International Core Journal of Engineering Volume 8 Issue 10, 2022
ISSN: 2414-1895 DOI: 10.6919/ICJE.202210_8(10).0033

Research on Credit Rating of E-commerce Sellers based on K-means Algorithm

-- Taking Taobao Children's Clothing Store as an Example

Jingjing Hu

Shanghai University, Shanghai, China

*jingjinghu@shu.edu.cn

Abstract

With the popularity of e-commerce, seller credit has become a decisive factor in transactions on e-commerce platforms. This article clarifies the current e-commerce sellers' credit score calculation rules and e-commerce sellers' credit grading methods. For the seller's e-commerce credit score, eight related indicators are selected: product quality, picture matching, reasonable pricing, service attitude, delivery speed, after-sales service, logistics speed, and packaging quality. Take Taobao.com children's clothing sellers as an example.We quantify the weights of the eight factors. Then, the K-means algorithm was used to grade the credit of the collected 48 stores on the Python platform.

Keywords

E-commerce Sellers; Credit Evaluation System; K-means Algorithm.

1. Introduction

Integrity is the most important factor that cannot be ignored in e-commerce platforms. The credit level of sellers affects the success of transactions and the trust of consumers. The existing Taobao seller's store credit score calculation mechanism is based on the evaluation given by the buyer to add or subtract the store's credit score accordingly. Among them, one point is added for positive reviews, no additional points are added for moderate reviews, and one point is subtracted from negative reviews. This rule has three obvious flaws: credit scores are not comparable and authentic, credit calculation methods are simple, indistinguishable, and credit evaluation has a small screening surface.

Yang et al.[1] proposed that the credit problem of e-commerce is caused by the information asymmetry between consumers and merchants. E-commerce platforms need to place supervision responsibilities in an important position, increase supervision, and ensure the interests of both parties. Based on the existing e-commerce platform scoring model, Cao et al.[2] constructed a hierarchy, and added three factors into the model: transaction amount weighting, buyer credit weighting, and transaction time weighting. Xu et al.[3] analyzed and proposed an improvement method for the e-commerce scoring mechanism, and constructed a new seller credit scoring model—a multi-factor correction model. Huang and Xiao[4] took Alibaba as an example to study how to use effective information to evaluate corporate credit as accurately as possible, build a credit model through AHP analytic hierarchy process, and use grey relational analysis. Chen[5] reconstructed the evaluation indexes of merchants by studying the defects in the merchants' scoring system under the existing C2C e-commerce model, and divided them into subjective indexes and objective indexes. Jiang and Yang[6] studied how to build a cross-border e-commerce credit evaluation system in the context of continuous data provision by big data. Yan and Chen[7,8] studied how to construct a credit evaluation

ISSN: 2414-1895 DOI: 10.6919/ICJE.202210_8(10).0033

system from the perspective of dynamic event transactions. Gu[9] analyzed the research on the credit problem of the seller service rating system from the perspective of game theory, gave Nash equilibrium solutions under each model, and finally proposed Taobao DSR (Detail Seller Rating) credit problem resolution mechanism in the three-party game between sellers, buyers and Taobao platform.

All standard paper components have been specified for three reasons: ease of use when formatting individual papers, automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and conformity of style throughout a conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type [3], within parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components [4], incorporating the applicable criteria that follow.

2. Data and Method

2.1 Data Collection

We collected the credit score ratings of 48 children's clothing Taobao stores, and divided the ratings into numbers. The higher the number, the lower the rating, as shown in Table 1. Then, using the random sampling method, 100 reviews are randomly selected from each store, and the favorable rate is calculated. In the selection of indicators, we have selected eight factors to consider: product quality, picture matching, reasonable pricing, service attitude, delivery speed, after-sales service, logistics speed, and packaging quality through research and reference to historical documents. The weights are quantified according to the favorable rate, as shown in Table 2, Table 3, and Table 4. Finally, we sort and number the stores from low to high according to the credit score. The data is shown in Table 5.

2.2 Data Processing

We use the range transformation method to standardize the data in Table 5. The transformation formula is as follows:

$$x^* = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Where x_{min} is the minimum value of the sample data, and x_{max} is the maximum value of the sample data.

2.3 Method

We take product quality, picture matching, reasonable pricing, service attitude, delivery speed, after-sales service, logistics speed, and packaging quality as credit scoring indicators, and use K-means clustering to perform category analysis on the data, where the clustering category k= 3.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202210_8(10).0033

Table 1. Store Credit Rating

Credit Score	Digital Rating
4-10	20
11-40	19
41-90	18
91-150	17
151-250	16
251-500	15
501-1000	14
1001-2000	13
2001-5000	12
5001-10000	11
10001-20000	10
20001-50000	9
50001-100000	8
100001-200000	7
200001-500000	6
500001-1000000	5
1000001-2000000	4
2000001-5000000	3
5000001-10000000	2
>10000000	1

Table 2. Product Quality, Picture Matching, Reasonable Pricing Score

Favorable Rate	Score
95%-100%	20
90%-95%	19
85%-90%	18
80%-85%	17
75%-80%	16
70%-75%	15
65%-70%	14
60%-65%	13
55%-60%	12
50%-55%	11
45%-50%	10
40%-45%	9
35%-40%	8
30%-35%	7
25%-30%	6
0%-25%	1-5

ISSN: 2414-1895 DOI: 10.6919/ICJE.202210_8(10).0033

Table 3. Logistics Speed, Delivery Speed and Service Attitude Score

Favorable Rate	Score
90%-100%	15
80%-90%	14
70%-80%	13
60%-70%	12
50%-60%	11
40%-50%	10
30%-40%	9
20%-30%	8
10%-20%	7
0%-10%	1-6

Table 4. After-sales Service and Packaging Quality Score

Favorable Rate	Score
90%-100%	10
80%-90%	9
70%-80%	8
60%-70%	7
50%-60%	6
40%-50%	5
30%-40%	4
20%-30%	3
10%-20%	2
0%-10%	1

ISSN: 2414-1895 DOI: 10.6919/ICJE.202210_8(10).0033

Table 5. Indicator Scores of 48 Stores

NO.	Digatal Rating	Product Quality	Reasonable Pricing	Picture Matching	Logistics Speed	Delivery Speed	Service Attitude	After-sales Service	Packaging Quality
1	1	15	15	15	14	12	13	9	9
2	2	16	16	16	10	15	13	8	10
3	2	14	16	15	9	14	12	8	9
4	3	16	16	16	11	14	13	8	7
5	4	17	16	16	12	13	13	8	8
6	4	16	7	11	13	14	14	7	8
7	5	17	18	17	8	14	11	8	5
8	5	15	15	15	12	14	13	8	8
9	5	13	13	13	13	13	13	7	7
10	5	15	15	15	13	13	13	7	9
11	6	15	15	15	12	13	13	6	7
12	6	19	19	19	14	14	14	10	10
13	7	17	17	17	12	12	12	9	8
14	7	18	17	17	14	12	13	9	8
15	8	14	15	14	10	13	12	8	8
16	8	18	9	13	13	13	13	9	9
17	8	14	16	15	11	11	11	8	9
18	9	19	18	18	14	14	14	9	9
19	10	17	18	17	12	12	12	8	9
20	10	13	14	13	10	13	12	6	7
21	10	17	17	17	12	13	13	8	7
22	10	18	19	18	13	14	14	9	8
23	10	19	19	19	14	13	14	9	9
24	11	18	19	18	13	13	13	8	9
25	11	18	17	17	14	13	14	9	9
26	11	19	16	17	14	13	14	9	9
27	11	18	18	18	13	12	13	9	9
28	12	18	15	16	14	13	14	9	9
29	12	17	19	18	14	13	14	9	9
30	12	16	16	16	14	13	14	9	9
31	13	15	15	15	12	14	13	8	8
32	14	20	20	20	15	14	15	10	10
33	15	17	19	18	14	13	14	7	9
34	15	17	15	16	10	13	12	8	9
35	15	19	18	18	13	14	14	8	9
36	15	18	18	18	13	13	13	9	9
37	16	19	10	14	14	12	13	10	9
38	16	17	16	16	14	13	14	9	9
39	16	16	14	15	14	13	14	8	9
40	16	13	15	14	10	12	11	7	6
41	16	18	18	18	14	14	14	8	9
42	17	18	19	18	14	12	13	9	9
43	17	18	19	18	14	13	14	9	9
44	18	19	20	19	15	13	14	10	10
45	19	18	18	18	14	13	14	9	9
46	19	19	19	19	14	10	12	8	9
47	20	20	20	20	15	15	15	10	10
48	20	20	20	20	14	13	14	9	10

ISSN: 2414-1895 DOI: 10.6919/ICJE.202210_8(10).0033

3. Result

3.1 Credit Rating Results of Platform

Table 6. Credit Rating Results of Taobao Platform

Clustering results		Number of stores
Category 1	1,2,3,4,5,6,7,8,9,10,11,12	12
Category 2	13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32	20
Category 3	33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48	16

The rating results of the platform are determined based on the long-term accumulated credit score of each store. The Taobao platform divides a total of 20 ratings for sellers. According to Table 1, we divided the 48 Taobao children's clothing sellers we studied into artificial credit rating, and divided them into 3 categories. Among them, those with a digital rating <= 6 are the first category, which represented a high credit rating; those with a digital rating of 7-14 are the second category means that the credit rating is average; the digital rating >= 15 is the third category, indicating that the credit rating is low. The specific classification is shown in Table 6.

3.2 Credit Rating Results of K-means Clustering

Table 7. Credit Rating Results of K-means Clustering

Clustering results		Number of stores
Category 1	6,16,37	3
Category 2	1,2,3,4,5,7,8,9,10,11,13,15,17,20,21,30,31,34,39,40	20
Category 3	12,14,18,19,22,23,24,25,26,27,28,29,32,33,35,36,38,41,42,41,44,45,46,47,48	25

Use Python to perform cluster analysis on the collected data of 48 children's clothing stores, where the number of clusters k is taken as 3, and the results are shown in Table 7.

4. Discussion and Conclusion

By comparing Table 6 and Table 7, it can be seen that there is a big difference between the credit rating results given by the platform and the K-means cluster analysis results, especially the stores numbered 1-12, the credit rating given by the platform is high, but in the cluster analysis results more than half of the stores are classified into the same category as the stores with a relatively average credit rating (numbered 13-32); for stores with an average credit rating on the platform (numbered 13-32), more than half are with a low credit rating (numbered 33- 48) in the same category.

The eight indicators we consider are all strongly related to credit evaluation. Therefore, compared with the credit rating given by the platform, the credit classification obtained through experiments is more scientific. Therefore, it can be concluded that Compared with simply defining the credit rating of e-commerce sellers based on the total score obtained by adding and subtracting points from good, medium and bad reviews, the credit rating of e-commerce sellers obtained by covering various indicators that consumers care about is better and has practical significance.

We put forward suggestions for the research: (1) The credit rating of Taobao sellers does not have an absolute guarantee, and consumers need to be cautious when making purchase decisions. (2) The e-commerce platform needs to improve the credit evaluation system, change the current model of single credit scoring standard and simple calculation, and improve the coverage of buyers' scoring content.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202210_8(10).0033

References

[1] Yang F, Wang A, Wu J, Tang L. Designing credit supervision mechanism in C2B2C e-commerce based on game theory[J]. Systems Engineering-Theory & Practice, 2017, 37(8):2102-2110.

[2] Cao J, Shi L, Zhang C, Zhao P. Research of improved algorithm in E-commerce credit evaluation mechanism [J]. Journal of Wuhan Institute of Metallurgical Management, 2018,28(03):10-12.

[3] Xu Q, Wang T, Jiang C, Yang S. The Creation and Validity Test of Seller's Credit Scoring Multifactorial Correction Model in E-commerce--A Case Study on Taobao[J]. Soft Science, 2017.

[4] Huang X, Xiao B. Research on enterprise e-commerce credit evaluation based on AHP grey relational analysis-taking Alibaba e-commerce enterprises as an example [J]. Shang, 2016(25): 130-132.

[5] Chen S. Research on the seller's credit evaluation system under the C2C e-commerce model [J]. Science and Technology Economic Market, 2013(09): 83-85.

[6] Jiang P, Yang L. Construction of cross-border e-commerce credit evaluation system under the background of big data [J]. Computer Knowledge and Technology, 2019, 15(18): 308-309.

[7] Yan B, Chen C. The seller's trust evaluation on electronic commerce based on event system theory[J]. Journal of Yanbian University(Natural Science Edition), 2017.

[8] Yan B. Construction of seller credit evaluation model of e-commerce platform from the perspective of dynamic transaction events [J]. E-Commerce, 2017(12):55-70.

[9] Gu H. Research on the credit problem of e-commerce platforms from the perspective of game theory--Taking Taobao seller service rating system as an example[J].China Business Theory,2016(13):66-69.