Research on design, production and sales factors of mobile phone based on online review analysis

Tianlu Zhu¹*, Zixuan Wang²

¹Department of mechanical engineering, north China electric power university, Baoding, Hebei province, 071000, China
²Department of economic management, north China electric power university, Baoding, Hebei province, 071000, China

*Corresponding author’s e-mail: 1939378494@qq.com

Abstract. Based on the relationship between mobile online review elements and mobile product stages, this paper extracted the online review data of four mobile phones on the online shopping platform. The SPSS logarithm was used for descriptive statistics and cluster analysis, and an indicator system of purchase decision factors based on mobile phone design, production and sales was established. Finally, relevant countermeasures and Suggestions are put forward for enterprises and shopping platform from the perspectives of mobile phone design, production and sales, by analyzing the index system and the consumer preference of the target user group.

1. Introduction
In recent years, with the effective combination of Internet technology and statistical analysis tools, mobile online reviews have become an important source of information for enterprises to optimize the design, production and sales of commodities. Such a large number of text messages containing consumer experience are of rich expressive significance, and because of their open group participation, they often have a strong influence on the word-of-mouth communication and sales volume of products and have a high research value. In this context, how to effectively use mobile online reviews to accurately extract information and then help mobile phone manufacturers to improve mobile phone design, standardize production process, formulate sales strategy, and finally improve the core competitiveness of enterprises has become an important topic facing the mobile phone industry.

2. Literature review
Generally speaking, the sample data of online reviews can present the required content according to the needs of researchers after the integration of information extraction and the processing of emotional analysis. Guo, Y (2016) used the potential Dirichlet analysis data mining method to extract valuable comments from visitors to quantify the hotel rating provided by customers and identify the key dimensions of customer service expressed by hotel tourists [1]. Pang B et al. (2002) compared and verified the applicability of SVM based on machine learning by taking the number of film reviews as a sample [2]. From a visual perspective, scholars such as Schtiuel Uckert, M (2014) reviewed and analyzed articles related to online reviews of tourism and hotels published in academic journals from 2004 to 2013, and divided their research topics into five themes: online purchase satisfaction, online management, online reviews, opinion mining and review motivation [3]. In terms of emotion analysis, Wei Shi (2013) completed the classification of sentence elements by building a spatial model based on
fuzzy emotion ontology, and further proposed an emotion analysis formula and verified its effectiveness[4]. Chunlin Qian et al (2019) established an improved collaborative filtering recommendation model based on the uncertainty theory for the uncertainty of relationship evaluation, and made innovations in the research algorithm [5-6]. To sum up, this paper aimed at the development strategy of the mobile phone industry, used the document frequency method to extract data, used emotional analysis to identify the needs, carried out cluster analysis on the types of consumer groups, and studied the characteristics of different consumer groups, and put forward Suggestions for improvement in the mobile phone industry from the three aspects of design, production and sales.

3. Research model and method

3.1. Sentiment analysis
There are usually fixed scoring rules for emotional analysis. In this study, 5-point likert scale was used to score negative sentences, and 1 to 5 points were assigned, corresponding to a gradually increasing degree of satisfaction.

3.2. Factor analysis
In order to improve the efficiency, the dimensionality was firstly reduced through factor analysis. The steps of factor analysis are as follows:
- Bartlett Test of Sphericity and kaiser-meyer-olkin were selected to Test the suitability of the comment data for factor analysis. The larger the bartlett sphericity test value and KMO is, the more suitable the variable is for factor analysis.
- Construct factor variables.
  \[ y_k = u_{1k}X_1 + u_{2k}X_2 + \cdots + u_{kk}X_k \]  
  \[ y_k \] represents the KTH principal component of the original variable, \( k = 1,2,\cdots,p \).
- Name and explain factor variables and calculate the factor score. The factor variable is expressed as a linear combination of the original variable by the formula, it is:
  \[ F_i = \beta_{i1}X_1 + \beta_{i2}X_2 + \cdots + \beta_{ip}X_p \]  
  And \( i = 1,2,\cdots,m \).

3.3. cluster analysis
Variables or samples are classified according to the degree of affinity between variables or samples, and the variables or samples with the highest similarity are gathered into the same class. The research steps are as follows:
- Standardize data
- Squared Euclidean distance was used to measure the degree of affinity between clustering objects:
  \[ \text{SEUCLID} = \sum_k (x_i - y_i)^2 \]  
  \( x_i \) represents the value of the first sample on the ith variable; \( y_i \) represents the value of the second sample in the ith variable.
- Clustering. Merge each object point and the nearest point to it into one class, repeating until all the points are in one class.

4. Experiment and result analysis
In this paper, online reviews of four mobile phones including XIAOMI 8, SAMSUNG GALAX S8, iPhone XS and HUAWEI Mate20 were selected as research samples, and data were selected from official websites of each mobile phone brand and two shopping platforms with the largest number of users -- TMALL and JD. Select the latest online comments from January 2019 to April 2019, eliminate irrelevant comments, and select 600 valid data.
4.1. data processing and conversion

According to the online comment data collected, the comments of ordinary consumers are relatively short and cannot involve the synthesis of these 17 types of evaluation indexes. Therefore, there are many missing values in the process of rating the comments. In this paper, linear interpolation method is adopted to supplement and improve the missing values for a series of analysis such as factor analysis and clustering analysis.

4.2. factor analysis based on principal component analysis

In this paper, SPSS17.0 was used to process the original data, and the results were shown in table 1. It was found that the KMO value was 0.809, indicating that the original variable was suitable for factor analysis according to the test standard. The statistical value of the Bartlett test was 136 and the adjoint value (sig value) was 0.000, far less than the significance level of 0.05. Therefore, the null hypothesis was rejected and the correlation between the original variables was considered suitable for factor analysis.

| Table 1. KMO and Bartlett’s Test |
|----------------------------------|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | .809 |
| Bartlett’s Test of Sphericity | Approx. Chi-Square | 3722.948 |
| df | 136 |
| Sig. | .000 |

In this paper, a total of three factors with eigenvalues greater than 1 were extracted. According to the output results of SPSS software, the information of common factors was obtained as shown in table 2. The table includes the initial eigenvalues, explanatory variances and cumulative explanatory variances of each factor. In this paper, three common factors were extracted. The cumulative explanatory variances were 85.781%, that is, the explanatory rates of these three factors to the original variables reached 85.781%, indicating that these three factors could reflect most of the original information.

| Table 2. Total Variance Explained |
|-----------------------------------|
| Initial Eigenvalues | Extraction Sums of Squared Loadings | Rotation Sums of Squared Loadings |
| Component | % of Total Variance | Cumulative % | Component | % of Total Variance | Cumulative % | Component | % of Total Variance | Cumulative % |
| 1 | 3.573 | 39.691 | 39.691 | 3.573 | 39.691 | 39.691 | 2.851 | 31.671 | 31.671 |
| 2 | 2.534 | 28.149 | 67.840 | 2.534 | 28.149 | 67.840 | 2.834 | 31.482 | 63.153 |
| 3 | 1.615 | 17.940 | 85.781 | 1.615 | 17.940 | 85.781 | 2.037 | 22.628 | 85.781 |

Table 3 is the factor load matrix after rotation, which reflects the load values of different categories in different factors. It can be seen from table 3 that factors of technology, material, quality and origin have higher loading on factor f1. The factors of system, screen, function, battery, software, appearance and trait have higher load on factor f2. Factors of product, sales, price, logistics, service and after-sales service have higher load on factor f3. According to the factor loading of different factors, the three factors are preliminarily named as technology quality factor, software configuration factor and product service factor, which are the clustering number.
### Table 3. Component Matrix

| Factor | Component |
|--------|-----------|
| 1      | Systematic Score (0.690) Screen Score (0.683) Functional Score (0.672) Battery Score (0.546) Software Score (0.733) Appearance Score (0.589) Trait Score (0.609) |
| 2      | Technical Score (0.849) Material Score (0.775) Quality Score (0.821) Origin Score (0.758) |
| 3      | Product Category Score (0.503) Sales Score (0.634) Price Score (0.838) Logistics Score (0.209) Service Score (0.821) After-sales Score (0.297) |

#### 4.3. Cluster analysis

In this study, variables were taken as objects for clustering and r-type clustering analysis was performed. Mobile phone purchase decision factors is obtained by system clustering, the variables of the clustering tree (figure 1).

#### 4.4. Establishment of index system of mobile phone purchase decision factors

According to the market segmentation theory and the above factor analysis and clustering analysis results, 17 types of evaluation indexes can be divided into three aspects: design, production and sales. AHP can be used to construct the factor index system of mobile phone purchase decision, as shown in figure 2.
From the perspective of the indicator system, the enterprise and shopping platform can attract users from different perspectives and improve the competitiveness of enterprises by targeting at three types of target user groups. For users who prefer design elements, improving the design of mobile phone products is the first priority. The specific measures can be realized through product diversification and product differentiation strategy. For example, the screen can be designed as comprehensive screen, water drop screen, bangs screen, etc., to meet the different preferences of users [7].

Factors of production quality preference of consumers are more in love with "black" science and technology, use of mobile phone products has the strict request, therefore, to keep improving in production, with the introduction of new technology, standardization of production processes, to the whole process of total quality management of products, mobile phone products to high-tech, standardization, precise positioning requirement.

The strategy for those who prefer sales factors should start from the perspective of marketing. Design factors can be summarized in 4P 'S strategies, including product strategy, price strategy, distribution strategy, and promotion strategy, improve product performance, adjust the right price, improve the quality of service, online marketing, promotion and personnel promotion, the combination of overall improve the competitiveness of the enterprises.

5. The conclusion
It has become an important task for the mobile phone industry to make full use of mobile online reviews to accurately extract information and improve the core competitiveness of enterprises. The shortcomings of this paper are as follows: only four mobile phone products with a high market share are selected without considering the sales reviews of some niche brands. In the future, samples need to be further expanded to reflect the more real situation of the mobile phone industry.

References
[1] Guo Y, Barnes S J, Jia Q. Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation[J]. Tourism Management, 2017, 59:467-483.
[2] Pang B, Lee L, Vaithyanathan S. Thumbs Up. Sentiment Classification Using Machine Learning Techniques [C]//Proceedings of the 2002 ACL Conference on Empirical Methods in Natural Language Processing. 2002: 79-86.
[3] Schtiuel Uckert, M. Hospitality And Tourism Online Reviews: Recent Trends And Future Directions. [J]. Journal Of Travel & Tourism Marketing, 2014,32:608-621
[4] Shi Wei, Hongwei Wang, Shaoyi He. Semantic analysis of Chinese online comments [J]. Journal of
Information Science And Technology, 2013, 32(08): 860-867.

[5] Chunlin Qian, Xingfang Zhang, Lihua Sun. Improved collaborative filtering recommendation model based on online comment sentiment analysis [J]. Journal of Shandong university (Engineering Science edition), 2019, 49(01): 47-54.

[6] Su Jiafu; Yang Yu; Yang Tao. Measuring knowledge diffusion efficiency in R&D networks. Knowledge management research and practice. 2018, 16(2): 208-219.

[7] Yang Jie, Su Jiafu; Song Lijun. Selection of Manufacturing Enterprise Innovation Design Project Based on Consumer’s Green Preferences. Sustainability 2019, 11(5), 1375