



Research on the influence of specific star rating of different products on customers' purchase decision

Di Jin, Kangtao Lv, Jiayi Lin

School of Science, Hangzhou Normal University, Hangzhou, Zhejiang, 311121, China

Abstract: With the rapid development and increasing popularity of online shopping, customers often collect products by comparing star rating and professional evaluation personnel. In order to obtain a higher purchase rate, merchants have to analyze the star rating, especially the influence of specific star rating (such as 1 star) on the purchase of users. First of all, through contingency table analysis and Pearson chi-square test for a specific comment text keywords and star rating on correlation analysis, we find the specific keywords of three products (such as a favorite and disappointed) is related to a particular star rating significant explanatory text reviews the consistency between the emotional tendencies and specific star rating, so users can overall attitude towards products with star rating similar to replace. Moreover, in order to analyze the effects of a particular star rating, we divide influence factors into the type of the user (normal users or professional assessment) and users' purchases (whether to buy the product), and different conditions of two kinds of influence factors are combined. Between the specific star rating and association rule mining, we find in the home appliance and articles for daily use, specific star rating (such as 1 star) and ordinary users and not buy the product has a strong affinity. Based on the above analysis, we combine the actual situation and put forward some Suggestions on the sales strategy.

Keywords: Specific star rating; Customer purchase decision; Contingency table analysis; Association rule mining.

1. Data preprocessing

First of all, we select three types of products: household appliances, baby products and household products for pre-processing, and eliminate redundant information and brands with too small sample size. Then, we conduct comparative analysis based on star rating and text comments to understand the overall situation of the three products and the specific situation of different brands in the products.

2. Analysis of significant correlations between specific keywords and star ratings in text comments

Words with specific emotional tendency often appear in the text comments of customers on products, and the emotional tendency of different review texts and the rating of different stars will influence our assessment of the final situation of products to a certain extent, thus affecting consumers' purchase. So is there a direct correlation between certain emotional tendencies and ratings in textual comments? To this end, we selected several representative keywords extracted from the comment text and analyzed whether there was a significant correlation between them and the corresponding ratings.

First of all, we divide the stars into high star rating (4 and 5 stars) and low star rating (1 and 2 stars). Then, we select two keywords with different emotional attitudes from the comment text, namely favorite and disappointed. We conduct contingency table analysis of the comment texts with these two keywords and their corresponding star ratings as well as Pearson's Chi-square test to analyze whether comments with different emotional attitudes are related to specific ratings [1].

Through analysis, we obtain the relationship between high and low star ratings and specific review texts as follows:

Table1 List of product A specific reviews and star ratings

Star ratings		Specific comments under low star ratings	
Low star ratings		favorite	disappointed
1 star	Percentage of rows	34.75%	65.25%
2 star	Percentage of rows	23.97%	76.03%
Star ratings		Specific comments under high star ratings	
High star ratings		favorite	disappointed
4 star	Percentage of rows	64.00%	36.00%
5 star	Percentage of rows	66.88%	21.25%
Pearson chi-square test for p values		0.014	

Table2 List of product B specific reviews and star ratings

Star ratings		Specific comments under low star ratings	
Low star ratings		favorite	disappointed
1 star	Percentage of rows	44.75%	55.25%
2 star	Percentage of rows	43.97%	56.03%
Star ratings		Specific comments under high star ratings	
High star ratings		favorite	disappointed
4 star	Percentage of rows	68.60%	31.40%
5 star	Percentage of rows	57.89%	42.11%
Pearson chi-square test for p values		0.038	

Table3 List of product C specific reviews and star ratings

Star ratings		Specific comments under low star ratings	
Low star ratings		favorite	disappointed
1 star	Percentage of rows	39.76%	60.24%
2 star	Percentage of rows	33.87%	66.13%
Star ratings		Specific comments under high star ratings	
High tar ratings		favorite	disappointed
4 star	Percentage of rows	71.60%	28.40%
5 star	Percentage of rows	67.79%	32.21%
Pearson chi-square test for p values		0.020	

From the contingency table analysis and Pearson chi-square test, we can find that between the review text of two keywords in three products and their corresponding star ratings, the p values < 0.05 . Reject the null hypothesis, it is believed that the different star ratings are significantly correlated with specific review texts, at the same time, from the percentage of specific reviews per star of each of the three products, the proportion of low star rating emotions tend to be "disappointed", while the high star rating emotions tend to be "favorite" .

3. Mining analysis of association rules for specific star rating

Through data access analysis, we find that whether reviews come from vine users and whether customers purchase the product have a certain impact on the review scores. Conversely, are average customers more likely to write some type of review to the unpurchased product after reading a series of particular star ratings (such as 1 star rating) ?

In order to analyze whether the specific star ratings can trigger more comments from ordinary users who have not bought the product at all, we conduct the association rule mining analysis on the three data sets^[2], and take the 1 star rating as an example. Firstly, we divided the mix of vine users and purchases into four categories:

Table4 Vine user and product purchase mix

Vine user	Product purchase	Combination type
Y	Y	A
Y	N	B
N	Y	C
N	N	D

Note: Below we are making professional assessors Vine users.

For the low star ratings (1 star ratings and 2 star ratings) of the three product data

sets, we calculate the proportion of users in the four combinations above, the details are as follows:

Table5 product A low star rating user information crosstab

User information			star_rating		summation
			1	2	
vine	N	Percentage of star_rating	76.8%	23.2%	100.0%
verified_purchase	Y	Percentage of star_rating	66.3%	44.7%	61.3%
	N		33.7%	55.3%	38.7%
combine	C	Percentage of star_rating	33.7%	55.3%	38.7%
	D		66.3%	44.7%	61.3%

Table6 product B low star rating user information crosstab

User information			star_rating		summation
			1	2	
vine	N	Percentage of star_rating	52.1%	47.9%	100.0%
verified_purchase	Y	Percentage of star_rating	15.4%	8.2%	12.0%
	N		84.6%	91.8%	88.0%
combine	C	Percentage of star_rating	84.6%	91.8%	88.0%
	D		15.4%	8.2%	12.0%

Table7 product C low star rating user information crosstab

User information			star_rating		summation
			1	2	
vine	Y	Percentage of star_rating	0.0%	0.6%	0.2%
	N		100.0%	99.4%	99.8%
verified_purchase	Y	Percentage of star_rating	72.2%	78.9%	74.8%
	N		27.8%	21.1%	25.2%
combine	B	Percentage of star_rating	0.0%	0.6%	0.2%

Next, we apply SPSS software to analyze the association rules between vine users, verified_purchase, combine and low star ratings. The analysis process is as follows:

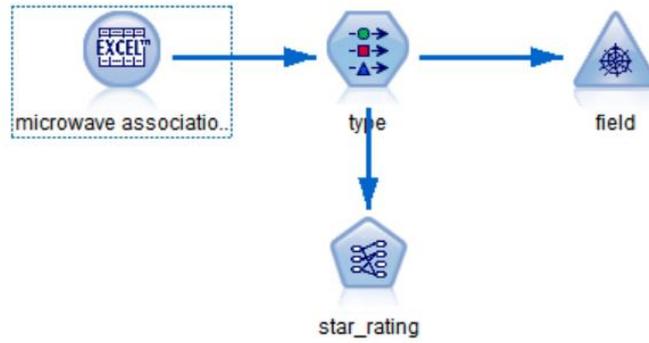


Figure1 Association rule analysis flow chart

By analyzing association rules, we obtain the associated network diagrams between vine users, verified_purchase, combine and low star ratings.

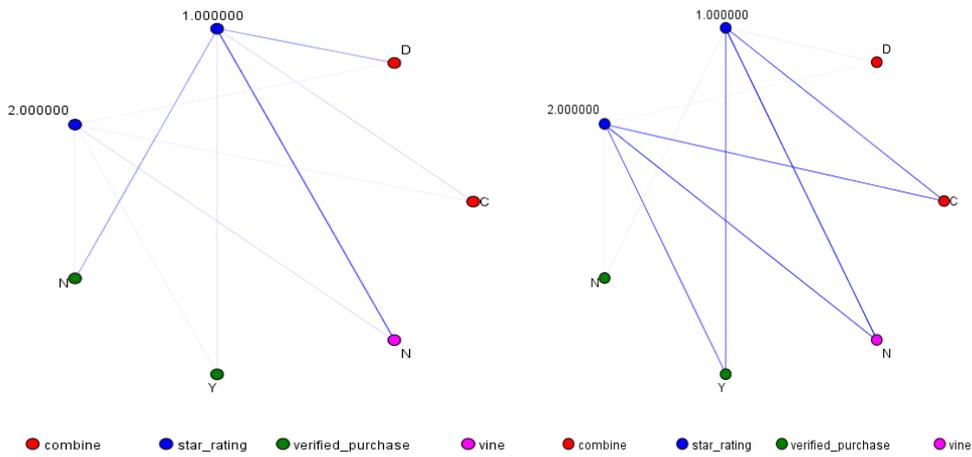


Figure2 Household appliances association network diagram

Figure3 Baby products association network diagram

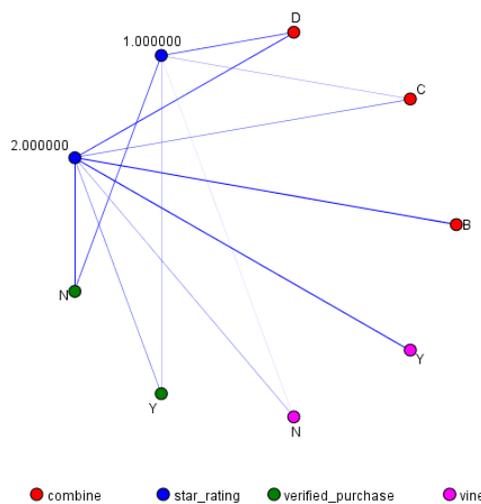


Figure4 Household products association network diagram

From the above three products associated network diagrams, most of the customers in the 1-star ratings of microwave ovens and hair dryers are average users or have

not bought the products or are the ordinary users who have not bought the products (combination type D). While from the Fig.1, the pacifiers 1-star evaluations mostly are ordinary customers or ordinary users who have bought the products (combination type C). Thus, microwave ovens and hair dryers in specific star ratings (such as 1 star ratings) do lead to more reviews, but the baby pacifiers in particular star ratings will not significantly cause more. So once the company place the microwave ovens and hair dryers on sale in the online marketplace, it is necessary to pay attention to the specific star ratings (such as 1 star) of the reviews, then adjust sales strategy in time.

4. The research conclusion

Through analysis, we find that there is a significant correlation between the specific comment keywords corresponding to each star of the three products and the specific star rating. Therefore, the analysis of users' emotional attitude towards the products can be approximately replaced by the star rating.

At the same time, in the household appliances and household products, specific star rating (such as 1 star) could indeed cause more comments, so related businesses in the two products should be pay attention to the comment of specific star rating, according to the comments on the content analysis the strengths and weaknesses of existing products in time, and looking for a professional evaluation personnel evaluation, give professional advice, thus to adjust product sales strategy, in order to gain more benefits.

References

- [1] liao qin, he zhifeng, Chen zhihong. Data mining and mathematical modeling. National defense industry press, 2010.02;
- [2] zhang yanping, zhou pengcheng. Analysis on the influence of online customer evaluation on online shopping decision [J]. Northern economics and trade, 2013 (07) : 71-72.