ORIGINAL ARTICLE

Product selection based on sentiment analysis of online reviews: an intuitionistic fuzzy TODIM method

Zhenyu Zhang[1] · Jian Guo[1] · Huirong Zhang[2] · Lixin Zhou[3] · Mengjiao Wang[4]

Abstract

Online reviews contain a great deal of information about consumers' purchasing preferences, which seriously affects potential consumers' purchasing decisions. Using the online review data to help customers make purchasing decisions has become a concern of customers, which has theoretical and practical application value. Therefore, a product selection model is presented based on sentiment analysis combined with an intuitionistic fuzzy TODIM method. Firstly, the product features are extracted by the Apriori algorithm based on online reviews. The sentiment orientation and intensity of the sentiment words for the product features are identified by the lexicon-based sentiment analysis approach. Next, the sentiment orientation of the product features is represented by an intuitionistic fuzzy value. Then the intuitionistic fuzzy TODIM method is used to determine the ranking results of the alternative products. Finally, the case study of mobile phone selection is given to illustrate the proposed approach. The results show that the proposed method considers the online reviews' sentiment orientation and intensity and the consumers' gain and loss in the purchasing product process and is more reasonable than the previous research.

Keywords Online review · Intuitionistic fuzzy set · Sentiment analysis · TODIM

Introduction

The rapid development of the Internet has brought great convenience to people's lives, and online consumer groups are increasing. Online reviews contain a great deal of information about consumers' purchasing preferences, which seriously affects potential consumers' purchasing decisions. It becomes an essential information source for consumers to make purchasing decisions and significantly impacts consumers' decision-making behavior [1–3]. However, due to the complexity of online reviews, consumers cannot effectively use online review data. Therefore, fully and effectively using the online review data to make it the basis of purchasing

decisions has become a concern of many scholars, merchants, and consumers.

In the existing research [18], the ranking product methods through online reviews include two parts: sentiment analysis and multi-attribute decision-making. The first part is to identify the sentiment orientation of online reviews by analyzing online reviews extracted from online platforms. The second part is to select the best alternative product considering selected criteria based on sentiment analysis. However, some product ranking methods based on online reviews only consider online reviews' positive and negative sentiment tendencies [4–9]. The sentiment tendency of each sentence is divided into positive or negative, and ignore the information that the sentiment orientation in online reviews is neutral, resulting in a loss of information in the product purchase decision process.

The sentiment orientations of online reviews are classified into positive, negative, and neutral to avoid the loss of online review information. Intuitionistic fuzzy set (IFS) includes membership, non-membership, and hesitation simultaneously, providing a useful tool to represent the positive, negative, and neutral sentiments in online review data. The IFS is widely used to describe sentiment orientation and sentiment intensity [18]. However, the existing research ignores

✉ Lixin Zhou
  zhoulixin1861@hotmail.com

[1] School of Automation, Nanjing University of Science and Technology, Nanjing 210014, China

[2] School of Labor Relationship, Shandong Management University, Jinan 250357, China

[3] Business School, University of Shanghai for Science and Technology, Shanghai 200093, China

[4] Department of Traffic and Prison Management, Shanghai Police College, Shanghai 200137, China

the customers' psychological behavior and gain and loss during purchasing. TODIM method is suitable for describing the psychological behavior of the customers in the product ranking process [32–35]. The main idea of the TODIM method is to compare the product feature value of each alternative product and obtain the gain and loss value, then calculate the dominance degree between every two alternative products and the overall prospect values of each product [36–39]. According to the overall prospect values, the alternative products are ranked.

Therefore, an online review-based product selection model combined with an intuitionistic fuzzy TODIM (IF-TODIM) method is developed. Firstly, the Apriori algorithm is used to extract the product features that customers focus on based on online reviews. Then the sentiment orientation and intensity of the sentiment words for the product features are identified by the lexicon-based sentiment analysis approach. The proportion of the sentiment orientations of the product features are represented by an intuitionistic fuzzy value (IFV). Finally, the IF-TODIM method is used to determine the final ranking results of the alternative products.

The rest of our work is organized as follows. "Related works" introduces some related works on the ranking selection. Considering the advantages of IFVs and intuitionistic fuzzy sets (IFSs) representing the sentiment orientations of product features, "Preliminaries" provides some concepts of IFVs and IFSs. "The IF-TODIM method for product ranking based on online review" develops a new IF-TODIM method for product selection based on online reviews. A case study is given to illustrate the effectiveness of the developed IF-TODIM method in "Case study". "Conclusion" takes some conclusions.

Related works

Recently, some scholars have concentrated on ranking products through online reviews [4–9, 18]. Zhang et al. [4] identified multiple important product features, then extracted sentences about each feature from online reviews, divided the online reviews into subjective and comparative reviews using a dynamic programming algorithm. The online reviews' sentiment orientation was determined to construct a weighted product graph and rank the products using an improved PageRank algorithm. Later, Zhang et al. [5] improved the algorithm by considering the importance of different reviews. The weight of each review was determined by the review's usefulness and time. Kang et al. [6] proposed a customer satisfaction analysis framework based on customer review mining analysis for product improvement decision making. Najmi et al. [7] calculated each product's score by both review and brand. The review score was derived from sentiment analysis and usefulness analysis, and the brand score

was calculated by an improved PageRank algorithm, and the products were ranked based on their combined scores. Li et al. [8] used the value function of prospect theory to determine the perceived value of alternative products based on consumers' expectations of product attributes and the sentiment orientation of product attributes in online reviews. Fan et al. [9] used the stochastic PROMETHEE-II method to determine product ranking based on online ratings.

Fuzzy set theory has been applied in the product ranking or recommendation to represent the uncertainty in the online review data [10–12]. Different forms of fuzzy sets have been used to represent product feature values, such as fuzzy set, hesitant fuzzy set (HFS), Pythagorean fuzzy set (PFS), interval type-2 fuzzy set (IT2 FS), and IFS. Peng et al. [13] calculated the similarity measures of words to cluster each product feature synonyms, then determined the important product features based on the total frequency of each product feature in the reviews. The subjective evaluation of experts was contributed to obtaining a fuzzy decision matrix of important product features, and finally, the products were ranked by the fuzzy PROMETHEE method. Zhang et al. [14] regarded different sentiment scores of product features as different membership values and integrated different sentiment scores by HFS. A product ranking method based on 2-additive fuzzy measures and Choquet integral was developed. Considering IT2 FS was more accurate than the traditional fuzzy set in representing the uncertainty, Bi et al. [15] represented the uncertainty of the product features' sentiment orientations using IT2 FS. Fu et al. [16] used deep learning models and K-means clustering algorithms to identify sentiment tendencies, considered the credibility of the number of online reviews for different products. Interval-valued PFS sets were used to represent product attribute values, and finally, the Heronian mean operator was used to integrate product attribute information to derive product ranking. To retain both the online review sentiment propensity and its probability, Liu and Teng [17] used probabilistic linguistic term sets (PLTSs). The PL-TODIM method was proposed for alternative products based on the new entropy measures and possibility degrees. The probability multivalued neutrosophic linguistic numbers (PMVNLNs) was developed by Ji et al. [18] to characterize online reviews and reflect the differences in positive (negative) information. Regret theory was combined with outranking methods to construct a review-based decision support model. Liang et al. [19] considered the randomness and ambiguity of online reviews and the interrelationship between product features in the decision support model and developed a linguistic intuitionistic normal cloud (LINC) model. Liang et al. [20] represented tourists' sentiment preferences by distributed linguistic according to the online reviews, developed a method for determining the ideal and minimum value solutions, and proposed a DL-VIKOR to rank the alternative hotels for tourists.

IFS has been widely used to describe sentiment orientation and sentiment intensity [21]. In the transforming process, the proportions of the positive, negative, and neutral sentiment orientations were transformed into the membership, non-membership, and hesitance values in IFVs, respectively. Therefore, the IFS has strong flexibility and practicality in the product ranking problem. Liu et al. [22] constructed a purchase decision model based on the IF-TOPSIS method,

hesitance value in IFV, $\mu_A(x_i) + \nu_A(x_i) + \pi_A(x_i) = 1$. The Hamming, Euclidean, and generalized distances between the two product features A and B are defined as follows.

$$d_h(A, B) = \frac{1}{2n} \sum_{i=1}^{n} (|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)|) \quad (1)$$

$$d_e(A, B) = \sqrt{\frac{1}{2n} \sum_{i=1}^{n} \left(|\mu_A(x_i) - \mu_B(x_i)|^2 + |\nu_A(x_i) - \nu_B(x_i)|^2 + |\pi_A(x_i) - \pi_B(x_i)|^2 \right)} \quad (2)$$

which focuses on product preference through similarity to the ideal solution. Liu et al. [23] ranked the alternative products using the combined intuitionistic fuzzy weighted average (IFWA) operator with the PROMETHEE II method. Çalı and Balaman [24] represented the online ratings of hotel customers by IFSs, and IF-ELECTRE was used to rank alternative hotels with VIKOR integration. Zhang et al. [25] calculated the feature weights considering the customers' attention and developed a product ranking model combining 2-additive fuzzy measures, non-linear programming, and Choquet integration.

Therefore, the main contributions of the developed IF-TODIM method for ranking products are as follows. Firstly, a new product selection method based on online reviews is proposed to consider the consumers' online reviews and psychological behavior. Secondly, the product features are exacted by the Apriori algorithm, which is different from the previous research. Thirdly, in the IF-TODIM method, new ranking methods of intuitionistic fuzzy values (IFVs) are developed to compare the gain and loss of each product feature. The objective weight values of product features are calculated by considering entropy measures. Fourthly, compared with the previous method, product ranking with the IF-TODIM method has advantages over the intuitionistic fuzzy TOPSIS (IF-TOPSIS) method.

Preliminaries

The IFVs have the advantage of representing the feature values of products. In the transforming process, the proportions of the positive, negative, and neutral sentiment orientations are transformed into the membership, non-membership, and hesitance values in IFVs, respectively. Therefore, some basic concepts of IFVs and IFSs are introduced.

Definition 1 [26]. Let $A = \{\langle x_i, \mu_A(x_i), \nu_A(x_i) \rangle | x_i \in X\}$ and $B = \{\langle x_i, \mu_B(x_i), \nu_B(x_i) \rangle | x_i \in X\}$ be two IFSs representing the feature values of products, where $\mu_A(x_i)$, $\nu_A(x_i)$ and $\pi_A(x_i)$ are the membership value, non-membership value , and

$$d_g(A, B) = \left[\frac{1}{2n} \sum_{i=1}^{n} (|\mu_A(x_i) - \mu_B(x_i)|^\lambda + |\nu_A(x_i) - \nu_B(x_i)|^\lambda + |\pi_A(x_i) - \pi_B(x_i)|^\lambda) \right]^{1/\lambda} \quad (3)$$

Definition 2 [27, 28]. Let $A = \{\langle x_i, \mu_A(x_i), \nu_A(x_i) \rangle | x_i \in X\}$ be an IFS, and the entropy measures can be defined as follows.

$$E_1(A) = \sum_{i=1}^{n} (1 - (\mu_A(x_i) + \nu_A(x_i))) = \sum_{i=1}^{n} \pi_A(x_i) \quad (4)$$

$$E_2(A) = \sum_{i=1}^{n} \left(1 - (\mu_A(x_i) + \nu_A(x_i))^k \right), k = 2, 3, \cdots, \infty \quad (5)$$

$$E_3(A) = \sum_{i=1}^{n} \left(1 - (\mu_A(x_i) + \nu_A(x_i)) \cdot e^{1-(\mu_A(x_i)+\nu_A(x_i))} \right) \quad (6)$$

$$E_4(A) = \sum_{i=1}^{n} \left(1 - (\mu_A(x_i) + \nu_A(x_i)) \cdot \sin\left((\pi/2)(\mu_A(x_i) + \nu_A(x_i))\right) \right) \quad (7)$$

The score measures of IFVs act as an important role in comparing the magnitude of alternative product feature values. Some new score measures considering the signed distance of IFVs are introduced as follows.

Definition 3. Let $a = <\mu_a, \nu_a>$ and $b = <\mu_b, \nu_b>$ be two IFVs representing the feature values of products, $\tilde{0} = <0, 1>$ and $\tilde{1} = <1, 0>$ are the worst and best evaluation values of the product features. Then, the new score measures R_h, R_e and R_g of IFVs are defined as follows.

$$R_h(a) = d_h\left(a, \tilde{0}\right) = \frac{1}{2}(|\mu_a - 0| + |\nu_a - 1| + |\pi_a - 0|) = \frac{1}{2}(1 + \mu_a + \pi_a - \nu_a) \quad (8)$$

$$R_e(a) = d_e\left(a, \tilde{0}\right) = \sqrt{\frac{1}{2}\left(|\mu_a - 0|^2 + |\nu_a - 1|^2 + |\pi_a - 0|^2\right)} = \sqrt{\frac{1}{2}\left(\mu_a^2 + \pi_a^2 + |1 - \nu_a|^2\right)} \quad (9)$$

$$R_g(a) = d_g\left(a, \tilde{0}\right) = \left(\frac{1}{2}\left(|\mu_a - 0|^\lambda + |v_a - 1|^\lambda + |\pi_a - 0|^\lambda\right)\right)^{1/\lambda}$$

$$= \left[\frac{1}{2}\left(\mu_a^\lambda + \pi_a^\lambda + |1 - v_a|^\lambda\right)\right]^{1/\lambda} \qquad (10)$$

Property 1. Let $a = \langle \mu_a, v_a \rangle$ be an IFV, where μ_a, $v_a \in [0, 1]$ and $0 \le \mu_a + v_a \le 1$, then $R(a) \in [0, 1]$.

Proof. If the IFV $a = \langle 0, 1 \rangle$, then

$$R_h(a) = d_h\left(a, \tilde{0}\right) = \frac{1}{2}(|0 - 0| + |1 - 1| + |0 - 0|) = 0,$$

$$R_e(a) = d_e\left(a, \tilde{0}\right)$$
$$= \sqrt{\frac{1}{2}\left(|0 - 0|^2 + |1 - 1|^2 + |0 - 0|^2\right)} = 0,$$

$$R_g(a) = d_g\left(a, \tilde{0}\right)$$
$$= \left(\frac{1}{2}\left(|0 - 0|^\lambda + |1 - 1|^\lambda + |0 - 0|^\lambda\right)\right)^{1/\lambda} = 0.$$

If the IFV $a = \langle 1, 0 \rangle$, then
$$R_h(a) = d_h\left(a, \tilde{0}\right) = \frac{1}{2}(|1 - 0| + |0 - 1| + |0 - 0|) = 1,$$

$$R_e(a) = d_e\left(a, \tilde{0}\right)$$
$$= \sqrt{\frac{1}{2}\left(|1 - 0|^2 + |0 - 1|^2 + |0 - 0|^2\right)} = 1,$$

$$R_g(a) = d_g\left(a, \tilde{0}\right)$$
$$= \left(\frac{1}{2}\left(|1 - 0|^\lambda + |0 - 1|^\lambda + |0 - 0|^\lambda\right)\right)^{1/\lambda} = 1.$$

Therefore, $R(a) \in [0, 1]$.

The corresponding ranking method of IFVs is defined as
(1) If $R(a) > R(b)$, then $a \succ b$;
(2) If $R(a) = R(b)$, then $a \sim b$.

The IF-TODIM method for product ranking based on online review

Problem description

The following symbols are used to represent collections and variables in the product selection problem.

$A = \{A_1, A_2, \ldots, A_n\}$: a collection of n alternative products, where A_i represents the i-th product, $i = 1, 2, \ldots, n$ and the consumers select the alternative product set A.

$F = \{f_1, f_2, \ldots, f_m\}$: a collection of m features, the products' features from the online reviews that the consumer focuses on, where f_j represents the j-th feature, $j = 1, 2, \ldots, m$.

$W = \{\omega_1, \omega_2, \ldots, \omega_m\}$: the weight vector of the features, where ω_j represents the weight of the feature f_j, $\omega_j > 0$ and $\sum_{j=1}^m \omega_j = 1$.

$Q = \{q_1, q_2, \ldots, q_n\}$: the collection of the number of online reviews for the alternative product A_i, where q_i means the number of online reviews about the alternative product A_i, $i = 1, 2, \ldots, n$.

$D_{ik} = \left\{d_{ik}^1, d_{ik}^2, \ldots, d_{ik}^m\right\}$: the online review collection of the alternative product A_i, where d_{ik}^j represents the k-th online review on the features f_j in the i-th alternative product, $i = 1, 2, \ldots, n, j = 1, 2, \ldots, m, k = 1, 2, \ldots, q_i$.

The problem is how to select alternative products A_1, A_2, $\ldots$, A_n based on online review D_{ik} and feature weight ω_j, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$, $k = 1, 2, \ldots, q_i$.

The flowchart of the product selection is shown in Fig. 1. The input information is the crawled online reviews of alternative products. The process includes two parts: sentiment orientation identification and product ranking based on the IF-TODIM method. In the first part, the Apriori algorithm is first used to identify the product features that customers focus on based on online reviews. The sentiment orientation and intensity of the sentiment words for the product features are identified by the lexicon-based sentiment analysis approach. The second part is to convert the sentiment orientation of the product features into an IFV and then use the IF-TODIM method to determine the final ranking results of the alternative products.

## Sentiment orientation identification of the online reviews

(1) Product feature extraction

A product feature extraction method based on online review data mining is introduced to extract the features of the alternative products that the consumers focus on from the online reviews. The process is described as follows.

First, the online review data is segmented, and the online review data after the segmentation is tagged. For the sake of accuracy and rationality, the ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System, http://ictclas.nlpir.org/) tool is used for word segmentation of online review data. The lexical marking is for nouns, verbs, adjectives, or verbs with noun functions and proper nouns to improve the accuracy of the search.

Secondly, the association rule transaction file is created using the part-of-speech tagging, and the frequent itemset is searched based on the association rule Apriori algorithm. Here, the minimum support value is 1%, and at the same time, more than three frequent items are not considered.

Fig. 1 The flowchart of the product selection
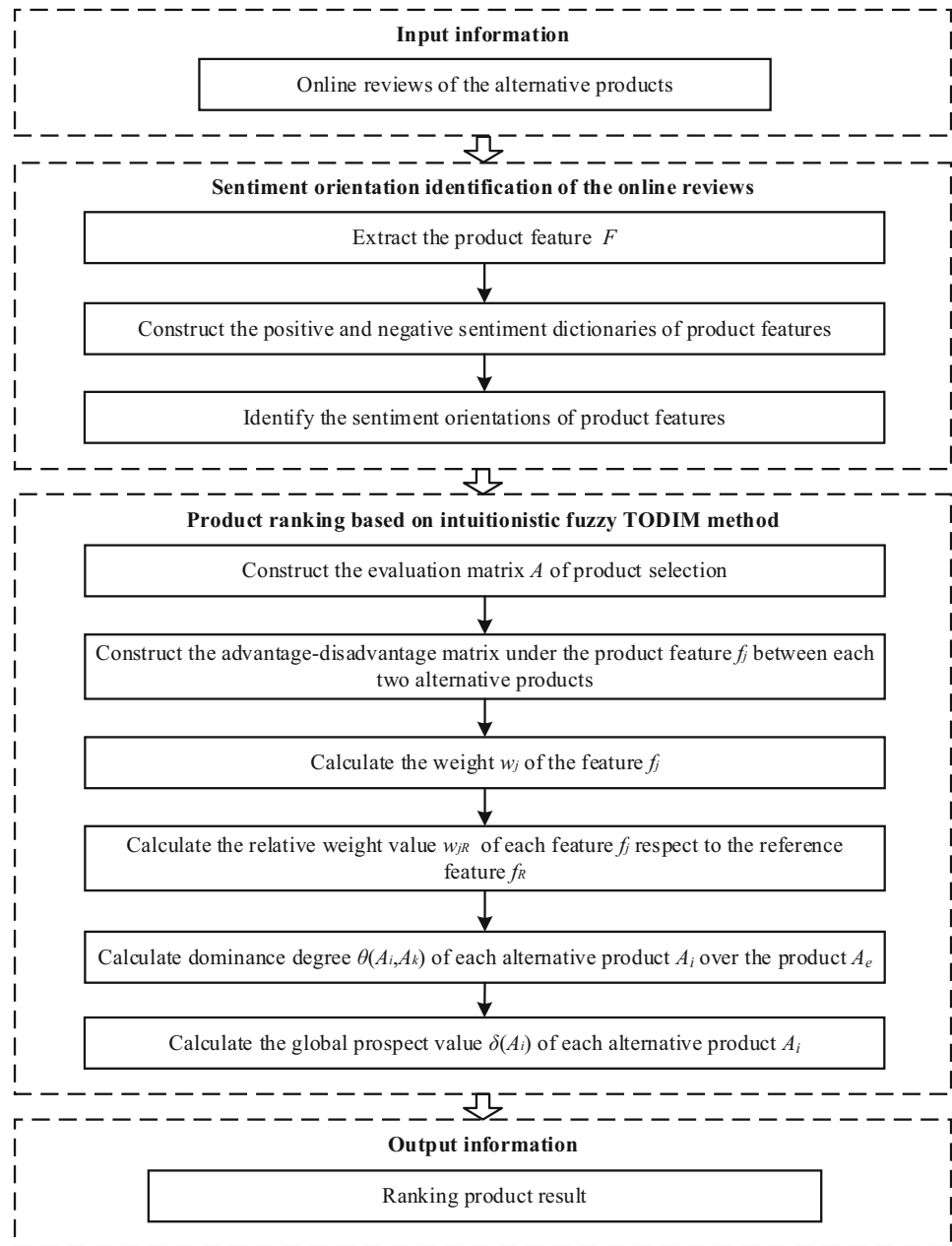
The frequent itemset is pruned and corrected according to the neighboring rules and independent support and formed into a product feature set F^{TF}.

Then, the common Chinese frequent item noun set F^{FF} of non-product features (such as some common product brands, colloquial zed nouns, and personal names) and the product feature set F^{SF} (containing single nouns) are constructed, and F^{TF} is filtered to form the final product feature set F, i.e. $F = F^{TF} - F^{FF} - F^{SF}$.

(2) Construct the positive and negative sentiment dictionaries of product features

Normally, different features have different positive or negative sentiment dictionaries. A word exhibits different sentiment orientations in the sentiment dictionaries of different features. For example, "high" is the negative sentiment word in the dictionary of the feature "price" and the positive sentiment word of the feature "pixel". Therefore, the positive and negative sentiment dictionary for each product feature should be constructed separately.

Firstly, according to the online review set after the part-of-speech tagging, the association rule object file for the feature f_j in the review is created. The frequent itemset F is searched based on the association rule Apriori algorithm to form the feature annotation set.

Assume that $\overline{W}_i^j$ represents a sentiment word of the feature f_j of the alternative product A_i, then $\overline{W}^j$ means all of the sentiment words of the feature f_j, $\overline{W}_i^j$ and $\overline{W}^j$ are defined as follows.

$$\overline{W}_i^j = \overline{W}_{i1}^j \cup \overline{W}_{i2}^j \cup \ldots \cup \overline{W}_{iq_i}^j, \\ i = 1, 2, \ldots, n, \, j = 1, 2, \ldots, m \quad (11)$$

$$\overline{W}^j = \overline{W}_1^j \cup \overline{W}_2^j \cup \ldots \cup \overline{W}_n^j, \, j = 1, 2, \ldots, m \quad (12)$$

In addition, let $\overline{W}_{HowNet}^+ = \left\{ \overline{W}_1^{HN+}, \overline{W}_2^{HN+}, \ldots, \overline{W}_{4566}^{HN+} \right\}$ and $\overline{W}_{HowNet}^- = \left\{ \overline{W}_1^{HN-}, \overline{W}_2^{HN-}, \ldots, \overline{W}_{4370}^{HN-} \right\}$ represent the positive and negative sentiment word sets in the HowNet sentiment dictionary. $\overline{W}_{HowNet}^+$ and $\overline{W}_{HowNet}^-$ include 4566 positive sentiment words and 4370 negative sentiment words, respectively. Assume that $\overline{W}_j^+$ and $\overline{W}_j^-$ are the positive and negative sentiment dictionary of the feature f_j, $\overline{W}_j^+$ and $\overline{W}_j^-$ are defined as follows.

$$\overline{W}_j^+ = \overline{W}_{HowNet}^+ \cap \overline{W}^j, \, j = 1, 2, \ldots, m \quad (13)$$

$$\overline{W}_j^- = \overline{W}_{HowNet}^- \cap \overline{W}^j, \, j = 1, 2, \ldots, m \quad (14)$$

To improve their accuracy, $\overline{W}_{HowNet}^+$ and $\overline{W}_{HowNet}^-$ need to make adjustments manually.

(3) Identify the sentiment orientations of product features

Each feature's positive, neutral or negative sentiment orientations of each review are calculated. The principle of identifying the sentence's sentiment orientation is as follows [18]. If the number of positive sentiment words in the sentence is greater than that of the negative sentiment, the sentiment is considered positive. If the number of negative sentiment words in a sentence is greater than that of positive sentiment words, the sentence's sentiment orientation is considered negative. If there are equal positive and negative sentiment words or no sentiment words in the sentence, then the sentence is considered neutral in its sentiment orientation. If there is a negative word in the sentence, the sentiment orientation of the sentence is reversed. The rules are shown as follows.

For each sentiment word set $\overline{W}_{ik}^j$ obtained by online reviews, $\overline{W}_{ik}^{j+}$ and $\overline{W}_{ik}^{j-}$ are the sets of positive and negative sentiment words, which are the intersections between $\overline{W}_{ik}^j$ and $\overline{W}_i^{j+}$ or $\overline{W}_i^{j-}$. Let $s_{ik}^j = \left(\alpha_{ik}^j, \beta_{ik}^j, \gamma_{ik}^j \right)$ express the sentiment orientation vector of the sentence D_{ik}^j, where $\alpha_{ik}^j, \beta_{ik}^j, \gamma_{ik}^j \in \{0, 1\}$ and $\alpha_{ik}^j + \beta_{ik}^j + \gamma_{ik}^j = 0$ or 1. If $\overline{W}_{ik}^j$

is an empty set, then $s_{ik}^j = (0, 0, 0)$. If the number of $\overline{W}_{ik}^{j+}$ is greater than the number of $\overline{W}_{ik}^{j-}$, then $s_{ik}^j = (1, 0, 0)$. If the number of $\overline{W}_{ik}^{j+}$ is less than the number of $\overline{W}_{ik}^{j-}$, then $s_{ik}^j = (0, 0, 1)$. If the number of $\overline{W}_{ik}^{j+}$ is equal to the number of $\overline{W}_{ik}^{j-}$ and $\overline{W}_{ik}^j$ is not an empty set, then $s_{ik}^j = (0, 1, 0)$. The sentiment orientations of the product features are calculated by the above rules.

Product ranking based on IF-TODIM method

(1) Transform the sentiment orientations of product features into IFVs

IFVs are a useful tool for representing the ambiguity and hesitation of products' features. IFVs can simultaneously reflect the like, neutral, and unlike of the online review [29]. Based on the theory of IFSs, online reviews of alternative products' sentiment orientations can be expressed simply and completely by IFVs [30].

In addition, most online reviews now have a click-and-click feature that makes it easy to understand the usefulness of each review. Therefore, more important weights are assigned to more praises, which are more useful reviews. Let X_{ik}^j be the importance of each review, and X_{ik}^j is determined by the number of likes and calculated as follows.

$$x_{ik}^j = e^{\frac{N_{ik}^j - \min(N_i)}{\max(N_i) - \min(N_i)}}, \, i = 1, 2, \ldots, n, \\ j = 1, 2, \ldots, m, \, k = 1, 2, \ldots, q_i \quad (15)$$

where N_{ik}^j is the number of likes on the k-th review of the product A_i, N_i is the set of the number of likes of the product A_i.

Let $q_{ij}^{pos}, q_{ij}^{neu}$ and q_{ij}^{neg} be the frequency of online reviews of positive, neutral, and negative sentiment orientations that characterize the alternative product, and $q_{ij}^{pos}, q_{ij}^{neu}$ and q_{ij}^{neg} are defined as follows.

$$q_{ij}^{pos} = \sum_{k=1}^{q_i} x_{ik}^j \alpha_{ik}^j, \, i = 1, 2, \ldots, n, \, j = 1, 2, \ldots, m \quad (16)$$

$$q_{ij}^{neu} = \sum_{k=1}^{q_i} x_{ik}^j \beta_{ik}^j, \, i = 1, 2, \ldots, n, \, j = 1, 2, \ldots, m \quad (17)$$

$$q_{ij}^{neg} = \sum_{k=1}^{q_i} x_{ik}^j \gamma_{ik}^j, \, i = 1, 2, \ldots, n, \, j = 1, 2, \ldots, m \quad (18)$$

Let $q_{ij}^{pos}, q_{ij}^{neu}$ and q_{ij}^{neg} represent the percentage of the alternative product features for positive, neutral, and negative sentiment orientations. The calculation formula is defined as

follows

$$p_{ij}^{pos} = \frac{q_{ij}^{pos}}{q_{ij}^{pos} + q_{ij}^{neu} + q_{ij}^{neg}}, \quad i = 1, 2, \ldots, n,$$
$$j = 1, 2, \ldots, m \quad (19)$$

$$p_{ij}^{neu} = \frac{q_{ij}^{neu}}{q_{ij}^{pos} + q_{ij}^{neu} + q_{ij}^{neg}}, \quad i = 1, 2, \ldots, n,$$
$$j = 1, 2, \ldots, m \quad (20)$$

$$p_{ij}^{neg} = \frac{q_{ij}^{neg}}{q_{ij}^{pos} + q_{ij}^{neu} + q_{ij}^{neg}}, \quad i = 1, 2, \ldots, n,$$
$$j = 1, 2, \ldots, m \quad (21)$$

Obviously, $p_{ij}^{pos} + p_{ij}^{neu} + p_{ij}^{neg} = 1, p_{ij}^{pos}, p_{ij}^{neu}, p_{ij}^{neg} \geq 0$.
Thus, based on the interpretation of the IFV [31], an IFV $a_{ij} = (\mu_{ij}, v_{ij}, \pi_{ij})$ can represent the percentages of positive, neutral, and negative sentiment orientations of product features, where $\mu_{ij} = p_{ij}^{pos}$, $v_{ij} = p_{ij}^{neg}$ and $\pi_{ij} = p_{ij}^{neu}$.

(2) IF-TODIM method for ranking products

Step 1: calculate the feature values (μ_{ij}, v_{ij}) in each alternative product and construct the decision matrix $A = (a_{ij})_{m \times n}$ of product selection.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$
$$= \begin{bmatrix} (\mu_{11}, v_{11}) & (\mu_{12}, v_{12}) & \cdots & (\mu_{1n}, v_{1n}) \\ (\mu_{21}, v_{21}) & (\mu_{22}, v_{22}) & \cdots & (\mu_{2n}, v_{2n}) \\ \cdots & \cdots & \cdots & \cdots \\ (\mu_{m1}, v_{m1}) & (\mu_{m2}, v_{m2}) & \cdots & (\mu_{mn}, v_{mn}) \end{bmatrix} \quad (22)$$

where a_{ij} represents the value of the criteria f_j in the alternative product A_i, all the values of a_{ij} are represented by IFVs.

Step 2: compare the feature values of each two alternative products by Eqs. (8)–(10) and construct the advantage-disadvantage matrix, where "A" or "D" means that A_i is larger or smaller than A_k.

Step 3: calculate the weight w_j of the feature f_j as follows.

Firstly, calculate the entropy E_{ij} of each product feature by Eqs. (4)–(7), and normalize the entropy by the following equation:

$$H = (h_{ij})_{m \times n} = \left(\frac{E_{ij}}{\max(E_{ij})} \right)_{m \times n},$$

$$i = 1, 2, \cdots, m; \; j = 1, 2, \cdots, n \quad (23)$$

Then, the entropy weights of each product feature are calculated as follows.

$$w_j = \frac{1}{n - \sum_{j=1}^{n} a_j} \times (1 - a_j) \quad (24)$$

where $a_j = \sum_{i=1}^{m} h_{ij}$.

Step 4: the feature with the largest weight value is regarded as the reference feature f_R. The relative weight value w_{jR} of each feature f_j over the reference feature f_R is calculated by Eq. (25).

$$w_{jR} = \frac{w_j}{w_R} \quad (25)$$

Step 5: the dominance degree $\vartheta(A_i, A_e)$ of the alternative product A_i over A_e is calculated by Eq. (26).

$$\vartheta(A_i, A_e) = \sum_{j=1}^{n} \phi_j(A_i, A_e) \quad (26)$$

where

$$\phi_j(A_i, A_e)$$
$$= \begin{cases} \sqrt{w_{jR} d(a_{ij}, a_{ej}) \Big/ \sum_{j=1}^{n} w_{jR}}, & if \; a_{ij} \succ a_{ej} \\ 0, & if \; a_{ij} \sim a_{ej} \\ -\frac{1}{\theta} \sqrt{\left(\sum_{j=1}^{n} w_{jR} \right) d(a_{ij}, a_{ej}) \Big/ w_{jR}}, & if \; a_{ij} \prec a_{ej} \end{cases}$$

where $d(a_{ij}, a_{ej})$ is calculated by Eqs. (1)–(3).

Step 6: the global prospect value $\delta(A_i)$ of each alternative product A_i is calculated by Eq. (27).

$$\delta(A_i) = \frac{\sum_{e=1}^{m} \vartheta(A_i, A_e) - \min_i \left\{ \sum_{e=1}^{m} \vartheta(A_i, A_e) \right\}}{\max_i \left\{ \sum_{e=1}^{m} \vartheta(A_i, A_e) \right\} - \min_i \left\{ \sum_{e=1}^{m} \vartheta(A_i, A_e) \right\}} \quad (27)$$

Step 7: rank the alternative products according to the global prospect values $\delta(A_i)$. The larger the value of $\delta(A_i)$ is, the better the alternative product A_i is.

Table 1 Partial display of mobile phone feature sentiment dictionaries $\overline{W}_j^+$ and $\overline{W}_j^-$

Features	Positive sentiment words	Negative sentiment words
Appearance	(beautiful), (comfortable), 惊(stunning), (exquisite), (good-looking)…	丑(ugly), (heavy), (thick), (small)…
Screen	腻(delicate), (comfortable), 惊(stunning), (likeable), 撼(shocking)…	(small), 扭(awkward), (uncomfortable), 爽(unpleasant)…
Photo	满(satisfied), 耐(durable), (nice)…	耗(power-hungry), (inadequate), 慢(slow)…
Battery	(perfect), (excellent), 腻(fine), 晰(clear), (beautiful)…	脏(dirty), (stuck), (uninteresting), 丑(ugly)…
Price	(cheap), (affordable), (worthwhile), (reasonable), (cost-effective)…	(expensive), (bad), (high)…
System	(smooth), 爽(cool), (fast), (convenient), 悍(powerful)…	顿(stuck), 慢(slow), 杂(complicated), (unstable)…

Case study

Decision-making process

Online reviews of five mobile phones from Jingdong Mall (https://www.jd.com/) are crawled. The five alternative mobile phones are iPhone X, Huawei P10, OPPO R11S, Mito T8, and VIVO X9. The crawler software Bazhuayu (http://www.bazhuayu.com/) is used to crawl 5000 reviews (1000 reviews per phone). After processing, 2000 reviews (400 reviews for each phone) are extracted from the review data set obtained. The mobile phone features extracted by the Apriori algorithm that customers focus on are $F = \{$Appearance, Screen, Photo, Battery, Price, System$\}$. The positive sentiment dictionary $\overline{W}_j^+$ and negative sentiment dictionary $\overline{W}_j^-$ of mobile phone feature are constructed by Eqs. (11)–(14) in Table 1. "/n /a" (Price/n is/v affordable/a) is taken as an example to express the process of the sentiment orientation. $\overline{W}_5^+ \cap \overline{W}_{51}^1 \neq \emptyset$ and $s_{51}^1 = (1, 0, 0)$. Thus, the sentiment orientation is positive.

The positive, neutral, and negative sentiment orientation numbers q_{ij}^{pos}, q_{ij}^{neu} and q_{ij}^{neg} of alternative mobile phones are calculated by Eqs. (15)–(18) and shown in Table 2.

The steps of ranking mobile phones by the IF-TODIM method are shown as follows.

Step 1: calculate the feature values (μ_{ij}, v_{ij}) of each alternative mobile phone by Eqs. (19)–(21) and construct the decision matrix $A = (a_{ij})_{m \times n}$. For example, the appearance value of IPONE X (A_1) is [0.757, 0.187], where

$$\mu_{11} = p_{11}^{pos} = \frac{q_{11}^{pos}}{q_{11}^{pos} + q_{11}^{neu} + q_{11}^{neg}}$$
$$= \frac{557.0}{557.0 + 41.2 + 137.6} = 0.757,$$

$$v_{11} = p_{11}^{neg} = \frac{q_{11}^{neg}}{q_{11}^{pos} + q_{11}^{neu} + q_{11}^{neg}}$$
$$= \frac{137.6}{557.0 + 41.2 + 137.6} = 0.187.$$

The intuitionistic fuzzy decision matrix of mobile phone selection is shown in Table 3.

Step 2: compare the feature values of each two alternative products by Eqs. (8)–(10) and construct the advantage-disadvantage matrix. For example, the score measures of IPHONE X (A_1) and HUAWEI P10 (A_2) under the attribute appearance (f_1) are

$$R_h(a_{11}) = d_h\left(a_{11}, \tilde{0}\right) = \frac{1}{2}\left(1 + \mu_{a_{11}} + \pi_{a_{11}} - v_{a_{11}}\right)$$
$$= \frac{1}{2}(1 + 0.757 + 0.056 - 0.187) = 0.813.$$

$$R_h(a_{21}) = d_h\left(a_{21}, \tilde{0}\right) = \frac{1}{2}\left(1 + \mu_{a_{21}} + \pi_{a_{21}} - v_{a_{21}}\right)$$
$$= \frac{1}{2}(1 + 0.964 + 0.024 - 0.012) = 0.988.$$

Therefore, $R_h(a_{11}) \prec R_h(a_{21})$. The score measure of A_1 under the attribute appearance (f_1) is smaller than A_2, represented by "D". The advantage-disadvantage matrix is shown in Table 4.

Step 3: calculate the weight w_j of the feature f_j.

Calculate the entropy e_{ij} of each mobile phone feature by Eq. (4) in Definition 2. For example, the entropy of a_{11} is $e_{11} = 1 - (\mu_A(x_i) + v_A(x_i)) = 1 - (0.757 + 0.187) = 0.056$. The entropy matrix E is as follows.

$$E = \begin{pmatrix} 0.056 & 0.017 & 0.050 & 0.046 & 0.014 & 0.019 \\ 0.024 & 0.030 & 0.028 & 0.049 & 0.020 & 0.065 \\ 0.005 & 0.018 & 0.008 & 0.027 & 0.031 & 0.006 \\ 0.005 & 0.063 & 0.026 & 0.030 & 0.019 & 0.051 \\ 0.006 & 0.021 & 0.009 & 0.086 & 0.011 & 0.032 \end{pmatrix}$$

Table 2 The sentiment orientation numbers q_{ij}^{pos}, q_{ij}^{neu} and q_{ij}^{neg} of each feature of alternative mobile phones

Features	IPHONE X (A_1)			HUAWEI P10 (A_2)			OPPO R11S (A_3)			MEITU T8 (A_4)			VIVO X9 (A_5)		
	pos	neu	neg	pos	neu	neg	pos	neu	Neg	pos	neu	neg	pos	neu	neg
Appearance	557.0	41.2	137.6	572.9	14.3	7.1	403.0	2.0	2.4	474.6	2.5	22.5	390.3	2.5	17.6
Screen	479.4	10.1	103.8	543.6	19.3	81.2	336.8	6.5	20.0	286.5	27.7	125.9	406.8	9.3	26.6
Photo	469.1	29.2	86.6	514.6	15.7	30.9	331.8	2.9	32.3	393.1	10.8	10.4	415.0	3.9	19.3
Battery	268.3	26.7	285.7	286.6	27.5	246.7	267.6	9.9	87.6	92.1	13.4	341.4	225.6	38.0	178.0
Price	120.7	7.5	405.9	335.0	11.2	214.9	160.0	11.4	195.6	311.9	8.4	122.1	421.3	4.8	10.0
System	479.2	11.2	96.9	480.3	36.5	44.3	346.0	2.2	23.4	298.6	22.7	124.4	381.6	14.1	44.4

Table 3 Intuitionistic fuzzy decision matrix of mobile phone selection

Features	Appearance (f_1)	Screen (f_2)	Photo (f_3)	Battery (f_4)	Price (f_5)	System (f_6)
IPHONE X (A_1)	[0.757, 0.187]	[0.808, 0.175]	[0.802, 0.148]	[0.462, 0.492]	[0.226, 0.760]	[0.816, 0.165]
HUAWEI P10 (A_2)	[0.964, 0.012]	[0.844, 0.126]	[0.917, 0.055]	[0.511, 0.440]	[0.597, 0.383]	[0.856, 0.079]
OPPO R11S (A_3)	[0.989, 0.006]	[0.927, 0.055]	[0.904, 0.088]	[0.733, 0.240]	[0.436, 0.533]	[0.931, 0.063]
MEITU T8 (A_4)	[0.950, 0.045]	[0.651, 0.286]	[0.949, 0.025]	[0.206, 0.764]	[0.705, 0.276]	[0.670, 0.279]
VIVO X9 (A_5)	[0.951, 0.043]	[0.919, 0.060]	[0.947, 0.044]	[0.511, 0.403]	[0.966, 0.023]	[0.867, 0.101]

Table 4 Advantage-disadvantage matrix under each feature f_j between two alternative mobile phones

	A_1/A_2	A_1/A_3	A_1/A_4	A_1/A_5	A_2/A_3	A_2/A_4	A_2/A_5	A_3/A_4	A_3/A_5	A_4/A_5
f_1	D	D	D	D	D	A	A	A	A	D
f_2	D	D	A	D	D	A	D	A	A	D
f_3	D	D	D	D	A	D	D	D	D	A
f_4	D	D	A	D	D	A	D	A	A	D
f_5	D	D	D	D	A	D	D	D	D	D
f_6	D	D	A	D	D	A	A	A	A	D

Then, the normalized entropy matrix is obtained by Eq. (23) as follows.

$$H = \left(h_{ij}\right)_{m \times n} = \begin{pmatrix} 1.000 & 0.304 & 0.893 & 0.821 & 0.250 & 0.339 \\ 0.369 & 0.462 & 0.431 & 0.754 & 0.308 & 1.000 \\ 0.161 & 0.581 & 0.258 & 0.871 & 1.000 & 0.194 \\ 0.079 & 1.000 & 0.413 & 0.476 & 0.302 & 0.810 \\ 0.070 & 0.244 & 0.105 & 1.000 & 0.128 & 0.372 \end{pmatrix}$$

Finally, the entropy weight can be calculated by Eq. (24) as $W = \left(w_j\right)_{n \times 1} = (0.076, 0.177, 0.122, 0.325, 0.110, 0.190)^T$.

Step 4: the feature with the largest weight value is regarded as the reference feature f_R. The relative weight value w_{jR} of each feature f_j over the reference feature f_R is calculated by Eq. (25). The relative weight values are shown in Table 5.

Step 5: the dominance $\vartheta(A_i, A_e)$ of the alternative mobile phone A_i over A_e is calculated by Eq. (26).

Here, assume that $\theta = 1$ [36], then the gain and losses $\phi_j(A_i, A_e)$ are calculated and shown in Table 6. For example,

$\phi_1(A_1, A_2) = -\frac{1}{\theta}\sqrt{\left(\sum_{j=1}^n w_{jR}\right)d(a_{ij} - a_{ej})\Big/ w_{jR}} = -\frac{1}{1}\sqrt{3.077 \times 0.1255 \Big/ 0.233} = -1.340.$

Table 5 The relative weight value w_{jR}

	Appearance	Screen	Photo	Battery	Price	System
w_j	0.076	0.177	0.122	0.325	0.110	0.190
w_{jR}	0.233	0.544	0.376	1.000	0.338	0.587

Then, the dominance degree of mobile phone A_i over mobile phone A_e is calculated by Eq. (26). For example, the dominance degree of mobile phone A_1 over A_2 is $\vartheta(A_1, A_2) = \sum_{j=1}^{6} \phi_j(A_1, A_2) = -1.340 - 0.499 - 0.857 - 0.480 - 1.691 - 0.590 = -5.457$. The dominance degree matrix is shown in Table 7.

Step 6: the global prospect value $\delta(A_i)$ of each alternative mobile phone A_i is calculated by Eq. (27). The global prospect values are $\delta(A_1) = \frac{\sum_{e=1}^{m} \vartheta(A_1, A_e) - \min\{\sum_{e=1}^{m} \vartheta(A_1, A_e)\}}{\max\{\sum_{e=1}^{m} \vartheta(A_1, A_e)\} - \min\{\sum_{e=1}^{m} \vartheta(A_1, A_e)\}} = \frac{-21.059 - (-21.059)}{1.408 - (-21.059)} = 0$, $\delta(A_2) = 0.785$, $\delta(A_3) = 0.913$, $\delta(A_4) = 0.220$ and $\delta(A_5) = 1$.

Step 7: rank the alternative mobile phones according to the global prospect values $\delta(A_i)$, the ranking result is $\delta(A_5) > \delta(A_3) > \delta(A_2) > \delta(A_4) > \delta(A_1)$.

The alternative mobile phones are sorted as VIVO X9 > OPPO R11S > Huawei P10 > Mito T8 > IPHONE X. In the case of priority price, system performance, and appearance, the optimal choice is VIVO X9. According to the online reviews of VIVO X9, most of the online reviews indicate that the system is fluent and the mobile phone is cost-effective. Most of the online reviews of IPHONE X are too expensive, resulting in a lower ranking.

Analysis of the effect of the parameter

The product selection based on online reviews involves the attenuation coefficient θ. The attenuation coefficient θ affecting the ranking result is analyzed by taking different values. When the attenuation coefficient $\theta = 1, 2, 3, 4$, the product ranking result calculated by the IF-TODIM method under different attenuation coefficients are shown in Table 8. From the result, the A_5 is always the best choice under different attenuation coefficients. Therefore, different attenuation coefficient values have no effect on the product ranking results.

Comparison analysis

The developed IF-TODIM method is compared with the IF-TOPSIS [22], IF-VIKOR [40], and IF-PROMETHEE [9] methods to illustrate the effectiveness.

(1) Comparison with IF-TOPSIS method

The main idea of the IF-TOPSIS method for ranking products is to normalize the original data matrix and determine the distance between the alternative products and the optimal or worst solution based on each attribute index's weight [20]. The relative closeness of each alternative product to the optimal solution is used as the evaluating basis. The steps of the IF-TOPSIS method are as follows.

Step 1: the mobile phone's positive ideal solution (PIS) A^+ and negative ideal solution (NIS) A^- are defined as follows.

$$A^+ = \left\{ a_{j+}, \max_{i=1}^{m} \langle a_{ij} \rangle; \ j = 1, 2, \ldots, n \right\} \quad (28)$$

$$A^- = \left\{ a_{j-}, \min_{i=1}^{m} \langle a_{ij} \rangle; \ j = 1, 2, \ldots, n \right\} \quad (29)$$

Then, the PIS and NIS of each mobile phone feature are shown in Table 9.

Step 2: Calculate the weighted distance from each alternative mobile phone A_i to the PIS A^+ and the NIS A^-.

$$D_i^+ = \sum_{j=1}^{n} \omega_j d(a_{ij}, a_{j+}) \quad (30)$$

$$D_i^- = \sum_{j=1}^{n} \omega_j d(a_{ij}, a_{j-}) \quad (31)$$

Step 3: Calculate the relative closeness (CI_i) of each alternative mobile phone A_i as follows.

$$CI_i = \frac{D_i^-}{D_i^+ + D_i^-}, \ i = 1, 2, \ldots, m \quad (32)$$

Here, the weighted Hamming distance between the alternative mobile phone A_i and the PIS A^+ or the NIS A^- represented by the IFSs is calculated. The ranking result calculated by the IF-TOPSIS method is shown in Table 10. The product ranking result is $A_3 > A_5 > A_2 > A_1 > A_4$. Namely, the best choice to buy the alternative mobile phone based on online reviews is OPPO R11S (A_3).

(2) Comparison with IF-VIKOR method

The IF-VIKOR method is developed by Yang et al. [40] to select the best compromise hotels.

Step 1: the PISs and NISs of mobile phones are shown in Table 9.

Table 6 Gain and loss matrix under each feature f_j between two alternative mobile phones

f_1

	A_1	A_2	A_3	A_4	A_5
A_1	0.000	−1.340	−1.383	−1.238	−1.245
A_2	0.234	0.000	−0.341	0.067	0.065
A_3	0.247	0.078	0.000	0.103	0.101
A_4	0.224	−0.511	−0.615	0.000	−0.128
A_5	0.225	−0.495	−0.601	0.018	0.000

f_2

	A_1	A_2	A_3	A_4	A_5
A_1	0.000	−0.499	−0.840	0.309	−0.817
A_2	0.155	0.000	−0.675	0.345	−0.646
A_3	0.274	0.227	0.000	0.413	0.069
A_4	−0.893	−1.023	−1.226	0.000	−1.210
A_5	0.266	0.216	−0.197	0.407	0.000

f_3

	A_1	A_2	A_3	A_4	A_5
A_1	0.000	−0.857	−0.748	−0.978	−0.934
A_2	0.221	0.000	0.083	−0.470	−0.371
A_3	0.205	−0.418	0.000	−0.629	−0.559
A_4	0.251	0.118	0.144	0.000	0.044
A_5	0.247	0.109	0.137	−0.290	0.000

f_4

	A_1	A_2	A_3	A_4	A_5
A_1	0.000	−0.480	−1.105	0.548	−0.536
A_2	0.240	0.000	−0.995	0.598	−0.239
A_3	0.559	0.504	0.000	0.782	0.498
A_4	−1.098	−1.199	−1.558	0.000	−1.222
A_5	0.252	0.078	−0.966	0.603	0.000

f_5

	A_1	A_2	A_3	A_4	A_5
A_1	0.000	−1.691	−1.296	−1.918	−2.374
A_2	0.382	0.000	0.289	−0.905	−1.666
A_3	0.289	−1.086	0.000	−1.414	−1.801
A_4	0.434	0.206	0.324	0.000	−2.289
A_5	0.539	0.380	0.455	0.319	0.000

f_6

	A_1	A_2	A_3	A_4	A_5
A_1	0.000	−0.590	−0.788	0.311	−0.570
A_2	0.278	0.000	−0.522	0.358	0.092
A_3	0.178	0.213	0.000	0.417	0.203
A_4	−0.865	−1.047	−1.170	0.000	−1.815
A_5	0.190	−0.303	−0.545	0.364	0.000

Step 2: calculate the S_i and R_i of alternative mobile phone A_i.

$$S_1 = 0.375, \ S_2 = 0.154, \ S_3 = 0.077, \ S_4 = 0.420, \ S_5 = 0.060.$$

$$R_1 = 1, \ R_2 = 0.494, \ R_3 = 0.704, \ R_4 = 1, \ R_5 = 0.366.$$

Step 3: assume that $v = 0.5$, and calculate the Q_i of alternative mobile phone A_i.

$$Q_1 = 0.062, \ Q_2 = 0.770, \ Q_3 = 0.710, \ Q_4 = 0, \ Q_5 = 1.$$

Step 4: obtain the ranking result of alternative mobile phones.

The ranking result of alternative mobile phones is obtained as $A_5 > A_2 > A_3 > A_1 > A_4$.

(3) Comparison with IF-PROMETHEE method

The IF-PROMETHEE method [23] is developed to support the consumers' purchase decisions.

Step 1: the priority index of A_i over A_j is shown in Table 7.

Step 2: the entering flow $\varphi^+(A_i)$ and exiting flow $\varphi^-(A_i)$ are calculated as

$$\varphi^+(A_1) = -5.265, \ \varphi^+(A_2) = -0.856, \ \varphi^+(A_3) = -0.137, \ \varphi^+(A_4) = -4.031, \ \varphi^+(A_5) = 0.352$$

$$\varphi^-(A_1) = 0.758, \ \varphi^-(A_2) = -2.353, \ \varphi^-(A_3) = -3.535, \ \varphi^-(A_4) = -0.470, \ \varphi^-(A_5) = -4.337$$

Step 3: the comprehensive outranking indices $\varphi(A_i)$ are

$$\varphi(A_1) = -6.023, \ \varphi(A_2) = 1.497, \ \varphi(A_3) = 3.399, \ \varphi(A_4) = -3.561, \ \varphi(A_5) = 4.689$$

Step 4: the ranking result of the alternative mobile phones is $A_5 > A_3 > A_2 > A_4 > A_1$.

(4) Discussion

To illustrate the effectiveness of ranking products based on the IF-TODIM method and online reviews, the product ranking results of the IF-TODIM method and the three other methods are shown in Fig. 2. The results show that the product ranking result by the IF-TODIM method is the same as the IF-PROMETHEE method and different from those of the two other methods. The best choice to buy the mobile phone obtained by the IF-TODIM, IF-VIKOR, and IF-PROMETHEE methods is A_5 (VIVO X9), while that of the IF-TOPSIS method is A_3 (OPPO R11S). A_1 (IPHONE X) and A_4 (Mito T8) are always the worst two choices. The main

Table 7 Dominance degree of mobile phone A_i over mobile phone A_e

	A_1	A_2	A_3	A_4	A_5
A_1	0.000	− 5.457	− 6.160	− 2.967	− 6.475
A_2	1.509	0.000	− 2.163	− 0.007	− 2.764
A_3	1.751	− 0.482	0.000	− 0.328	− 1.488
A_4	− 1.947	− 3.457	− 4.101	0.000	− 6.620
A_5	1.718	− 0.015	− 1.717	1.422	0.000

Table 8 Product ranking results under different attenuation coefficients

θ	Ranking result
1	$A_5 > A_3 > A_2 > A_4 > A_1$
2	$A_5 > A_3 > A_2 > A_4 > A_1$
3	$A_5 > A_3 > A_2 > A_4 > A_1$
4	$A_5 > A_3 > A_2 > A_4 > A_1$

Table 9 The PIS and NIS of each mobile phone feature

	PIS	NIS
f_1	[0.989, 0.006]	[0.757, 0.187]
f_2	[0.927, 0.055]	[0.651, 0.286]
f_3	[0.949, 0.025]	[0.802, 0.148]
f_4	[0.733, 0.240]	[0.206, 0.764]
f_5	[0.966, 0.023]	[0.226, 0.760]
f_6	[0.931, 0.063]	[0.670, 0.279]

Table 10 The ranking result calculated by the IF-TOPSIS method

	D_i^+	D_i^+	CI_i	Product ranking
A_1	0.240	0.134	0.359	4
A_2	0.136	0.238	0.637	3
A_3	0.064	0.403	0.864	1
A_4	0.292	0.082	0.219	5
A_5	0.078	0.297	0.793	2
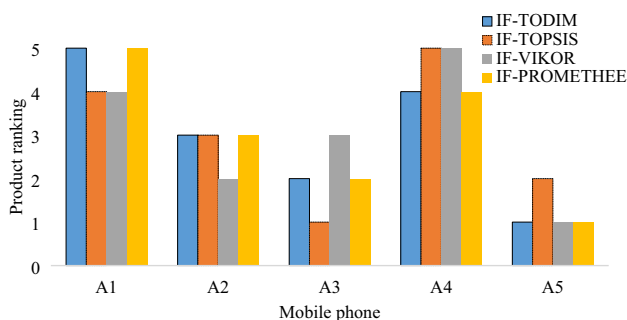
Fig. 2 Product ranking results by the developed IF-TODIM method and other methods

reason for the different results is that the IF-TODIM method considers the gain and loss of each mobile phone feature and prospect value in the product ranking process. VIVO X9 has some advantages in the attribute of price, and other features reappraise from all the features. The ranking result by the IF-TODIM method is closer to the actual situation. The customers are fully rational in purchasing mobile phones under the IF-TOPSIS and IF-VOKOR method. Customers are non-fully rational in the purchase decision process. The IF-TOPSIS and IF-VOKOR method is not reasonable for the ranking product. Therefore, the IF-TODIM method based on online reviews is more reasonable than the IF-TOPSIS and IF-VIKOR method.

Conclusion

In this paper, a new analytical method for ranking products is presented. The main idea of ranking product method through

online reviews and IF-TODIM is as follows. Firstly, the Apriori algorithm is used to identify the product features based on online reviews. Then the sentiment orientation and intensity of the sentiment words for the product features are identified by the lexicon-based sentiment analysis approach. Next, the sentiment orientation of the product features is converted into an IFV, and then the IF-TODIM method is used to determine the ranking results of the alternative products.

The proposed method fully considers consumers' subjective needs and different sentiment orientations (positive, neutral, and negative) for each product feature. The IFVs are used to fully reflect the different sentiment orientations of online reviews, which is more elaborate than previous studies and makes up for the lack of consideration of the neural sentiment orientation. In addition, the gain and loss of each mobile phone feature in the product ranking process are also considered. The obtained result is closer to the actual purchase needs of consumers. In general, the degree of membership, non-membership, and hesitation in IFV provides an effective way to solve the problem of product ranking. The proposed method has operability and practical application value and provides a new decision-making technology to solve the problem of product purchase decision-making using online review data in the current era of big data.

The developed method provides a convenient tool to give recommendations for purchasing products, and the decision support system needs to improve. In addition, the emojis and photos in the online review data are neglected during the data pre-processing process. In future work, it is necessary to study the product ranking method combing with emojis and photos.

Data availability The data used to support the findings of this study are included within the article.

Declarations

Conflict of interest We declare that we do have no commercial or associative interests that represent a conflict of interests in connection with this manuscript. There are no professional or other personal interests that can inappropriately influence our submitted work.

Research involving human participants and/or animals This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Naragund GH, Santhosh Kumar KL, Majumdar J (2015) Development of decision making and analysis on customer reviews using sentiment dictionary for human-robot interaction. Int J Adv Res Comput Commun Eng (IJARCCE) 4(8):387–391
2. Zhang Z, Zhang H, Zhou L, Li Y (2021) Analyzing the coevolution of mobile application diffusion and social network: a multi-agent model. Entropy 23(5):521
3. Zhou L, Lin J, Li Y, Zhang Z (2020) Innovation diffusion of mobile applications in social networks: a multi-agent system. Sustainability 12(7):2884
4. Zhang K, Narayanan R, Choudhary AN (2010) Voice of the customers: mining online customer reviews for product feature-based ranking. WOSN 10:11–11
5. Zhang K, Cheng Y, Liao WK, Choudhary A (2011, August) Mining millions of reviews: a technique to rank products based on importance of reviews. In: Proceedings of the 13th international conference on electronic commerce, pp 1–8
6. Kang D, Park Y (2014) Review-based measurement of customer satisfaction in mobile service: sentiment analysis and VIKOR approach. Expert Syst Appl 41(4):1041–1050
7. Najmi E, Hashmi K, Malik Z, Rezgui A, Khan HU (2015) CAPRA: a comprehensive approach to product ranking using customer reviews. Computing 97(8):843–867
8. Li MY, Zhao XJ, Zhang L, Ye X, Li B (2020) Method for product selection considering consumer's expectations and online reviews. Kybernetes 50(9):2488–2520
9. Fan ZP, Xi Y, Liu Y (2018) Supporting consumer's purchase decision: a method for ranking products based on online multi-attribute product ratings. Soft Comput 22(16):5247–5261
10. Zhang Z, Li J, Sun Y, Lin J (2019) Novel distance and similarity measures on hesitant fuzzy linguistic term sets and their application in clustering analysis. IEEE Access 7:100231–100242
11. Wu S, Lin J, Zhang Z (2020) New distance measures of hesitant fuzzy linguistic term sets. Phys Scr 96(1):015002
12. Zhang Z, Lin J, Miao R, Zhou L (2019) Novel distance and similarity measures on hesitant fuzzy linguistic term sets with application to pattern recognition. J Intell Fuzzy Syst 37(2):2981–2990
13. Peng Y, Kou G, Li J (2014) A fuzzy PROMETHEE approach for mining customer reviews in Chinese. Arab J Sci Eng 39(6):5245–5252
14. Zhang D, Wu C, Liu J (2020) Ranking products with online reviews: a novel method based on hesitant fuzzy set and sentiment word framework. J Operat Res Soc 71(3):528–542
15. Bi JW, Liu Y, Fan ZP (2019) Representing sentiment analysis results of online reviews using interval type-2 fuzzy numbers and its application to product ranking. Inf Sci 504:293–307
16. Fu X, Ouyang T, Yang Z, Liu S (2020) A product ranking method combining the features–opinion pairs mining and interval-valued Pythagorean fuzzy sets. Appl Soft Comput 97:106803
17. Liu P, Teng F (2019) Probabilistic linguistic TODIM method for selecting products through online product reviews. Inf Sci 485:441–455
18. Ji P, Zhang HY, Wang JQ (2018) A fuzzy decision support model with sentiment analysis for items comparison in e-commerce: The case study of http://PConline.com. IEEE Trans Syst Man Cybern Syst 49(10):1993–2004

19. Liang R, Wang JQ (2019) A linguistic intuitionistic cloud decision support model with sentiment analysis for product selection in E-commerce. Int J Fuzzy Syst 21(3):963–977

20. Liang X, Liu P, Wang Z (2019) Hotel selection utilizing online reviews: a novel decision support model based on sentiment analysis and DL-VIKOR method. Technol Econ Dev Econ 25(6):1139–1161

21. Fan ZP, Li GM, Liu Y (2020) Processes and methods of information fusion for ranking products based on online reviews: an overview. Information Fusion 60:87–97

22. Liu Y, Bi JW, Fan ZP (2017) A method for ranking products through online reviews based on sentiment classification and interval-valued intuitionistic fuzzy TOPSIS. Int J Inf Technol Decis Mak 16(06):1497–1522

23. Liu Y, Bi JW, Fan ZP (2017) Ranking products through online reviews: a method based on sentiment analysis technique and intuitionistic fuzzy set theory. Information Fusion 36:149–161

24. Çalı S, Balaman ŞY (2019) Improved decisions for marketing, supply and purchasing: mining big data through an integration of sentiment analysis and intuitionistic fuzzy multi criteria assessment. Comput Ind Eng 129:315–332

25. Zhang D, Li Y, Wu C (2020) An extended TODIM method to rank products with online reviews under intuitionistic fuzzy environment. J Operat Res Soc 71(2):322–334

26. Szmidt E, Kacprzyk J (2000) Distances between intuitionistic fuzzy sets. Fuzzy Sets Syst 114(3):505–518

27. Szmidt E, Kacprzyk J (2001) Entropy for intuitionistic fuzzy sets. Fuzzy Sets Syst 118(3):467–477

28. Chen TY, Li CH (2010) Determining objective weights with intuitionistic fuzzy entropy measures: a comparative analysis. Inf Sci 180(21):4207–4222

29. Xu Z, Zhao N (2016) Information fusion for intuitionistic fuzzy decision making: an overview. Inf Fusion 28:10–23

30. Atanassov KT (1989) More on intuitionistic fuzzy sets. Fuzzy Sets Syst 33(1):37–45

31. Xu Z (2007) Intuitionistic fuzzy aggregation operators. IEEE Trans Fuzzy Syst 15(6):1179–1187

32. Liu P, Zhang P (2020) Normal wiggly hesitant fuzzy TODIM approach for multiple attribute decision making. J Intell Fuzzy Syst 39(1):627–644

33. Lu J, Wei C (2019) TODIM method for performance appraisal on social-integration-based rural reconstruction with interval-valued intuitionistic fuzzy information. J Intell Fuzzy Syst 37(2):1731–1740

34. Deng X, Gao H (2019) TODIM method for multiple attribute decision making with 2-tuple linguistic Pythagorean fuzzy information. J Intell Fuzzy Syst 37(2):1769–1780

35. Huang YH, Wei GW (2018) TODIM method for Pythagorean 2-tuple linguistic multiple attribute decision making. J Intell Fuzzy Syst 35(1):901–915

36. Zhang Z, Lin J, Zhang H, Wu S, Jiang D (2020) Hybrid TODIM method for law enforcement possibility evaluation of judgment debtor. Mathematics 8(10):1806

37. Gomes L, Lima M (1992) TODIM: Basics and application to multicriteria ranking of projects with environmental impacts. Found Comput Decis Sci 16(4):113–127

38. Zhang Z, Zhao X, Qin Y, Si H, Zhou L (2021) Interval type-2 fuzzy TOPSIS approach with utility theory for subway station operational risk evaluation. J Ambient Intell Hum Comput. https://doi.org/10.1007/s12652-021-03182-0

39. Gomes L, Lima M (1992) From modeling individual preferences to multicriteria ranking of discrete alternatives: a look at prospect theory and the additive difference model. Found Comput Decis Sci 17(3):171–184

40. Yang Z, Gao Y, Fu X (2021) A decision-making algorithm combining the aspect-based sentiment analysis and intuitionistic fuzzy-VIKOR for online hotel reservation. Ann Operat Res. https://doi.org/10.1007/s10479-021-04339-y