applied sciences

MDPI

Article

Identification of Usefulness for Online Reviews Based on Grounded Theory and Multilayer Perceptron Neural Network

Jiani Hou * and Aimin Zhu

School of Management, Shenyang University of Technology, Shenyang 110870, China
* Correspondence: 13998270823@163.com

Abstract: In the context of the continuous development of e-commerce platforms and consumer shopping patterns, online reviews of goods are increasing. At the same time, its commercial value is self-evident, and many merchants and consumers manipulate online reviews for profit purposes. Therefore, a method based on Grounded theory and Multi-Layer Perceptron (MLP) neural network is proposed to identify the usefulness of online reviews. Firstly, the Grounded theory is used to collect and analyze the product purchasing experiences of 35 consumers, and the characteristics of the usefulness of online reviews in each stage of purchase decision-making are extracted. Secondly, the MLP neural network classifier is used to identify the usefulness of online reviews. Finally, relevant comments are captured as the subject and compared with the traditional classifier algorithm to verify the effectiveness of the proposed method. The experimental results show that the feature extraction method considering consumers' purchase decisions can improve the classification effect to a certain extent and provide some guidance and suggestions for enterprises in the practice of operating online stores.

Keywords: online review; the usefulness of online reviews; multilayer perceptron neural network; Grounded theory

check for updates

Citation: Hou, J.; Zhu, A. Identification of Usefulness for Online Reviews Based on Grounded Theory and Multilayer Perceptron Neural Network. *Appl. Sci.* **2023**, *13*, 5321. https://doi.org/10.3390/app13095321

Academic Editor: Christos Bouras

Received: 8 March 2023
Revised: 2 April 2023
Accepted: 12 April 2023
Published: 24 April 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, with the popularity of mobile Internet technology and the severe background of the COVID-19 pandemic, more and more consumers have changed their traditional shopping modes and chosen to complete the basic needs of life online. At the same time, as e-commerce platforms facilitate consumers' purchase decisions, users can leave their feelings and comments on the use of products or services in the form of online comments at the post-purchase stage [1]. With the change in consumption patterns, the number of orders and corresponding online comments on e-commerce platforms is steadily on the increase. At the same time, the commercial value of e-commerce platforms and merchants and the reference value of shopping decisions for consumers are also self-evident. Faced with such redundant growth, it is difficult for consumers to find valuable information and advice from massive online reviews [2].

Due to the large amount of online review information, lack of standardization, wide audience, and other characteristics, consumers will be affected by them in the whole purchase decision-making process [3,4]. In addition, due to the uneven authenticity and quality of review information, e-commerce platforms lack unified evaluation standards, and manual screening of online reviews for consumers to make decisions is not practical, which greatly increases costs and reduces efficiency [5]. Therefore, how to enable consumers to quickly screen useful online review information in each stage of purchase decision-making has become an urgent practical problem to be solved so that the online review system can better serve the platform, enterprises, and consumers.

The identification of online reviews is mainly based on whether they are useful for consumers to make decisions. However, in the real shopping situation, merchants often

employ professional "brushes" to carry out fake shopping behavior for the purpose of increasing sales. The purpose of merchants is only to increase the sales of commodity links and the praise written by "brushes" after confirming receipt of goods. In the process of fake shopping, it is difficult for "brushes" to operate according to the shopping process of real consumers. Moreover, due to the lack of experience and feelings of real consumers about the goods, the online comments written by them cannot truly reflect the real attributes and shopping experience of the goods. Therefore, the identification of online comments should start not only from the comment text itself but also from the whole process of consumer purchase to identify online comments that do not conform to the real purchase behavior of the order.

In summary, a method is proposed to identify the usefulness of online reviews by using Grounded theory and an MLP neural network to consider consumer purchasing behavior. Firstly, based on the Grounded theory, a semi-structured interview is conducted on the real shopping experiences of 35 interviewees, and the usefulness of online reviews is characterized by open coding, spindle coding, and selective coding from each stage of consumers' purchase decision making. Secondly, the captured data are used to test and train the model. Through multiple sets of horizontal and longitudinal experiments, the optimal feature combination that best matches the MLP neural network classifier model is found and compared with other traditional classifiers so as to verify the superiority of the classification effect of the model.

The remainder of the paper is organized as follows: Section 2 reviews the literature on the influencing factors and recognition algorithms of the usefulness of online reviews. Section 3 introduces the process of identifying the usefulness of online reviews' considering consumers' purchase decisions based on Grounded theory. Section 4 carries out the experimental design and verifies the validity and interpretability of the proposed method through multiple sets of horizontal and longitudinal comparison tests. Section 5 gives the experimental conclusion and briefly introduces the shortcomings and future prospects of the paper.

2. Literature Review

2.1. Research on the Influencing Factors of the Usefulness of Online Reviews

The usefulness of online comments refers to the extent to which viewers of e-commerce platforms think the comments are valuable and helpful after reading them [6]. At present, the influencing factors for the usefulness of online reviews have been analyzed from various perspectives. For example, Lee et al., Malika et al., and Ketron et al. studied the usefulness of online reviews from the perspective of grammatical features and found that online reviews with high-quality grammatical structure can make consumers have higher credibility and a stronger purchase desire. On the other hand, reviews of low quality are less credible and make consumers less inclined to purchase [7–9]. Cheng et al., Ullah et al., and Gang et al. studied the usefulness of online reviews from the perspective of semantic features, used text mining and sentiment analysis methods to study Airban users' focus on online review content, and found that consumers like to comment on the basis of existing reviews. Location, convenience, and service are the most important factors for consumers in reviews [10–12]. Guo et al., Weimer et al., and Li et al. studied the usefulness of online reviews from the perspective of genre characteristics and used four styles, namely credibility, readability, substantiality, language acquisition, and word counting, to predict the usefulness of reviewers, which had a better effect than the benchmark model [13–15].

Existing studies all start with online reviews themselves and study the features in their content that can be used to measure the usefulness of online reviews as identifying factors. However, few studies judge the usefulness of online reviews from the perspective of consumers' purchasing behaviors by identifying the features of fake shopping behaviors. The purchase decision-making process of online consumers can be divided into several stages: demand generation, information collection, product evaluation, purchase decision, and post-purchase behavior [16]. Existing studies have shown that online reviews have

varying degrees of influence on all aspects of the decision-making process of consumers [17]. Grounded theory is a qualitative research method that generalizes and summarizes unknown phenomena through the cyclic methods of data collection, conceptual abstraction, and comparative analysis, which is suitable for summarizing the identification characteristics of the usefulness of online reviews. In conclusion, based on the Grounded theory, the feature extraction method is selected considering consumers' purchase decision-making behavior to construct an index of the usefulness of online reviews, which fits the real online shopping situation of consumers and can improve the recognition effect of the model.

2.2. Research on the Algorithm of Online Review Usefulness Recognition

The identification of the usefulness of online reviews is fundamentally a binary problem, and some research have used machine learning methods to identify it. Lee et al. classified the usefulness of online hotel reviews by the quality of reviews, the emotion of reviews, and the characteristics of reviewers and selected four classification models, DT, RF, PCA+LGR, and PCA+SVM, to conduct empirical research on different data sets. Through different analytical perspectives, it can be found that RF classification technology has the best predictive performance [18]. Ou et al., different from previous studies, used a statistical regression model to analyze the influencing factors of the usefulness of online reviews and chose a support vector machine model to train attractive attribute classifiers [19]. Aiming at the problem that the existing opinion mining technology ignores the quality of comments, Chen and Tseng proposed a method of first extracting representative comment features and then classifying the information quality of online comments using support vector machines [20]. Later, some scholars found that MLP neural networks performed better than traditional classifiers in text classification. For example, Hu et al. and Lee et al. used the multi-layer perceptron neural network to predict the usefulness of online reviews from the perspective of product information, review features, and text features and argued that the MLP neural network was better than the linear regression analysis method in the calculation of the minimum mean square error [21,22]. In conclusion, the core problem of this paper is to construct a classifier based on useful and useless samples of tags using supervised learning and then predict untagged comments as useful or useless comments. Based on the above literature review, a multi-layer perceptron neural network is selected as the classifier model, which identifies the usefulness of online reviews and compares them with other traditional classifiers to prove their superiority.

To sum up, most of the existing studies on the usefulness of recognizing online comments only take the length of comments, the number of comments, the degree of emotion of comments, and other factors as identification characteristics based on the comment text itself. However, in the real situation, merchants' "brushing" behavior for the purpose of increasing sales volume or consumers' large product reviews driven by the interest of "cash back on good reviews" cannot be identified from the perspective of the identification characteristics above. Therefore, a feature construction method is proposed, considering the purchasing decision-making process of consumers. Based on the Grounded theory, the product purchasing experiences of 35 consumers are analyzed, and the features that can effectively identify the usefulness of online reviews are selected from the five stages of demand generation, information collection, product evaluation, purchase decision, and post-purchase behavior, respectively. Finally, the MLP neural network classification model is used to identify the usefulness of online reviews, and the optimal feature combination set is obtained by a multi-group comparison test, which is tested on other traditional classifiers. By crawling more than 6000 online comments from JD.com (accessed on 5 November 2022) as the subject data set, the feasibility and effectiveness of the proposed method are verified. The experimental results show that the feature construction method based on Grounded theory considering consumer decision-making processes proposed can effectively improve the classification effect of the MLP neural network. It provides theoretical support and technical guarantees for e-commerce business operations and consumer decision-making.

3. A Method to Identify the Usefulness of Online Reviews

The fundamental problem of identifying the usefulness of online reviews is a binary one, which can be divided into two categories from the perspective of consumers: useful and useless. Most scholars use machine learning methods such as Neural Networks, Naive Bayes, and Support Vector Machines to identify the usefulness of online reviews, which are quantitative methods. In this paper, qualitative and quantitative methods are used to analyze it, and the identification features are constructed using qualitative methods and Grounded theory. The classifier is constructed using quantitative methods, and the specific research framework is shown in Figure 1.

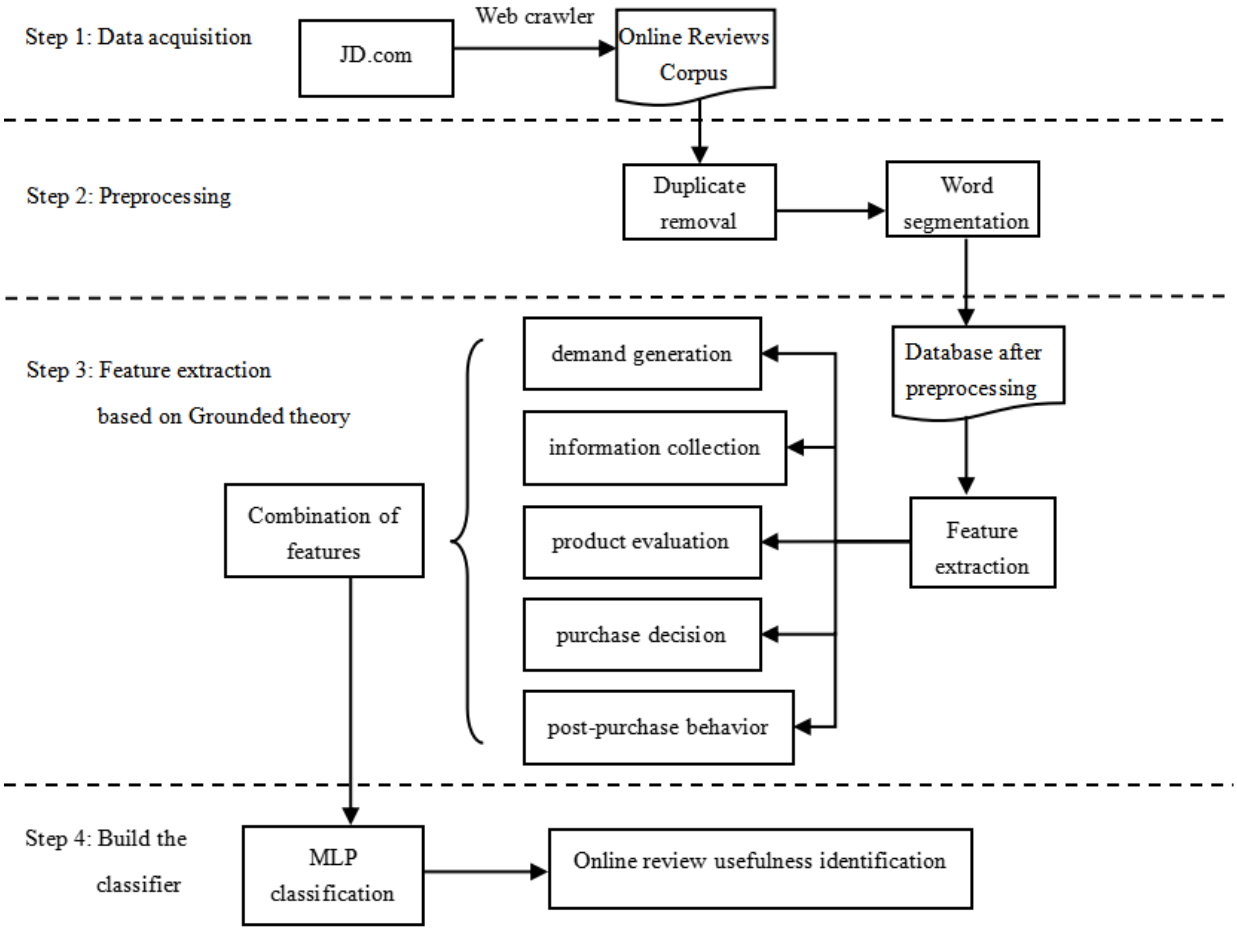


Figure 1. Research methodology framework.

The specific steps of the research method proposed are as follows: Firstly, in the stage of data acquisition and preprocessing, the online review database of related products crawled from JD.com (accessed on 5 November 2022) is established and the data preprocessing is carried out. Secondly, in the feature engineering stage, feature extraction is carried out from each stage of the consumer decision-making process based on the Grounded theory. Through the semi-structured interview data of 35 consumers, feature extraction is carried out on the usefulness of online reviews from the stages of demand generation, information collection, product evaluation, purchase decision, and post-purchase behavior. The extracted features are combined to obtain multiple feature representations of the usefulness of online reviews. Finally, multiple feature combinations are trained on the MLP classifier, the optimal feature combination solution is obtained through comparative analysis, and a reasonable interpretation is obtained in the real situation.

3.1. TF-IDF Method

Among the natural language processing methods for text data, the existing research mainly uses the word bag method and the TF-IDF method [23]. The word bag method refers to the word segmentation representation of the text data, and then the number of occurrences of each word in the text is counted to obtain the word-based feature representation [24]. However, meaningless modal particles exist in many text data in real situations, but the bag of words method cannot identify them. Therefore, some scholars add a reverse file frequency index based on the core idea of the bag of words method so that the model has a better recognition effect [25]. Therefore, the TF-IDF method was selected to perform feature representation on the crawled online review data, and it was used as the baseline system to verify the effectiveness of the proposed method in the comparative analysis.

3.2. Feature Extraction Based on Grounded Theory

In the stage of data preprocessing, the online comment text should be processed with duplicate removal, word segmentation, and stop words removal, and the pre-processed online comment text data set can be obtained. In the stage of constructing the feature representation of online reviews, 35 experienced online shoppers are interviewed through semi-structured interviews. The interview outline is shown in Table 1. Secondly, based on the Grounded theory, open coding, spindle coding, and selective coding are carried out for the semi-structured interview content [26], and feature selection and feature combination are carried out for the coding results considering the purchasing decision-making process of consumers [27]. The decision-making process of consumers is divided into several stages: demand generation, information collection, product evaluation, purchase decision, and post-purchase behavior [16]. Therefore, the whole semi-structured interview process is also centered on the decision-making process of consumers. A total of 29 (P_1–P_{29}) complete interview texts are collected in this study, while the rest were incomplete or not completed as required. After the interview, the audio recordings were transcribed into 29 interview texts with more than 20,000 words. Through data coding analysis of the above interview texts, the feature selected in the final stage of demand generation is keyword retrieval number (N_{kr}); in the information collection stage, the selected characteristic is the number of comments viewed (N_{cv}); in the product evaluation stage, the selected characteristics are the number of other products to be viewed in the store where the target located (N_{PSP}) and the number of similar products of the target to be viewed (N_{PP}); in the purchase decision stage, the selected characteristic is the number of days the product was added to the cart (D_p); in the post-purchase behavior stage, the selected characteristics are the days between receiving goods and commenting (D_{rc}), the days between first comments and additional comments (D_{FA}), the number of comment to be viewed (N_B), and the number of likes on comments (N_G). The details are shown in Figure 2.

Table 1. Interview outline.

Number	Interview Questions
1	How many hours a week do you spend reading reviews on e-commerce platforms? What is the frequency?
2	Could you tell me something about your usual shopping process?
3	How do you select search keywords for products?
4	What shopping experience would you most like to see in a product review?
5	What metrics do you look for when reviewing reviews?
6	Do you compare the target products you want to buy? How does it compare?
7	Have you ever collected or added the products you finally purchased?
8	How long do you comment on the product firstly and additionally after receiving it?

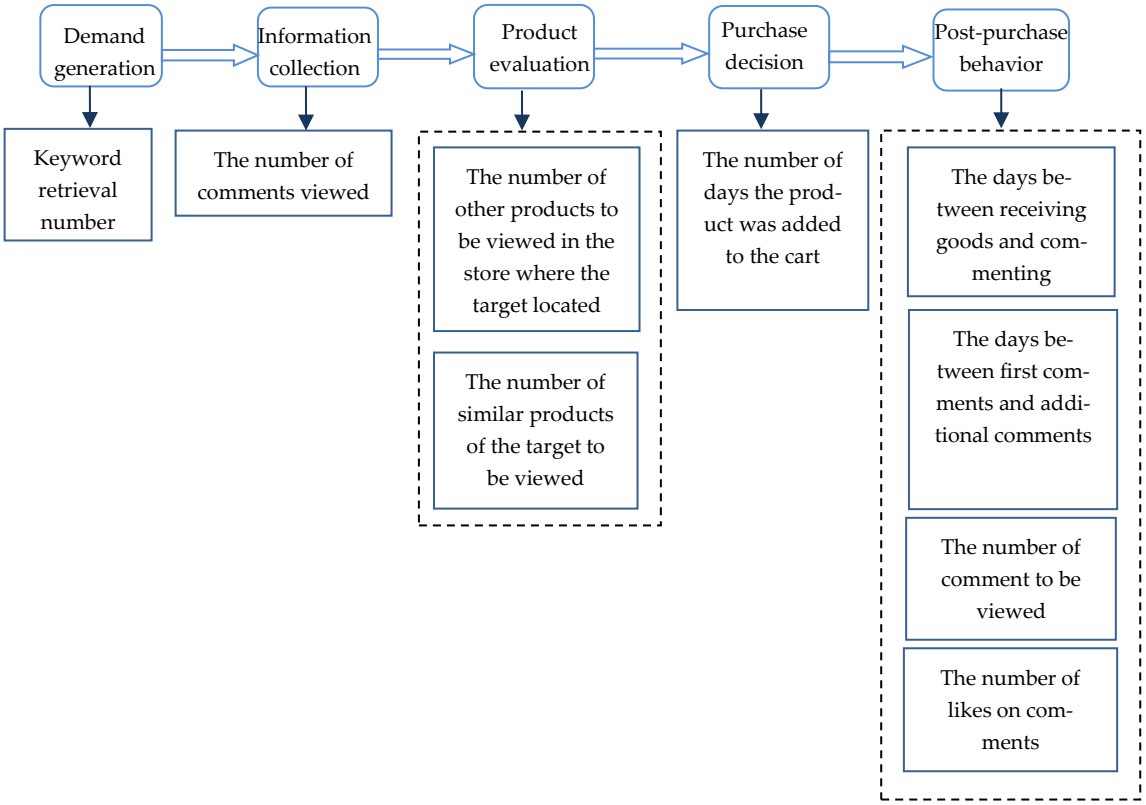


Figure 2. Consumer purchase decision-making process and feature classification situation.

(1) N_{kr}: refers to the keyword retrieval number. According to the interview record of Question 3 in P_2, "*Generally, when I want to buy seasonal clothes, I will search several keywords, such as '2022 new women's clothing', '2022 winter extra thick', and 'blouse loose for women', to find clothes suitable for me*";

(2) N_{cv}: refers to the number of comments viewed. P_{14} indicates that he will browse the reviews of the target products that he wants to buy, especially for the reviews that contain the product features that he/she is interested in;

(3) N_{psp}: refers to the number of other products to be viewed in the store where the target is located. The interview record for P_{16} for question 2 records that "*when I find something I want to buy, I will click it to see the reviews and the stores, especially when I buy clothes. I will also check whether the style of the store is suitable for me and the reviews of other products in this store*";

(4) N_{pp}: refers to the number of similar products from the target to be viewed. P_{21} indicates that the price of the same product may be inconsistent in different stores, so browsing and comparison will be carried out in the selection of products;

(5) D_p: refers to the number of days the product was added to the cart. P_{10} 's interview record for question 7 states that "*I always buy some things with appropriate prices or possible needs, and it is more cost-effective to buy them together when there is a huge promotion activity*";

(6) D_{rc}: refers to the days between receiving goods and commenting. P_{24} indicates that generally, after confirming the receipt of goods, products will be reviewed after a period of experience;

(7) D_{fa}: the days between the first comment and additional comments. According to the interview record for question 8 in P_{27}, "*Sometimes I review and sometimes I don't. Sometimes I will make additional comments after a period of experience when I find that there is a big difference between my experience and that at the beginning*";

(8) N_b: refers to the number of comments to be viewed. The interview record of P_{29} on question 5 records that "*There are too many fake comments now, but relatively speaking, there are reference value for the more popular ones*";

(9) N_g: refers to the number of likes on comments. The interview record of P_8 for question 5 records that "*It must be the number of likes of comments. Usually, I will also give a like to comments that are relevant and valuable for reference*".

Due to the different nature of the selected features, the data are different. In order to improve the recognition efficiency of the classifier, the normalization of each feature index should be carried out. The specific formula is shown in Formula (1), where *original_value* represents the original value of the feature data, min represents the minimum value of all sample data under the feature, max represents the maximum value of all sample data under the feature, and *value* represents the feature value after normalization processing.

$$value = \frac{(original_value - \min)}{(\max - \min)} \tag{1}$$

3.3. MLP Neural Network

In recent years, the neural network model has been widely used in natural language processing due to its hidden layer, which can automatically combine features and alleviate data scarcity [28]. A neural network is a classification model composed of a large number of interconnected neurons. Due to its strong adaptive learning ability, it can make decisions on complex problems with multiple inputs and outputs without relying on research objects [29]. The MLP neural network adds one or more fully connected hidden layers between the output layer and the input layer and transforms the hidden layer output through an activation function, thus avoiding the situation where the linear, non-fractal data cannot be recognized. So it is more suitable for classification and recognition of the usefulness of online comments.

As shown in Figure 3, the MLP neural network model is composed of an input layer, a hidden layer, and an output layer, respectively, and the neurons of each layer are transmitted layer by layer to each neuron of the next layer, forming a fully connected layer. First of all, the sentence vector representation is input as the input layer, and an m neuron transmits to each neuron of the hidden layer, respectively. Relu function is used to activate it in the hidden layer and transmit it to each neuron in the output layer through the n neuron. The specific formula of the hidden layer is as follows:

$$Y_1 = f(W_1 X_1 + b_1) \tag{2}$$

where X_1 is the input layer, W_1 is the connection coefficient, and b_1 is the bias. In addition, the Relu function is chosen as the activation function of the hidden layer, and its specific formula is as follows:

$$f(x) = \max(0, x) \tag{3}$$

Finally, the output result of Formula (3) is taken as the input of the output layer, and the Softmax function is used to carry out regression operation in this layer. The specific formula of the output layer is as follows:

$$Y_2 = (W_2 X_2 + b_2) \tag{4}$$

where X_2 is the hidden layer; W_2 is the connection coefficient, and b_2 is the bias. Y_2 is the final output; it is a two-dimensional text category representation where -1 means the online comment is useless and 1 means the online comment is useful.

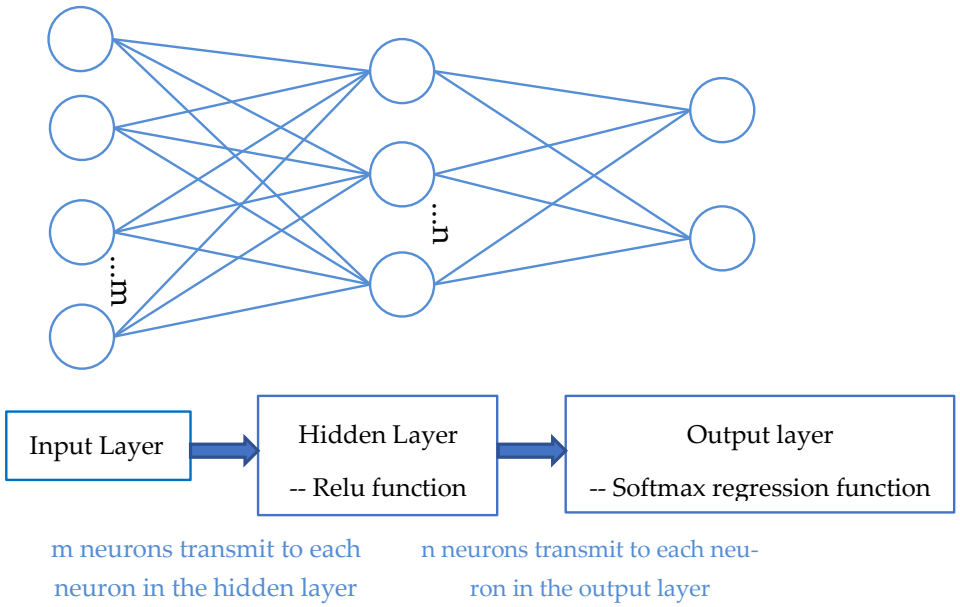


Figure 3. MLP neural network model structure.

As shown in Figure 3, the MLP neural network model is composed of m neurons in the input layer, n neurons in the hidden layer, and two neurons in the output layer. The independent variable X_i is a m dimensional vector and the dependent variable Y is a two-dimensional vector. The overall formula of the model is as follows:

$$Y = G\{W_2[f(W_1X + b_1)] + b_2\} \quad (5)$$

where G is the Softmax regression function of the output layer; W_1 and W_2 are the connection function; and b_1 and b_2 are the bias.

4. Experimental Design

In this section, the method proposed will be analyzed and explained in detail using the online review data of mobile phone products. Firstly, the specific method proposed is introduced, and then it is compared with the existing methods of online review usefulness identification so as to illustrate the effectiveness of the proposed method.

4.1. Construction of Training Set and Test Set

Python is used to crawl relevant online comment data from JD.com (accessed on 5 November 2022), a total of 6215 comments. Each piece of data contains the characteristic information of each purchasing stage of the consumer mentioned above.

4.1.1. Automatic Annotation Setting of Training Set

Since the machine learning classifier needs to divide the data into a training set and a test set, the training set is used to train and the test set is used to test, so as to verify the feasibility and accuracy of the method. In the problem of identifying the usefulness of online comments, its essence is a binary problem, which only needs to judge whether the online comment is useful or not. Therefore, the usefulness of online comments is recorded as +1, and the useless comments as −1. For the training set data, the determined classification results should be given to the machine to learn. In previous studies, some scholars annotated the training set data with artificial labeling, which is slightly less objective [30,31]. However, some of the literature regards the number of likes on comments as the degree of usefulness of online comments [32]. Although this index cannot be summarized simply, it is interpretable in the automatic labeling and recognition of training sets. Therefore, the number of likes on comments is chosen as the index for automatic

labeling of training set text. In addition, for the selection of the threshold for the number of likes, 100 pieces of data randomly selected from the data were marked with the category of manual labeling on usefulness, adjusted the results of automatic labeling under different thresholds, and compared them with the results of manual labeling. The fluctuation of the accuracy rate is shown in Figure 4. It can be seen that with the increase of the number of likes threshold, the accuracy of the automatic labeling results increases first and then decreases. It reaches its peak value when the number of likes threshold is 6. Therefore, the number of likes indexes for automatic labeling in the training is set at 6, which can make the automatic labeling classification results accurate and reliable.

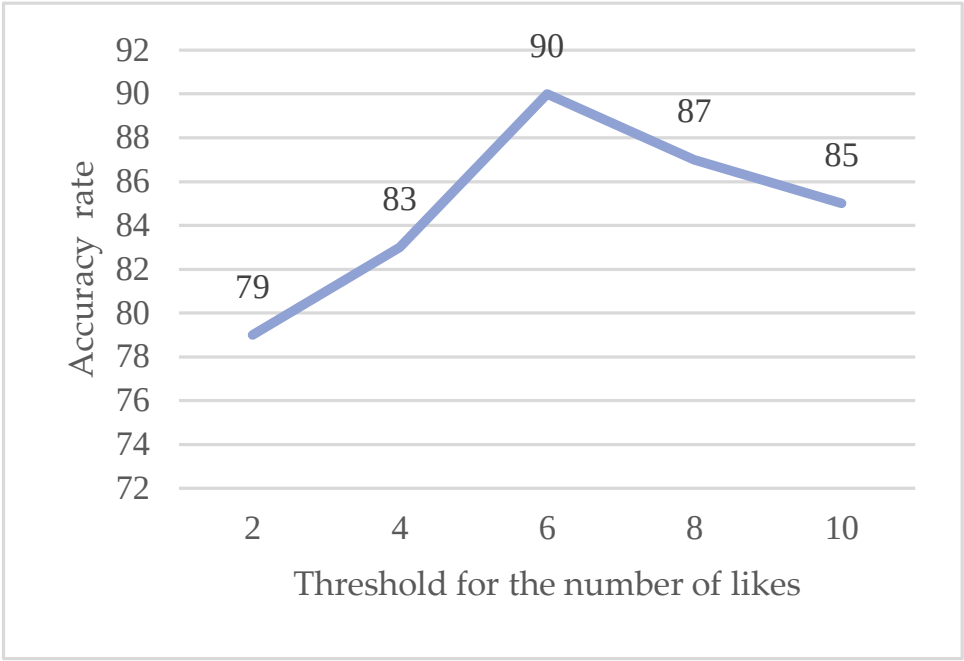


Figure 4. Fluctuation of accuracy under different thresholds.

4.1.2. Selection of Training Set and Test Set

In the process of classifier model training, the proportion and selection of the training set and data set affect the classification effect of the model [33]. In previous studies, most scholars divided the original database by 1:9, selected one of them as the training set, and used the trained classification model to classify and identify the usefulness of online reviews with 9 of them [34]. However, the random selection of the training set and test set is contingent to some extent. In order to make the model not be affected by the randomness of the training set and test set and have a better classification effect, the original data are divided into 10 parts in equal proportion; one part is set as the training set in turn, and the other part is set as the test set. The average value of 10 times the classifier training results is taken as the final classification result.

4.2. Evaluation Indicators

Aiming at the evaluation problem of an online review usefulness recognition method, most scholars choose to use the accuracy rate, recall rate, and their harmonic average value to conduct an all-round evaluation of the classification model, and the evaluation results can honestly reflect the advantages, disadvantages, and feasibility of the recognition method [35]. Therefore, the evaluation indexes selected are the useful review accuracy rate, the useful review recall rate, the F_1 value, and the overall review accuracy rate. The useful review accuracy rate ($Precision_uf$) refers to the proportion of useful comments among the data identified as useful; the useful review recall ($Recall_uf$) refers to the percentage of useful comments that are identified as useful; and the F_1 value refers to the harmonic

average of the accuracy rate and recall rate. Overall accuracy (*Precision*) refers to the ratio of all correctly identified data in the database. The specific calculation formulas are shown in Equations (6)–(9).

$$Precision_uf = \frac{A}{A + C} \tag{6}$$

$$Recall_uf = \frac{A}{A + B} \tag{7}$$

$$F_1 = \frac{2 * Precision_uf * Recall_uf}{Precision_uf + Recall_uf} \tag{8}$$

$$Precision = \frac{A + D}{A + B + C + D} \tag{9}$$

where A is the number of original useful comments correctly identified; B is the number of original useful comments misidentified; C is the number of original useless comments incorrectly identified; and D is the number of original useless comments correctly identified. The specific data division is shown in Figure 5.

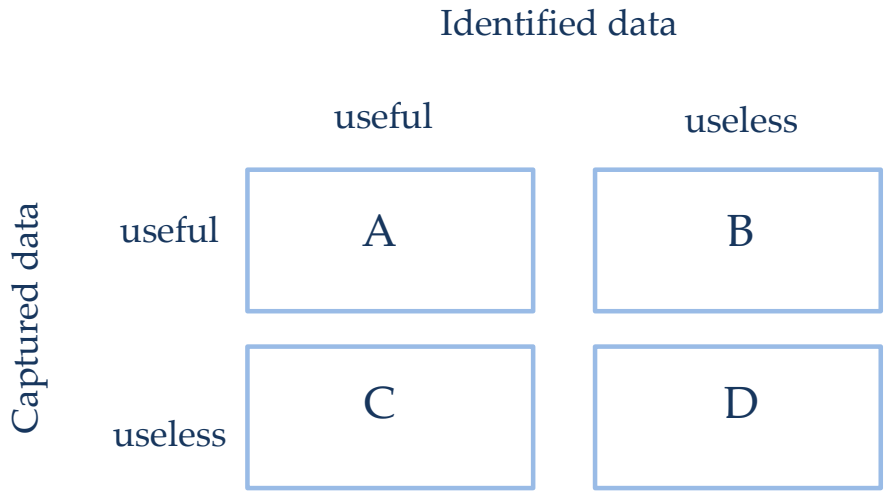


Figure 5. Division of data.

4.3. Result Analysis

4.3.1. Classification Results of MLP Model Based on Different Feature Combination

Firstly, the MLP classification model trained by the TF-IDF feature extraction method is taken as the baseline system. Secondly, the features of demand generation (DG), information collection (IC), product evaluation (PE), purchase decision (PD), and post-purchase behavior (PB) in consumer purchase decision making are trained separately, and the combined features of the whole consumer purchase decision making process (PDP) are trained to obtain multiple different MLP classifier models. The feasibility and effectiveness of identifying the usefulness of online reviews from the perspective of consumers' purchase decision-making process are verified by comparative analysis. In addition, in order to judge the necessity of each index in each stage and the degree of impact on the usefulness of online reviews, for the purchase stage with multiple indexes, it is necessary to identify the decline in model recognition results after removing a single index. That is, for the two stages of product evaluation and post-purchase behavior, the relevant six indicators are removed one by one, and the impact of each indicator on the usefulness of online reviews is judged by comparing the decline degree of the model identification results. The specific experimental results are shown in Table 2.

Table 2. Recognition results of the MLP neural network classifier.

Feature Combination	Precision_uf	Recall_uf	F_1	Precision
TF-IDF	0.647	0.546	0.592	0.599
PDP	0.916	0.864	0.889	0.896
DG	0.792	0.732	0.761	0.779
IC	0.821	0.769	0.794	0.805
PE	0.896	0.735	0.808	0.847
PE-N_{psp}	0.875	0.724	0.792	0.832
PE-N_{pp}	0.885	0.728	0.799	0.84
PD	0.845	0.723	0.779	0.805
PB	0.912	0.851	0.88	0.889
PB-D_{rc}	0.894	0.847	0.87	0.876
PB-D_{fa}	0.893	0.832	0.861	0.87
PB-N_b	0.802	0.769	0.785	0.796
PB-N_g	0.813	0.771	0.791	0.799

As shown in Table 2:

(1) As a baseline system, the effectiveness (F_1) of the MLP neural network model trained by the TF-IDF feature extraction method to identify the usefulness of online reviews is 59.2%. However, the MLP neural network model proposed is to train the multi-stage combined features from the consumer purchase decision-making process. The recognition effect (F_1) reached 88.9%, compared with an increase of nearly 30%. The results show that the feature extraction method, considering the purchasing decision-making process of consumers, significantly improves the classification effect of the MLP classifier. The usefulness of online reviews can be identified not only from the comments themselves but also from the whole purchasing decision-making process of consumers, and the usefulness of online reviews can be identified through the characteristic indicators in each stage;

(2) Both the classification effect of the classification model with the indexes of each stage trained separately and the other features of the multi-index stage trained after the removal of a single index decrease to a certain extent. Among them, by analyzing the recognition effects of classifier models at different stages of consumer decision-making, it can be seen that the influence of each stage on the usefulness of online reviews, in descending order, is: post-purchase behavior, product evaluation, purchase decision, information collection, and demand generation. In the product evaluation stage, N_{psp} (the number of other products to be viewed in the store where the target is located) has a greater impact on the identification results than N_{pp} (the number of similar products to be viewed). In the post-purchase behavior stage, the degree of influence on the recognition results of the usefulness of online comments, in descending order, is: N_b (the number of comments to be viewed), N_g (the number of likes on comments), D_{fa} (the days between the first comment and additional comments), and D_{rc} (the days between receiving goods and commenting).

In order to further optimize the recognition effect of the classification model and study the necessity of each index, each refinement index is added to the feature representation in turn according to its influence degree, from large to small, on the classification results. The recognition effect of the nine classifiers obtained is shown in Table 3.

Table 3. Classification results of the MLP neural network classifiers based on different feature combinations.

Feature Combination	Precision	Recall	F_1	Accuracy
N_B	0.913	0.764	0.832	0.843
N_B, N_G	0.91	0.768	0.833	0.85
N_B, N_G, N_{PSP}	0.912	0.782	0.842	0.862
N_B, N_G, N_{PSP}, D_{FA}	0.909	0.79	0.845	0.869
N_B, N_G, N_{PSP}, D_{FA}, N_{PP}	0.902	0.824	0.861	0.874
N_B, N_G, N_{PSP}, D_{FA}, N_{PP}, N_{cv}	0.911	0.832	0.87	0.881
N_B, N_G, N_{PSP}, D_{FA}, N_{PP}, N_{cv}, D_{rc}	0.914	0.86	0.886	0.89
N_B, N_G, N_{PSP}, D_{FA}, N_{PP}, N_{cv}, D_{rc}, D_p	0.914	0.872	0.893	0.9
N_B, N_G, N_{PSP}, D_{FA}, N_{PP}, N_{cv}, D_{rc}, D_p, N_{kr}	0.916	0.864	0.889	0.892

As can be seen from Table 3, when the feature combination is N_B (the number of comment to be viewed), N_G (the number of likes on comments), N_{PSP} (the number of other products to be viewed in the store where the target locates), D_{FA} (the days between first comments and additional comments), N_{PP} (the number of similar products of the target to be viewed), N_{cv} (the number of comments viewed), D_{rc} (the days between receiving goods and commenting), D_p (the number of days the product was added to the cart), the MLP classifier model has the best recognition effect, with the F_1 value reaching 89.3%; when the feature combination is the feature of the whole process of consumer purchase decision, the F_1 value decreases to 88.9%. It can be seen that the addition of N_{kr} (keyword retrieval number) is not conducive to the model's recognition of the usefulness of online reviews. In the decision-making process of the consumer, the stage of demand generation is far from the writing stage of online comments after the final deal is reached, which may be affected to some extent, resulting in deviation. However, except for the demand generation stage, the addition of each feature index in other stages all improves the classification and recognition effect of the model, which can prove the effectiveness of the feature extraction method proposed in this paper considering the purchasing decision-making process of consumers.

4.3.2. Comparison between the Proposed Method and the Classical Method

In order to further verify the recognition effectiveness of the proposed method, it is compared with the traditional classifier model Naive Bayes and Support Vector Machine (SVM). The feature representation of TF-IDF method , the optimal feature combination (N_B, N_G, N_{PSP}, D_{FA}, N_{PP}, N_{cv}, D_{rc}, and D_p) and the all feature combination (N_B, N_G, N_{PSP}, D_{FA}, N_{PP}, N_{cv}, D_{rc}, D_p, and N_{kr}) based on Grounded theory considering purchase decision-making process are all considered as the feature representation to train SVM classifier model, Naive Bayes classifier model and MLP neural network classifier, respectively. The specific experimental results are shown in Figure 6.

As shown in Figure 6:

(1) In the selection of feature extraction methods, the overall recognition effect of the three classifiers trained by the TF-IDF method is poor, and they only perform well in the recall rate, while the *Precision_uf*, F_1 value and *Precision* their scores are low compared with those trained by the other two feature representation methods considering consumers' purchase decisions. Since the TF-IDF feature representation method only starts from the perspective of text word frequency and the usefulness of online reviews in real situations is affected by many factors, it can be seen that the feature extraction method carried out in this paper from each stage of consumers' purchase decisions has certain effectiveness;

(2) In the comparison of the three feature extraction methods, the optimal feature combination has the best performance, and its recognition effect on the three different classifier models is excellent. For the SVM classifier model, the F_1 value based on optimal feature combinations is between 4.3% and 15.4% higher than those based on all feature combinations and the TF-IDF method. For the Naive Bayes classifier

model, the F_1 value based on optimal feature combination is between 1% and 5.4% higher than those based on all feature combinations and the TF-IDF method. For the MLP neural network classifier model, the F_1 value based on the optimal feature combination is 0.4% and 30.1% higher than those based on all feature combinations and the TF-IDF method. It can be concluded that the selection of the optimal feature combination significantly improves the recognition effect of the classifier.

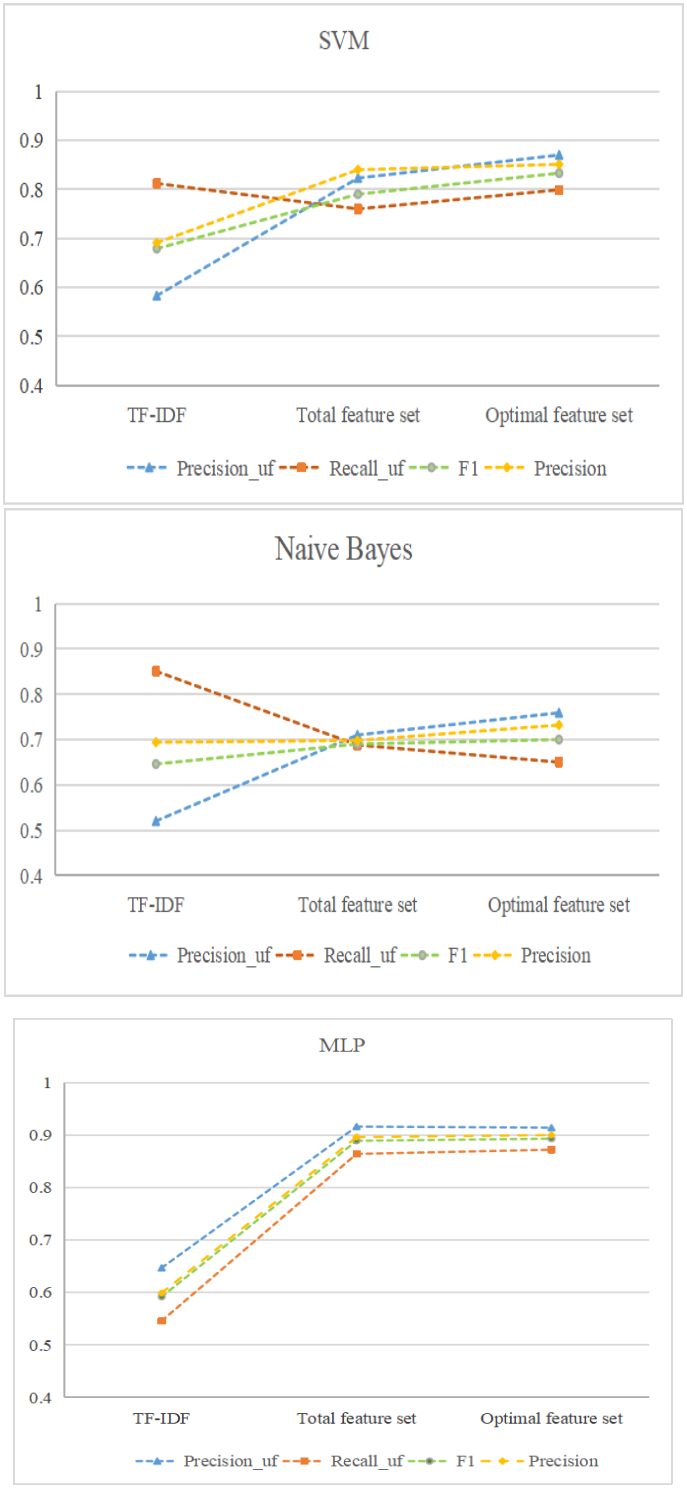


Figure 6. Comparison of the recognition effect under different feature representations of three classifiers.

The recognition effect of the three classifier models under the optimal feature representation will be further analyzed below.

As can be seen from Figure 7, in terms of the selection of classifier models, the MLP classifier has the best overall recognition effect, and all indexes are higher than the other two traditional classifiers. The F_1 value of the SVM classification model based on the optimal feature combination is 83.3%, while the F_1 value of the MLP classifier is 6% higher than that of the SVM classifier, reaching 89.3%. Because the SVM classifier is also very suitable for the analysis of short text data, the classification effect is not as outstanding as that of the MLP classification model, but it still meets the basic standard. The F_1 value of the Naive Bayes classification model based on the optimal feature combination is only 70%, and the F_1 value of the MLP classifier is 19.3% higher than that of the Naive Bayes classifier. Although the optimal feature combination improves the classification effect of the model, the selection of the classifier is still particularly important. So it can fully reflect the matching degree of the optimal feature combination representation and the MLP classifier model, and it also proves the effectiveness of the recognition results.

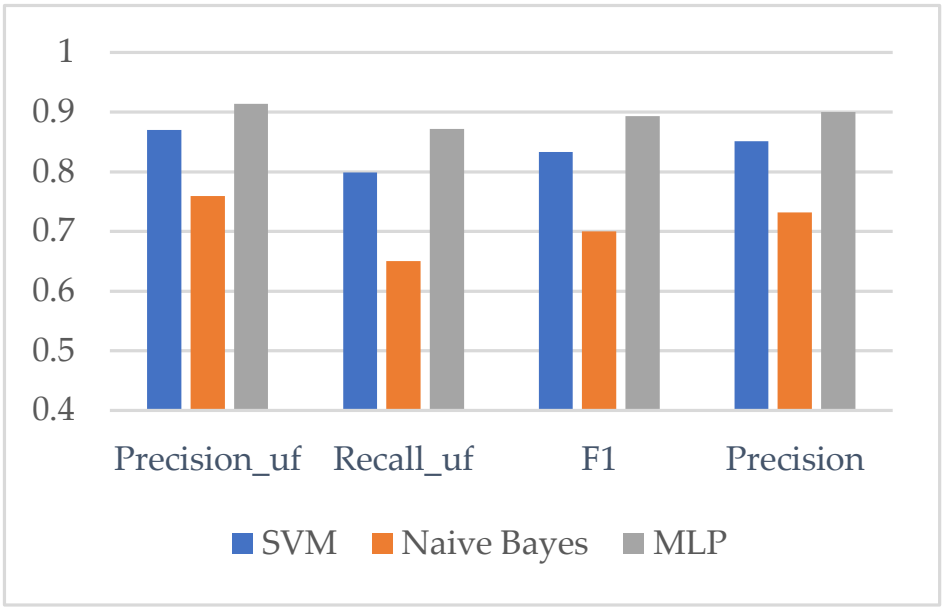


Figure 7. Comparison of the recognition effect under the optimal feature representation of three classifiers.

5. Conclusions

With the booming development of e-commerce platforms, the number of orders and online reviews have exploded. At the same time, the problem of identifying the usefulness of online reviews needs to be solved. Therefore, it is focused on how to efficiently identify the usefulness of online reviews. Aiming at the feature extraction stage, a feature extraction method is proposed based on Grounded theory, which considers the purchasing decision-making process of consumers. The feature extraction method represents the usefulness of online reviews in five stages: demand generation, information collection, product evaluation, purchase decision, and post-purchase behavior. At the classifier selection stage, the MLP neural network is selected as the classifier to identify the usefulness of online reviews. Through multiple sets of horizontal and longitudinal comparison experiments, the influence degree and necessity of the proposed features on the usefulness of online reviews are analyzed longitudinally so as to obtain the best combination of features, which are respectively trained on the three classifiers for horizontal comparison. The experimental results show that the feature extraction method, considering the consumer purchase decision process, can effectively improve the recognition effect of the classifier. Through

the training results of the optimal feature combination on the three classifiers, it is shown that the feature construction method based on the Grounded theory considering the consumer purchase decision-making process with the MLP model has high adaptability and interpretability, and the recognition effect has improved to a large extent.

The method proposed in this paper still has some shortcomings and needs to be perfected in future work. First of all, the paper studies the identification problem of the usefulness of online reviews, which is only a classification problem. It can also further conduct a quantitative analysis of the usefulness of online reviews. Secondly, for feature selection, online comments not only have text data but also data in the form of pictures, which can also be added to the idea of image information features in future research. For the selection of models, there are still other neural network models that can analyze the semantic aspects of short texts, and the classification effect can be verified in future experiments. Finally, there is also room for research on the application of the usefulness of online reviews. For example, the impact of the usefulness of online reviews on product sales and other practical problems facing enterprises remain to be solved.

Author Contributions: Methodology, J.H.; Data curation, J.H.; Writing—original draft, J.H.; Writing—review & editing, A.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The research data in this paper are obtained by crawling JD.com (accessed on 5 November 2022) website.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shin, S.; Chung, N.; Zheng, X.; Koo, C. Assessing the Impact of Textual Content Concreteness on Helpfulness in Online Travel Reviews. *J. Travel Res.* **2019**, *58*, 579–593. [CrossRef]
2. Kostyra, D.; Reiner, J.; Natter, M.; Klapper, D. Decomposing the effects of online customer reviews on brand, price, and product attributes. *Int. J. Res. Mark.* **2016**, *33*, 11–26. [CrossRef]
3. Purnawirawan, N.; Pelsmacker, P.D.; Dens, N. Balance and sequence in online reviews: How perceived usefulness affects attitudes and intentions. *J. Interact. Mark.* **2012**, *26*, 244–255. [CrossRef]
4. Chen, H.N.; Huang, C.Y. An investigation into online reviewers' behavior. *Eur. J. Mark.* **2013**, *47*, 1758–1773. [CrossRef]
5. Cheung, C.M.K.; Lee, M.K.O. What drives consumers to spread electronic word of mouth in online consumer-opinion platforms. *Decis. Support Syst.* **2012**, *53*, 218–225. [CrossRef]
6. Liu, X.; Wang, G.A.; Fan, W.; Zhang, Z. Finding Useful Solutions in Online Knowledge Communities: A Theory-Driven Design and Multilevel Analysis. *Inf. Syst. Res.* **2020**, *31*, 731–752. [CrossRef]
7. Lee, S.G.; Trimi, S.; Yang, C.G. Perceived Usefulness Factors of Online Reviews: A Study of Amazon.com. *J. Comput. Inf. Syst.* **2018**, *58*, 344–352. [CrossRef]
8. Malik, M.; Hussain, A. An analysis of review content and reviewer variables that contribute to review helpfulness. *Inf. Process. Manag.* **2018**, *54*, 88–104. [CrossRef]
9. Ketron, S. Investigating the effect of quality of grammar and mechanics (QGAM) in online reviews: The mediating role of reviewer crediblity. *J. Bus. Res.* **2017**, *81*, 51–59. [CrossRef]
10. Cheng, M.; Xin, J. What do Airbnb users care about? An analysis of online review comments. *Int. J. Hosp. Manag.* **2019**, *76*, 58–70. [CrossRef]
11. Rullah, R.; Zeb, A.; Kim, W. The impact of emotions on the helpfulness of movie reviews. *J. Appl. Res. Technol.* **2015**, *15*, 359–363.
12. Ren, G.; Hong, T. Examining the relationship between specific negative emotions and the perceived helpfulness of online reviews. *Inf. Process. Manag.* **2019**, *56*, 1425–1438. [CrossRef]
13. Zhang, W.; Xu, H.; Wan, W. Weakness finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. *Expert Syst. Appl.* **2012**, *39*, 10283–10291. [CrossRef]
14. BSparks, A.; So, K.K.F.; Bradley, G.L. Responding to negative online reviews: The effects of hotel responses on customer inferences of trust and concern. *Tour. Manag.* **2016**, *53*, 74–85. [CrossRef]
15. Guo, B.; Zhou, S. What makes population perception of review helpfulness: An information processing perspective. *Electron. Commer. Res.* **2016**, *17*, 585–608. [CrossRef]

16. Weimer, M.; Gurevych, I. Predicting the perceived quality of web forum posts. In Proceedings of the RANLP 2007, Bororets, Bulgaria, 27–29 September 2007; pp. 643–648.
17. Li, S.T.; Pham, T.T.; Chuang, H.C. Do reviewers' words affect predicting their helpfulness ratings? Locating helpful reviewers by linguistics styles. *Inf. Manag.* **2019**, *56*, 28–38. [CrossRef]
18. Lee, P.J.; Hu, Y.H.; Lu, K.T. Assessing the helpfulness of online hotel reviews: A classification-based approach. *Telemat. Inform.* **2018**, *35*, 436–445. [CrossRef]
19. Ou, W.; Huynh, V.N.; Sriboonchitta, S. Training attractive attribute classifiers based on opinion features extracted from review data. *Electron. Commer. Res. Appl.* **2018**, *32*, 13–22. [CrossRef]
20. Chen, C.C.; Tseng, Y.D. Quality evaluation of product reviews using an information quality framework. *Decis. Support Syst.* **2011**, *50*, 755–768. [CrossRef]
21. Hu, H.; Li, P.; Chen, Y. Biterm-based multilayer perceptron network for tagging short text. In Proceedings of the 2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Siem Reap, Cambodia, 15–17 July 2015; pp. 212–217.
22. Lee, S.; Choeh, J.Y. Predicting the helpfulness of online reviews using multilayer perceptron neural networks. *Expert Syst. Appl.* **2014**, *41*, 3041–3046. [CrossRef]
23. Qader, W.A.; Ameen, M.M.; Ahmed, B.I. An Overview of Bag of Words; Importance, Implementation, Applications, and Challenges. In Proceedings of the 2019 International Engineering Conference (IEC), Erbil, Iraq, 23–25 June 2019; pp. 200–204.
24. Ngo-Ye, T.L.; Sinha, A.P. The influence of reviewer engagement characteristics on online review helpfulness: A text regression model. *Decis. Support Syst.* **2014**, *61*, 47–58. [CrossRef]
25. Yu, C.T.; Salton, G.M. On the construction of effective vocabularies for information retrieval. *ACM SIGIR Forum* **1973**, *9*, 48–60.
26. Chen, R.Y.; Zheng, Y.T.; Xu, W. Secondhand seller reputation in online markets: A text analytics framework. *Decis. Support Syst.* **2018**, *108*, 96–106. [CrossRef]
27. Kokkodis, M.; Ipeirotis, P.G. Reputation transferability in online labor markets. *Manag. Sci.* **2016**, *62*, 1687–1706. [CrossRef]
28. Rout, J.K.; Singh, S.; Jena, S.K. Deceptive review detection using labeled and unlabeled data. *Multimed. Tools Appl.* **2017**, *76*, 3187–3211. [CrossRef]
29. Zhang, W.; Du, Y.; Yoshida, T.; Wang, Q. DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network. *Inf. Process. Manag.* **2018**, *54*, 576–592. [CrossRef]
30. Noekhah, S.; Salim, N.B.; Zakaria, N.H. A novel model for opinion spam detection based on multi-iteration network structure. *Adv. Sci. Lett.* **2018**, *24*, 437–1442. [CrossRef]
31. Le, Q.V.; Mikolov, T. Distributed representations of sentences and documents. In Proceedings of the 31st International Conference on Machine Learning, Beijing, China, 21–26 June 2014; Volume 23, pp. 1188–1196.
32. Zhang, H.; Ji, D.; Ren, Y.; Zhang, H. Finding deceptive opinion spam by correcting the mislabeled instances. *Chin. J. Electron.* **2015**, *24*, 52–57.
33. Zhang, Y.; Lin, Z. Predicting the helpfulness of online product reviews: A multilingual approach. *Electron. Commer. Res. Appl.* **2018**, *27*, 1–10. [CrossRef]
34. Malik, M.; Hussain, A. Helpfulness of product reviews as a function of discrete positive and negative emotions. *Comput. Hum. Behav.* **2017**, *73*, 290–302. [CrossRef]
35. Zheng, X.; Zhu, S.; Lin, Z. Capturing the essence of word-of-mouth for social commerce: Assessing the quality of online e-commerce reviews by a semi-supervised approach. *Decis. Support Syst.* **2013**, *56*, 211–222. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.