Research Background

With technological advancement and social development, the use of the Internet has gradually become widespread around the world. The Internet now has developed to provide a platform for uses ranging from completing daily demands to conducting research. With respect to completing daily demands, the Internet has provided possibility for online shopping. Given the fact that people nowadays have overwhelmed schedules and heavy workloads due to the fast pace of our society, more and more people prefer to shop online instead of going to department stores and supermarkets in person. However, online shopping possesses deficiencies and inconvenience despite its advantages. Shortcomings like being unable to see the products in person have become the greatest worry among customers as they may risk purchasing low-quality products due to lack of key information presented online. On the other hand, producers also suffer from the worry of selling their products. As a result, determining what characteristics of products are crucial to sales volume is the main challenge for online companies. To solve this problem, we choose a specific kind of product——cellphone——to analyze what kinds of cellphones have the highest sale volume.

Current Research Status

Current Research mainly focuses on several key factors which are considered to influence sales volume. Jie Zhang and Jianan Zhong [Jie Zhang, Jianan Zhong. Research of promotion’s influence to customers’ purchasing behaviors.]conducted research to analyze how sale promotion influences customers’ minds and predict the purchase choices of customers. Gang Du and Zhenyu Huang [Gang Du, Zhenyu Huang. Prediction of customers’ purchasing behaviors in the Big Data environment.] employed the Teradata platform to build decision-making tree model to predict customers’ purchasing behaviors, further improving the efficiency and accuracy of prediction. Zhanbo Zhao, Luping Sun, and Meng Sun [Zhanbo Zhao, Luping Sun, and Meng Sun. Research of comparison between factors in C2C influencing page view and sales volume.] discovered that factors influencing page view and sales volume are substantially different. To be more specific, price, scale, reputation and insurance have significant influence on page view and sales volume. Zhihai Hu, Dandan Zhao and Yi Zhang [Zhihai Hu, Dandan Zhao and Yi Zhang. Empirical research of online comments’ influence on sales volume.] employed sales of skin care products on Taobao as an example to analyze the influence of online comments to sales volume. The aforementioned researches mainly explored certain factors influencing sales volume, but lacked generality. Therefore, online sellers were unable to determine the influential order of all these factors.

With respect to the research methods, current researches mainly employed three methods: Grey Relational Analysis, C2C Model, and BP Neural Network Fitting. As for Grey Relational Analysis, Fatao Wang employed Grey Relational Analysis to determine the main factors for the development of online shopping. Naicong Hou, Xu Zhang, Enjun Zhang [Naicong Hou, Xu Zhang, Enjun Zhang. Grey Relational Analysis of online reputation and sales volume——movie data as an example[J], marketing, 2015(02):28-30] presented reputation as the most influential factor of purchase. Xiao Shi [Xiao Shi. Research of influential factors of online sales based on Grey Relational Analysis.] conducted a quantitive research of the interrelation of sales and price, comment rate, popularity with the utilization of Grey Relational Analysis. As for C2C model, Youzhi Xue and Yongfeng Guo [Youzhi Xue and Yongfeng Guo, Research of the competing strategy of C2C electronic seller: analysis based on Taobao[A], e-commerce, 2012(15),129-140] employed Tobit model to discover that customers valued more on price and delivery fee. Jingsha Fu [Jiangsha Fu. Determining factors of online sales based on C2C.] created a quantify model of influential factors. As for BP Neural Network Fitting, Yanli Ma built an evaluating system including refund rate, descriptions and online comments. All these aforementioned methods are theoretically capable of analyzing the influence of certain factors on sales volume, but are lack of practicality.

In conclusion, current researches have failed to analyze influential factors in a systematic and comprehensive way, and they have failed to reveal specific characteristics that achieve higher sales volume. Therefore, our research results fill in this academic blank by offering a clear view into the characteristics that cellphones with high sales volume have and applying our results to predicting sales volume.

Research purpose and significance

Since online sellers constantly worry about ways to promote sales volume, we conduct research in the hope of offering a practical solution by determining which characteristics contribute to improving sales volume. Our research purposes can be summarized as below:

* To conduct qualitative research to have a general understanding of the characteristics that contribute to high sales volume.
* To conduct quantitative research to rank factors that are considered to have influence on sales volume.
* To determine specific characteristics within each factor that contribute to highest sales volume.
* To predict the sales volume of cellphones with a given characteristic.

Our research results will be of great reference and help to online cellphone sellers by offering clear explanation of what kinds of cellphones have the highest sales volume. Online cellphones sellers can consequently adjust their products according to our research results to achieve higher sales volume.

Research method and train of thinking

After gathering data of information about product selling in AliExpress, we extract useful and relevant data concerning different influential factors for further research. Next, we come to the data procurement to reduce the number of independent factors. In this process, we apply three different method——Grey Relational Analysis, Principal Component Analysis, and Entropy of Information. The Grey Relational Analysis fails to reduce the number of influential factors, while the other two methods effectively complete the goal. Then we apply the results from data procurement for modeling. In the modeling process, we apply results from Principal Component Analysis to Analytic Hierarchy Process, KNN, and Linear Regression. At this point, we have reached the conclusion of the rank of different independent factors. Furthermore, we conduct optimization to each model. We optimize KNN by Bayes Distinction and Linear Regression by Principal Component Regression, while we optimize Entropy of Information to BP Neural Network Fitting. Afterwards, we employ the BOOST algorithm to synthesize the three methods and reach the conclusion that which characteristic contribute to the highest sale volume. Finally, we practice the application of our research results by predicting future sales conditions.