mathematics

MDPI

Article

A Large-Scale Reviews-Driven Multi-Criteria Product Ranking Approach Based on User Credibility and Division Mechanism

Wenzhi Cao, Xingen Yang and Yi Yang *

School of Frontier Crossover Studies, Hunan University of Technology and Business, Changsha 410205, China; c.wenz.hutb@gmail.com (W.C.); yangxg02@163.com (X.Y.)
* Correspondence: yangshijiazu@my.swjtu.edu.cn

Abstract: Massive online reviews provide consumers with the convenience of obtaining product information, but it is still worth exploring how to provide consumers with useful and reliable product rankings. The existing ranking methods do not fully mine user information, rating, and text comment information to obtain scientific and reasonable information aggregation methods. Therefore, this study constructs a user credibility model and proposes a large-scale user information aggregation method to obtain a new product ranking method. First, in order to obtain the aggregate weight of large-scale users, this paper proposes a consistency modeling method of text comments and star ratings by mining the associated information of user comments, including user interaction information and user personalized characteristics information, combined with sentiment analysis technology, and then constructs a user credibility model. Second, a double-layer group division mechanism considering user regions and comment time is designed to develop the large-scale group ratings aggregation approach. Third, based on the user credibility model and the large-scale ratings aggregation approach, a product ranking method is developed. Finally, the feasibility and effectiveness of the proposed method are verified through a case study for automobile ranking and a comparative analysis is furnished. The analysis results of the application case of automobile ranking show that there is a significant difference between the ranking results obtained by the ratings aggregation method based on the arithmetic mean and the ranking results obtained by this method. The method in this study comprehensively considers user credibility and group division, which can be reflected in user aggregation weights and the group aggregation process, and can also obtain more scientific and reasonable decision results.

Keywords: multi-criteria; online reviews; user credibility; product ranking

MSC: 90B50

Citation: Cao, W.; Yang, X.; Yang, Y. A Large-Scale Reviews-Driven Multi-Criteria Product Ranking Approach Based on User Credibility and Division Mechanism. *Mathematics* **2023**, *11*, 2952. https://doi.org/10.3390/math11132952

Academic Editors: James Liou and Artūras Kaklauskas

Received: 15 May 2023
Revised: 20 June 2023
Accepted: 27 June 2023
Published: 1 July 2023

1. Introduction

With the increasing popularity and development of the Internet, e-commerce and review platforms have emerged. These platforms allow consumers to post online reviews about their experiences and opinions on various product criteria, including quality and functionality. Online reviews offer valuable information to consumers who lack expert knowledge of the products they wish to purchase and help inform their purchase decisions. However, the abundance of online review data and the fake reviews posted by malicious users make it challenging for consumers to determine the authenticity of reviews and make informed purchase decisions. Thus, it is crucial to identify credible reviews from the vast amount of review data available. This study focuses on the credibility of online product reviews.

To date, numerous experts and scholars have conducted extensive research on the credibility of online product reviews, and they applied their findings to evaluate information credibility across different review platforms and social networks [1,2]. These studies

involve developing credible models for online reviews to validate their reliability. Researchers examine the factors that influence the credibility of reviews and investigate the distribution patterns of ratings. For instance, Verma et al. proposed a credibility model for online reviews by analyzing the influencing factors of content, communicator, context, and consumer [3]. They also explored credibility variables associated with these factors and established a causal relationship between the variables and the credibility of online reviews by exploring 22 propositions. Banerjee et al. proposed a theoretical model of reviewer credibility based on dimensions such as positivity, engagement, experience, reputation, competence, and social connections [4]. They employed robust regression to determine the significance of these factors. Furthermore, Sun et al. developed a reputation rating method based on user rating bias and rating characteristics [5]. They discovered that the ratings provided by reliable users exhibit a peak distribution, whereas those provided by malicious users are substantially biased. On the other hand, scholars have presented research on the credibility of reviews from review text, ratings, user information, etc. Xiang et al. analyzes the reliability of review data from the aspects of review text semantic features, sentiment, and ratings [6]. Meel et al. put forward a holistic view of how the information is being weaponized to fulfil the malicious motives and forcefully make a biased user perception about a person, event, or firm [7]. In addition, review text and ratings are also key factors in the study of the credibility of reviews. For instance, Hazarika pointed out that there is an inconsistency between the review text and the rating in product reviews [8]. Almansour et al. proposed to build a system by fusing review text, ratings, and sources [9]. Lo et al. studied the credibility of reviews from the consistency of review text and ratings [10]. However, these studies have rarely introduced the latest text analysis tools to analyze the sentiment of texts. Additionally, traditional statistical and mathematical approaches are not sufficiently precise and efficient in reviewing texts. In this study, we employ the latest text analysis tool to develop a model based on sentiment analysis to assess the usefulness of comments.

Sentiment analysis (SA) technology is an important means of obtaining emotional tendencies in large-scale comments. SA is the process of gathering and analyzing people's opinions, thoughts, and impressions regarding various topics, products, subjects, and services [11]. SA involves analyzing text or speech using computer techniques to determine the sentiment or emotional state within the text [12,13]. The latest text analysis model, which can be pretrained on large-scale text data, can be fine-tuned to address sentiment analysis tasks [14]. Yang et al. proposed a new SA model, SLCABG [15], and the related experimental results show that the model can effectively improve the performance of text SA. At the same time, SA is also used to analyze large-scale review sets. For instance, Haque et al. used a supervised learning method on a large-scale Amazon dataset to polarize it and achieve satisfactory accuracy [16]. Guo et al. identified the key dimensions of customer service voiced by hotel visitors using a data mining approach, latent Dirichlet analysis (LDA) [17], and the related set included 266,544 online reviews for 25,670 hotels located in 16 countries. Thus, this research employs the latest text analysis model, BERT, to conduct SA on comment text, and quantify the SA of the comment text to five levels corresponding to user ratings. This study constructs a sentiment score acquisition method for text comments based on manually annotated sentiment training libraries and combined with the BERT model.

Currently, there are numerous Internet-based review platforms exclusively for automobile brands, which offer rich and standardized data, providing information resources for useful research. Therefore, this research focuses on automobiles as the research subject and develops a model of review usefulness. Research on ranking decisions for automotive products is divided into two primary areas of investigation, namely rating-driven ranking decisions and text review-driven ranking decisions. In the research on rating-based ranking decisions, distributed linguistic term sets remain the core representation tool for transforming rating information. PROMETHEE-II and TODIM are extended to the linguistic term set environment to propose product ranking approaches [18,19]. In the research on rank-

ing decisions based on text reviews, by combining the sentiment ratings and star ratings based on the output of the DUTIR sentiment dictionary, scholars developed a PageRank algorithm product ranking technique based on a directed graph model [20]. Additionally, scholars have addressed the accuracy problem of sentiment intensity recognition by developing a ranking decision approach based on ideal solutions and introducing two interval type fuzzy sets [21]. Scholars used sentiment analysis techniques to output five types of sentiment ranking by considering the advantages of probabilistic linguistic term sets in characterizing sentiment tendencies and their distribution forms. They combined these sentiment rankings with TODIM and evidence theory to construct a related product ranking decision approach [22]. The above studies primarily aggregate group wisdom knowledge in large-scale online reviews from a statistical viewpoint, using fuzzy sets, linguistic term sets, and other representational methods. However, they have not fully combined the current advanced text analysis techniques to analyze review texts, or considered word-of-mouth credibility and aggregation weights of heterogeneous individuals, and the inconsistency between text reviews and star ratings, as well as reviewer information disclosure, to conduct research. Thus, the current identification of false reviews is not precise enough. On the other hand, the current research does not consider the aggregation of large-scale ratings from group users, which is easily affected by large-scale fake reviews. To address this issue, this study proposes the construction of a user credibility model, and applies the text analysis model to analyze the user credibility weight of each review during the process of aggregating the wisdom knowledge of group online reviews. A group user score aggregation method is also built to calculate the comprehensive score of automobile brands. Based on this analysis, an automobile ranking decision method is developed.

In summary, this study examines how to weaken the influence of fake reviews and extract real and credible reviews for product ranking, and proposes a user credibility model based on the consistency of review sentiment orientations and ratings to solve the problem of difficult automobile ranking decisions. Compared with existing approaches, this approach examines the credibility of reviews in terms of online review text content, performs sentiment analysis on the review text, uses the high accuracy of the text analysis model, quantifies the sentiment intensity of each review text, and further analyzes the user disclosure information to compute user credibility weights. The contributions of this approach can be summarized as follows.

(1) A user weight model based on user disclosure information is constructed, which includes authentication information, interaction information, and driving information. Then, the sentiment analysis techniques and expert knowledge are used to measure the degree of consistency between ratings and text comments, and a comprehensive user weight calculation model is developed.
(2) The large-scale group ratings aggregation approach based on user region and comment time division is proposed, and a product ranking method is developed.

2. Problem Description and Data Description

Many consumers encounter difficulties in selecting the appropriate automobile for themselves because of their lack of professional experience and knowledge about automobiles. To address this issue, this research proposes a review credibility model based on user disclosure information and consistency to examine the aggregation of automobile reviews, and compute and rank the overall rating of each automobile brand. The notation defined below is employed to denote the aggregation and variables for this problem.

$X = \{x_1, x_2 \dots x_n\}$: n represents the number of alternative target automobiles, and x_i represents the i-th target automobile, $i = 1, 2, \dots, n$.

$A = \{a_1, a_2, \dots, a_8\}$: The data analysis demonstrates that there are eight automobile criteria and a_j is the j-th criteria of the automobile, $j \in [1, 8]$, corresponding to {space, power, control, electricity/fuel consumption, comfort, exterior, interior, value for money}.

$U = \left\{u_i^1, u_i^2, \ldots, u_i^{K_i}\right\}$: u_i^k represents the k-th user who commented on the target automobile x_i, and K_i represents the number of users who commented on the target automobile x_i.

$T = \left\{t_{ij}^1, t_{ij}^2, \ldots, t_{ij}^{K_i}\right\}$: t_{ij}^k represents the text of a comment made by user e_i^k on a criterion a_j of target automobile x_i. In this study, a user can only make one comment on an automobile in the dataset. Therefore, the number of comments is equal to the number of users making comments, $k = 1, 2, \ldots, K_i$.

$S = \left\{s_{ij}^1, s_{ij}^2, \ldots, s_{ij}^k\right\}$: s_{ij}^k represents the star rating of user e_i^k for a criterion a_j of the target automobile x_i, $k = 1, 2, \ldots, K_i$.

Finally, the automobile's overall rating is determined from the above dataset by computing the mapping function: $G = F(X, A, T, S)$, G represents the composite rating and F represents the mapping function.

3. A User Credibility Model Based on Consistency and User Disclosure Information

In this section, a user credibility model based on consistency and user disclosure information is developed to compute user credibility weights, as shown in Figure 1.

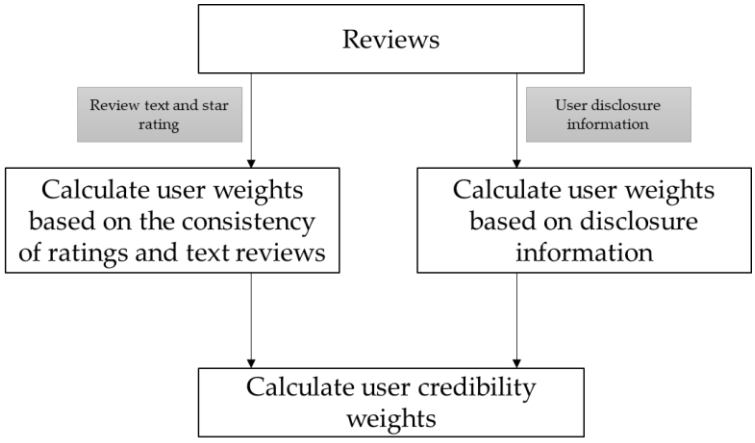

Figure 1. User credibility weights calculation flow chart.

The target automobile $X = \{x_1, x_2 \ldots x_n\}$ is filtered according to constraints such as price, budget and model, and relevant comments are crawled on the Autohome.com forum, using Python crawlers, which comprise comment text, ratings, and user disclosure information. Python and SQL tools are e employed to eliminate illegal comment data, and the data are normalized and stored in the format shown in Table 1.

Table 1. Comment text and rating scale.

Automotive Indicators	Space (a_1)	Power (a_2)	Control (a_3)	Consumption (a_4)	Comfort (a_5)	Exterior (a_6)	Interior (a_7)	Vfm (a_8)
Comment text	t_{i1}^k	t_{i2}^k	t_{i3}^k	t_{i4}^k	t_{i5}^k	t_{i6}^k	t_{i7}^k	t_{i8}^k

3.1. User Weights Based on the Consistency of Ratings and Text Reviews

Step 1. Building automobile text review sentiment dataset

(1) Method construction

Step 1.1 Obtaining text review training set

The automobile reviews are crawled on the Autohome.com forum using Python crawlers, and prepossessing is performed to remove garbled characters, missing information, and other data that do not meet the specifications. At the same time, according to the types of automobiles currently on the market, they are divided into new energy automobiles and gasoline automobiles, and the review data are screened according to the score

distribution to make the score distribution even. Finally, the data matrix $\overline{M} = (D_h)_{H_0 \times 1}$ is obtained.

Step 1.2 Obtaining text review training set with expert sentiment values

The obtained automobile review sample library is uploaded to the built automobile review labeling system to ensure that the data labeling conforms to the specification. At the same time, in order to obtain accurate data, L experts were hired to mark the comment text. Each review text is annotated once by L experts. The experts formulate the rating rules through discussion, and then mark the emotional strength of the comment text according to the rules, as illustrated in Figure 2. Finally, the data matrix $M = \left(D_h^l\right)_{H_0 \times L}$ is obtained.

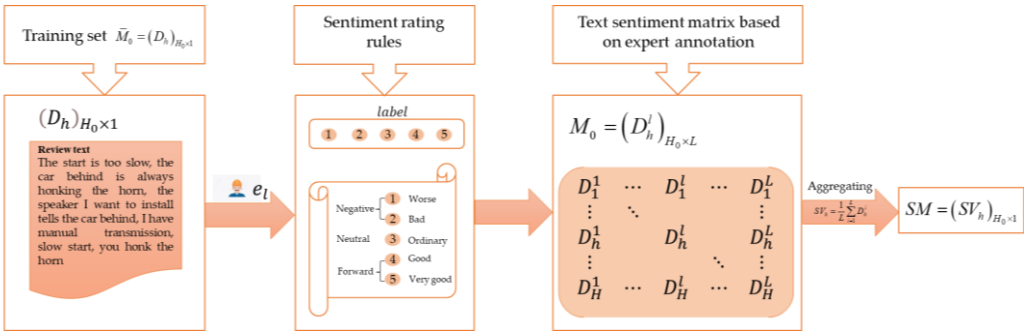

Figure 2. Automobile review data expert labeling flow diagram.

Step 1.3 Aggregating of expert sentiment values for text review training set

The variance of the marked sentiment score of each comment is calculated and a threshold set to filter the data, so as to maintain the stability of the comment data. Then, the average value of the expert sentiment values SV_h for text review D_h is calculated as the emotional strength of the label.

$$SV_h = \frac{1}{L}\sum_{l=1}^{L} D_h^l \tag{1}$$

(2) Method execution

Step 1.1 Obtaining text review training set

Using crawlers to obtain more than 4000 new energy automobile reviews and more than 4000 gasoline automobile reviews in the forum, a total of more than 8000 automobile review data were obtained. Review data including missing information and garbled characters were removed through preprocessing, and finally 7361 automobile review data were obtained.

$$\overline{M}_0 = (D_h)_{7361 \times 1} \tag{2}$$

Step 1.2 Obtaining text review training set with expert sentiment values

Eight experts were hired to discuss and formulate the sentiment rating rules for automobile review texts; they logged in to the annotation system to annotate the above-mentioned obtained review texts. Finally the matrix M_0 of the text emotional annotations of the eight experts was obtained.

$$M_0 = \left(D_h^l\right)_{7361 \times 8} \tag{3}$$

Step 1.3 Aggregating of expert sentiment values for text review training set

First, the variance of the sentiment labeling rating of eight experts for each comment was calculated, and the labeled data with a variance greater than 2 were removed. Then, the sentiment annotation matrix of the above-mentioned automobile review text was aggregate. Finally, 6563 automobile review text sentiment annotation data were obtained to form a

sentiment analysis model training dataset. This included eight automobile attributes, where each attribute has a comment text. Finally 52,504 texts were obtained. The distribution of sentiment intensity is shown in Table 2.

$$SV_h = \frac{1}{8}\sum_{l=1}^{8} D_h^l$$
$$SM = (SV_h)_{6563\times 1} \tag{4}$$

Step 2. Training text review sentiment values based on BERT sentiment analysis model

Table 2. Distribution of sentiment analysis model training dataset.

Sentiment Rating	Dataset
1	9888
2	9200
3	10,952
4	12,216
5	10,248

Step 2.1 Building the automobile review sentiment analysis model based on BERT

In this study, the deep learning framework pytorch was used to build the automobile review sentiment analysis model. The main process is shown as the Figure 3.

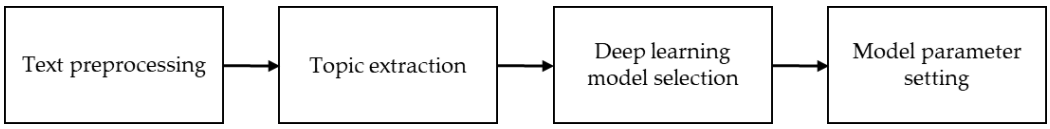

Figure 3. Sentiment analysis model construction flow diagram.

The process was as follows: First, remove stop words, stemming, and other preprocessing of the automobile review text. Then extract the topics in the comment text, and then select an efficient deep learning model according to the short text processing effect. This study used the BERT model to build a sentiment analysis model.

The overall framework of the model is a stack of encoders with multiple layers of transformers, with a single-layer structure, as illustrated in Figure 4. $E_1, E_2, \ldots, E_N$ is the embedded word after the embedding process, *Trm* represents the transformer encoding layer of the host, and $T_1, T_2, \ldots, T_N$ represents the word encoding after the multilayer process. The affective intensity prediction process includes embedding, multi-head self-attention, feedforward, and layer normalization.

At the same time, the model parameters are preliminarily set according to the length and data volume of the automobile review text data. The parameter settings of the model are shown in Table 3.

Table 3. BERT model parameters.

Parameters	Name	Value
Max_length	Maximum text length	512
Epoch	Training batches	5
Batch_size	Number of batch gradients down	16

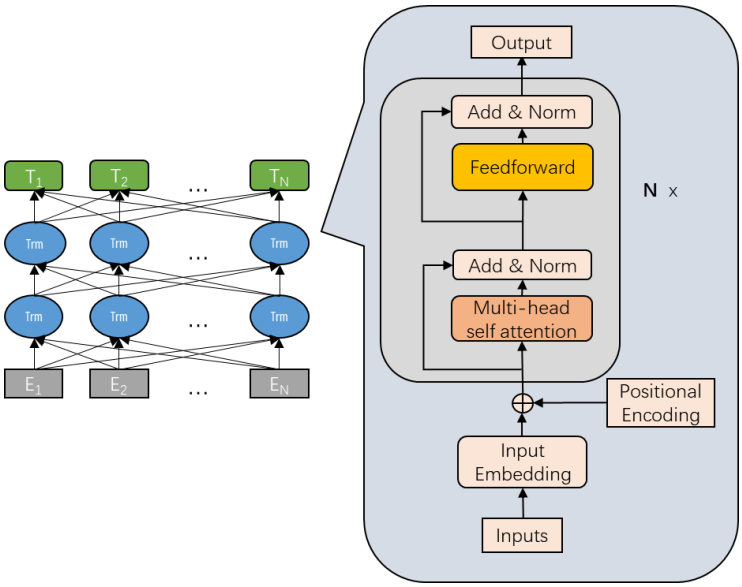

Figure 4. BERT sentiment analysis model structure diagram.

Step 2.2 Training text review sentiment values

In this study, the aforementioned prepared training dataset was employed to train the BERT model based on the pytorch framework to obtain a sentiment analysis model with eight automobile features. Figure 5 shows the process of sentiment analysis model training. Table 4 shows the accuracy of the model.

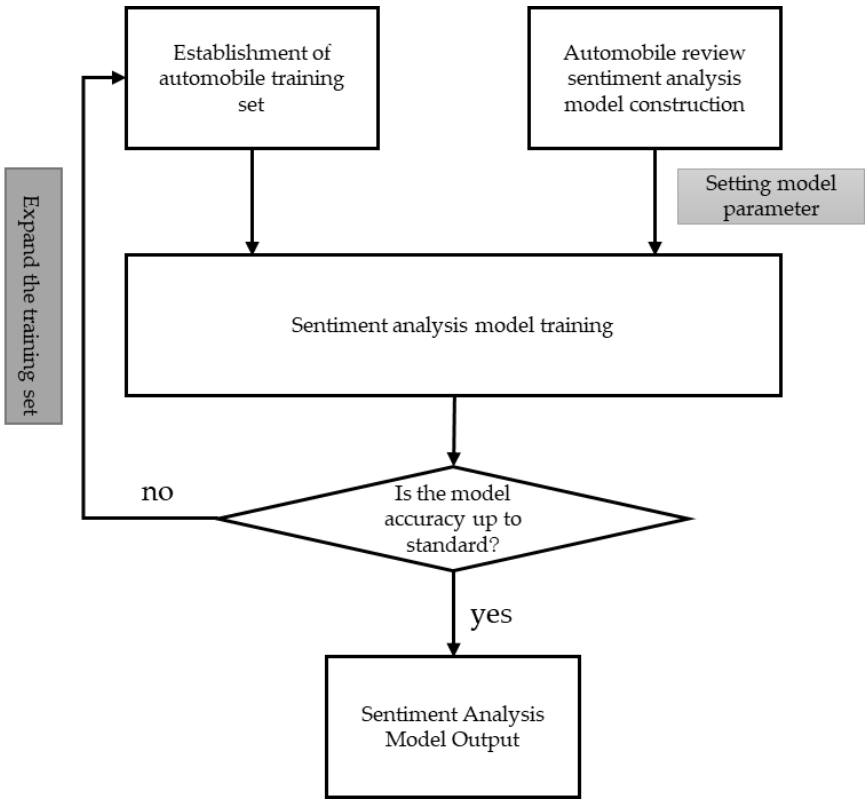

Figure 5. Flow chart of automobile review sentiment analysis model training.

Table 4. Sentiment analysis model accuracy.

Automobile Properties	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
Model accuracy	89%	81%	80%	76%	76%	87%	77%	85%

Step 3. Predicting text review sentiment values

To quantitatively predict the sentiment intensity of all comments, the trained sentiment analysis model was employed. The model predicts an emotional intensity based on the input utterances using the above module, as illustrated in Figure 6. The top part [CLS] is the identification of the beginning of the text, which contains the information of the entire sentence, but has no real meaning. The output of all other positions will be biased by placing more weight on the weight of their position, so the first place is output. Then, this is processed by the linear classifier module to predict a label.

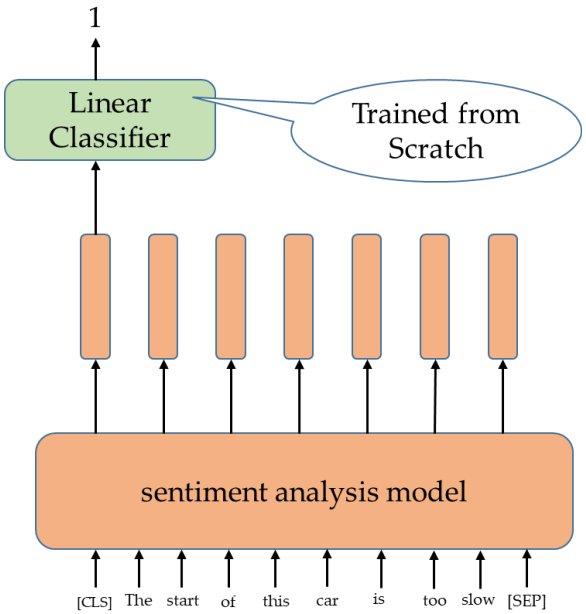

Figure 6. Sentiment intensity prediction.

The sentiment intensity S of the a_j ($j \in [1, 8]$) indicator review text of user e_i^k for an automobile x_i is obtained by examining the review text t_{ij}^k with the sentiment analysis model. This yields the sentiment intensity prediction matrix $S_p = \left\{ s_{pij}^1, s_{pij}^2, \ldots, s_{pij}^k \right\}$ for all reviews of the automobile x_i. The granularity s_{pij}^k is the same as the user star rating s_{ij}^k, which takes on a range of values $\{1, 2, 3, 4, 5\}$.

Step 4. User weights based on the consistency of ratings and text reviews

Generally, fake reviews are characterized by inconsistencies between the star rating s_{ij}^k and the sentiment intensity s_{pij}^k of their review texts. Thus, this section proposes approaches to compute the consistency weights of the two factors. The higher the weight, the higher the degree of consistency and the more credible the review. The computation process is as follows:

$$r_{ij}^k = \frac{1}{4}\left(4 - \left|s_{ij}^k - s_{pij}^k\right|\right) \tag{5}$$

The consistency weight matrix is obtained for online reviews of automobiles, $R = \left[r_{ij}^k\right]$, $r_{ij}^k \geq 0$.

3.2. User Weights Based on Disclosure Information

User disclosures include whether they are authenticated, their interaction index (number of replies, number of likes, and number of views), as illustrated in Table 5, and their daily travel rate, expressed as $I = \{I_1, I_2, I_3\}$. User disclosures provide a side view of the credibility of reviews.

Table 5. User disclosure information form.

Automotive Series	User	Certified (I_1)	Interaction Index (I_2)			Usage Rate (I_3)	
			Views	Points	Replies	Daily Travel Rate	Mileage
x_i	e_i^k	Yes/No	p_1	p_2	p_3	q_1	q_2

Step 1. Certification indicators

Users who post by word of mouth on the platform can be categorized as certified and non-certified owners. Certified owners are users who have purchased the automobile. The platform's automobile owner certification requires uploading personal information such as certified automobile models and driving licenses. This information is audited by the platform. In contrast, uncertified owners may be users who have not purchased the automobile in question. Reviews from certified owners are more credible. The weights $w_{iI_1}^k$ are computed, as illustrated in the following formula:

$$w_{iI_1}^k = \begin{cases} 1, Certified \\ 0.5, Non-certified \end{cases} \tag{6}$$

Step 2. Travel rate indicator

The trip rate weighting is determined by combining the daily trip rate and mileage. Research suggests that many indicators of an automobile require sufficient mileage to test its performance. Thus, the higher the usage of the automobile, the deeper the user's experience of the automobile's performance and the more credible the reviews they publish. The usage rate of an automobile can be computed based on its daily driving rate and mileage driven. Based on statistics, the daily driving rate of the automobile is around $15 \sim 78$ km/d, and this study divides the interval accordingly to compute the daily driving rate weights, as shown in Table 6. Furthermore, statistics on mileage posted by word of mouth in automobile forums indicate that mileage is concentrated at $0 \sim 1000$ km. The higher the mileage, the lower the number of published word-of-mouth entries, according to which the following intervals are divided, as illustrated in Table 7.

Table 6. Daily travel rate weights rule.

Daily Travel Rate (km/d)	[78,+∞)	[57,78)	[36,57)	[15,36)	[0,15)
q_1	1.0	0.8	0.6	0.4	0.2

Table 7. Mileage weights rule.

Mileage (km)	[8000,+∞)	[5000,8000)	[3000,5000)	[1000,3000)	[0,1000)
q_2	1.0	0.8	0.6	0.4	0.2

The usage weights $w_{iI_2}^k$ can be computed by summing up the travel rate weights and the mileage weights as follows, where μ represents the automobile usage weight computation parameter:

$$w_{iI_2}^k = \mu q_1^{ik} + (1-\mu) q_2^{ik} \tag{7}$$

Step 3. Interactive indicators

In this study, word-of-mouth entries posted on automobile forums are viewed, liked, and replied to by other users, and the ratio of the sum of the three to the length of posting is called the interaction index. A higher interaction index indicates that the review is more recognized and considered more credible. To assign weights to the interaction index, the following intervals were computed, as illustrated in the following formula:

$$w_{iI_3}^k = \begin{cases} 1.0, I_3 \in [15199, \infty] \\ 0.8, I_3 \in [5424, 15199] \\ 0.6, I_3 \in [1769, 5424] \\ 0.4, I_3 \in [674, 1769] \\ 0.2, I_3 \in [0, 674] \end{cases} \tag{8}$$

Step 4. The weight of k-th user e_i^k for automobile x_i is given as follows based on the above three indicators:

$$f_i^k = \frac{1}{3}\left(w_{iI_1}^k + w_{iI_2}^k + w_{iI_3}^k\right) \tag{9}$$

3.3. User Comprehensive Weight Calculation

Fusing the consistency weight of online reviews with the user disclosure weight yields a user credibility weight for reviews c_{ij}^k. The formula is as follows, where μ is the parameter for computing the credibility weight of online reviews:

$$c_{ij}^k = \mu r_{ij}^k + (1-\mu) f_{ij}^k \tag{10}$$

where $f_{ij}^k = f_k^i (j = 1, 2, \cdots, 8)$.

4. Large-Scale Ratings Aggregation Based on Group User Division

This section proposes a multi-criteria ratings aggregation method for group users to address the issue of weakening the role of user reputation weights in large groups of users, thus weakening the impact of false reviews on the overall rating. The approach first divides users into multiple sets based on their purchase location. Then, all user sets are divided into sets based on their purchase time. User ratings are then computed using user reputation weights. Finally, the aggregation of all user sets is distributed to compute the overall rating.

Step 1. Group division method

Step 1.1 Group division method based on user geography

As the users of the platform are automobile owners from different regions of the country, their experience and needs of the automobile may vary. Therefore, the users of each automobile x_i purchase u_i^k are divided into eight collections based on geography. The geographical divisions of China are set up as

$$D = \{D_d | d = 0, 1, \ldots, 7\} = \{\text{No region, Northeast, North, Central, East, South, Southwest, Northwest}\},$$

The seven sets of users by geography are represented as

$$\cup_{d=0}^{7}\left\{u_i^{D_d}\right\}, u_i^{D_d} = \left\{u_i^{dk}\middle| k = 1, 2, \ldots, |D_d|\right\},\ d = 0, 1, \ldots, 7.$$

Step 1.2 Group division method based on the time of user comments

Time is a crucial factor that should not be overlooked, and reviews are even more time sensitive, with different references at different times. Additionally, the automobiles themselves are being updated and the purchase of services and prices are constantly changing. Thus, the study of reviews must also be approached according to different time periods. This study divides the collection based on geographical divisions using years as the re-

search step, and the difference between the earliest and latest reviews, n, as the research quotient. The time collection is $T = \{T_t | t = 1, 2, \ldots, n\}$. The set of users divided by geography and time can be expressed as $\cup_{d=1}^{7} \cup_{t=1}^{n} \left\{ u_i^{D_d T_t} \right\}$, $u_i^{D_d T_t} = \left\{ u_i^{dtk} \middle| k = 1, 2, \ldots, |D_d T_t| \right\}$, $d = 0, 1, \ldots, 7, t = 1, 2, \ldots, n$, where $|D_d T_t|$ represents the number of users who commented during time period t in region d. The group division structure is shown in Figure 7.

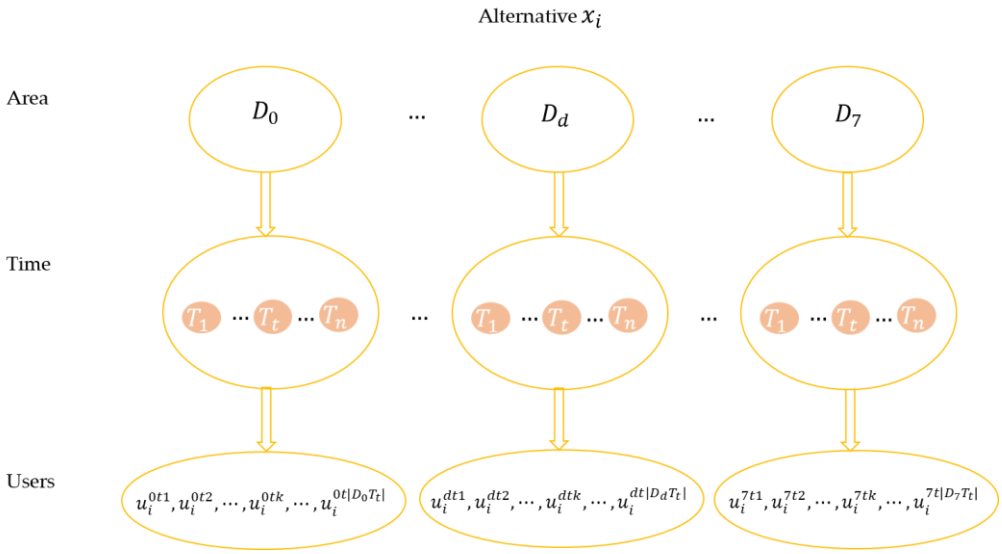

Figure 7. Group user division flowchart.

Each user is assigned a weight based on the set division of users, and the final credibility weight for each comment is u_{ij}^{dtk},$k = 1, 2, \ldots, |D_d T_t|$, which denotes the credibility weight of the a_j ($j \in [1, 8]$) indicator for the automobile x_i by the k-th user in year T_t under a geographical region D_d. This is then normalized to give c_{ij}^{dtk}.

$$c_{ij}^{dtk} = \frac{u_{ij}^{dtk}}{\sum_{k=1}^{|D_d T_t|} u_{ij}^{dtk}} \tag{11}$$

Step 2. Calculating the overall user rating

To further mitigate the impact of false reviews, the original star rating s_{ij}^{k} of each online review was arithmetically averaged with the predicted sentiment intensity s_{pij}^{k} of its text t_{ij}^{k} to obtain a new rating $\overline{S}_{ij}^{k}$ for each review on each criteria of the automobile.

$$\overline{S}_{ij}^{k} = \lambda s_{ij}^{k} + (1 - \lambda) s_{pij}^{k} \tag{12}$$

where $\overline{S}_{ij}^{k}$ denotes user u_i^k rating of indicator a_j for automobile x_i.

Step 3. Aggregating group user ratings

After the aforementioned group segmentation, the rating corresponding to user u_i^{dtk} is $\overline{S}_{ij}^{dtk}$,$\left\{ \overline{S}_{ij}^{dtk} | k = 1, 2, \ldots, |D_d T_t| \right\}$. The explanation of the related parameters is shown in Table 8.

Table 8. Explanation of formulas $\overline{S}_{ij}^{dtk}$.

Parameter	Meaning
d	The d-th purchase region
t	The t-th time period
k	The k-th user in the user set of the t-th time period in the d-th region
i	The i-th alternative automobile
j	The j-th feature of the automobile
$\overline{S}$	Composite score of text sentiment strength and raw rating

Finally, the rating is multiplied by the credibility weight and summed to obtain the final rating $\overline{S}_{ij}$.

$$\overline{S}_{ij} = \frac{1}{8}\sum_{d=0}^{7}\left(\frac{1}{n}\sum_{t=1}^{n}\left(\sum_{k=1}^{|D_dT_t|} c_{ij}^{dtk} \times \overline{S}_{ij}^{dtk}\right)\right) \tag{13}$$

where $k = 1, 2, \ldots, |D_dT_t|$, and $\overline{S}_{ij}$ represents the combined rating of all users of the a_j indicator for automobile x_i.

Step 4. Aggregating multi-criteria ratings

Finally, the eight indicators were aggregated to find the computed composite rating $\overline{S}_i$ for automobile x_i. w_j is the weight of criteria a_j.

$$\overline{S}_i = \sum_{j=1}^{8} w_j \overline{S}_{ij} \tag{14}$$

5. Product Ranking Methods

Step 1. Collect the data and structure it to obtain the comment dataset of the automobile.
Step 2. Obtain the user weights.

Step 2.1 Calculate the user weights based on information disclosure.
Step 2.2 Calculate the user weights based on the consistency of ratings and text reviews.
Step 2.3 Obtain the user comprehensive weight.

Step 3. Aggregate large-scale ratings.

Step 3.1 Group division based on user geography and comment time.
Step 3.2 Calculate the overall user rating based on ratings and emotional analysis value of text comments.
Step 3.3 Aggregate group user ratings.
Step 3.4 Aggregate multi-criteria ratings.

Step 4. The ranking results for alternative target automobiles $x_i (i = 1, 2, \ldots, n)$ are obtained based on the final overall ratings $\overline{S}_i (i = 1, 2, \ldots, n)$

$$x_{\sigma(i)} \succ x_{\sigma(i+1)} (i = 1, 2, \ldots, n)$$

where $\overline{S}_{\sigma(i)} \succ \overline{S}_{\sigma(i+1)} (i = 1, 2, \ldots, n)$.

6. Application of the Method

The popularity of e-commerce has resulted in the development of numerous review platforms. As a large commodity, the market for automobiles is huge, and this has resulted in the emergence of specialized automobile review platforms, such as AutoZone. These platforms offer reviews of almost all automobiles and provide comprehensive information, making them one of the most crucial sources of information for consumers. However, many consumers are plagued by false reviews because of their lack of automobile-related knowledge, making it challenging for them to make an informed choice. Thus, this section

is based on user disclosure information and a consistency user-credibility model analysis approach to assist consumers in making informed purchasing decisions.

Step 1. Determining product sets, criteria sets, and data acquisition

As shown in Table 9, six alternative target automobiles were selected based on consumers' budgets and models. All criteria of each automobile brand were analyzed, with the computation process detailed below.

A Python crawler was written to crawl the review data of the corresponding target automobile in AutoZone as of 30 December 2022, as each automobile brand was released at a different time, and thus its review count was different. The crawled data were structured using the Python program. Table 9 shows the review data obtained after removing illegal review data, such as garbled codes and null values.

Table 9. Alternative target automobile brands.

Automobile Model	Number of Reviews	Overall Rating	Review Time	Price (in RMB)
GAC Toyota-Hanlanda (x_1)	3762	4.51	2006–2022	26.88–34.88
Dongfeng Nissan-Xuan Yi (x_2)	2689	4.53	2016–2022	9.98–17.49
SAIC Volkswagen-ToucanL (x_3)	3329	4.52	2018–2022	19.90–28.38
Dongfeng Honda-XR-V (x_4)	2392	4.86	2015–2022	13.29–15.29
SAIC-Volkswagen-Polo (x_5)	3647	4.71	2014–2022	9.09–12.49
FAW-Volkswagen-Golf (x_6)	3756	4.86	2014–2022	12.98–22.98

Step 2. Use a user credibility model based on consistency and user disclosures to obtain user credibility weights.

Step 2.1 User weights based on the consistency of ratings and text reviews

Step 2.1.1 Sentiment analysis based on text comments

The trained sentiment analysis model was employed to predict the sentiment intensity of the preprocessed online review text for each automobile. A new rating, denoted s^k_{pij}, was obtained for each of the eight automobile features of each online review, as illustrated in Table 10.

Table 10. Predicted emotional intensity data table for Dongfeng Nissan-Xuan Yi (x_2).

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
0	4	4	3	4	4	5	3	4
1	5	3	3	5	2	4	2	5
2	4	4	5	5	5	5	5	4
3	4	4	4	3	5	5	4	4
4	5	4	3	3	5	5	5	5

Step 2.1.2 Consistency weights based on ratings and text

After obtaining the predicted sentiment intensity s^k_{pij} of all the review texts, a consistency analysis was conducted with their corresponding original star ratings to determine a consistency weight r^k_{ij}. Table 11 illustrates the data for the consistency weighting r^k_{ij} component of the Dongfeng Nissan-Xuan Yi (x_2).

Table 11. Consistency weights of ratings and texts for Dongfeng Nissan-Xuan Yi (x_2).

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
0	4	3	3	4	4	3	4	4
1	3	3	4	3	4	4	4	3
2	3	3	3	3	3	3	4	1
3	4	2	4	4	4	4	3	3
4	3	4	4	1	4	4	4	4

Step 2.2 Calculation of weights based on user information disclosure

The characteristic information weights of all review publishers for each automobile were computed. The authentication weight (I_1), usage rate weight (I_2), and interaction index weight (I_3) were computed using the computation approach proposed in the previous section. The parameters $\mu = 0.3$ for computing automobile usage rate weight were set, and aggregation was used to obtain the characteristic information weight f_i^k for each review publisher, as illustrated in Table 12 for the partial data of Dongfeng Nissan-Henyi.

Table 12. Information on indicators for calculating the weighting of the Nissan-Xuan Yi user feature.

No.	Certified	Purchase Time	Comment Time	Mileage	p_1	p_2	p_3	I_1	I_2	I_3	f_i^k
0	NO	2017-08	2017-08	1500	2035	0	0	0.5	0.6	0.3	1.4
1	YES	2016-10	2017-08	8788	1919	0	0	1.0	0.6	0.6	2.2
2	NO	2017-08	2017-08	998	673	0	1	0.5	0.4	0.2	1.1
3	NO	2017-08	2017-09	760	2529	0	0	0.5	0.6	0.2	1.3
4	NO	2018-01	2018-01	33	134,664	166	95	0.5	1.0	0.2	1.7

Step 2.3 Combined user weighting calculation

Combining user feature information weights and review consistency weights, and setting $\mu = 0.7$, yields a credibility weight c_{ij}^k for each review about all automobile criteria for each automobile brand, as illustrated in Table 13.

Table 13. Combined weighting of users of Dongfeng Nissan- Xuan Yi (x_2).

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
0	3.19	2.49	2.49	3.19	3.19	2.49	3.19	3.19
1	2.79	2.79	3.49	2.79	3.49	3.49	3.49	2.79
2	2.34	2.34	2.34	2.34	2.34	2.34	3.04	0.94
3	3.07	1.67	3.07	3.07	3.07	3.07	2.37	2.37
4	2.46	3.16	3.16	1.06	3.16	3.16	3.16	3.16

Step 3. Aggregation of group user ratings

Step 3.1 Group division

Step 3.1.1 Grouping based on user geography

According to the purchase area of the reviews divided into eight collections, the Python location function was used to achieve regional collection division, and the number of reviews in each collection was determined, as illustrated in Table 14.

Table 14. Dongfeng Nissan—Xuan Yi (x_2) regional breakdown set.

Area	Number of Reviews	Area	Number of Reviews
No region	91	East	836
Northeast	210	South	408
North	240	Southwest	240
Central	438	Northwest	210

Step 3.1.2 Group segmentation based on user purchase time periods

In addition to the regional set division, each regional set was again divided into sets by observing the distribution of users' time to purchase an automobile, as illustrated in Table 15.

Table 15. Dongfeng Nissan—Xuan Yi (x_2) Regional—Time Group Segmentation Set.

Area	2016–2018	2019	2020–2022
No region	1	39	51
Northeast	25	44	141
North	32	138	70
Central	43	172	223
East	96	386	354
South	70	231	107
Southwest	16	106	118
Northwest	13	48	165

Step 3.1.3 Combined user weighting normalization

Table 16 shows the results of normalizing the combined weights of users after dividing each set.

Table 16. Set internal confidence weights normalized.

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
1	0.082	0.097	0.055	0.082	0.101	0.075	0.071	0.078
2	0.080	0.095	0.075	0.056	0.099	0.073	0.089	0.099
3	0.086	0.100	0.058	0.110	0.082	0.079	0.094	0.105
4	0.094	0.108	0.088	0.094	0.113	0.108	0.101	0.113
5	0.062	0.078	0.102	0.085	0.081	0.100	0.073	0.059
6	0.112	0.081	0.105	0.088	0.084	0.102	0.096	0.107
7	0.093	0.085	0.087	0.070	0.089	0.085	0.100	0.089
8	0.112	0.081	0.105	0.112	0.107	0.102	0.076	0.107
9	0.097	0.111	0.113	0.074	0.116	0.089	0.104	0.093
10	0.089	0.103	0.105	0.113	0.085	0.103	0.097	0.085
11	0.091	0.062	0.107	0.115	0.042	0.083	0.099	0.065

Step 3.1.4 Calculation of the overall user rating

(1) Rating and text emotional intensity combined

The average of the raw ratings of each comment and the sentiment intensity obtained from the text sentiment quantification was computed to obtain the rating $\overline{S}_{pij}^{k}$ for each comment. Table 17 presents the average ratings for some of the comments.

Table 17. Mean ratings of Dongfeng Nissan- Xuan Yi (x_2) ratings and textual sentiment intensity.

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
0	5.0	2.5	4.5	5.0	4.0	4.5	3.0	4.0
1	3.5	3.5	3.0	3.5	4.0	4.0	4.0	4.5
2	3.5	3.5	4.5	4.5	4.5	3.5	4.0	3.5
3	5.0	3.0	4.0	3.0	5.0	5.0	2.5	4.5
4	3.5	4.0	3.0	3.5	5.0	3.0	3.0	4.0

(2) Overall user rating calculation

Based on each review's rating and its user reputation weighting to compute its composite rating, Table 18 illustrates the composite ratings computed for a collection of group users.

Table 18. Overall rating of Dongfeng Nissan- Xuan Yi (x_2).

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
0	0.068	0.102	0.098	0.047	0.095	0.041	0.104	0.096
1	0.097	0.039	0.052	0.109	0.036	0.041	0.055	0.051
2	0.077	0.080	0.077	0.045	0.093	0.095	0.041	0.065
3	0.075	0.050	0.058	0.054	0.102	0.048	0.080	0.057
4	0.099	0.083	0.069	0.111	0.096	0.098	0.105	0.097

Step 3.1.5 Aggregation of group user ratings

(1) Group aggregation by user purchase time

Aggregation is based on a collection of time-of-purchase users to compute an overall rating, as shown in Table 19.

Table 19. Rating matrix for aggregation of Dongfeng Nissan- Xuan Yi (x_2) by the time interval in the Northeast.

Time Interval	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
2016–2018	4.49	3.46	3.65	3.91	4.36	4.06	3.93	4.46
2019	4.45	3.89	3.96	4.16	4.50	4.62	4.26	4.34
2020–2022	4.68	4.09	4.49	4.47	4.66	4.74	4.57	4.65

(2) Group aggregation by user geography

A composite rating is computed based on the aggregation of the set of users in the area of purchase, as illustrated in Table 20. Each serial number represents a geographical group.

Table 20. Dongfeng Nissan- Xuan Yi (x_2) geographical aggregation rating matrix.

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
0	4.38	4.09	4.14	4.38	4.68	4.63	4.03	4.49
1	4.45	3.99	4.12	4.22	4.53	4.57	4.22	4.49
2	4.54	4.02	4.12	4.26	4.51	4.61	4.34	4.51
3	4.50	3.98	4.15	4.27	4.52	4.54	4.30	4.48
4	4.47	3.92	4.16	4.24	4.52	4.53	4.25	4.40
5	4.49	3.76	4.07	4.18	4.48	4.48	4.22	4.40
6	4.45	3.78	4.17	4.28	4.44	4.51	4.16	4.40
7	4.54	3.81	4.03	4.18	4.50	4.47	4.25	4.48

(3) Aggregation of multi-geographical ratings

The multi-criteria composite ratings were obtained by aggregating the aggregated ratings of all regional user groups. Table 21 illustrates the multi-criteria composite ratings for the six automobile brands.

Table 21. Multi-criteria automobile composite rating table.

Automobile Criteria	x_1	x_2	x_3	x_4	x_5	x_6
a_1	4.68	4.48	4.65	4.43	3.86	3.65
a_2	4.26	3.92	4.42	4.22	4.08	4.30
a_3	4.25	4.12	4.42	4.32	4.54	4.40
a_4	4.00	4.25	3.89	4.14	4.16	3.92
a_5	4.54	4.52	4.09	3.51	4.08	3.87
a_6	4.68	4.54	4.66	4.67	4.72	4.61
a_7	4.03	4.22	3.99	3.73	3.92	4.21
a_8	4.37	4.46	4.28	4.12	4.42	3.96

Step 3.2 Aggregation of multi-criteria ratings

A multi-criteria rating aggregation approach was implemented, and consumers set all weights $w_j = 0.125$, $j = 1, 2, \ldots, 8$ to obtain the overall ratings of the automobile brands. Table 22 illustrates the calculation results of the comprehensive score when λ in Formula (12) takes different values, and their corresponding rankings for the six automobile brands. And Figure 8 shows the changing trend. Figure 6 demonstrates the pattern of the composite ratings obtained from the three different methods.

Table 22. Overall ratings for alternatives with different ratios (λ) of ratings to text sentiment scores.

Automobile Brand	Original (l_1)	Ranking	λ = 0 (l_2)	Ranking	λ = 0.5 (l_3)	Ranking	λ = 0.2(l_4)	Ranking	λ = 0.8 (l_5)	Ranking
x_1	4.51	6	4.15	1	4.35	1	4.23	1	4.48	1
x_2	4.53	4	4.14	2	4.31	2	4.21	2	4.42	2
x_3	4.52	5	4.13	3	4.30	3	4.20	3	4.41	3
x_4	4.86	1	3.93	5	4.14	5	4.02	5	4.27	5
x_5	4.71	3	4.02	4	4.22	4	4.11	4	4.34	4
x_6	4.86	1	3.91	6	4.11	6	3.99	6	4.25	6

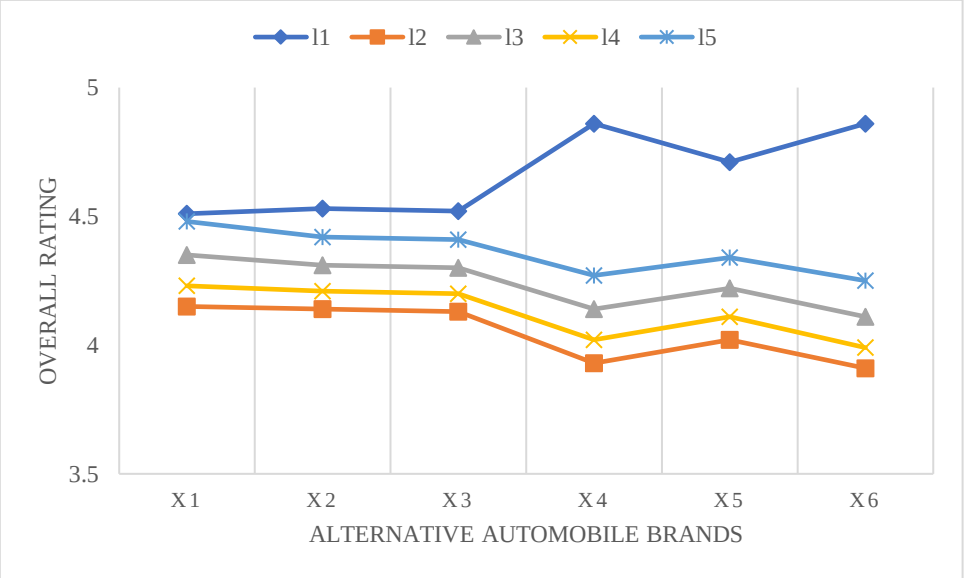

Figure 8. Ranking results for all alternatives.

Step 4. Obtain the ranking results

The composite ratings of each automobile brand were obtained and ranked based on the aforementioned composite rating computation, and the findings are presented in Table 23.

Table 23. Overall scoring ranking table.

Case1	$x_4 = x_6 > x_5 > x_2 > x_3 > x_1$
Case2	$x_1 > x_2 > x_3 > x_5 > x_4 > x_6$
Case3	$x_1 > x_2 > x_3 > x_5 > x_4 > x_6$
Case4	$x_1 > x_2 > x_3 > x_5 > x_4 > x_6$
Case5	$x_1 > x_2 > x_3 > x_5 > x_4 > x_6$

The ranking order of the automobiles was changed by performing a consistency analysis of the ratings with the review text and the fusion of the text feature information to compute the overall rating of the automobiles, which differed from the original rating of the automobiles. Six automobiles were ranked with the original overall rating of

$x_4 = x_6 > x_5 > x_2 > x_3 > x_1$. The ranking results of the overall rating and sentiment analysis overall rating computed using the above approach were $x_1 > x_2 > x_3 > x_5 > x_4 > x_6$. The change in rating and brand ranking demonstrates that fake reviews affect the overall rating and ranking of an automobile brand. Users can employ this analysis to choose an automobile brand that suits their needs for each automobile criterion by simply adjusting the weight W of each automobile criteria. For instance, if they prefer an automobile with comfortable space, Toucan L is the recommended choice; if they prioritize low fuel consumption, Dongfeng Nissan- Xuan Yi is a suitable choice.

This study employs a text sentiment analysis approach as the foundation to examine the consistency between ratings and review text while fusing review text features, which can well verify the authenticity of each review. This study's experimental findings demonstrate that the proposed analysis approach effectively mitigates the impact of false reviews, allowing consumers to obtain comprehensive ratings of automobile brands devoid of the influence of false reviews, as well as criteria-specific ratings for each brand. Therefore, consumers can personalize the selection of automobile brands based on their needs.

7. Conclusions

Considering that user feature information and content feature information can also reflect the credibility of reviews, this study also calculates the weight of user feature information and content feature information of each review. Considering the inconsistency between online review texts and the corresponding star ratings, this study uses a deep learning model to analyze the sentiment of online review texts, predict the sentiment intensity rating of each text, and compare it with the corresponding star ratings given by users to obtain the consistency weight. By combining objective and subjective factors, the feasible weight of each review can be more accurately calculated. In recalculating the composite ratings, this study split the reviews of an automobile from multiple sets by the location and time of purchase and calculated the composite ratings for each set within the set, taking full account of the impact of the location and time of purchase on the credibility of the reviews. Compared with similar existing studies, the research process, methods, and results of this paper are more interpretable and enlightening. In particular, in terms of user credibility, the proposed approach closely explores the personality characteristics disclosed by users. However, there are still limitations in this study; for example, the product quality complaint data are not considered, and the information is not comprehensive enough. Missing values for future user reviews are also not considered. This research is aimed at automobiles, and further research is needed on how to deal with other fields. In the future, further research and attempts will be made regarding the consideration of user personalized preferences in the ratings aggregation process, and the integration of product quality complaint data provided by users for product rankings.

Author Contributions: Methodology, X.Y. and Y.Y.; Software, X.Y.; Validation, W.C.; Investigation, Y.Y.; Writing—original draft, W.C.; Visualization, Y.Y.; Supervision, Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used to support the findings of the study are available from the corresponding author upon request. The author's email address is yangshijiazu@my.swjtu.edu.cn.

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

References

1. Esposito, C.; Galli, A.; Moscato, V.; Sperlí, G. Multi-criteria assessment of user trust in Social Reviewing Systems with subjective logic fusion. *Inf. Fusion* **2022**, *77*, 1–18. [CrossRef]
2. Moran, G.; Muzellec, L. eWOM credibility on social networking sites: A framework. *J. Mark. Commun.* **2017**, *23*, 149–161. [CrossRef]
3. Verma, D.; Dewani, P.P. eWOM credibility: A comprehensive framework and literature review. *Online Inf. Rev.* **2021**, *45*, 481–500. [CrossRef]

4. Banerjee, S.; Bhattacharyya, S.; Bose, I. Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business. *Decis. Support Syst.* **2017**, *96*, 17–26. [CrossRef]
5. Sun, H.-L.; Liang, K.-P.; Liao, H.; Chen, D.B. Evaluating user reputation of online rating systems by rating statistical patterns. *Knowl.-Based Syst.* **2021**, *219*, 106895. [CrossRef]
6. Xiang, Z.; Du, Q.; Ma, Y.; Fan, W. A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tour. Manag.* **2017**, *58*, 51–65. [CrossRef]
7. Meel, P.; Vishwakarma, D.K. Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Syst. Appl.* **2020**, *153*, 112986. [CrossRef]
8. Hazarika, B.; Chen, K.; Razi, M. Are numeric ratings true representations of reviews? A study of inconsistency between reviews and ratings. *Int. J. Bus. Inf. Syst.* **2021**, *38*, 85–106. [CrossRef]
9. Almansour, A.; Alotaibi, R.; Alharbi, H. Text-rating review discrepancy (TRRD): An integrative review and implications for research. *Future Bus. J.* **2022**, *8*, 3. [CrossRef]
10. Lo, A.S.; Yao, S.S. What makes hotel online reviews credible? An investigation of the roles of reviewer expertise, review rating consistency and review valence. *Int. J. Contemp. Hosp. Manag.* **2019**, *31*, 41–60. [CrossRef]
11. Wankhade, M.; Rao, A.C.S.; Kulkarni, C. A survey on sentiment analysis methods, applications, and challenges. *Artif. Intell. Rev.* **2022**, *55*, 5731–5780. [CrossRef]
12. Yadav, A.; Vishwakarma, D.K. Sentiment analysis using deep learning architectures: A review. *Artif. Intell. Rev.* **2020**, *53*, 4335–4385. [CrossRef]
13. Zhang, L.; Wang, S.; Liu, B. Deep learning for sentiment analysis: A survey. *WIREs Data Min. Knowl. Discov.* **2018**, *8*, e1253. [CrossRef]
14. Xu, H.; Liu, B.; Shu, L.; Yu, P.S. BERT post-training for review reading comprehension and aspect-based sentiment analysis. *arXiv* **2019**, arXiv:1904.02232.
15. Yang, L.; Li, Y.; Wang, J.; Sherratt, R.S. Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access* **2020**, *8*, 23522–23530. [CrossRef]
16. Haque, T.U.; Saber, N.N.; Shah, F.M. Sentiment analysis on large scale Amazon product reviews. In Proceedings of the 2018 IEEE International Conference on Innovative Research and Development (ICIRD), Bangkok, Thailand, 11–12 May 2018; pp. 1–6.
17. Guo, Y.; Barnes, S.J.; Jia, Q. Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tour. Manag.* **2017**, *59*, 467–483. [CrossRef]
18. Fan, Z.P.; Xi, Y.; Liu, Y. Supporting consumer's purchase decision: A method for ranking products based on online multi-criteria product ratings. *Soft Comput.* **2018**, *22*, 5247–5261. [CrossRef]
19. Wu, Q.; Liu, X.; Qin, J.; Wang, W.; Zhou, L. A linguistic distribution behavioral multi-criteria group decision making model integrating extended generalized TODIM and quantum decision theory. *Appl. Soft Comput.* **2021**, *98*, 106757. [CrossRef]
20. Guo, C.; Du, Z.; Kou, X. Products Ranking Through Aspect-Based Sentiment Analysis of Online Heterogeneous Reviews. *J. Syst. Sci. Syst. Eng.* **2018**, *27*, 542–558. [CrossRef]
21. Bi, J.W.; Liu, Y.; Fan, Z.P. Representing sentiment analysis results of online reviews using interval type-2 fuzzy numbers and its application to product ranking. *Inf. Sci.* **2019**, *504*, 293–307. [CrossRef]
22. Liu, P.; Teng, F. Probabilistic linguistic TODIM method for selecting products through online product reviews. *Inf. Sci.* **2019**, *485*, 441–455. [CrossRef]