Summer 7-5-2020

Brand User Attention Model Based on Online Text Reviews: An Empirical Study of New Energy Automobile Brands

Minjuan Chai
School of Economics and Management, Zhejiang Sci-tech University, Hangzhou 310018 China

Jianhong Luo
School of Economics and Management, Zhejiang Sci-tech University, Hangzhou 310018 China, jh_luo@zstu.edu.cn

Jiachen Lan
School of Economics and Management, Zhejiang Sci-tech University, Hangzhou 310018 China

Brand User Attention Model Based on Online Text Reviews: An Empirical Study of New Energy Automobile Brands

Minjuan Chai Jianhong Luo Jiachen Lan*
School of Economics and Management, Zhejiang Sci-tech University, Hangzhou 310018 China

Abstract: Accurately grasping the rules of user behavior and market changes and timely adjusting decisions and strategies are the ways for brand development and innovation. In this paper, we proposed a brand user attention model based on online text review analysis. First of all, we collected and preprocessed the user comment text from the online forum. Secondly, through the LDA topic model and LDAvis visual analysis, the potential topics of user reviews were extracted, and a multi-dimensional feature analysis model was constructed to reveal the users' attention features of brand products. Finally, took the new energy automobile brands as an example, the users' attention features for the different new energy automobile brands were explored and the empirical study was carried out. This study found that the brand user attention model based on online text analysis can effectively extract the characteristics that brand users care about, obtain valuable business insight, and provide support for managers' decision-making.

Keywords: brand user attention, text analysis, LDA topic model, empirical analysis, new energy automobiles

1. INTRODUCTION

Brands are consumers' evaluation and cognition of the enterprise and its production and sales of products or services and are an important basis for clarifying the company's positioning, formulating production and marketing strategies, and enhancing market competitive advantages. Analyzing the brand user's attention and discovering the potential consumption characteristics of the user will help the company to accurately grasp the user behavior and market changes in the fierce market competition, to make better decisions. To get a more comprehensive and in-depth understanding of the differences between the products and services of each brand and the importance that the user places on product features, the user's perception opinion of the product should be studied first.

Nowadays, the public is more inclined to use social media to exchange information about products and brands, listen to multiple views and opinions[1], so social media contains important business information that brands and sellers need. At present, there are mainly three types of social media platforms: online communities and forums, blogs, and social networks. They can allow the public to participate in brand interaction more extensively and deeply[2]. The information getting from online forums is more real, more timely, and more comprehensive. Analyzing the actual textual content of social media platform can better understand the consumers' interest for marketing[3]. Compared with conducting offline surveys, collecting review data from online forums is more controllable, which can not only clearly and accurately reflect the user's true perception and evaluation of the product, but also effectively reduce the cost of obtaining experimental data. Therefore, this article uses the comments of online forums to study users' opinions on brand products.

This paper proposed a brand user attention model and construction method based on text analysis to analyze the characteristics of users' attention to brand products. In this process, first, user review data was collected from online forums and preprocessed. Second, latent topic models were established through LDA (Latent Dirichlet Allocation, LDA) topic analysis and LDAvis visual analysis methods, and users' attention features were extracted from the latent topic models. Then the user's attention and its evolution characteristics

* Corresponding author: Jianhong Luo, email: jh_luo@zstu.edu.cn.

were analyzed through the multi-dimensional feature analysis model. Finally, the main new energy automobile brands were taken as an example of empirical analysis to verify the effectiveness of the method proposed in this paper.

2. RELATED WORKS

Users' attention to the brand is very important for brand innovation and further development of the industry, it reflects the effect of brand building. At present, the rapid development of information technology provides good support for the brand user attention analysis. More and more scholars have invested in the research of promoting brand building by collecting and analyzing brand users' comments. Tu Haili, Tang Xiaobo and Xie Li[4] built a user demand mining model based on the online comment data and got the satisfaction of users' demand for the brand product or service attributes, and then put forward the direction of brand improvement. Wei Jiahua, Fan Lili and Lin Xi[5] thought that private brand purchase intention research is an effective way to understand customers and product development, and built a scale of the influence of product characteristics on private brand purchase intention. Through empirical analysis, they found that product characteristics had a strong influence on purchase intention. Through these research results, it can be seen that brand research is of great significance for the further promotion of brand construction, including planning, design, publicity and management. Analyzing user needs by using brand user attention data is an important direction of brand study.

In the study of consumers' purchase behavior, most scholars adopt the method of the questionnaire. Zhao Yuxiang, Liu zhouying and Xu Weihan[6] designed the questionnaire according to the Kano model classification method, analyzed the data collected by the questionnaire, obtained the final decision weight ranking of the elements of the public science platform's game, and put forward corresponding countermeasures and suggestions according to the research results; When Gao Junbo, An Bowen and Wang Xiaofeng[7] studied the potential influential topics in online forums, they put forward a method of clustering influential words, which can extract important topic information timely and accurately for users and forum managers, and find hot issues in the forums. It can be seen that at present, most scholars choose to use the questionnaire survey method to obtain the user's ideas and opinions, and then study the consumer's purchase behavior. There are more and more studies at home and abroad to obtain consumer information through the social media text mining method. It has important theory and practice to use online comments of online forums to study consumer behavior and provide data-based decision support for managers.

In recent years, many scholars have analyzed online social media based on the LDA topic model and established relevant models, such as microblog, online forum, etc., and achieved certain results. In China, Luo Jianhong et al[8] used LDA topic model to explore the user's repost pattern of enterprise social media content; Chatao Chen et al[9] inferred forum topic and user interest by using LDA topic model; Wang Zhenhuang et al[10] proposed a microblog topic visualization system based on the topic model. This system uses a variety of interconnected views and rich interaction methods. It supports users to analyze and explore the results of the topic model, and can effectively help users analyze microblog topics; Xiong Huixiang et al[11] proposed a tag generation method based on the LDA topic model, which can describe the user's microblog features more accurately. In foreign countries, Malek hajjem et al[12] used LDA topic model and other methods to improve twitter topic analysis, which improved the measurement of topic consistency; Daniel Ramage et al[13] extended LDA topic model, and reflected the content of twitter from the content, style, status and social characteristics of posts, and gave two results of information consumption-oriented tasks. It can be seen that LDA is more practical and practical for topic analysis of online social media text, and LDA has great potential in this field. Many scholars have expanded their research based on the original LDA topic model. Therefore, this paper puts forward a brand user attention model based on LDA topic model, which analyzes user behavior from multiple perspectives, aiming to help brands, sellers and forum operators accurately understand the consumption characteristics of the users, summarize the connotation of users' attention about brand products, and make more accurate decisions.

3. BRAND USER ATTENTION MODEL BASED ON TEXT ANALYSIS

The comment text of the online forum truly reflects the users' evaluation of the product. The flow chart of the proposed brand user attention model based on text analysis is shown in Figure 1.

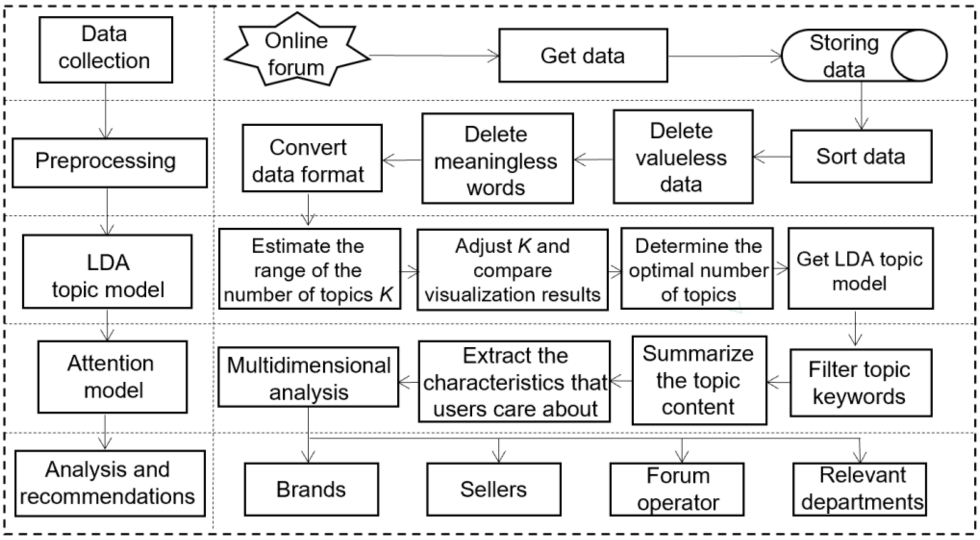

Figure 1. Brand User Attention Model

3.1 Data collection and preprocessing

According to the characteristics of the online forum and the specific needs of data, we collected online comment texts related to brands, including user name, user category, user registration time, comment time, title, comment content, gender, user location and other attributes.

In order to ensure the validity of the experimental data, the text collected from the forum needs to be preprocessed. Firstly, delete incomplete data. Secondly, delete meaningless words in the data set, such as stop words, links, special characters and other words that need to be ignored. Thirdly, transform the processed data into a data format that can be used for LDA topic modeling.

3.2 LDA topic modeling and visualization

After data collection and data preprocessing, we use the LDA topic model and LDAvis to analyze massive text data and find potential topics from the data set. The LDAvis tool developed by Carson Sievert and Kenneth E. Shirley is used to realize the interactive LDA topic model based on Web pages[14], so as to show the topical differences between brands intuitively.

LDA is a text topic generation model. It was first proposed by Blei et al in 2003[15]. LDA includes three levels: document, topic and word. Its basic idea is that the topic of a document is mixed. Each document represents a probability distribution of some topics, and each topic represents a probability distribution of some words.

LDA is also known as a three-layer Bayesian probability model, and its generation process can be represented in Figure 2. Among them, α and β represent the prior distribution hyperparameters of the document—topic probability θ and topic—word probability distribution φ respectively, Z represents a topic, W represents the word of the text, N represents the total number of words in each document, M represents the collection of the whole document, and K represents the total number of topics in the document[16].

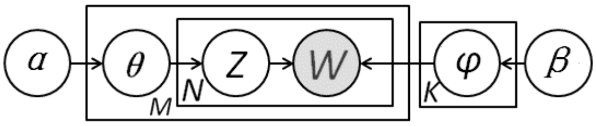

Figure 2. Bayesian network diagram of LDA model

In 2014, Carson Sievert and Kenneth E. Shirley put forward LDAvis, which is a method combining R language and D3. The LDAvis model can map the results of topic recognition to a two-dimensional space based on a multidimensional scaling algorithm, and then reveal the topic-topic, topic-word associations[17]. LDAvis thinks that the words selected according to the probability of the word item belonging to the topic are not the best result. At the same time, LDAvis puts forward a new basis for selecting the words representing the topic: the correlation between the words and the topic. According to the correlation between the words and the topic, the words belonging to the topic are reordered according to the weight, and the top-weighted words are selected to represent the topic[18].

In this paper, LDA is used to analyze the text data and get the result of the topic classification. Before determining the optimal number of topics, we estimate the range of topic number according to experience and classification, and continuously adjust the number of topics K at the speed of increment of 1. By observing the LDAvis visualization results of topics corresponding to each K value, the results were compared in terms of the practical significance, the topics simplification and the differences between topics, and finally, we got the optimal number of topics in the data set. Then an appropriate LDA topic model is obtained.

3.3 Brand user attention model

Users' cognition and evaluation of the brand mainly focus on the product characteristics of the brand, so scholars study the brand product characteristics to promote the in-depth analysis and improvement of the brand building. Li Tianjiao et al[19] think that the product features of pure electric automobiles mainly include the following: comfort performance, driving performance, appearance and interior decoration, safety performance, information interaction, energy consumption, charging convenience, etc. Therefore, to build a brand user attention model, we need to classify each topic content according to the product characteristics, and summarize the Characteristics that User Care about (CUC) from the resulted topics.

The topics in the LDA topic model can be numbered as T_1, T_2, ..., T_K. We use the topic model to get the distribution of each word in the topic, select the words with larger weight as the topic keywords, and summarize the content of each topic. According to the connotation of the product features, the meaning of each topic is extracted as the CUC of the product, with the numbers of C_1, C_2, ..., C_h.

After extracting brand CUC based on LDA topic model, we need to further study the collected data and analyze the characteristics of brand user attention and its evolution. The multi-dimensional analysis of data can find insights from multiple perspectives, make up for some aspects that cannot be involved in independent dimensions, to get more meaningful results. Therefore, three dimensions of the time, region and gender are analyzed respectively. By comparing research and analysis, we can further discover the changing law of consumers' attention to product characteristics.

3.3.1 CUC time dimension analysis

Using the "user comment time" attribute of the data, summarize the brand's comment months of each CUC, and take the number of reviewers per month as the measurement index. And the histogram is used to show the distribution of the number of reviewers in each CUC in each month of a year.

$$Q_{mc} = \sum_{i=1}^{n} P_{mci} \tag{1}$$

Among them, m is the month ($1 \le m \le 12$), c is one of the CUC, Q_{mc} is the total number of reviewers in c in the m month, P_{mci} is the number of reviewers of the i^{th} keyword in c in the m month.

3.3.2 CUC regional dimension analysis

Using the "user location" attribute, the number of reviewers in each province or municipality of each CUC is summed. Taking the number of reviewers obtained by statistics as the measurement index, select the top five provinces or municipalities directly under the central government with the number of reviewers ranking first, and draw the bar chart according to the category of CUC.

$$Q_{cr} = \sum_{i=1}^{n} P_{cri} \tag{2}$$

Among them, c is one of the CUC, r is the region, Q_{cr} is the total number of reviewers in r region in c, P_{cri} is the number of reviewers in r region of the i^{th} keyword in c.

3.3.3 CUC gender dimension analysis

By using the attribute of "user gender", the number of male reviewers and female reviewers under each CUC of each new energy automobile brand are counted, and the ratio of male to female is taken as the measurement variable of the attention value of gender dimension, and the bar chart is shown.

$$R = \frac{\sum_{i=1}^{n} P_{cim}}{\sum_{i=1}^{n} P_{cif}} \tag{3}$$

Among them, R is the proportion of men and women, c is one of the CUC, P_{cim} is the number of male reviewers of the i^{th} keyword in c, P_{cif} is the number of female reviewers of the i^{th} keyword in c.

3.4 Analysis and suggestions

Through the analysis of the time, region and gender dimensions of the brand CUC, we can find the evolution characteristics when the users' attention to the brand products. Combined with the characteristics and positioning of the brand itself, the brand user attention model was proposed in this paper. By analysis with this brand user attention model, it can put forward practical and effective suggestions for the brand, the seller, the forum operator, the policymaker and other relevant departments, and also provides suggestions for the brand and the industry development.

4. EMPIRICAL ANALYSIS

4.1 Data collection and preprocessing

Based on the above methods, this paper took the new energy automobile brands as an example, and made an empirical analysis of the brand user attention of Tesla Model X and BYD Yuan new energy automobiles. This paper selected the most influential forum in the field of automobile - "home of automobile" forum as the source of original data[20], and collected online comment data of Tesla Model X and BYD Yuan in the forum, including user name, user category, user registration time, comment time, title, comment content, gender, user location and other attributes. The collected data of BYD Yuan is from January 1, 2018 to December 31, 2018; As Tesla Model X has less data, in order to improve the reliability of experimental data, Tesla's data time range has been extended from January 1, 2017 to December 31, 2018. In this experiment, 11975 pieces of data about the Tesla Model X and 19902 pieces of data about the BYD Yuan were collected.

4.2 LDA topic modeling and visualization

As Section 3.2 described, after comparing the LDAvis visualization results of topics corresponding to each K value, we got the optimal number of topics is 4 for both two data sets. The results of the Tesla Model X (left) and the BYD Yuan (right) are shown in Figure 3.

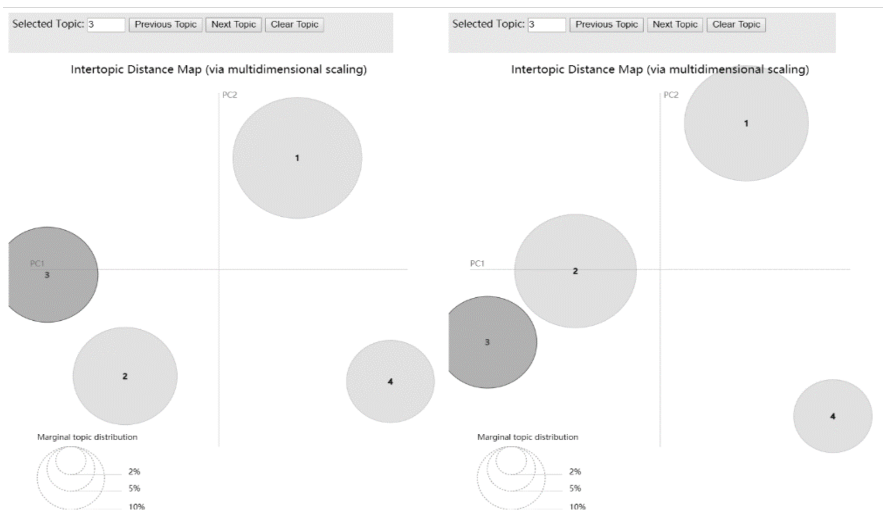

Figure 3. Tesla Model X (left) and BYD Yuan (right) topic model

In Figure 3, each topic is represented by one circle, the size of the circle represents the weight of each topic in the data set, and circles are sorted and numbered by size. We can click the circle and get the detail topic information including keywords list, etc.

4.3 Brand user attention model

According to the above results, the top ten words in each topic for Tesla Model X and BYD Yuan are list in Table 1 and Table 2.

For convenience, Tesla Model X is represented as T_X, BYD Yuan is represented as T_Y. T_{X1}, T_{X2}, T_{X3} and T_{X4} are respectively the four topic numbers of Tesla Model X, and T_{Y1}, T_{Y2}, T_{Y3} and T_{Y4} are respectively the four topic numbers of BYD Yuan. The topic meaning of each topic is summarized and refined according to the top ten weight words for each topic. Then each topic meaning is described as topic content extraction in Table 1 and Table 2.

Table 1. Tesla Model X topic content extraction

Brand name	Topic number	Top ten weight words for each topic (ordered by weights)	Topic content extraction
Tesla Model X （X）	T_{X1}	感觉(feel)、体验(experience)、喜欢(like)、座椅(seat)、科技(technology)、内饰(interior)、加速(accelerate)、鹰翼门(eagle wing gate)、感受(feeling)、模式(mode)	high-tech and experience
	T_{X2}	充电(charge)、驾驶(drive)、自动(automatic)、公里(kilometer)、辅助(auxiliary)、安装(installation)、功能(function)、朋友(friend)、交付(deliver)、销售(sales)	charging performance and endurance mileage
	T_{X3}	充电(charge)、公里(kilometer)、电量(electric quantity)、服务区(service area)、速度(speed)、出发(set off)、行程(itinerary)、里程(mileage)、到达(arrivals)、北京(Beijing)	self-driving travel and long-distance travel experience
	T_{X4}	轮毂(wheel hub)、细节(detail)、效果(effect)、碳纤维(carbon fiber)、原厂(original)、完美(perfect)、产品(product)、施工(construction)、卡钳(calipers)、改色(change color)	auto parts and manufacturing materials

Table.2 BYD Yuan topic content extraction

Brand name	Topic number	Top ten weight words for each topic (ordered by weights)	Topic content extraction
BYD Yuan（Y）	T_{Y1}	设计(design)、空间(space)、感觉(feel)、方向盘(steering wheel)、天窗(Sunroof)、后排(Back row)、不错(not bad)、功能(function)、座椅(seat)、全景(panoramic)	design sense and in-car features
	T_{Y2}	喜欢(like)、感觉(feel)、新能源(new energy)、老婆(wife)、提车(mention cars)、销售(sales)、孩子(child)、买车(buy a car)、时间(time)、家里(at home)	the joy of buying a car and benefits to family life
	T_{Y3}	生活(life)、喜欢(like)、风景(scenery)、出发(set off)、感觉(feel)、地方(local)、媳妇(wife)、时间(time)、景区(scenic)、拍照(Take a picture)	family self-driving travel experience
	T_{Y4}	充电(charge)、公里(kilometer)、续航(endurance)、电池(battery)、里程(mileage)、安装(installation)、电量(electric quantity)、服务区(service area)、km、车辆(automobile)	charging, endurance and infrastructure

As can be seen from the topic content extraction of Tesla Model X and BYD Yuan, the topics result of Tesla Model X and the topics of BYD Yuan show the characteristics of users care about on the two different new energy automobile brands. It depicts the elements of new energy automobiles that users may consider before and after purchase the automobiles. To further understand the characteristics or elements requirements of users, the new energy automobile CUCs framework is constructed according to the above topic modelling results. The CUCs framework includes six categories of CUC: technology interaction, energy consumption, long-distance performance, accessory material, appearance and interior decoration, and short-distance performance. Each CUC is numbered as C_1, C_2, ..., C_6. The proportion of each CUC for each brand is list in Table 3 and Table 4. It means the different proportions of the characteristics of each brand that users pay attention to.

Table 3. The CUC of Tesla Model X

CUC number	C_1	C_2	C_3	C_4
CUC	technology interaction	energy consumption	long-distance performance	the material of accessories
Proportion（%）	40	24	21	15
Corresponding topic number	T_{X1}	T_{X2}	T_{X3}	T_{X4}
Topic content extraction	high-tech and experience	charging performance and endurance mileage	self-driving travel and long-distance travel experience	auto parts and manufacturing materials

Table 4. The CUC of BYD Yuan

CUC number	C_5	C_6	C_3	C_2
CUC	exterior and interior	short-distance performance	long-distance performance	energy consumption
Proportion（%）	36	32	26	6
Corresponding topic number	T_{Y1}	T_{Y2}	T_{Y3}	T_{Y4}
Topic content extraction	design sense and in-car features	joy of buying a car and benefits to family life	family self-driving travel experience	charging, endurance and infrastructure

It is clear that although the number of topics of the two brands is the same, but the proportion of each CUC is different for the two brands. Therefore, the brand user attention model is effective to mine the CUC of different brands and can give more insight for the studied brands.

4.3.1 CUC time dimension analysis

According to formula (1), the "user comment time" attribute of the experimental data is analyzed, and

Figure 4 is the monthly number of reviewers of Tesla Model X.

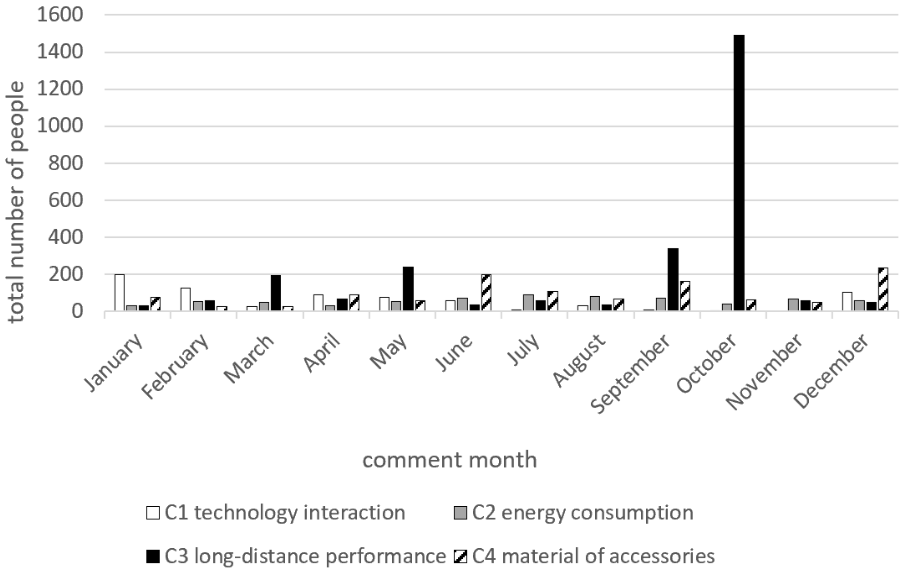

Figure 4. Monthly number of reviewers of Tesla Model X

It can be seen that users talk about technology interaction more in winter and the long-distance performance in spring. In early summer, early autumn and early winter, users pay more attention to the material of accessories of new energy automobiles. It is also clear that Tesla Model X users were active in October, and the number of reviewers on long-distance performance increased significantly, which may be related to seasons, holidays and national subsidy policies.

4.3.2 CUC regional dimension analysis

According to formula (2), we analyzed the "user location" attribute of the experimental data and got the CUC regional dimension analysis results of the two brands. Figure 5 and Figure 6 show the distribution of the top five provinces in terms of the number of reviewers of Tesla Model X and BYD Yuan towards different CUC.

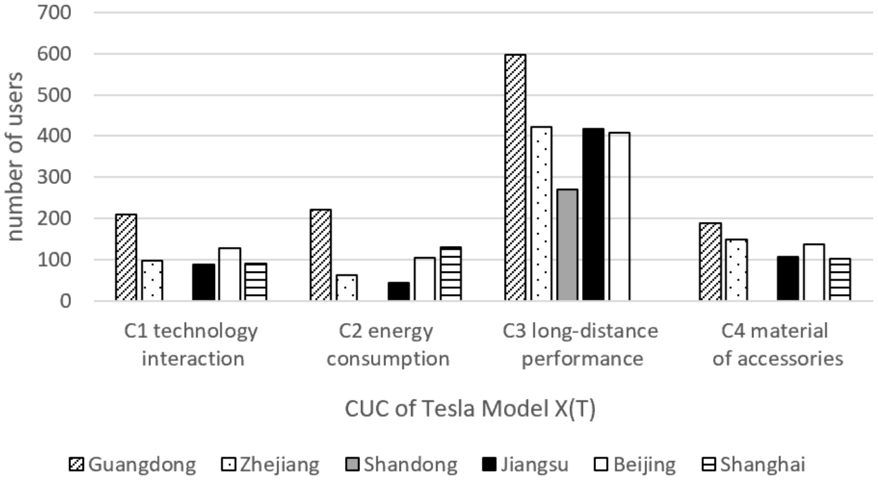

Figure 5. The distribution of the top five provinces of Tesla Model X

It is found that four CUCs including energy consumption, long-distance performance, the material of accessories and technology interaction have attracted more attention. Guangdong, Jiangsu, Zhejiang, Shandong, Shanghai and Beijing are the top five provinces or municipalities with the largest number of reviewers in each CUC. Tesla is targeted at high-income people, who also generally live in the above provinces or municipalities. First-tier cities usually limit the flow of fuel automobiles, so new energy automobiles have entered the choice of people living in big cities. The province with the largest number of reviewers under each CUC is Guangdong, so Guangdong has the largest sales market.

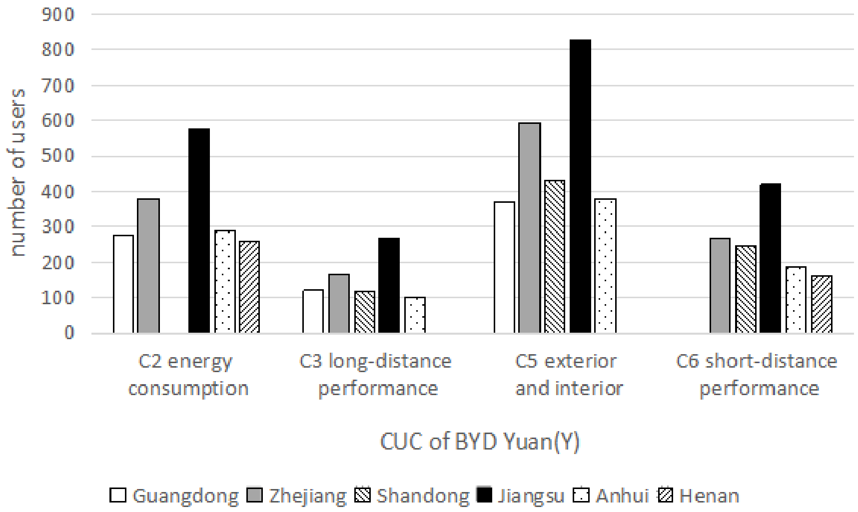

Figure 6. The distribution of the top five provinces of BYD Yuan

As can be seen from Figure 6, users of BYD Yuan mainly pay attention to energy consumption, exterior and interior, short-distance performance and long-distance performance. Jiangsu, Zhejiang, Shandong, Guangdong, Henan and Anhui are the top five provinces in terms of the number of users. According to the Statistics Bulletin released by each province in 2018, the experimental results are consistent with the actual economic strength of each province. Jiangsu and Zhejiang are the first and second provinces in each CUC, which shows that BYD Yuan has a large market in Jiangsu and Zhejiang.

4.3.3 CUC gender dimension analysis

Finally, according to the gender characteristics, the influencing factors of new energy automobile CUC are analyzed by formula (3). Figure 7 is a statistical chart of the gender ratio of each CUC of Tesla Model X and BYD Yuan.

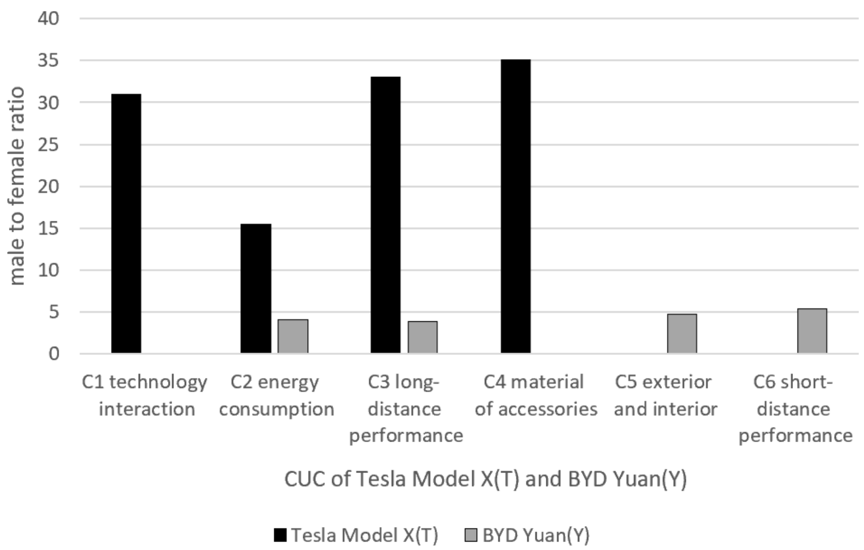

Figure 7. The gender ratio of CUC of the two different brands

From the figure, for Tesla Model X, the ratio of men to women for the CUC of the material of accessories is the largest, and that of energy consumption is the smallest. Relatively speaking, men generally know about cars and pay more attention to characteristics such as auto parts, car structure and materials. For BYD Yuan, the ratio of men and women for the CUC of long-distance performance is the smallest, while that of short-distance performance is the largest. Therefore, when buying a car or commenting online, women are more inclined to share a specific experience of self-driving with their family and friends, and pay more attention to the emotional experience of the car, while men are more concerned about the practicality of the car, which is in line with the characteristics that women are more emotional than men. At the same time, the gender ratio of Tesla Model X and BYD Yuan has a large difference. Therefore, the gender ratios of different brands are different, specifically, the gender ratio of high-price new energy automobiles is higher than that of low-price new energy automobiles.

4.4 Analysis and suggestions

4.4.1 Opinions on different brand user attention

This paper proposed a new energy automobile user attention model based on text analysis. The attention model is mined and constructed from the perspective of six CUCs. The different brand has different performance in the same dimension.

Specifically, the differences in users' attention of different brands are as follows:

- Compared with mid-low-end, the public is more concerned about the high-end new energy automobiles' technology interaction and the material of accessories, and more concerned about the short-distance performance and exterior interiors of mid lower end new energy automobiles.
- In terms of the time dimension, users of high-end new energy automobiles are active during holidays, such as national day, while the user behavior of mid-low-end new energy automobiles is mainly affected by the seasonal change and national subsidy policy.
- In the regional dimension, the characteristics of different brands show that the vast majority of their consumers are from economically developed regions, but some consumers of mid-low-end new energy automobiles are from middle economic level regions.
- In the gender dimension, the proportion of men and women in high-end new energy automobiles is far greater than that in mid-low-end new energy automobiles.

4.4.2 Suggestion

LDA topic model makes the topic of online reviews intuitive, and the extraction of new energy automobile user attention model shows the multidimensional distribution of attention, reveals the user's attention to new energy automobiles comprehensively. It is conducive to brand makers, sellers and discussion operators to grasp the mechanism of users' attention, to formulate their operation and management plans, and can also help related government departments to put forward effective policies and measures.

Brands should apply the known CUCs to production, sales and operation. First of all, through the user attention model, they can fully understand the user's attention to the product, promote product function optimization. Secondly, the revealed model meaning should be transformed into actual operating plans. According to the user's car purchase needs, they could produce ads that fit the public's psychology and convey the corporate spirit consistent with the public's emotions. Finally, expanding the sales channels is a good way to let the public more contact with the brand, understand the brand.

As a department that directly contacts with customers, the seller plays an important role in product introduction. After knowing the user's attention to new energy automobiles, we can establish sales guidelines, conduct training for sellers, and formulate detailed sales plans. When communicating with customers, the sellers shall focus on the points based on the characteristics of customers' time, region and gender, and the seller should meet the needs of customers for product characteristics and provide corresponding services to users.

Through the CUC multi-dimensional analysis of new energy automobiles, forum operators can first guide forum members to publish posts that cater to the user's attention, grasp hot topics and adjust topics in time.

Secondly, they could use the brand's competitive advantage to improve the core competitiveness of the forum, make the forum as broad as possible, and help more users solve problems.

Since 2010, in order to encourage the public to buy new energy automobiles, the state has continuously introduced policies to promote the development of the new energy automobile industry[21]. When formulating relevant policies of the new energy automobile industry, relevant departments can fully consider the potential consumption characteristics of customers, and put forward targeted optimization measures according to different users' attention characteristics under different time and space conditions, so that the new energy automobile industry can develop better and faster, and the energy structure can be further optimized and improved.

4.5 Analysis summary of new energy automobile user attention model

The user attention model was proposed and empirical studied with two new energy automobile brands of Tesla and BYD. Through the visual analysis and the topic modelling, the CUC framework was found, which includes six characteristics. Combined with multi-dimensional analysis, the user's attention to the product was deeply explored. The business insights obtained from the model and analysis have important value, which can help brand makers, sellers, forum operators and relevant departments to carry out brand developing and promote relevant work. The model proposed in this paper is applicable to the field of new energy automobiles, and can help to understand the market situation of different brands. The user attention model can be extended to more fields and more brands in the future, and can provide strong support for managers' decision-making, and promote brand construction of various industries.

5. CONCLUSION

With the rapid development of Internet economy, users' attention to product characteristics has become a necessary reference for brand growth and innovation. Obtaining and analyzing consumers' opinions and evaluations of the brand plays an important role in the development of the brand and even the whole industry. This paper proposed a brand user attention analysis method based on text analysis, which collects online comment texts of forums, carries out visual analysis with LDA topic modelling, extracts CUC based on the topic model, and constructs multi-dimensional characteristics analysis model, so as to put forward constructive suggestions. The empirical study shows that this method can get the characteristics that users pay attention to about brands, which can help decision-making for relevant departments.

Based on the text analysis, the brand user attention model makes users' cognition and evaluation of brand products or services clearer. This paper also analyzed the changing rules of users' attention, and put forward constructive suggestions for brand makers, sellers and relevant departments to help establish a brand-building method based on text analysis.

The brand user attention model based on text analysis proposed in this paper is still a basic research model and there is still some room for improvement. In the follow-up research, we can increase the dimension of attention analysis, explore the important business information hidden in CUC, and make the decision-making more comprehensive and accurate. When the experimental data is more abundant, it can analyze the situation and trend of the brand sales market, help risk management and make timely and accurate adjustments when necessary. At the same time, the brand user attention model is applied and improved in the actual application, so as to further improve the accuracy and effectiveness of decision support.

ACKNOWLEDGEMENT

This research was supported by the Natural Science Foundation of Zhejiang Province (No. Y18G020017, No. LZ18G010001), the National Nature Science Foundation of China (No. 71501172, No.71471165), and Zhejiang Social Science Planning Project (No. 17NDJC258YB).

REFERENCES

[1] Kilgour M, Sasser S L, Larke R. (2015). The social media transformation process: curating content into strategy. Corporate Communications: An International Journal, 20(3), 326-343.

[2] Zhang Shusen, Liang Xun, Qi Jinshan. (2017). A review on role identification methods in social networks. Chinese Journal of Computers, 40(3), 649-673(in Chinese)

[3] LUO Jianhong, PAN Xuwei, ZHU Xiyong. (2015). Identifying digital traces for business marketing through topic probabilistic model. Technology Analysis & Strategic Management, 27(10): 1176–1192

[4] Tu Haili, Tang Xiaobo, Xie Li. (2015). Research on user needs mining model based on online reviews. Journal of the China Society for Scientific and Technical Information, 34(10), 1088-1097(in Chinese)

[5] Wei Jiahua, Fan Lili, Lin Xi. (2013). An empirical study on the influence of product characteristics on private brand purchase intention. Commercial Times, (12), 34-36(in Chinese)

[6] Zhao Yuxiang, Liu Zhouying, Xu Weihan. (2019). Investigating gamification elements of citizen science platform based on Kano model. Library & Information, 39(03), 38-049(in Chinese)

[7] Gao Junbo, An Bowen, Wang Xiaofeng. (2008). Study on potential influence topic in on-line community. Journal of Computer Applications, 28(1), 140-142(in Chinese)

[8] Luo Jianhong, Pan Xuwei, Zhu Xiyong. (2017). Discovery of repost patterns by topic analysis in enterprise social networking. Aslib Journal of Information Management, 69(2), 158-173

[9] Chen Chaotao, Ren Jiangtao. (2017). Forum latent Dirichlet allocation for user interest discovery. Knowledge-Based Systems, 126, 1-7

[10] Wang Zhenhuang, Chen Siming, Yuan Xiaoru. (2018). Visual analysis for microblog topic modeling. Journal of Software, 29(4), 1115-1130(in Chinese)

[11] Xiong Huixiang, Ye Jiaxin. (2018). Microblog tags generation based on LDA theme model. Information Science, 36, 7-12(in Chinese)

[12] Hajjem M, Latiri C. (2017). Combining IR and LDA topic modeling for filtering microblogs. Procedia Computer Science, 112, 761-770

[13] Ramage D, Dumais S, Liebling D. (2010, May). Characterizing microblogs with topic models. In Fourth international AAAI conference on weblogs and social media

[14] Sievert C, Shirley K. (2014, June). LDAvis: A method for visualizing and interpreting topics. In Proceedings of the workshop on interactive language learning, visualization, and interfaces (pp. 63-70)

[15] Luo Jianhong, Pan Xuwei, Wang Shixiong, Huang Yujing. (2019). Identifying target audience on enterprise social network. Industrial management & data systems, 119(1), 111-128

[16] Wu Kai, Wang Ying. (2015). The initial exploration on microblogger knowledge discovery with user mention relations. Library & Information, 159(02), 123-127(in Chinese)

[17] Liu Ziqiang, Xu Haiyun, Yue Lixin, Fang Shu. (2019). Research on core technology topic identification based on Chunk-LDAvi. Library and Information Service, 63(9), 73-84(in Chinese)

[18] Sun Guochao, Xu Shuo, Qiao Xiaodong. (2015). Review on visualization of topic models. Technology Intelligence Engineering, 1(6), 51-61(in Chinese)

[19] Li Tianjiao, Ma Jun. (2017). Positioning method for research on characteristics of pure electric vehicle products. Auto Industry Research, (1), 15-21(in Chinese)

[20] Yang Jie, Nice. (2016). Research on factors affecting consumer purchase of new energy vehicles. Shanghai Management Science, (3), 56-63(in Chinese)

[21] Zhang Lei, Qin Quande. (2018). China's new energy vehicle policies: Evolution, comparison and recommendation. Transportation Research Part A: Policy and Practice, 110, 57-72