Volume: 14 Issue: 04 Pages: 5509-5516 (2023) ISSN: 0975-0290

Research on Influencing Factors of Garment Commodity Bad Reviews Based On LDA Theme Model

RenZhiyu
Email:1921746300@qq.com

ABSTRACT

The rise of e-commerce platform makes more consumers buy clothing goods online. However, due to the specific transaction process of online shopping, consumers are prone to generate various complaints and dissatisfaction, and thus conduct negative comments. Therefore, it is of great significance for apparel e-commerce to improve the quality of goods and services and improve user satisfaction to find out the reasons for consumers' dissatisfaction with clothing products from a large number of bad reviews. The article takes the poor evaluation data of affordable clothing brands and luxury clothing brands on the JD platform as the research object, conducts word frequency statistics and LDA theme model analysis on the poorly reviewed texts, and believes that the main factors affecting the poor evaluation of clothing products are Commodity Quality, Customer Service, Merchant Integrity, and Logistics Speed. In addition, by comparing and analyzing the difference evaluation data of the two types of clothing brands, it can be found that the most important influencing factor of affordable clothing brands is Commodity Quality, while the most important factor of the difference evaluation of luxury clothing brands is Customer Service.

Keywords -**bad reviews, influencing factors, text mining, LDA topic model.**

Date of Submission: Dec 2, 2022 Date of Acceptance: Dec 27, 2022

I. INTRODUCTION

With the advent of the digital era, the Internet has brought many conveniences to people's lives. Various e-commerce platforms, including Tmall, Taobao, JD, Amazon, etc., are emerging. Online shopping has become an important channel for people to purchase goods, and the number of online shoppers is growing. According to the 47th Statistical Report on the Development of China's Internet released by CNNIC, as of December 2020, the number of Internet users in China was 989 million, and the Internet penetration rate reached 70.4%, including 782 million online shopping users, an increase of 72.15 million over March 2020, accounting for 79.1% of the total Internet users[1]. It can be seen from the data released by the National Bureau of Statistics that the expenditure on clothing consumption per capita in China will increase year by year from 2018 to 2019[2]. The residents' consumption level continues to improve, which predicts that the proportion of consumers in clothing consumption will become larger and larger in the future.

At the same time, with the rise of the Web 2.0 concept characterized by advocating personalization, user interaction has been reflected[3]. Consumers' access to relevant product or service information has also changed from traditional word of mouth to online comments. The online shopping behavior of consumers is deeply affected by the network and online comments, and online shopping comments have become the main source of information for consumers to buy goods. However, due to various problems existing in the online trading process, the bad reviews of clothing products are increasing day by day. Most consumers make online shopping decisions by browsing product reviews, and too many bad reviews will reduce consumers' willingness to buy. In view of this, it is necessary to dig deeply into the influencing factors contained in the poor evaluation of clothing products. Clothing e-commerce merchants and online shopping platforms can improve their service quality based on these factors to improve their competitiveness.

II. LITERATURE REVIEW

Online comment is a kind of unstructured or semi-structured data with messy content but important value[4]. It refers to that after consumers purchase and use goods, they publish opinions about the goods or businesses to public groups or organizations on the Internet[5]. In the case of asymmetric online shopping information, it is the basis for consumers to make purchase decisions[6]. Online comments can be divided into positive comments and negative comments according to the polarity of comments (Valence)[7]. Negative comments (negative comments) are more useful than positive comments[8]. Consumers believe that negative online comments are more accurate and have higher diagnostic power than positive and neutral comments[9]. Therefore, negative comments have greater reference value for consumers' purchase decisions[10]. At present, the research on poor evaluation is mainly divided into two aspects:

(1) Explore the impact of negative comments on consumers' purchase intention. On the basis of the theory of conformity effect and attribution, Lu Haixia used binomial logistic regression to test the hypothesis and concluded that the number of negative comments and the quality of negative comments will have a negative impact on

consumers' shopping. In addition, if the quality of negative comments is high, the number of negative comments will reduce the possibility of consumers' buying[11]. Wang Yang and others believed that the negative emotional tendency contained in the negative comment information, the timeliness of the comment itself, and the risk perceived by potential consumers by the negative comment can all promote the negative comment to have a stronger negative effect, thus having a significant negative impact on the purchase intention[12]. Kim[13], based on the attribution emotion behavior model, studied the impact of attribution of customer perceived errors on customers' negative emotions and willingness to restructure.

(2) Mining the influencing factors hidden in the negative comments. Li Ming found out the reasons for consumers' dissatisfaction through a specific analysis of the content of book reviews in online bookstores, and classified the content into six main aspects. At the same time, he put forward his own suggestions on book publishing, sales and readers[14]. In addition, Bi Datian used the LDA theme model to mine and analyze the merchants' negative comments on Taobao, and obtained important factors that affect consumers' negative comments, thus providing countermeasures and suggestions for merchants' marketing activities[15]. Cao Jun took the user's negative comments on "Baidu takeout" and "Meituan takeout" as the research object, and by using Word2vec tools to cluster the characteristic vocabulary of negative comments, he explored the influencing factors of users' negative comments on takeout, and put forward suggestions on takeout businesses' operation according to the influencing factors [16]. Sun Qingyun crawled the comment data on the MOOC platform, extracted the theme of the comment text with negative emotional tendency by using the LDA model, explored the important factors that affect students' behavior of making negative comments, and gave the corresponding countermeasures and suggestions [17]. Mehta et al. studied the factors affecting hotel consumer satisfaction during the COVID-19 by improving the new measurement scale, and found 12 topics that were discussed most. The main reasons for dissatisfaction were staff, service, rooms, cleanliness, slowbooking speed, and hotel response to the epidemic[18].

Nowadays, although there are researches on the influencing factors of poor reviews on shopping platforms, there is a lack of in-depth exploration from the perspective of clothing commodities. Therefore, based on the current shopping environment and research status, this paper takes the clothing category as an example to conduct text mining and analysis of consumers' bad comments on the online shopping platform, identify and summarize the main factors of bad comments, which has reference value and guiding significance for the improvement of service quality of online shopping platforms and clothing e-commerce merchants.

III. DATA ACQUISITION OF ONLINE BAD REVIEWS OF APPAREL PRODUCTS

3.1 Data selection

Today, there are many e-commerce shopping platforms, such as JD, Taobao, Tmall and Maigou. Because the quantity and quality of online comments on these platforms are different, it is very important to choose the most appropriate platform among many shopping platforms. This paper finally chooses JD shopping platform as the source platform of clothing product's poor evaluation data, for the following reasons:

(1) By calculating the Baidu weight, Alexa ranking and PR value of shopping websites, we can see that JD ranks second in the ranking list of shopping websites (Table 1), which shows that JD website has high user loyalty and large daily average traffic, so we can get more reliable data.

Table1. Ranking summary of shopping websites

Websites	Baiduweight	Alexa ranking	PR value	Websitescore	Ranking
Alibaba	9	36	7	4738	1
Jingdong	9	20	7	4669	2
Maigou	9	3113	5	4554	3
Taobao	8	6	7	4304	4

(2)This paper takes online consumer clothing negative comments as the research data sample, and through observing the negative comments system of each shopping platform, it is found that JD has a relatively convenient negative comments interface, and it is easier to crawl data, which will make the negative comments data crawled in this paper more accurate.

Based on the above analysis, among the negative reviews of many shopping platforms, the online negative reviews of JD shopping platform have good typicality, and the influencing factors of consumer dissatisfaction can be excavated from the text reviews. Therefore, this paper selects the online negative comments of JD shopping platform as the research object.

3.2 Selection of clothing brands

For the research on the influencing factors of the poor evaluation of clothing products, it is necessary to consider not only the gender of consumers, but also the consumption capacity of consumers. Therefore, this paper selects 29 clothing commodity brands with different price points on the JD platform as the source of the data for the difference review, including 9 affordable brands and 20 luxury brands. See Table 2 and Table 3 for details. By selecting a brand with a large number of bad comments, we can get the reasons for bad comments with high quantity and quality, thus improving the differentiation of the analysis results.

Table 2. Affordable clothing brands

Sima	Dandy	Jeep
Taiping bird	Home of Hailan	Ngggn
Uniqlo	Zara	Pierre Cardin

Table 3. Luxury clothing brands

Gucci	Hazzys	Lime Flare	Versace
Ralph Lauren	Denton Hepburn	Aigle	Descent
Aec'Teryx	Kolon Sport	Jack Wolfskin	Kenzo
Montbell	Ami Paris	Canada Goose	Coach
TheNorthFace	Tory Burch	Ralph Lauren	Burberry

Volume: 14 Issue: 04 Pages: 5509-5516 (2023) ISSN: 0975-0290

3.3 Data crawling

According to the data needs of this article, we crawled the negative comments of typical clothing brands on the JD shopping platform, and manually cleaned them to get 9748 valid data. Python can collect data through various extension libraries to achieve easier and more convenient data crawling than other languages. Therefore, Python is selected to crawl the negative comments in this paper. The crawling process is shown in Figure 1.

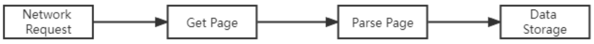

Figure 1. Data crawling flow chart

3.4 Word frequency statistics

After the word segmentation of the crawling poor comment text, it is very necessary to count the frequency of words appearing in the poor comment, from which we can roughly identify the influencing factors of consumers' poor comment willingness. This paper extracts the high frequency words of poor comment of the first 40 clothing products, and the word frequency statistics are shown in Table 4.

Table 4. Statistics of Word Frequency of Clothing Commodities

Word	quality	customer service	garbage	negative comment	business	fade	Fake goods	size	Bad	color
Frequency	1681	872	857	677	504	497	446	444	438	417
Word	express	work	return goods	Fabric	picture	shopping	Thread end	Worthless	delivery	brand
Frequency	407	392	385	342	323	318	282	270	261	239
Word	pilling	commodity	Very bad	disappointment	size	texture of material	logistics	No	service	Price
Frequency	229	225	221	207	199	197	191	187	187	183

From the above table, we can see that the high frequency words of bad evaluation of clothing commodity brands are mostly the words of clothing commodity quality, customer service, fake goods, merchants, express delivery, returns, etc. It can be inferred that the influencing factors of bad evaluation of clothing commodities are roughly related to quality, service, merchant behavior, and logistics.

However, word frequency statistics is used to count the number of times a word appears in the text. Therefore, the influencing factors of bad comments analyzed based on word frequency statistics are not comprehensive enough, and we need to use the LDA topic model to further explore, because the LDA topic model gives the topic of each document in the form of probability distribution, and uses it as the topic clustering or text classification, so as to obtain more accurate influencing factors of bad comments.

IV. ANALYSIS OF POOR APPAREL REVIEW FACTORS BASED ON LDA THEME MODEL

4.1 LDA thememodel

LDA theme model was proposed by David Blei et al. in 2003. Its full name is Latent Dirichlet Allocation theme model[19].It is an unsupervised topic machine learning algorithm based on the bag of words model (BOW)[20], which is usually used to find basic topics in the corpus. The main idea is to find a topic and its words from a document, and these words are distributed according to a certain probability. When the LDA topic model is applied in practice, the processed data is usually put into the model and the appropriate number of topics is set. Two output results can be obtained, namely, document topic distribution and topic word distribution. In the document topic distribution, lines represent each document and lists represent each topic. In the topic word distribution, lines represent each word in the corpus, and lists represent each topic.

4.2 Determination of the optimal number of topics

In this paper, the method of perplexity is used to select the optimal number of topics. In information theory, the degree of confusion is used to measure the quality of a probability distribution or probability model prediction sample. It can also be used to compare two probability distributions or probability models. [21]The formula is as follows:

$$\text{perplexity}\,(D_{\text{test}}) = \exp\left\{\frac{-\sum_{d=1}^{M}\log(p(w_d))}{\sum_{d=1}^{M}N_d}\right\} \quad (4.1)$$

Where, M -- the size of the test corpus

Nd -- Text size of the d (number of words)

Use Python language to process the confusion of text. The subject number confusion line is shown in Figure 2.

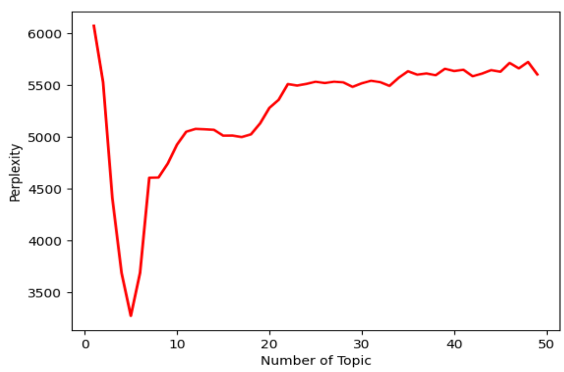

Figure 2. Topic number-confusion diagram

The stronger the model generation ability is, the smaller the perpetuity value is. In the above figure, with the increase of K value, the training confusion gradually decreases. When k belongs to (1,5), the curve drops sharply; When k is greater than 5, the curve rises; When k=5, the smaller the permeability value is, so 5 is the best value of K.

4.3 LDA topic modeling

This paper uses Python language to analyze LDA topic model, generates corpus after word segmentation of text data, builds dictionaries, calculates text vectors using TF-IDF, and finally performs LDA topic model fitting.

Given that the number of topics is 5, the LDA theme model is used to model the processed data, and the theme word distribution of the poor comments on clothing products is obtained, as shown in Table 5. The table not only contains the information about topics and words, but also shows the probability distribution of each word under each topic.

Table 5. Distribution of theme words of clothing brands

Topic1	Probability	Topic2	Probability	Topic3	Probability	Topic4	Probability	Topic 5	Probability
hair	0.185	degree	0.109	customer service	0.07	shopping	0.067	garbage	0.11
fall	0.172	return goods	0.086	thread end	0.061	fake goods	0.066	not good	0.081
pilling	0.057	logistics	0.068	service	0.055	experience	0.06	effect	0.066
really	0.038	that's slow	0.06	satisfied	0.053	retreat	0.053	pilling	0.056
disappointment	0.033	sell	0.056	after sales	0.049	thread end	0.051	fade	0.052
garbage	0.032	no	0.054	goods	0.046	buy	0.04	Wear	0.05
speechless	0.03	speechless	0.048	attitude	0.044	want to	0.038	thread end	0.046
too	0.029	regret	0.044	bad	0.036	physical store	0.029	fall	0.038
negative comment	0.028	express	0.042	service attitude	0.033	consumer	0.029	glimpse	0.034
quality	0.027	buy	0.034	retreat	0.032	fingding	0.025	size	0.031

By observing the data results, it can be found that the change of the probability value of the occurrence of the feature words of each topic presents an obvious long tail effect. The feature words with low tail probability value have very little impact on the semantic representation of the topic, so they can be ignored[17]. According to the distribution of theme words of clothing brands in Table 8, theme 1 and theme 5 are about the quality of clothing products, theme 2 is about logistics speed, theme 3 is about customer service, and theme 4 is about the credibility of online stores.

4.4 Visual analysis

In order to observe the topic-word distribution more intuitively, the results are analyzed visually. In this paper, the PyLDAvis package in Python language is used to visually analyze the results. Figure 3 shows the distribution of the subject-words in the negative comments on clothing commodities.

After analyzing the visualization results of pyLDAvis, the bubble distribution on the left indicates different topics, and the first 30 feature words in the topic are on the right. Since the number of topics set= 5, there are only 5 bubbles. The frequency of each topic is reflected by the size of the bubble, so the size and number of the bubble indicate the frequency of the topic. You can choose to view a specific topic by hovering the mouse over the bubble on the left. After selection, the right panel will display the vocabulary related to this topic. Parameters in the upper right corner λ It represents the relevance of a word to the subject. If λ the closer to 1, the more frequent words under the topic are more relevant to the topic; If λ the closer to 0, the more special and unique words under the theme are more relevant to the theme. This time, we are considering λ= 1.

In Figure 3, when the bubble number is 1, when the bubble number is 1, the correlation degree and the overall frequency of the words appearing in Topic 1 can be directly observed. Among them, the words with the highest overall frequency (the longest blue bar) are quality, and the words with the highest correlation (the red bar) in Topic 1 are wool, drop, fake, quality, etc. The above words are all related to the quality of goods.

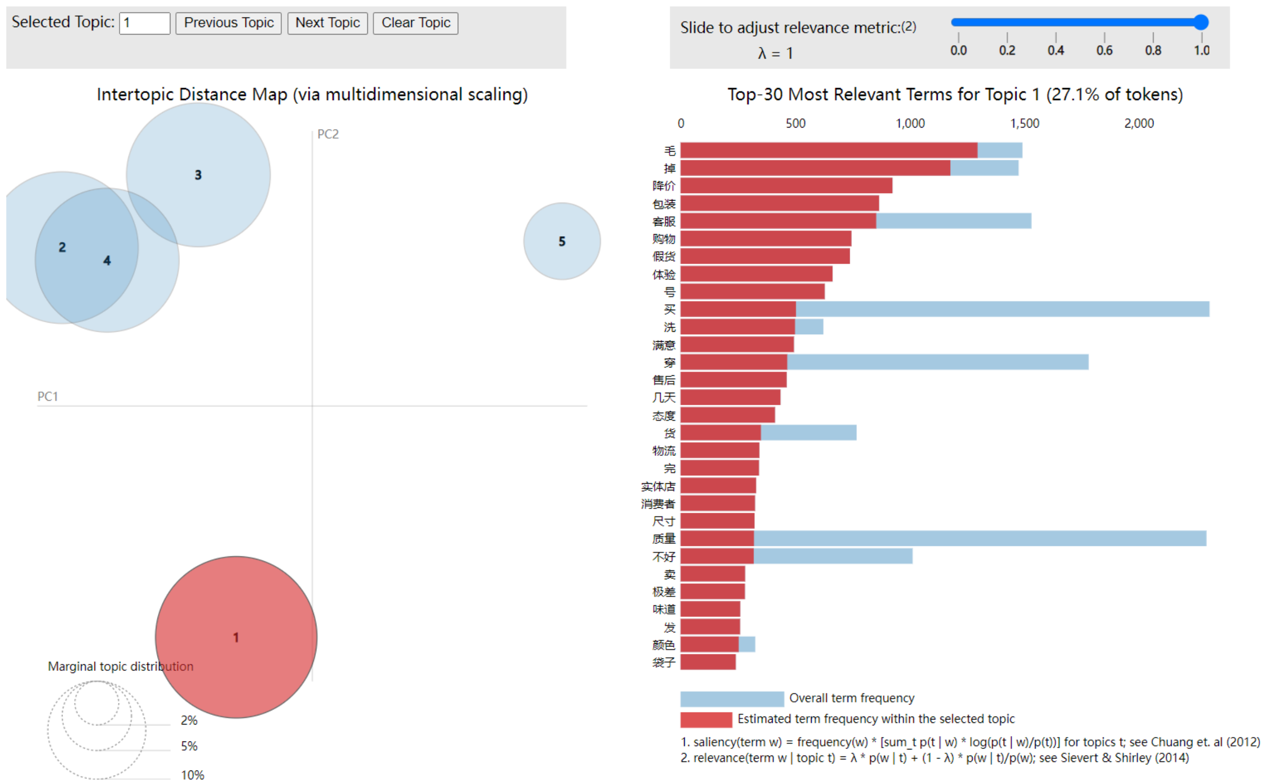

Figure 3. The subject-word visualization of clothing brand negative comments

In view of this, according to the above data and chart analysis, it can be concluded that the factors that consumers have bad comments on clothing products include product quality, customer service, business integrity and logistics speed.

V. Comparative Analysis of the Factors of Poor Evaluation Based on Clothing Types

5.1 Classification of clothing commodities

Clothing commodities can be classified into men's clothing, women's clothing and neutral clothing according to gender; According to age, it can be divided into children's clothing and adult clothing; Classified by price, it can be divided into affordable clothing and luxury clothing; According to the style, it can be divided into leisure clothing, professional clothing and sports clothing.

In this paper, we choose to classify clothing commodities according to price, and select luxury clothing brands and affordable clothing brands. From this perspective of classification, we can explore the reasons for consumers' negative comments on the two types of clothing brands, get the corresponding factors affecting the negative comments, and then put forward corresponding suggestions for improvement on the two types of clothing e-commerce, so as to improve consumer satisfaction.

5.2 Data analysis and results

(1)Word frequency statistics. Table 6 and Table 7 respectively show the first 30 high-frequency words of the two categories of clothing brands. It can be seen that the high-frequency word with the most negative comments on affordable clothing is quality, up to 1586 words. The words describing quality, including garbage, color fading and counterfeit goods, are also high in frequency; Customer service is the most frequently used word for luxury clothing, which also includes a series of words related to customer service, such as service and after-sales. Therefore, it can be inferred that the most important reason for poor evaluation of affordable clothing brands is quality, while the biggest factor for poor evaluation of luxury clothing brands is customer service.

Table 6. High frequency words of affordable clothing brands

Word	quality	garbage	customer service	negative comment	fade	fake goods	not good	colour	size	business
Frequency	1586	800	663	608	480	410	410	408	406	378
Word	feel	work	express	Fabric	return goods	not worth it	thread end	deliver goods	pilling	very bad

Volume: 14 Issue: 04 Pages: 5509-5516 (2023) ISSN: 0975-0290

Frequency	375	363	361	332	309	256	243	231	222	208
Word	size	texture of material	disappointment	logistics	no	bad	commodity	zipper	cloth	discrepancy
Frequency	198	194	184	181	179	167	160	152	151	150

Table 7. High frequency words of luxury clothing brands

Word	customer service	business	quality	return goods	negative comment	commodity	garbage	packing	fake goods	brand
Frequency	210	126	102	76	72	65	62	62	57	56
Word	price reduction	shopping	inner gall	price	experience	size	express	feel	service	work
Frequency	54	53	53	51	48	46	46	46	42	41
Word	for the first time	after sales	thread end	one piece	fake goods	tag	not good	exchange goods	reason	attitude
Frequency	39	39	39	37	37	34	33	33	32	31

(2)LDA theme model. Analyze the LDA topic mining algorithm for the poor comment text data, set the number of topics=10, and use the LDA topic model to model the processed data. Table 8 and Table 9 show the topic word distribution of the poor reviews of affordable clothing brands and luxury clothing brands. Due to too much data, only the first ten words of the first four topics are displayed here. The table not only contains the information of topics and words, but also shows the probability distribution of each word under the theme.

Table 8. Topic word distribution of affordable brands (partial)

Topic1	Probability	Topic2	Probability	Topic3	Probability	Topic4	Probability
too bad	0.106	fall	0.21	colour	0.092	not good	0.153
fake goods	0.071	hair	0.171	discrepancy	0.087	goods	0.112
logistics	0.067	wash	0.067	material object	0.077	bad	0.059
shopping	0.059	Wear	0.051	picture	0.076	service	0.059
disappointment	0.057	quality	0.034	return goods	0.075	business	0.057
satisfied	0.053	difference	0.027	a lot	0.067	slow	0.052
quality	0.051	say	0.025	code	0.056	hair	0.048
number	0.049	Hairiness	0.023	go through	0.053	one piece	0.036
feel	0.048	not good	0.023	consumer	0.043	misplaced	0.033
not worth it	0.04	deliver goods	0.018	don't buy	0.041	black	0.032

Table 9. Theme word distribution of luxury brands (part)

Topic1	Probability	Topic 2	Probability	Topic 3	Probability	Topic 4	Probability
business	0.152	after sales	0.094	customer service	0.159	service	0.145
fabric	0.042	return goods	0.063	ask	0.053	not good	0.088
not worth it	0.037	negative comment	0.056	gift	0.051	price reduction	0.052
a tiny bit	0.029	price difference	0.044	size	0.045	satisfied	0.038
brand	0.025	refund	0.04	thread end	0.034	this house	0.033
no way	0.024	retreat	0.035	say	0.031	a little	0.032
money	0.02	reason	0.032	design	0.026	service	0.032

						attitude	
deformation	0.015	deliver goods	0.027	store	0.024	just bought	0.03
Stall goods	0.014	say	0.026	range	0.022	flaw	0.028
size	0.014	brand	0.025	someday	0.018	customer service	0.025

According to the distribution of theme words of affordable brands in Table 8, theme one is about logistics, theme two is about quality, theme three is about after-sales, and theme four is about service. However, there are quality problems with each theme. It can be seen that among affordable clothing products, the biggest factor for consumers to make bad comments is the quality factor, including fake goods, hair loss, color fading, pilling, fabric and other problems.

Table 9 shows the theme word distribution of luxury brands. The first theme is mainly about merchants, the second theme is about commodity after-sales, the third theme is about customer service, and the fourth theme is about service. It can be seen from the above that the biggest factor for consumers to make bad comments is the service factor. Whether it is the business, after-sales, customer service or service, it is the internal service included in the goods and is an additional supplement to the goods. Of course, the distribution of subject words also includes quality, logistics and other issues.

In general, according to the above data analysis, we can draw a conclusion that the biggest factor for consumers to make a poor evaluation of affordable clothing brands is quality, and the biggest factor for consumers to make a poor evaluation of luxury clothing brands is service.

VI. CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusion

In this paper, we first use python language to capture the clothing bad comment data of the online platform, preprocess the data and make statistics of word frequency, then use LDA theme model to extract the theme of the cleaned bad comment text data, and visually process the extracted results. Finally, we draw the following conclusions: the reasons for consumers to make bad comments on clothing products are basically similar, including product quality, customer service, business integrity Logistics speed. However, there are differences in the main factors that consumers give bad comments on clothing products at different price levels. For affordable clothing brands, the main factor that consumers give bad comments is the quality of clothing products, while for luxury clothing brands, the main factor that consumers give bad comments is customer service.

6.2 Recommendations

In view of the above conclusions, this paper puts forward the following suggestions for service improvement of clothing e-commerce on the online shopping platform in order to improve consumer satisfaction.

(1) Improve the quality of clothing products

From the key words extraction and word frequency statistical analysis of clothing poor reviews, it can be seen that quality is the word with more frequency. Quality is also the main reason for poor evaluation of affordable clothing products. In real life, people pay special attention to the quality of clothing commodities, because quality is the core of clothing commodities. As long as the quality is up to standard, most consumers are likely to recognize the store, not only recommend it to people around them, but also become repeat customers of the store. Therefore, online merchants should first ensure the quality of their products, and improve the quality related problems raised by customers in their poor comments, such as discoloration, shedding, thread shedding, fabrics, etc.

(2) Improve service level

The words customer service, service, service attitude and after-sales appear more frequently in the word frequency statistics, and the biggest factor influencing the poor evaluation of luxury clothing brands is service. Therefore, businesses and households must be careful about this problem, and the quality of service will affect the future development of the store. Imagine the following picture: I bought an ill fitting dress on a shopping platform and wanted to return or replace it, but the customer service didn't seriously deal with the problem. This will definitely leave a negative impression on consumers, cause customers' dissatisfaction with the store, and lead to bad comments or even more serious consequences. Therefore, it is very important to improve the service level. The following measures can be taken: First, improve the service attitude of customer service, and do not be careless or indifferent; Secondly, we should deal with the problems of consumers in a timely manner, for example, we should immediately feed back the return and replacement problems they raised; Finally, protect customers' consumer rights and avoid the situation of "no responsibility once sold".

(3)Improve logistics efficiency

In the word frequency statistics and subject word extraction, there are more words related to express logistics. This shows that logistics is also a key issue. Many businesses think that it is OK to deliver the goods within the specified time. As for the reason of slow or broken delivery on the road, it has nothing to do with them. This idea is wrong and is easy to cause after-sales disputes. Therefore, businesses should face up to the logistics factor. We can take measures to reduce the impact of express factors, such as choosing a good express company and dealing with logistics related problems in a timely manner.

(4)Improve the credibility of merchants

The word frequency of false publicity and fakes is high. In addition, in luxury brands, it can be found that there is a high probability of poor reviews of fakes, and consumers often feel cheated due to false publicity by

Volume: 14 Issue: 04 Pages: 5509-5516 (2023) ISSN: 0975-0290

merchants. Therefore, businesses should improve their own credibility. When conducting discount activities, they should not break their promises, temporarily regret, or cheat consumers. This will reduce the trust of the store in the hearts of customers, and it is easy for this store to sell fake goods.

REFERENCES

[1] China Netcom . CNNIC released the 47th Statistical Report on China's Internet Development [EB/OL]2021-02-03[2022-03-12].http://www.cac.gov.cn/2021-02/03/c_1613923423079314.htm.

[2] Prospective Industry Research Institute. Analysis of the current situation and development trend of the clothing retail industry in 2019. The proportion of per capita clothing consumption expenditure of residents is decreasing year by year [EB/OL]2019-07-31[2022-03-12]. https: //www. qianzhan.com/analyst/detail/220/19073145402346.html.

[3] Wu Zhenghong. The trend of socialized, fan-based and content-based operation -- the new model of clothing marketing in the era of e-commerce 2.0 [J]. China Textile, 2016 (08): 98-100

[4] Fan Z P , Che Y J , Chen Z Y . Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis[J]. Journal of Business Research, 2017, 74(May):90-100.

[5] Hennig-ThurauT , Malthouse E C , Friege C , et al. The Impact of New Media on Customer Relationships[J]. Journal of Service Research, 2010, 26(3):311-330.

[6] Yin Guo-peng . What online reviews do consumers think are more useful—— The effect of social factors [J] Managing the World, 2012 (12): 10.

[7] Chevalier J A , Mayzlin D . The Effect of Word of Mouth on Sales: Online Book Reviews[J]. Journal of Marketing Research (JMR), 2006, 43(3):345-354.

[8] Casalo L V , Flavian C , Guinaliu M , et al. Avoiding the dark side of positive online consumer reviews: Enhancing reviews' usefulness for high risk-averse travelers[J]. Journal of Business Research, 2015, 68(9):1829-1835.

[9] Rohini A ,Burnkrant R E , Rao U H . Consumer Response to Negative Publicity: The Moderating Role of Commitment[J]. Journal of Marketing Research, 2000.

[10] Jabr W , Zheng Z E . Know Yourself and Know Your Enemy: An Analysis of Firm Recommendations and Consumer Reviews in a Competitive Environment[J]. Mis Quarterly, 2014, 38(3):635-654.

[11] Lu Hai-xia, Wu Xiao-ding, Su Li-xun. On The Effect of Negative Online Customer Review on Purchase Behavior [J] . Beijing Sociology, 2014 (05): 102-109.

[12] Wang Y, Wang WJ, Liu ZY. A study on the influence of online negative review information on potential consumers' purchase intention[J]. Intelligence Science, 2018, 36(10):8.

[13] Kim Y , Chang Y , Wong S F , et al. Customer attribution of service failure and its impact in social commerce environment[J]. International Journal of Electronic Customer Relationship Management, 2014, 8(1/2/3):136.

[14] Li M. Analysis of poor book reviews in online bookstores and its insights[J]. Library Journal, 2012, 32(09): 58-61+94.

[15] Bi DAT, Chu QH, Cao Ran. A study on the factors influencing consumers' willingness to give bad reviews based on text mining[J]. Intelligence Theory and Practice, 2020, 43(10): 137-143.

[16] Cao Jun,Wang Hu. Research on factors influencing poor reviews of takeaway users - based on text reviews and Word2vec [J]. Modern commerce industry,2017,38(2):55-56.

[17] Sun QY, Liu JH. A study on the influencing factors of MOOC poor evaluation intention based on text mining[J]. Statistics and Management,2021,36(9):105-112.

[18] Mehta M P , Kumar G , Ramkumar M . Customer expectations in the hotel industry during the COVID-19 pandemic: a global perspective using sentiment analysis[J]. Tourism Recreation Research, 2021:1-18.

[19] GangadharanV , Gupta D . Recognizing Named Entities in Agriculture Documents using LDA based Topic Modelling Techniques[J]. Procedia Computer Science, 2020, 171:1337-1345.

[20] Yang Fan. Research on Online Reviews Clustering Based on LDA Topic Model[D]. Dalian University of Technology.

[21] Guo T , Wu S Q , Shi Y C , et al. Research On Online Review Based On LDA Subject Model[J]. International Journal of Advanced Networking and Applications - IJANA, 2020(03).