



Business Information Mining Technology of Social Media Platform Based on PageRank Algorithm

Ying Liu^(✉)

School of Business, Jinggangshan University, Ji'an 343009, China
liuyingzhende@126.com

Abstract. The existing methods of business information mining are flexible and cannot effectively mine the business information of media platform. In order to summarize and manage the mass business information effectively, a research method of business information mining based on PageRank algorithm for social media platform is proposed. Different from the existing methods, it innovatively optimizes the business information evaluation algorithm of social media platforms, increases the flexibility of information mining, and realizes the business information mining of social media platforms. The experiment proves that the technology of business information mining based on social media platforms can effectively summarize and manage a large amount of business information.

Keywords: PageRank algorithm · Social media · Business information · Data mining

1 Introduction

The commercial information platform is a special type of management information platform, which provides support for market-oriented operation activities in enterprise activities. Since market competition requires constant innovation in a mature economy and society, and innovation requires scientific decision-making methods and an adequate information base, the value of such platforms as ancillary tools continues to increase, and the modalities for providing assistance continue to evolve and evolve, integrating more and more advanced theories and technologies with the times [1]. Typical areas of business activities in which ancillary support is provided include:

Market operations, specifically including macro-market analysis, analysis and evaluation of product sales strategies, evaluation of business value of research and development plans for new products, analysis and evaluation of market trends of products, evaluation and optimization of product portfolios and pricing schemes, etc.

Customer Management, such as Customer Relationship Management, New Customer Marketing Strategy Analysis and Evaluation, Customer Channel Performance Analysis and Contribution Analysis [2].

Product R&D management, such as product line portfolio evaluation and optimization, product line profitability calculation and analysis, product modification program income evaluation, new product risk and value comprehensive calculation and analysis.

Field of financial investment management of enterprises, such as long-term investment planning and optimization of production resources, risk analysis of investment projects of enterprises, analysis of optimal allocation of short-term production resources, value assessment and operation analysis of intangible assets, analysis of enterprise operation cost structure and elements, etc. [3].

In the modern business information service environment, information is becoming more and more intensive and information-based. If enterprises want to develop steadily in the increasingly fierce market competition, they must extract the information which is beneficial to the business operation after analyzing a large amount of business data deeply, so as to improve their own decision-making ability.

Madhumathi S and the others introduced that the work flow model can be implemented with the data mining in the E-commerce platforms. It helps the product/project manager in several ways. The multiple queries have figured out and those are solved here. The data and reviews are generated automatically. The text are generated with web crawler and stored in database as a raw data. The data are cleaned with Natural Language Processing methods and algorithms [4]. Fauzan A and the others use the association rules with rapid miner software, data mining approach, and predictive analysis that contains various data exploration scenarios. The study provides important evidence for adopting data mining methods in the industrial sector and their advantages and disadvantages. The business information mining of social media platform is realized based on chevron Pacific Indonesia platform [5]. The above methods do not evaluate the business information according to the web page ranking, and the flexibility of the business information mining is poor, and the effective media platform business information mining results cannot be obtained.

Data mining, as a tool that can analyze and check a large amount of business data, can help the business enterprise to extract valuable information from the updated and accumulated data. It can collect and analyze a large amount of data in the business environment based on the established business objectives of the business enterprise, and then extract the key data and information from it, so as to provide effective information services for the scientific business decision-making activities. Different from existing methods, this paper innovatively introduces data mining technology to conduct business information mining on social media platforms. Design the business information management system of social media platform; PageRank-based data mining algorithm evaluates business information according to web page ranking to realize business information mining of media platform.

2 Social Media Platform Business Information Mining

2.1 Business Information Management System for Social Media Platforms

Data mining based on PageRank, as a professional data processing technology, has been widely used in commercial information services, among which Data mining is also widely used in the information services of commercial media platforms, mainly in

the risk assessment of media platforms, product supervision and the establishment of commercial competitive intelligence platforms. Some experts have a popular explanation for data mining, that is, to explore knowledge in databases. But there are also some experts who regard it as only one of the most basic and vital links in the process of knowledge discovery. The process of knowledge discovery is shown in Fig. 1.

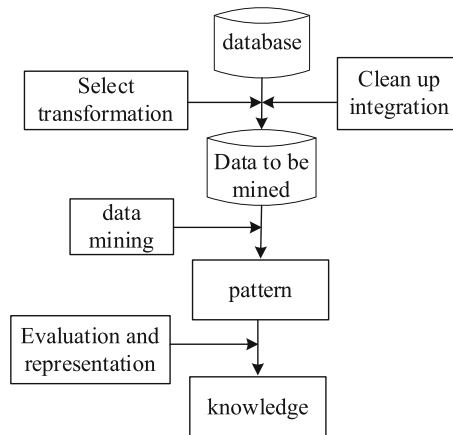


Fig. 1. Steps to turn mining data into feature information

As shown in Fig. 1, the result of data mining is the ability to interact with users using a knowledge base. Through the intuitive model to the user, or stored in the knowledge base as a new knowledge for the user to save. But in today's big data era, "data mining" than "knowledge discovery in the database" as a new concept is more popular. Therefore, this paper chooses terminology for data mining. In addition, the data mining process of the customer marketing response rate prediction model in this paper is basically in accordance with the above process and method steps to study the operation [6]. If we can realize the marketization and normalization of the "mortgage of control" mechanism and reduce the difficulty of its application, we can greatly reduce the cost of the debtor's recourse to funds and at the same time provide a mechanism for high-quality enterprises to send their own quality signals. The direction of business information business data management is shown in Fig. 2.

E-commerce can form cross-validation with other information of enterprises according to the historical data mastered, which is a powerful supplement to the scientific nature of product information management. When you do not have a social media business of your own, mine the available energy data into energy information to provide information services to businesses or social media organizations [7]. For example, when a company wishes to obtain credit from a social media platform, it can provide a report on whether the company's historical product data and historical information match for a fee. Similarly, social media institutions can be provided with services to monitor and analyse the electricity consumption of their leaders [8]. In this model, one is only responsible for the data he provides, and the information he analyzes, and is paid accordingly, without

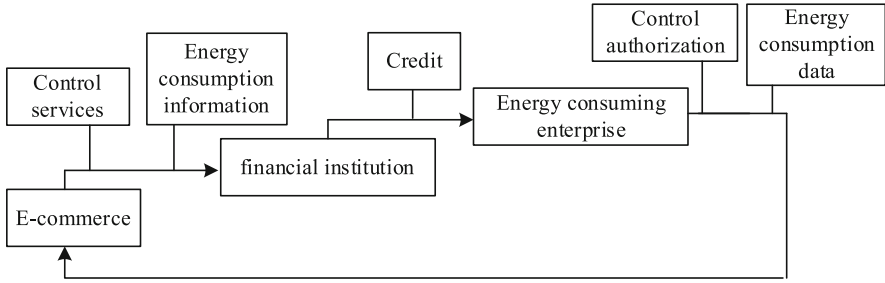


Fig. 2. Business information business data management direction

the social media risk itself. The overall architecture of the business model is shown in Fig. 3.

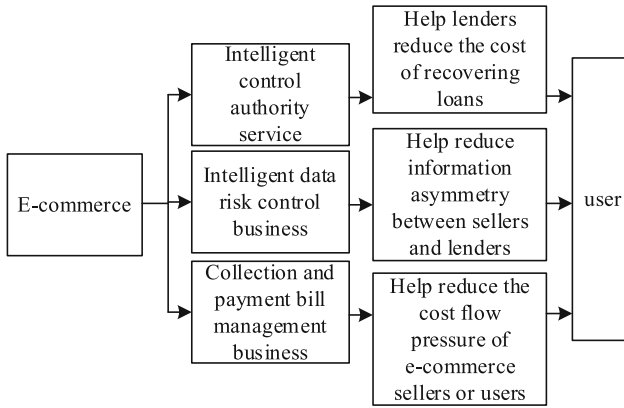


Fig. 3. Overall architecture of the business model

In this architecture, the main business relationships with users are divided into three parts: energy flow, information flow and cost flow. The information flow refers to the relationship between the user and the user due to the information generated by the transaction between the user and the user, and the energy consumption behavior of the user. Energy-use data risk control business shall be carried out around information flow, and help narrow the information asymmetry between debtors and credits. Receipt and payment billing management business shall be carried out around expense flow, and the optimized allocation of cash resources may be realized.

2.2 PageRank Based Business Information Evaluation Algorithm

Data mining is a process of extracting the unknown data and information which may exist in the data by some methods in the irregular, large quantity, unclear and random data. Data mining also has many other terms, different fields are different, for example, data

analysis, decision knowledge, knowledge discovery and so on. Knowledge discovery in the general artificial intelligence of the term, in fact, is equivalent to the database of data mining meaning. Data mining is the final step of data processing in business intelligence system. Data mining technology is good at finding hidden patterns of data sets by analyzing the correlation between attributes. There are two kinds of data mining patterns: supervised and unsupervised. Supervised data mining requires the user to specify the target's properties and a set of input properties. Based on the PageRank algorithm, each physical node is equivalent to an independent PageRank individual in the background of big data, and the scheduling of the associated nodes depends on PageRank algorithm. Set α_0 to represent the position information of the PageRank and α_1 to represent the position information of the information source. Using the above physical quantities, the selected parameters of the PageRank algorithm can be represented as follows:

$$i = \int_0^{\infty} \frac{w^2 |\alpha_1 - \alpha_0| \times q}{\lambda \dot{u}} dw \quad (1)$$

Among them, w represents pheromone concentration coefficient, q represents optimal foraging path selection conditions, t represents the average round-trip time, \dot{u} represents the behavior vector of ant colony. Mining coefficients are often expressed as $\dot{\xi}$, which can increase gradually with mining time, and are not affected by other physical quantities except PageRank algorithm. Set $|t|$ to represent the execution time of a mining cycle, the simultaneous formula (1) can represent the K-means of big data as:

$$x = \frac{\sum_m^{m'} i(\dot{\xi} + g) \cdot |j|}{\mu |t|} \quad (2)$$

Among them, m represents the lower bound mining condition of structured big data, m' represents the upper bound mining condition of structured big data, g represents the basic application privilege of PageRank algorithm, j represents the mining execution vector of large data parameters, and μ represents the execution load constant term of node organization. At this point, the calculation of all parameter vectors is completed.

Further exploiting the support and hyper-chaining relationship, the tradeoff between the two is that the final result will directly affect the result of the algorithm by giving the user a step-size parameter of the index page, which is more acceptable, even if the user's search efficiency loss is not significant. In this case, the definition of the distance with the step-size parameter needs to be compared:

$$d = \frac{p(b|a)}{D(x \cap y/a \cap b) \times ||T - p(a \cup b)||} \quad (3)$$

Expressed as item sets x and y , the two item sets are independent of each other and do not have duplicate attributes, which can be represented as: $x \cap y = \emptyset$. The personalization rules between set X and set Y must meet the following conditions: the personalization rules of two sets have certain universality and interesting. For example: $T = \{T_1, T_2, \dots, T_m\}$ is a collection of items, related data D is a collection of database

information, where each information d is a collection of items, making $d \subseteq I$. Each message has a glyph, called an d_i . Let a be an itemset, $a \in d$. Personalized association rules have $a \Rightarrow b$ implication, in which $a \subset I$, $b \supset I$, and $a \cap b = \emptyset$.

Based on the above association rules, the support and confidence of personalized characteristic data are calculated. If the $a \Rightarrow b$ is established in the information set D , it has the support of characteristic attribute. The percentage of the information set D that contains $a \cup b$ is also known as support, or $p(a \cup b)$. If the personalized association rule $a \Rightarrow b$ is established in the information set D and has confidence c , the information set D contains information a and all the information ratios of b , which is also called confidence $p(b|a)$.

There are three main algorithms in the standard data mining model: Bayesian algorithm, PageRank algorithm and PageRank algorithm. Unsupervised data mining techniques do not need to have predictable attributes. Before the research of data source, it is necessary to ensure the content of user's demand for data, which requires the clear characteristics of operational data and analytical data. Because the traditional business platform is mainly used in the daily transaction processing of enterprises, the data stored in the database accord with the characteristics of operational data. The data stored in order to meet the requirements of business intelligence analysis and mining, basically accord with the characteristics of analytical data. The difference between the two is shown in Table 1.

Table 1. Characteristics of operational data and analytical data

Operational data features	Characteristics of analytical data
Indicates the dynamic status of business processing	Indicates the static state of business processing
It is correct at the moment of access	Represents past data
It can be updated. It is updated by the entry personnel entering transactions	It is not updatable, and the access rights of end users are often read-only
Handle details	More attention is paid to the conclusive data, which is comprehensive or refined
The operation requirements can be known in advance, and the system can be optimized according to the expected workload	The operation requirements are not known in advance, and Yongyuan does not know what the user will do next
There are many transactions, each affecting a small part of the data	There are a small number of queries, and each query can access a large amount of data
Application oriented, supporting daily operation	Analysis oriented, support management requirements
Users don't have to understand the database, they just input data	Users need to understand the database to draw meaningful conclusions from the data

The basic logical structure of data warehouse includes star structure and snowflake structure. Compared with the two structures, the star structure contains fewer layers

of links for information retrieval than the snowflake structure, is easier to manage, and queries are more efficient; moreover, the star structure based on the star model facilitates the definition of data entities from the perspective of supporting decision makers, and the entities just reflect the core content (theme-related content), and the star model diagram is easy to understand, read and query. The data warehouse star structure diagram is shown in Fig. 4.

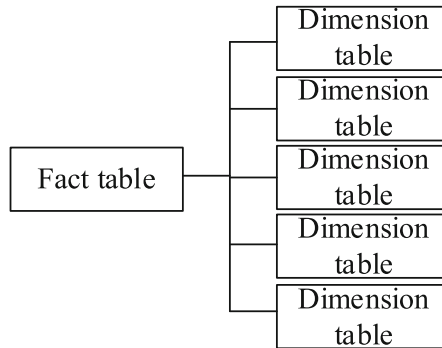


Fig. 4. Data warehouse star structure diagram

Based on the collected and accumulated customer information, the commodity database is used to analyze and mine the customer data, and to predict the customer's trust and desire to buy the product. Database is the process of information leading marketing, the process of information driving sales, and the process of information promoting efficiency. Database is a normal trend, is the trend of the customer as the center of the enterprise, is the trend of the data platform as the business philosophy, is to maintain customer resources to achieve sales goals to expand the trend of sales. Compared with the traditional marketing, the database has significant differences in marketing model, marketing channels, determining goals and many other aspects, as shown in Table 2.

From the macro point of view, the database can accurately predict and feedback the trading market in real time, from the micro point of view, the database can accurately plan the inventory customer's contribution to the enterprise. Therefore, an excellent database can promote the enterprise to understand the real needs of customers, and then interpret the ultimate customer goal in an all-round way, provide a platform for enterprises, humanization, in light of the marketing plan, reduce the operating costs of enterprises, accelerate the marketing efficiency of enterprises, and increase the adhesion degree of customers.

In addition, PageRank algorithm is increasing exponentially in recent years. Using database can effectively help enterprises to intercept effective new information, provide professional theoretical basis for follow-up strategy planning and new product R&D, and improve the competitive advantage of their own products.

Business intelligence refers to that an enterprise makes use of the existing internal and external database data as the source data, through a series of standardized operations such as cleaning, extracting, converting and loading, and then summarizes and

Table 2. Comparative analysis of traditional marketing and database

Measurement standard	Traditional marketing	Database marketing
Marketing model	One way communication marketing mode, passive acceptance of products and advertising information by the audience	Interactive marketing attracts the audience to actively participate in the whole process by designing various scenes
Delivery channel	Television, newspapers, advertisements and other mass media	Personal letters, e-mail, mobile phones, etc.
Target audience	—Communicate to many and face the public	One to one communication, accurate face to each customer and meet customer needs
Deliver content	It is strongly restricted by time, layout and cost	Convey rich advertising information, with less time and space constraints
Customer relationship	Unable to grasp every piece of feedback information in hand, the controllability is low	Through reliable feedback analysis, we can understand the psychology of the audience and establish a customer relationship database to maintain close contact with the audience
Cost	Facing the public, the general price is relatively expensive, especially the price of TV media is higher	Every advertising fee is spent on potential customers to avoid waste of supply and high cost performance
Sustainability	The repeatability is low, and the process is still complex	Develop new service items and simplify the purchase process, bringing the possibility of repeated purchase

stores them to the data warehouse, and then analyzes and processes the data in the data warehouse by adopting online analytical processing, data mining and other information technologies so as to convert them into the information and knowledge that users need to be able to understand, and transmit the potential characteristics and forecast results to the decision-makers so as to assist decision-making and improve performance. Therefore, business intelligence is not a new information technology, it is just a combination of data warehouse, online analytical processing and data mining. The business information feature mining management architecture is shown in Fig. 5.

To sum up, the ultimate goal of business information intelligence is to transform the predicted data and decision information into knowledge successfully by cleaning, processing and analyzing the information flow of all kinds of data sources.

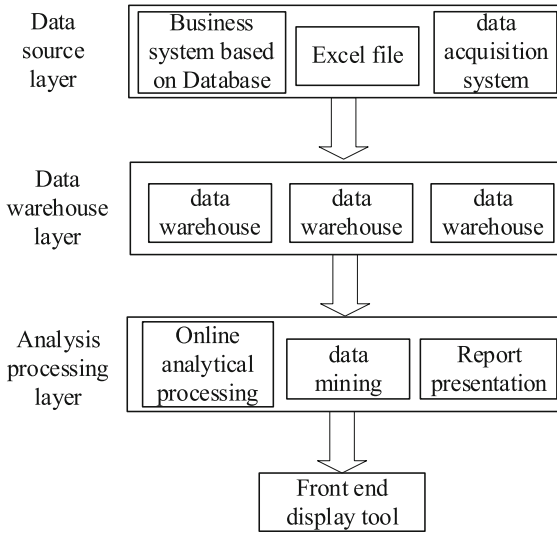


Fig. 5. Business information feature mining management architecture

2.3 Design the Realization of Business Information Mining in Media Platform

In the process of data mining, it is necessary to determine the mining object, prepare the data, build the model, analyze the results of data mining and apply knowledge. Data mining is very useful for finding and describing hidden patterns in a particular cube. Because the data in a cube grows rapidly, it is very difficult to find information manually. SQL Server2015’s Data mining algorithm allows automatic pattern finding and interactive analysis During interactive analysis mining, the specific data mining processing flow is shown in Fig. 6.

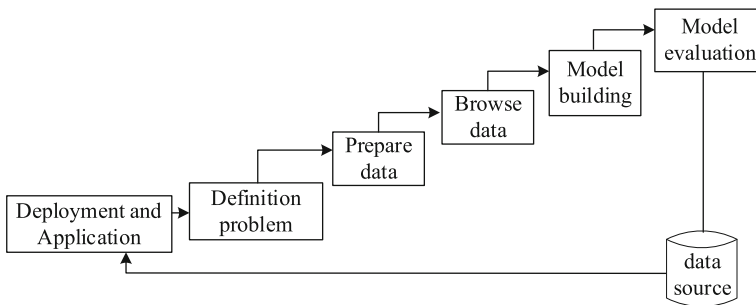


Fig. 6. Business information platform data collection and processing process

In terms of information security, the framework mainly integrates the existing operation platform and the security control mechanism on the database platform to guarantee the data security and operation security, protects the product sales business data by

encryption at the database level, and invokes the identity authentication service to confirm that each business has the correct principal user when the user enters the business view or functional navigation unit for the first time. The list of internal variables in the metrics calculation is also restricted to the subject matter of the business leader or the principal user of the calculation through security rules. The architecture of the marketing support platform is shown in Fig. 7.

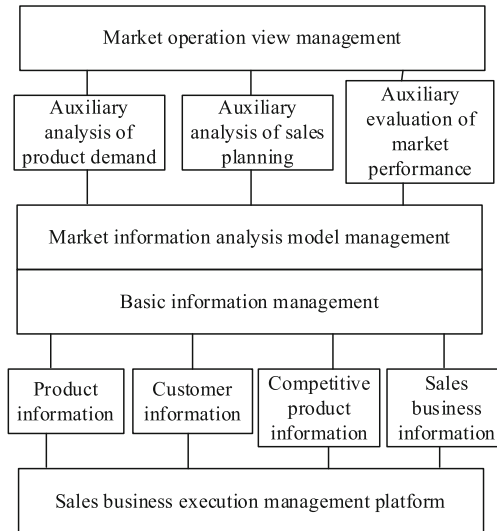


Fig. 7. Structure of business-to-information media platform

By assigning explicit fields to entities in the overall database scheme and implementing reference relationships between entities through external attributes, the logical design in the previous section is transformed into a relational data model, so that the database can be deployed on the current mainstream database management platform. The following are some examples of data tables that describe the field attributes. The product list is shown in Table 3, and each item represents one model of product, including basic information about the competitive products of the Company and other manufacturers. The sales record table is shown in Table 4.

This feature has performance advantages in some measurement tasks that require processing a large number of data sets to determine more reliable results. In the design scheme, the parallel model objects access different input data, and the output results are stored in different temporary data tables, so there is no interaction between them. This simple running logic is good for performance and also good for proper programming. Calls to the following algorithms can be executed as a combination of parameters, each of which is executed as a separate input item, or executed in parallel based on a multithreaded mode, to speed things up, as shown in Fig. 8.

Therefore, the enterprise should improve its information ability and perfect its internal information technology system. Now we are in an information age with rapid expansion and development. Overwhelming information is either actively or passively accessed

Table 3. Product information management table

Field	Field description	Field type	remarks
UQN	Product identification	CHAR	Primary key
DIN	Manufacturer identification	CHAR	Primary key and all zeros indicate the company
PQ	Product type	CHAR	_____
PSQ	Product subclass	CHAR	Can be empty
PME	Product model	VARCHAR	_____
PTK	The first time the product is on the market	DATETIME	_____
QOD	Quality grade	CHAR	_____
SOU	Current sales status	CHAR	Trial sale, sale and shutdown
CUG	Market share of similar products	NUMERIC	%
AOI	Current cumulative sales	INT	Element
AIU	Current cumulative selling expenses	INT	Element
AYU	Current remaining inventory	INT	_____

Table 4. Sales record table

Field	Field description	Field type	Remarks
OIN	Record number	INT	Primary key
CIN	Product identification	CHAR	Primary key
CPU	Time record	DATETIME	_____
CY	Sales batch	INT	_____
XIN	Selling price	INT	Element
CI	Customer number	INT	Foreign key
XQYM	Batch sales expenses	INT	Promotion expense items
CIYK	Contract amount	INT	Element
CIYX	Contract No	CHAR	Foreign key

by people. Therefore, selective access to information is very necessary. A successful

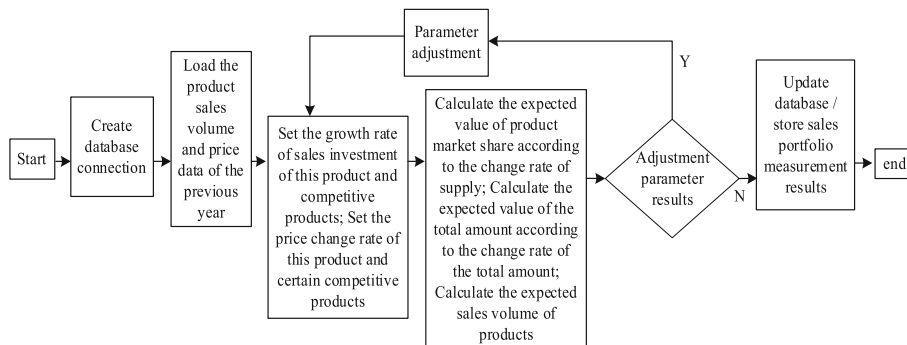


Fig. 8. Business platform information mining management process

analysis begins with the collection and screening of information, begins with the establishment of a pagerank network, and then calculates the probability of a conclusion in each pagerank network based on the quantized information. Finally, the correct conclusion can be reached only after repeated deliberation. Therefore, enterprises should filter purposefully, try their best to quantize every information, and put the quantized information into the PAGERANK network for analysis and collation, and summarize the possible logical relations. Only by combining the pagerank network with quantitative information can a variety of possibilities be analysed to arrive at a more factual conclusion; secondly, with the development of science and technology and economic integration, the increasing interdependence of production technology and economy among enterprises, the trend of resource exchange and complementarity is becoming more and more evident. Therefore, enterprises shall seek truly influential partners on the basis of resource sharing, mutual benefit and reciprocity, and through the combination of the strong and the strong, learn from each other to the maximum extent possible to make use of data and resources, improve their own products, promote the development of technologies, realize the upgrading of key components and technologies, effectively avoid homogenized competition, and enhance core competitiveness. Finally, enterprises must have a deep understanding of user needs, research appropriate products and services with users as the center, and form differentiated competition. Therefore, enterprises need to establish management process and decision-making system based on customer demand management and product planning, transform demand into products with differentiated competitive advantages, and quickly respond to the explicit and implicit demands of users, so as to grasp the market development trend and enterprise rhythm and maintain high profitability.

3 Analysis of Experimental Results

Collect the information, information level, cooperative innovation ability and enterprise performance of a national chain retail enterprise from a social media network platform, and evaluate it with a five level evaluation scale. When developing the website ranking evaluation data mining module, Microsoft SQL Server 200 is selected as a tool for

building a data warehouse; The data mining module developed by Microsoft Visual Basic6.0 is used to process the relevant data of business information mining on the social media platform, and the processing results are saved in the SQL Server database. The resulting data reached 5.29G. The reliability and validity of the collected data were tested to verify the stability and centralization of the data. In this paper, Cronbach’s alpha coefficient method is selected for reliability test. Simulation experiment environment is: Win10 operating system, Intel i5-9400F processor and NVIDIA GeForce GTX1070 graphics card. The model was trained and tested in the Python programming language. By comparative analysis, the Cronbach’s alpha coefficient of each variable is more than 0.7, the stability of data is better, the reliability is higher. In addition, the fitness of the model is tested by confirmatory factor analysis. The standardized factor load of each variable was > 0.5, AVE > 0.5, CR > 0.7, which accorded with the test standard, and the polymerization validity of the model was high. The AVE of each variable was higher than the value of correlation coefficient of this factor and other factors by the discernibility validity test, and the discernibility validity was good. The correlation test results and model fit indices are shown in Tables 5 and 6.

Table 5. Correlation test

Variable	Information acquisition	Information mining	Information creation	Technical cooperation	Organization cooperation	Strategic cooperation	Enterprise performance
Information acquisition	0.828	0.552**	0.482**	0.352**	0.342**	0.358**	0.335**
Information mining	0.552**	0.742	0.741**	0.582**	0.625**	0.575**	0.595**
Information creation	0.482**	0.742**	0.765	0.602**	0.633**	0.599**	0.628**
Technical cooperation	0.358**	0.582**	0.602**	0.825	0.509**	0.565**	0.671**
Organization cooperation	0.342**	0.623**	0.635**	0.509**	0.752	0.606**	0.652**
Strategic cooperation	0.381**	0.576**	0.599**	0.565**	0.618**	0.755	0.655**
Enterprise performance	0.335**	0.595**	0.625**	0.671**	0.652**	0.655**	-

Table 6. Model fitting index

Fitting index	Absolute fitting index			Relative fitting index			Parsimony fit index	
	X2/df	GFI	RMSEA	NFI	TLI	CFI	PGFI	PNFI
Standard value	<4	>0.9	<0.09	>0.9	>0.9	>0.9	>0.6	>0.6
Actual value	1.513	0.832	0.052	0.838	0.952	0.963	0.723	0.771

The attributes of the indicators shown in Table 6 are explained as follows: the absolute fitting index is the measure to which a model interprets the various relationships found by the sample data. The relative fitting index is a measure of how well a model fits the data. The parsimony fit index refers to the overall fit of the proposed model to the sample data.

In the phase of test implementation, professional testers verify the correctness of function implementation item by item based on the pagerank method according to the planned test cases and sample data, as well as the confirmed function requirement use cases. Using a sensitivity curve, the larger the area under the curve, the higher the accuracy, and the smoother the curve, the better. Business information mining flexibility detection results are shown in Fig. 9.

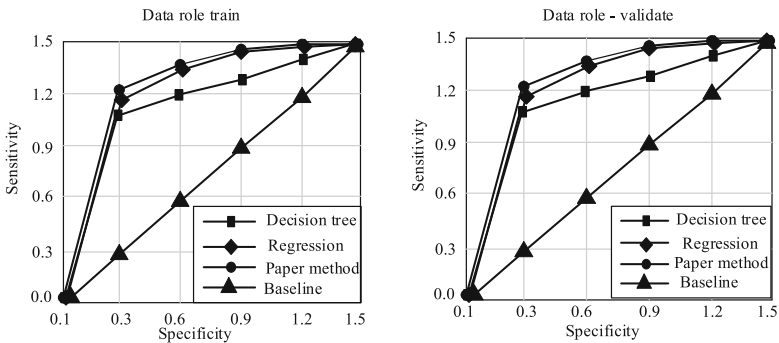


Fig. 9. Business information mining flexibility detection results

As a whole, the curve smoothness of the traditional method is not good, which shows that the stability of the traditional method is worse than the other two models. The PageRank algorithm is better than regression, but the two algorithms are very close and have no obvious advantages or disadvantages. Testing is done according to the planned priority. In addition to verifying that the expected response exists under the correct input conditions, it also verifies the stability in the case of exceptions such as permissions and incomplete data. According to Table 7, the phased testing tasks are summarized.

Applying the above method to specific functions, the test cases and test contents of each module are determined and implemented. Tables 8 and 9 are the module test items and results summary in the implementation phase.

Based on the above research results, the proposed PageRank -based business information mining technology for social media platform has high practicability and accuracy in the practical application.

Table 7. Phased data mining test tasks

Development stage	Main tasks related to testing
Requirement analysis	Determine the test cases according to the functional requirements analysis report, and preliminarily determine the test conditions, priorities and contents of each case
Software design	Determine the implementation difficulty of each function according to the design scheme, and then determine the test conditions, main test tasks, data samples and test plan of each test case
Programming implementation	Complete the source program audit and program debugging of the program unit
Program test	Perform item by Item Verification of test cases according to the test plan

Table 8. Basic information mining module test summary

Module function grouping	Test case	Test conclusion
Basic information query	Query and statistics of product information; Query and statistics of product cost information; Information inquiry and statistics of competitive shopping malls; Sales business information query and statistics	Adopt
Basic information analysis	User defined indicator statistical analysis; Multidimensional view generation; User defined index correlation analysis	Adopt
Business data import	Import actual business data from the company's sales business processing platform and generate records according to the database structure of the system	Adopt

Table 9. Product requirement information mining module test summary

Module function grouping	Test case	Test conclusion
Parameter estimation of measurement model	Estimation and setting of market aggregate change rate, price change rate and other parameters	Adopt

(continued)

Table 9. (continued)

Module function grouping	Test case	Test conclusion
Analysis and calculation of product sales and short-term demand for sales	Demand calculation based on total amount analysis method and demand calculation based on sales factor combination method; Demand calculation based on modified sales factor combination algorithm	Adopt
Analysis and calculation of product sales volume and short-term demand	Decomposition calculation of trend component, volatility component and periodicity component of sales volume and sales volume time series; Sales volume and sales volume time series extrapolation prediction calculation; Simple estimation calculation	Adopt
Other auxiliary functions	Query of calculation results; View generation; Report export	Adopt
	Report format conversion	Adopt

4 Conclusion

This paper proposes the business information mining technology based on PageRank, and creatively introduces the data mining technology to design the business information management system of the social media platform; Based on PageRank data mining algorithm, the business information is evaluated according to the ranking of the web page, and the business information mining of the media platform is completed. The experimental results show that this method has high practicability and accuracy in the practical application of social media platform. With the continuous application of data mining in business information services, I believe that the potential of data mining will be greatly stimulated, and the application of data mining in business information services will be further promoted. However, due to the limited conditions, the data mining efficiency of this method has some shortcomings. Future research will continue to use data mining technology to improve the efficiency of business information mining.

References

1. Madhumathi, S., Gomathi, R.: Data mining in Ecommerce platforms for product managers. Res. J. Eng. Technol. Int. Peer-reviewed J. Eng. Technol. **12**(1), 1–7 (2021)
2. Asrin, F., Saide, S., Ratna, S.: Data to knowledge-based transformation: the association rules with rapid miner approach and predictive analysis in evergreen IT-business routines of PT Chevron Pacific Indonesia. Int. J. Sociotechnol. Knowl. Dev. **13**(4), 141–152 (2021)
3. Su, J., Xu, R., Yu, S., Wang, B., Wang, J.: Idle slots skipped mechanism based tag identification algorithm with enhanced collision detection. KSII Trans. Internet Inf. Syst. **14**(5), 2294–2309 (2020)

4. Su, J., Xu, R., Yu, S., Wang, B., Wang, J.: Redundant rule detection for software-defined networking. *KSII Trans. Internet Inf. Syst.* **14**(6), 2735–2751 (2020)
5. Un, N.: On the tradeoffs of internet personal information and its governance mechanism. *Hebei Law Sci.* **38**(07), 96–113 (2020)
6. Zhang, Z., Shang, Y., Xu, Y.: Impact of IT-business alignment on business model innovation: based on data empowerment perspective. *Sci. Technol. Manage. Res.* **41**(12), 1–8 (2021)
7. Wang, C., Yin, S., Liu, W., et al.: High utility itemset mining algorithm based on improved particle swarm optimization. *J. Chin. Comput. Syst.* **41**(12), 1–8 (2021)
8. Ang, Z., Chen, J.: Sustainable business model of platform context: logic and implementation. *Sci. Sci. Manage. S. & T. (Mon.)* **42**(02), 59–76 (2021)