risk assessment of e-commerce transactions based on data mining, Modern Electronics Technique 5(2017)75.