

Research on the E-commerce Platform Performance and Green Supply Chain based on Data Mining and SVM

Ying Xu¹, Xuemei Zhang¹ and Hong Zhang^{2,*}

¹ Business School, Sias International University, Xinzheng City, 451150, China

¹ School of Electronic Commerce, Ningbo DaHongYing University, Ningbo City, 315175, China

² Department of Economics and Management, Changsha University, Changsha City, 410022, China

Corresponding author: Hong Zhang, 1445061952@qq.com

Abstract

In the network environment, supply chain management has greatly reduced the product development cycle, reduce the inventory. With the continuous development of information technology, e-commerce logistics platform has become the main factor affecting the development of logistics industry. In this paper, the authors research on the E-commerce platform performance and green supply chain based on data mining and SVM. The green supply chain considers the environmental problems in every link of the supply chain, and promotes the coordinated development of economy and environment. The result shows that the most critical factor that affects the satisfaction of consumer to B2C e-commerce platform is the accurate, complete and reliable logistics service.

Keywords: Data mining, E-commerce platform, Green supply chain, SVM algorithm

1. Introduction

Since the beginning of twenty-first Century, the global environmental problems are getting worse and worse, which makes people pay more and more attention to the study of environmental problems and the sustainable development of society. And national environmental laws and regulations are more and more strict, the country's green barriers will continue to strengthen, enterprises are under pressure from all aspects of environmental protection[1-2]. This makes the enterprise must consider the issue of environmental management, environmental management and participate in many enterprises only on the downstream, in order to produce green products, to achieve maximum economic and social benefits, so we must put environmental management into the supply chain management, a new method of supply chain management, supply chain management[3]. Green supply chain management, referred to as GSCM, it considers the various aspects of the supply chain environmental issues, focusing on environmental protection, promote the coordinated development of economy and environment. The emergence and development of electronic commerce is the result of economic globalization and the innovation of network technology[4-5]. It completely changes the original logistics, information flow and capital flow in the supply chain. It makes full use of the resources, improves the efficiency, reduces the cost and improves the service quality. E-commerce is emphasized in the computer network environment of commercial applications, not only the combination of hardware and software, but the buyers, sellers, vendors and partners in the application of Internet combination[6]. With the advantages of globalization, high efficiency, low cost, high efficiency, high selectivity and so on. The wide range of e-commerce, the use of an unprecedented network of customers, vendors,

suppliers and employees linked together. So as to realize the ideal mode of small batch, multi varieties, zero inventory, instant manufacturing and delivery, and greatly improve the speed of the transaction[7].

In the network environment, supply chain management can greatly shorten the product development cycle, reduce inventory, transportation cost can be further reduced, but it mainly consider how to achieve supply chain optimization and coordination of the whole, how to improve the service level and rapid response, and the impact of supply chain management in the implementation process on the environment the products and green issues are often not enough attention, and these problems are restricting the further optimization of supply chain management[8-9]. Sustainable development should be an important consideration, not only concerned about the short-term interests of the enterprise needs. Therefore, in the electronic commerce environment, considering the environmental management of the green supply chain management is the causes of many enterprises attach importance[10]. Green supply chain management emphasizes the overall and a long-term interest, focus on all aspects of the environment, reflects the company's green image, is the new development trend of supply chain management in e-commerce environment, the rapid development of the green supply chain management, more can be used effectively and quickly.

2. SVM algorithm and Data Mining

2.1. Support Vector Machine

The support vector machine method is a learning theory of VC dimension theory and structural risk minimization principle on the basis of statistics, according to the limited sample information between model complexity and learning ability to seek the best compromise, in order to obtain the best generalization ability. First, the linear separable problem is analyzed. For linearly separable training sets:

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad (1)$$

Hypothesis existence discriminant function:

$$f(x) = \text{sgn}((\omega \cdot x) + b) \quad (2)$$

Linear case, the optimal classification of the idea of the super plane can be described in figure 1.

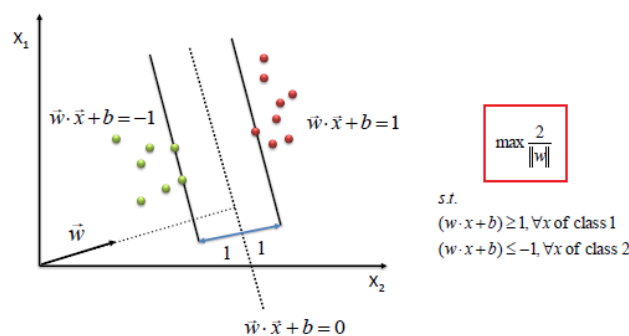


Figure 1. The Optimal Separating Hyper Plane

In order to maximize the classification of the two types of samples, it is needed to maximize the distance, that is, structural optimization problem:

$$\min \quad \frac{1}{2} \|\omega\|^2 \quad (3)$$

$$s.t. \quad y_i (\omega \cdot x + b) - 1 \geq 0$$

When the linearity is not available, there are some sample points which are not satisfied, then the formula is introduced into the:

$$\min \quad \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^n \xi_i \quad (4)$$

$$s.t. \quad y_i (\omega \cdot x + b) \geq 1 - \xi_i$$

The solution can be obtained by solving the equation:

$$\omega = \sum_{i=1}^n \alpha_i y_i x_i \quad (5)$$

$$b = y_i - \sum_{i=1}^n y_i \alpha_i (x_i \cdot x_j) \quad (6)$$

Then we get the linear discriminant function:

$$f(x) = \text{sgn} \left(\sum_{i=1}^n \alpha_i y_i (x_i \cdot x_j) + b \right) \quad (7)$$

2.2. Data Mining

Data mining (DM), also known as knowledge discovery in databases (Knowledge Discovery in Database, KDD), is a non trivial process to extract valid, novel, potentially useful and ultimately understandable patterns from large amounts of data in the. It combines database system, artificial intelligence, statistics, machine learning, information science, etc., is a new interdisciplinary field of application. Data mining technology has played a more and more important role in the decision support activities in various industries. A large number of data mining tools have emerged in the early stage of solving the problem of independent data mining software, to solve the common problem, set a variety of data mining algorithms in one of the horizontal data mining system, so far has emerged a lot of data mining tools. Data mining is a process of data analysis and knowledge acquisition, which is divided into steps, multi angle data analysis and knowledge acquisition. In order to combine the data mining process with the specific application and development process, the key step of the commercial development is to establish a unified process standard. This will help to form a record of work experience can effectively unified system, to strengthen the project plan and project management; the whole process will help beginners to successfully complete the data mining; for each step of the detailed planning and design process standards, good control and reduce the cost of the project. The standardization of data mining is still in the primary stage, so far, there is no uniform standard. According to the current standards of the problem solving methods and the different focus, the data mining standards are divided into four categories.

- 1) **Process standards:** defining data mining model generation, use and deployment of process standards, such as Fayy and ad CRISP-DM process standards, etc..

- 2) **Interface standard:** for the convenience of the customer application program calls for specific programming language and system data mining API interface, such as SQL/, MM, JDM, etc..
- 3) **Language standards:** data mining problem definition for problem description, knowledge discovery and expression of data mining language standards. Development of data mining platform and application program is with unified language standard specification. Similar with the SQL language, has designed a data mining query language (such as DM, QL, MSQL and Mine Rule), data mining (such as PM, ML language, CWM for DM) and the general data set, query definition and manipulation in one language (such as OLE, DB mining for DM).
- 4) **Network standards:** to solve the problem of distributed and remote data mining on the network data mining Web standards, such as, L for Analysis XM, Space Data, Web Semantic, etc..

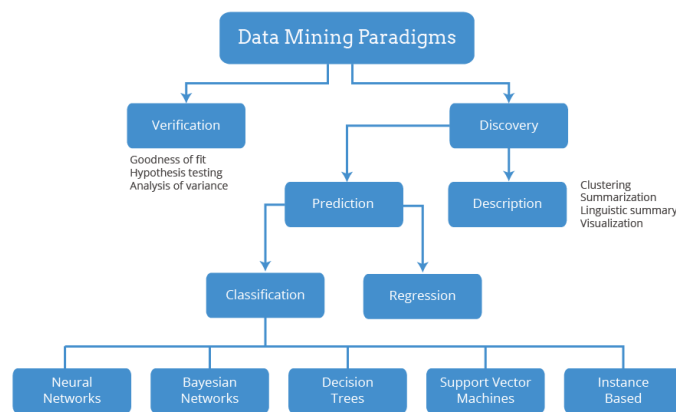


Figure 2. Data Mining

3. Electronic Commerce Platform and Green Supply Chain Management

3.1. Electronic Commerce Platform

With the development of mobile Internet, mobile Internet shopping is convenient, social to promote mobile online shopping consumer increasingly convenient, integrated platform based on social based on strong communication function also with information sharing, mobile shopping, payment, financial and other integrated services, rich B2C transactions and the scope of the market. As a result, B2C has become the main driving force of China's online shopping market, various types of B2C e-commerce platform as a comprehensive shopping mall platform, shopping platform, franchise enterprise network marketing platform, group purchase platform will gradually play its main role. Because the Internet itself cannot touch, virtual and anonymous, there were a lot of problems make B2C e-commerce market rapid development, due to the customer perception of online products and services there is a big uncertainty, resulting in online shopping overall customer satisfaction degree is relatively low. Consumers B2C e-commerce complaints two most followed by the arrival of slow return, which accounted for 18% and 8%. These all belong to the logistics service. Logistics service quality has caused a great impact on the consumer experience of consumers. According to the electronic commerce shopping data statistics, B2C e-commerce users of the logistics link satisfaction is only 75.1%, thus, the logistics service quality is an important link in the process of B2C e-commerce consumers concerned.



Figure 3. E-commerce Logistics Platform

3.2. Green Supply Chain Management

The traditional supply chain management is a chain, is a kind of integrated management, it starts from raw material suppliers, terminating in a user. It is on the whole supply chain organizations, between the departments of logistics, capital flow and information flow planning, coordination and control, the traditional e-commerce supply chain management process as shown in Figure 1, including time, quality, cost, service four goals, the final task is to achieve the enterprise benefit maximization. Traditional supply chain management is not designed to take into account the problem of product recycling, ignoring the impact on the environment. Green supply chain management compared with the traditional supply chain management, a multi functional environment, it will be from raw material suppliers, manufacturers and distributors through, finally user information flow, logistics, service and capital flow, integrated into a complete supply chain, using the method of system management, and the requirements of enterprises will be the idea of sustainable development into the implementation of green management of the entire supply chain. Green supply chain management will be produced in the production process of waste, waste and transport, warehousing, sales process and the damage caused by the user to be eliminated by the product recycling. In this way, the green supply chain management will be the disposal and recycling of waste products and raw materials after the process of using expanded to the entire supply chain management, increase the length of the supply chain, at the same time because of the reverse logistics, supply chain from the traditional chain into a ring.



Figure 4. Electricity Supplier Supply Chain

(1) Procurement management

Green procurement refers to as far as possible to the environment and ecology of the environment and no harm to the environment or small products or services. Through the source control, in the procurement of materials to consider whether it is environmentally friendly materials. Is green purchasing, it can have a positive impact on the environmental performance of enterprises in all aspects, to meet the public demand for environmentally friendly products, but also can reduce the whole cost, bring economic benefits and competitive advantage for enterprises. Green procurement is the most important part of the selection of excellent suppliers, suppliers must first have ISO14001 environmental certification standards, environmental protection enterprises, so as to establish a complete system of green supply chain.

(2) Green design

Green design is to give full consideration to the product function, quality, cost and development cycle in the stage of conception, advocate and no waste Recyclable design technology, 3R (Reduce, Reuse, Recycling) principle into the phase of product development of products on the environment and resources in the whole life cycle, the minimum consumption of products the impact on the environment and resources.

(3) Green production

Green production and cleaner production, which requires enterprises to cleaner production technology, process improvement through continuous improvement and management, instead of toxic materials and other methods to strengthen internal management, reduce waste and pollutants generated from the source, improve resource utilization, reduce the generation and emission of pollutants, in order to reduce the harm to the environment and human. The aim of green manufacturing is to reduce the impact of the product on the human and the environment throughout the entire production cycle, including the final disposal of the raw material from the raw material to the final disposal of the product.

(4) Green marketing

Green marketing is a new stage in the development of marketing. It is a new mode of thinking and operation of the new marketing mode of thinking and caring for the ecological level of the sustainable development. It refers to the supply chain enterprises in market research, product development, product pricing and promotional activities in the marketing process, to maintain the ecological balance and the importance of environmental protection green theory as a guide, the development of enterprises and the interests of consumers and society consistent. Green marketing is the core of the environmental impact of product marketing, is an extension of traditional marketing, pay more attention to the global environmental protection.

(5) Green Logistics

Green logistics is to improve the logistics system from the angle of environment, the formation of a logistics system with friendly environment, and by inhibiting logistics cause harm to the environment, which can promote the economic and social life of the healthy development of the logistics system. Materials and products in the transport process of the impact on the environment is a problem that can not be ignored, resulting in the transport process of the energy consumption and exhaust gas produced, and the possible leakage of materials and products, which will affect the environment. Successful green supply chain management must have green logistics, so that materials and products in the logistics process to reduce the harm caused to the environment, and to make full use of logistics resources.

(6) green recovery

Green recovery is the process of logistics activities from users to manufacturers or suppliers. It refers to the product is discarded, the product and parts of the recycling process, so that products or components to be recycled or recycled to reduce environmental pollution, improve resource utilization. Through the green recycling, can save raw materials, reduce production costs; reduce the pollution of the natural law.



Figure 5. Green Supply Chain

4. Empirical Analysis

4.1. Questionnaire Design

The main content of this study is the influence of logistics service quality of B2C e-commerce platform, this study adopts the method of questionnaire to collect data for ever B2C e-commerce platform consumer behavior groups. The investigation of B2C electronic business platform of logistics service quality and customer loyalty, we use Likert scale, so that consumers scoring for the problems described in accordance with their true feelings, true views in order to be able to close to the consumer. Formal sample collection, this study used convenience sampling method with special questionnaire website, designed the network questionnaire, the final total of 350 copies of questionnaires, obtained 295 valid questionnaires, the effective questionnaire was 84.1% overall.

Table 1. Sample Statistics

Variable	Option	The number	Percentage
Gender	Male	142	48.2%
	Female	153	51.8%
Age	20-25	67	22.7%
	25-30	146	49.1%
	>30	82	29.2%
Occupation	Student	196	66.5%
	Clerk	73	24.7%
	Liberal professions	26	8.8%

4.2. Reliability Analysis

In order to purify the measurement items of potential variables, this paper selects the total correlation coefficient (Item-Total Correlation Corrected, referred to as CITC) to eliminate the non - consistent measurement items. For the CITC value of less than 0.5 and can be added to increase the Cronbach series of numerical items to be deleted. Through the operation of SPSS18.0 statistical software, statistical results are shown in the following table:

Table 2. Cronbach's Alpha Reliability Test

Variable	Problem item	CITC	Remove the α value after the title	α coefficient
Reliability	R1	.423	.795	.802
	R2	.640	.724	
	R3	.712	.768	
Flexibility	F1	.546	.572	.621
	F2	.552	.596	
Time character	T1	.662	.737	.793
	T2	.603	.743	
	T3	.583	.706	
Customer satisfaction	CS1	.614	.745	.802
	CS2	.667	.714	
Customer trust	CT1	.646	.702	.828
	CT2	.713	.753	

As shown in Table 2, all items of the total correlation (CITC) was more than 0.5, the α coefficient of reliability, flexibility and time character reached 0.802, 0.621, 0.793, the internal consistency of the items is ideal. By using SPSS18.0 software, we test the reliability of the B2C business platform of logistics service quality, the KMO and Bartlett test of sphericity results shown in table 3.

Table 3. KMO and Bartlett's Test

Kaiser-Miyer-Olkin		.804
Bartlett Sphericity test	Approximate chi square	154.129
	Degree of freedom	210
	Significant	.000

From table 3 it can be seen that the B2C e-commerce platform of logistics service quality of the 12 items were obtained KMO value is 0.804, ranging from 0.8 to 0.9, said the logistics service quality in all issues very suitable for factor analysis, Bartlett Sphericity test results very significant correlation very good. The results can be seen in Table 4, there are 6 eigenvalues, and logistics service quality factors were extracted from the 6 factors, these 6 factors can explain the B2C e-commerce platform of logistics service quality 61.284%.

Table 4. Total Variance of Logistics Service Quality

Ingredients	Total	Variance	Accumulate
1	6.142	45.271	32.104
2	3.051	11.571	43.564
3	2.178	9.842	49.423

4	1.649	5.651	54.468
5	1.195	4.742	59.125
6	1.052	4.281	61.284

4.3. Multiple Regression Analysis

From table 5, we can see that the customer satisfaction to the logistics service quality of the regression model to adjust the R square value is 0.614 which indicates that the model has reached 50.7%, the customer satisfaction factor has a good explanation ability. The test value of DW is 2.157 in the range of 1.5 to 2.5 and is very close to 2, which indicates that there is no sequence correlation between the error term and the error term is independent. so the regression model can be judged. Through the above analysis, we can know that the regression model of customer satisfaction through the test.

Table 5. Abstract of Model

Model	R	R2	Adjust the R 2	The standard estimate error	Durbin-Watson
1	.614	.534	.493	.686	2.157

Table 6. Regression Coefficient

Variable	Non Standard coefficient		Standardized coefficient	t	Sig.
	B	Standard error	Beta		
constant	.642	.272		1.158	.051
R	.116	.059	.274	4.458	.000
F	.424	.045	.368	5.948	.003
T	.028	.027	.019	1.328	.000
CS	.202	.055	.014	1.239	.000
CT	.136	.049	.128	-1.921	.027

From table 6 , we can get that B2C e-commerce platform in the right way to provide customers with reliable logistics services, so that customers feel at ease, do not worry about risk, will significantly improve the customer satisfaction of the consumer experience of the platform. In addition, B2C e-commerce platform can quickly respond to orders, provide quick return service, accurate and timely information related to the goods to the customer feedback is also an important factor affecting customer satisfaction on the platform. In contrast, B2C e-commerce platform for the diversity of logistics services and the cost of logistics services to the extent of the impact of customer satisfaction is weak, and even less obvious.

5. Conclusions

The traditional industrial linear economic development, one-sided emphasis on economic growth, ignoring the compatibility of human development and natural environment, caused by environmental burdens, resources to support the situation, each enterprise faces a major challenge to the coordinated development of production and environment. At the same time, consumers require producers to provide products must meet the requirements of green products, the enterprises to carry out green management, production of green products, and to produce the waste recycling, reduce waste pollution of the environment, improve the utilization rate of resources. Therefore, the idea of

sustainable development in the production and operation of enterprises is particularly important. Effective environmental management is a powerful weapon to improve core competence. Green supply chain management is such a kind of economic and environmental win-win management model, it is a new strategic mode of enterprise development, is the inevitable trend of supply chain management in our country. And this model in the e-commerce environment, can make the supply chain enterprises to further improve efficiency, optimize resource utilization, achieve rapid market response, reduce operation time delay, improve customer service and seek to expand new market opportunities, and finally get the optimal economic and social benefit.

Acknowledgments

This work was financially supported by project of Humanities and Social Sciences Research Project of Ministry of Education, China (14YJC630154); ZheJiang Nature Science Foundation of China (LY13G030031); ZheJiang Philosophy & Society Science Foundation of China(16NDJC094YB); NingBo Nature Science Foundation of China (2014A610174); Ningbo Soft Science Foundation (2016A10059); Changsha Science and Technology Project: The design and implementation of the SMEs accounting original document information system based on cloud services.

References

- [1] D. Yao and Q. Liu, "Competitive pricing of mixed retail and e-tail distribution channels", *Omega*, vol. 33, (2005), pp. 235-247.
- [2] A. Enders and T. Jelassi, "The converging business models of Internet and bricks-and retailers", *European Management Journal*, vol. 18, (2000), pp. 542-550.
- [3] K. Anand, "Context-general and Context-specific Determinants of Online Satisfaction and Loyalty for Commerce and Content Sites", *Journal of Interactive Marketing*, vol. 24, no. 3, (2010), pp. 222-238.
- [4] A. Niklas and S. Fredrik, "Electronic commerce, marketing channels and logistics platform-a wholesaler perspective", *European Journal of Operational Research*, vol. 144, no. 2, (2003), pp. 270-279.
- [5] C. Druehl and E. Porteus, "Price competition between an Internet firm and a bricks and mortar firm", Working Paper, (2001), pp. 24-30.
- [6] J. Kim, "The role of etail quality, e-satisfaction and e-trust in online loyalty development process", *Journal of Retailing and Consumer Services*, vol. 16, no. 4, (2009), pp. 239-247.
- [7] S. Park and H. Keh, "Modeling hybrid distribution channels: a game-theoretic analysis", *Journal of Retailing and Consumer Services*, vol. 10, (2003), pp. 155-167.
- [8] K. Cattani and W. Gilland, "pricing strategies for a manufacturer adding a direct channel that competes with the traditional channel", *Production and Operations Management*, vol. 15, (2006), pp. 40-56.
- [9] Z. Huang and M. Benyoucef, "From e-commerce to social commerce: A close look at design features", *Electronic Commerce Research and Applications*, vol. 12, no. 4, (2013), pp. 246-259.
- [10] D. Jutla and P. Bodorik, "Developing internet e-commerce benchmarks", *Information Systems*, vol. 24, no. 6, (1999), pp. 475-493.