IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 1, 2021, accepted March 16, 2021, date of publication March 22, 2021, date of current version March 31, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3067844

Sentiment Analysis Technique and Neutrosophic Set Theory for Mining and Ranking Big Data From Online Reviews

IBRAHIM AWAJAN[1], **MUMTAZIMAH MOHAMAD**[1], **AND ASHRAF AL-QURAN**[2]
[1]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Besut Campus, Kuala Terengganu 22200, Malaysia
[2]Preparatory Year Deanship, King Faisal University, Al Hofuf 31982, Saudi Arabia

Corresponding author: Ibrahim Awajan (si2150@putra.unisza.edu.my)

This work was supported in part by the Center for Research Excellence and Incubation Management (CREIM), and in part by the Faculty of Informatics and Computing at the University Sultan Zainal Abidin (UniSZA).

ABSTRACT Recently, a huge amount of online consumer reviews (OCRs) is being generated through social media, web contents, and microblogs. This scale of big data cannot be handled by traditional methods. Sentiment analysis (SA) or opinion mining is emerging as a powerful and efficient tool in big data analytics and improving decision making. This research paper introduces a novel method that integrates neutrosophic set (NS) theory into the SA technique and multi-attribute decision making (MADM) to rank the different products based on numerous online reviews. The method consists of two parts: Determining sentiment scores of the online reviews based on the SA technique and ranking alternative products via NS theory. In the first part, the online reviews of the alternative products concerning multiple features are crawled and pre-processed. A neutral lexicon consists of 228 neutral words and phrases is compiled and the Valence Aware Dictionary and sEntiment Reasoner (VADER) for sentiment reasoning is adapted to handle the neutral data. The compiled neutral lexicon, as well as the adapted VADER, are utilized to build a novel adaptation called Neutro-VADER. The Neutro-VADER assigns positive, neutral, and negative sentiment scores to each review concerning the product feature. In this stage, the novel idea is to point out the positive, neutral, and negative sentiment scores as the truth, indeterminacy, and falsity memberships degrees of the neutrosophic number. The overall performance of each alternative concerning each feature based on a neutrosophic number is measured. In the second part, the ranking of alternatives is being evaluated through the simplified neutrosophic number weighted averaging (SNNWA) operator and cosine similarity measure methods. A case study with real datasets (Twitter datasets) is provided to illustrate the application of the proposed method. The results show good performance in handling the neutral data on the SA stage as well as the ranking stage. In the SA stage, findings show that the Neutro-VADER in the proposed method can deal successfully with all types of uncertainties including indeterminacy comparable with the traditional VADER in the other methods. In the ranking stage, the results show a great similarity and consistency while using other ranking methods such as PROMETHEE II, TOPSIS, and TODIM methods.

INDEX TERMS Sentiment analysis, VADER, online reviews, neutrosophic set, ranking product, simplified neutrosophic number, SNNWA operator, and decision making.

I. INTRODUCTION

The developments of e-commerce technology and platforms of the era of big data has rapidly emerged. To date, many social media networks and e-trade websites have provided platforms for consumers to select their products and post their online product reviews. The purchasing decision of expected customers is influenced to a certain degree by online reviews. That is, before making the purchasing decisions, potential customers can read and evaluate these reviews which help them make an appropriate purchasing decision. However, online reviews are not available for customers to read and then to assess what they desire to buy. This is mainly because of the diversity and the great number of these reviews. Thus, the importance of way to rank the desired products is clearly presented to help potential consumers make informed

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wang.

decisions. Recently few studies have used SA and MADM methods together to rank alternative products through online reviews [1]–[7].

SA adopts natural language processing which mainly focuses on pointing out the writer's moods, opinions, attitudes, and feelings from the different texts about different topics [8]. The sentiment orientation of each online review is automatically identified in the SA technique. Then the performance of alternative products and their features are analyzed.

Decision-making process consists of the evaluation of alternatives and the choice of the most preferable from them. MADM refers to rank alternatives or to select the best choice based on multiple attribute evaluation values of the different alternatives. MADM firstly proposed by [9] is one of the most significant research topics in decision theory. Both the theory and the methods of the MADM have been used in management, society, economy, military and engineering, and other fields.

However, most of the previous studies [1]–[7] on the combination of SA and MADM, do not handle neutral or indeterminate data. The reviews with neutral sentiment orientations represent the hesitant or uncertain evaluations of consumers concerning products, should not be ignored because they are also valuable for the potential consumer to make a reasonable decision. As a variation of fuzzy and intuitionistic fuzzy numbers, a neutrosophic number is a valuable data form to represent information with hesitance, neutrality, and uncertainty. Thus, based on the sentiment scores of reviews with positive, neutral, and negative sentiment orientations, neutrosophic numbers can be constructed to represent the performance of alternative products concerning product features.

For the first time, neutrosophy has been integrated with sentiment analysis [10] as an attempt to find a method that could solve the uncertainties arising from the discursive analysis. Then, it has been used in the sentiment analysis of tweets [11] and speech sentiment analysis [12].

This paper proposes a novel method that combines SA technique, MADM, and NS theory to rank the alternatives of products determined by the potential customers according to a set of product features and recommend the most preferable alternative. This paper focuses on English online reviews because English is the most represented language online.

The contribution of this research paper points out the following. Firstly, this paper proposes a comprehensive approach, which considers multiple factors of product selection such as product features, sentiment orientations of online reviews, product feature weights, and posted time of reviews. Secondly, VADER has been adapted to handle the neutral data. Thirdly, a neutral lexicon consists of 228 neutral words and phrases is compiled and then integrated into the adapted VADER to develop a Neutro-VADER. Fourthly, neutrosophic logic is utilized in SA and MADM processes simultaneously to rank the alternative products through real online reviews. Lastly, a comprehensive comparative study of the proposed method and other decision-making methods is conducted to validate the results in the decision process.

II. RELATED WORKS

There are three research areas related to this paper, namely SA, ranking products based on online reviews and, NS theory are discussed in this section.

A. SENTIMENT ANALYSIS

The processes of extracting, analyzing, and identifying opinions or emotions within a text are known as SA or opinion mining [13]. Whether the text is a document, a paragraph, a sentence, or an entire paragraph [14]. SA terminology and opinion mining first appeared in 2003 [15]. Given the importance of SA, has enjoyed a huge of research activity. Studies on SA have not been limited to the English language, but included many languages such as Arabic [16], [17], Chinese [18], [19], Indian [20], [21], and studies on other languages have been conducted in this direction. The following discussion is about three different aspects of SA.

1- Classification of text polarity: SA is a method for analyzing text to express polarity. Whether this classification is binary classification (positive or negative) [22], [23], or multi-class classification [24], [25].

[26], [27], are among the earliest studies in this domain. Pang *et al.* [26] found the effectiveness of applying support vector machine and naïve Bayes on movie reviews for the sentiment classification task. By two-class classification, Vargas-Calderón et al [28] measured the extent of the positive impact by including the Real Academia Espanola de la Lengua (RAE's) dictionary definitions in Word2Vec's neural network (NN) training sentences. A crawler was built to obtain such definitions. Khan *et al.* [29] proposed a multilingual framework applied on seven benchmark datasets for polarity classification with two different aspects of contributions. Tellez *et al.* [30] presented a simple method called enhanced SA and polarity classification (ESAP), to be a multilingual framework easy to implement and use in Twitter. Pang and Lee [31] dealt with the rating-inference issue, by expanding the prediction from two-class classification (positive or negative) into a star rating scale (3 or a 4 star). Kim and Hovy [32] applied various classifiers to identify the sentiment of text and holder of the text concerning a given topic. Lei *et al.* [33], proposed a sentiment-based rating prediction method (RPS) to improve prediction accuracy in recommender systems. Where the researchers proposed a social user sentimental measurement method for calculating user's sentiment towards different products and various items. The results shown the sentiment can characterize user preferences.

2-Levels of the text in sentiment analysis: The SA can be performed at three different levels of the text: Document-level [26], [34], [35], sentence-level [36]–[38], and aspect-level [39]–[41].

In document level. Bollegala *et al.* [40] proposed a method to perform cross-domain SA using a sentiment-sensitive thesaurus (SST). Blitzer *et al.* [41] extended the structural correspondence learning (SCL). Where the pivots had selected using the mutual information between a feature (bigrams or unigrams) and the domain label. Li *et al.* [44] proposed the SCL method for cross-lingual sentiment classification. This method used large amounts of monolingual data as well as a small dictionary to learn meaningful one-to-many mappings for pivot words. On the level of the sentence, González-Ibánez *et al.* [45] attempted to find a solution for the problem of automatic sarcasm detection in Twitter messages. The researchers studied the effect of both pragmatic and linguistic features of tweets in order to achieve the automatic separation of sarcastic messages into positive and negative ones. To sarcasm identification, Tsur *et al.* [46] proposed a semi-supervised algorithm entitled SASI. This algorithm recognizes sarcastic sentences through product reviews. Appel *et al.* [37] presented a hybrid approach at the sentence level to estimate the semantic orientation polarity and its intensity for sentences. This is achieved by using a sentiment lexicon enhanced with the assistance of SentiWordNet, natural language processing essential techniques, as well as fuzzy sets. Fu *et al.* [47] presented a model using rhetorical structure theory for text parsing. This approach able to make the representations concerning the relations between segments of text, which can lead to improved semantic representations of the text.

In the aspect level, Moghaddam and Ester [48] presented an unsupervised method for aspect extraction from unstructured reviews using known aspects. This is useful in cases where the customer satisfaction of services and products is well understood [49]. Farhadloo and Rolland [50] proposed a method to identify the aspects using cluster analysis. The authors performed a sentiment classification of text fragments into positive, neutral, and negative sentiments using score representation. Where aspect identification was based on a bag of nouns (BON).

3- Classification methods: The existing classification methods which are proposed in the literature are categorized into two groups: The machine learning approach, and the Lexicon-based approach [51], [52].

In the machine learning approach, Abdul aziz and Starkey [53] presented a method known as contextual analysis that can perform sentiment analysis without using any linguistics resources. Bastı*et al.* [54] demonstrated that contemporary machine learning techniques are viable data analysis tools for the critical assessment of initial public offerings valuation. The decision tree in [55] is one of the most popular inductive learning algorithms which is used in the supervised sentiment classification approach. In the same line, the support vector machine (SVM) as a machine learning technique that is based on statistical learning theory [56] is one of the most famous supervised classification methods according to the research report [57], [58]. This machine learning technique has high precision in text classification, which makes this technique famous in sentiment classification [59]. On the other hand, SVM is not probably suitable for large datasets classification [60]. EL abdouli *et al.* [61] presented a way to the SA on Twitter data written in multilingual. The tweets were classified into positive and negative classes using the symbols of emotion. The classification applied in this way is based on naïve Bayes classifier.

In the lexicon-based approach, one of the categories of lexicons approach is based on dictionary-based [62]–[65]. A set of hand-picked sentiment selected (features) are created and then expanded using a thesaurus or tools like WordNet [66]. Liu's lexicon [67] is a fully established example of this approach. Qiu *et al.* [68] proposed a method to assign polarities to newly discovered sentiment words in a domain. They [69] also presented a technique based on bootstrapping propagation and a few aspects like using of word seeds. A scoring algorithm was used to determine the polarity of tweets. Pandarachalil *et al.* [70] used the lexicon method to analyze Twitter sentiments. The polarity of tweets is estimated through three sentiment lexicons (SenticNet, SentislangNet, and SentiWordNet). Saif *et al.* [71] presented a lexicon-based approach to analyze the opinions on Twitter called SentiCircle. A proposed approach can detect the sentiment for both entity-level and tweet-level.

B. RANKING PRODUCTS BASED ON ONLINE REVIEWS

For the comparison of online products review, a prototype system depending on the web has been done by [72]. Natural language processing has been employed to read reviews automatically. For determining the polarity of reviews, Rajeev and Rekha [73] applied naïve Bayes classification. Also, they focused on both extracting the reviews of product features and assigning the polarity of those features. They applied their experiments on E-Commerce site: Flipkart, on mobile phone reviews to prove that the system helps the customers to choose the appropriate product.

To deal with the problem of online product ranking, Najmi *et al.* [4] proposed an approach to solving the problem of online product ranking by combining the SA of reviews, product aspect analyzer, review usefulness analysis, and product brand ranks. The authors applied their experiments on two main categories, TVs and Cameras using English online reviews (contains 197 products and 56,368 reviews from Amazon). Experiments show progress in enhancing the ranking process for new product releases when using the brand rank in the field of information retrieval for ranking online products.

Wang *et al.* [73] presented an econometric model that integrates between the SA and econometric models. This model helps in ranking product aspects through online reviews. They have taken into consideration inspecting the effect of opinion changes on sales volume to detect the aspect ranking. They gave a case study using online reviews in English. They collected their experimental corpora from Amazon to follow 386 digital cameras. Results showed that the aspect weight for digital cameras outperformed HAC (High Adjective Count) and TF-ID algorithms.

Guo *et al.* [7] proposed a ranking method that combines objective and subjective sentiment values. This method has been applied based on different aspects of alternative products through online reviews. This paper calculated the objective sentiment value of the product by determining the weights of aspects with the Latent Dirichlet Allocation (LDA) topic model. This study forms a directed graph model to fuse heterogeneous online reviews. The authors conducted a case study using Chinese online reviews. Besides, the authors proved that experimental results have a strong correlation with actual sales ranking by applying the Spearman coefficient.

In the same year, Kumar J and Abirami [74] proposed a framework to solve the problem of ranking alternative products (through reviews) based on aspects of products. They used the OpinRank review dataset in English, which contains 42,230 reviews to identify the aspects and opinion words, by applying a Harel-Koren fast multiscale layout. The products had been ranked based on positive and negative ranks through applying a Spearman's rank correlation coefficient based opinion ranking method. For the task of aspect-based sentiment classification, they performed the supervised learning methods to the sentiment classification using naïve Bayes, maximum entropy, and support vector machine.

To solve the problem of ranking alternative products through online reviews, Wu and Zhang [5] presented a method that combined SA techniques, intuitionistic fuzzy set theory, and MADM. This research has taken into account the attention degree to point out the weight of every feature. In this paper, the weight vector has been calculated by combining the frequency and attention degree of every feature. To obtain the dataset, the authors used the web crawler that has been written in Python to crawl reviews of related products from online shopping websites. Chinese online reviews have been used in the case study.

Liu *et al.* [1] also proposed a method based on SA technique and intuitionistic fuzzy set theory to rank products through online reviews. This process mainly contains two phases: First, identifying sentiment orientations of the online reviews through the SA technique. Second, ranking the alternative products through the intuitionistic fuzzy set theory. This study presented a way to convert the identified sentiment orientations into intuitionistic fuzzy numbers. According to intuitionistic fuzzy numbers and feature weights suggested by the consumer, an overall intuitionistic fuzzy number of every suggested alternative product is calculated by using the intuitionistic fuzzy weighted averaging operator. After that, the alternative products are ranked by PROMETHEE II.

C. NEUTROSOPHIC SET THEORY

Fuzzy set [75] and intuitionistic fuzzy set theories [76] can handle only uncertain and incomplete data but not neutral data that exists usually in real situations. To handle all types of uncertainty including neutrality, Smarandache [77] originally gave a concept of a NS from a philosophical point of view which is a part of neutrosophy. The words neutrosophy and neutrosophic were introduced by Smarandache in his 1998 book [78]. Etymologically, "neutrosophic" (noun) means knowledge of neutral thought, while "neutrosophic" (adjective), means having the nature of, or having the characteristic of neutrosophy. NS is characterized by a truth-membership function, indeterminacy-membership function, and falsity membership function, denoted by T, I, F, respectively, where the indeterminacy-membership function is independent of truth-membership function and falsity-membership function. The ranges of the functions T, I, and F are subsets of the real standard or nonstandard interval $]^-0, 1^+[$. To constrain them in the real standard interval $[0, 1]$ for convenient science and engineering applications, single-valued NS and its operators were introduced by [79] as the subclasses of the NSs.

Next, Ye [80] introduced a simplified NS which includes a single-valued NS, and proposed a MCDM method using the aggregation operators and cosine similarity measure for simplified NSs. However, Peng *et al.* [81] verified that in some cases, the simplified NSs' operations Ye [80] may be impractical. Therefore, they re-defined the operations and the aggregation operators for simplified neutrosophic numbers and established a multi-criteria group decision-making method based on the proposed operators. Later on, simplified NS and its variations have been applied extensively in MADM methods of operational research such as PROMETHEE [82] ELECTRE [83], TOPSIS-based QUALIFLEX [84], COPRAS [85], EDAS [86], TODIM [87], and others.

III. BASIC CONCEPTS

This section recapitulates the concepts of SA, neutrosophic, and simplified NSs. This section also provides an overview of some relevant concepts.

A. SENTIMENT ANALYSIS

SA is known as the process of computationally identifying and categorizing opinions expressed in any single piece of text, especially to point out whether the writer's opinion to a specific topic, product, etc. is positive, negative, or neutral.

SA techniques are classified into two main categories, namely, SA techniques based on machine learning and lexicon-based SA techniques. In turn, the SA techniques based on machine learning can also be divided into three subclasses, namely, (1) SA techniques based on supervised machine learning, (2) SA techniques based on unsupervised machine learning, and (3) SA techniques based on semi-supervised machine learning. Meanwhile, the lexicon-based SA techniques are divided also into two other subclasses, i.e., (1) dictionary-based SA techniques and, and (2) corpus-based SA techniques. The corpus-based SA techniques are mainly used to solve the problem of searching and finding opinion words with context-specific orientations. The basis of the techniques is to find the syntactic patterns as well as a seed list of various opinion words. The basis of dictionary-based SA techniques is to construct the SDs. In the dictionary-based SA techniques, a small set of sentiment words is manually determined at the beginning. After that, new sentiment words are found out and consequently, the number of the sentiment words is increased by looking for both the synonyms and antonyms of the sentiment words in the already known Corpora. The iterative process is carried out until no new sentiment words can be further found, and the SDs are composed in the light of the obtained sentiment words.

1) TWITTER SENTIMENT ANALYSIS

SA is usually applied to the data that is collected from the Internet and various social media platforms such as Facebook, Twitter, Google+, and several blogs.

Twitter is one of the most popular microblog platforms in which users can publish their thoughts and opinions. Twitter SA is an application of SA on Twitter data (tweets) which extracts meaningful information from tweets. Twitter SA also gives results in terms of percentage sentiment on a particular scale.

Now a day's Twitter is considered a good resource for getting user's opinions, hence SA of the tweets can be considered as an effective way of calculating public opinion for the business market as well as many other topics [88].

2) VADER SENTIMENT ANALYSIS

VADER is a model used for text SA. Vader model had been introduced in 2014 to deal with social media texts, movie reviews, and product reviews. VADER model is sensitive to both polarity (positive/negative) and intensity (strength) of emotion [89].

VADER has many advantages over other well-known methods of SA, including acting well on social media type text with no need for any training data. Furthermore, VADER is fast enough to be used online with streaming data, and it is not affected by a speed-performance trade-off. Moreover, VADER can detect sentiment from slangs and emojis that constitute a real important part of the social media environment.

VADER sentiment analyses are mainly based on certain points:

- *Punctuation:* For example, the use of an exclamation mark (!) raises up the magnitude of the intensity without changing the semantic orientation. The following tweet can be considered an example: "The food here is good!", it is noted that it is more intense than the following tweet "The food here is good." and an increase in the number of (!), leads to an increase in the magnitude consequently.
- *Capitalization:* In the case of using upper case letters to emphasize a sentiment-relevant word in the presence of other non-capitalized words, this consequently increases the magnitude of the sentiment intensity. For instance, "The food here is GREAT!" shows more intensity than "The food here is great!"
- *Degree Modifiers (Intensifiers):* Degree modifiers affect the sentiment intensity in two ways either increasing or decreasing the intensity. For example, "The service here is extremely good" is more intense than "The service here is good", On the other hand, "The service here is marginally good" does reduce the intensity.
- *Conjunctions:* Using conjunctions like "but" denotes a change in sentiment polarity, with the sentiment of the text following the conjunction being dominant. "The food here is great, but the service is horrible" has mixed sentiment, with the latter half directing the general rating.

B. SCRAPESTORM

ScrapeStorm is a new generation of web scraping software. ScrapeStorm is developed by the former Google search technology team and based on artificial intelligence technology. It is used to extract data and to clean the extracted data during the extraction process, which is a product tailored for academic research, data analysis, non-programming products, sales, and e-commerce, as well as finance.

C. NEUTROSOPHIC SETS

To begin with, the definition of NS, followed by the definition of simplified NS are recalled.

Definition (1) [77]: Let U be a universe of discourse. A neutrosophic set N in U is defined as:

$N = \{< u; T_N(u); I_N(u); F_N(u) >; u \in U\}$, where $T_N(u)$ is the truth membership function, $I_N(u)$ is the indeterminacy membership function, and $F_N(u)$ is the falsity membership function. $T_N(u)$, $I_N(u)$ and $F_N(u)$ are standard or nonstandard subsets of $]^-0; 1^+[$ respectively, that is $T; I; F : X \to]^-0; 1^+[$ and there is no restriction on the sum of $T_N(u)$, $I_N(u)$ and $F_N(u)$, therefore $^-0 \leq T_N(u)+I_N(u)+F_N(u) \leq 3^+$.

However, applying NSs to practical problems is difficult. Therefore, Ye [80] reduced NSs of nonstandard intervals into a kind of simplified NSs of standard intervals as follows.

Definition (2) [80]: Let U be a space of points (object) with a generic element u in U. A simplified NS S in U is defined as:

$S = \{< u; T_S(u); I_S(u); F_S(u) >; u \in U\}$, where $T_S(u)$ is the truth membership function, $I_S(u)$ is the indeterminacy membership function, and $F_S(u)$ is the falsity membership function. $T_S(u)$, $I_S(u)$ and $F_S(u)$ are singleton subintervals/subsets in the real standard interval [0, 1], that is $T_S; I_S; F_S$: $U \to [0, 1]$ and $0 \leq T_S(u) + I_S(u) + F_S(u) \leq 3$.

This paper uses the simplified NS whose membership values are singletons in the real standard interval [0, 1]. Thus, each simplified NS can be described by three real numbers in the real unit interval [0, 1].

IV. RESEARCH METHODS AND MATERIAL

This section presents the problem of ranking products through online reviews and the method for ranking products through online reviews as follows.

A. THE PROBLEM OF RANKING ALTERNATIVE PRODUCTS THROUGH ONLINE REVIEWS

We first formulate the problem of ranking products through online reviews, and then briefly describe a resolution process for it.

1) DESCRIPTION OF THE PROBLEM OF RANKING PRODUCTS THROUGH ONLINE REVIEWS

Consider a consumer who wants to purchase a product such as a phone. First of all, the consumer identifies several suitable suggested alternative phones. Although the consumer can identify alternative phones, the consumer is hesitant to choose the most suitable phone due to insufficient experience and knowledge. To determine the most suitable phone for the consumer among alternative phones, the consumer selects several candidate phone features. Besides, determining feature weights according to personal preferences.

To support the consumer purchase decision, the online reviews of the alternative products concerning the multiple phone features are collected from a website rich in related reviews. The issue addressed in this paper is how to rank the alternative products based on the online reviews and the feature weights determined by the consumer. The problem of ranking alternative products through online reviews mentioned above is clearly shown in Fig. 1.

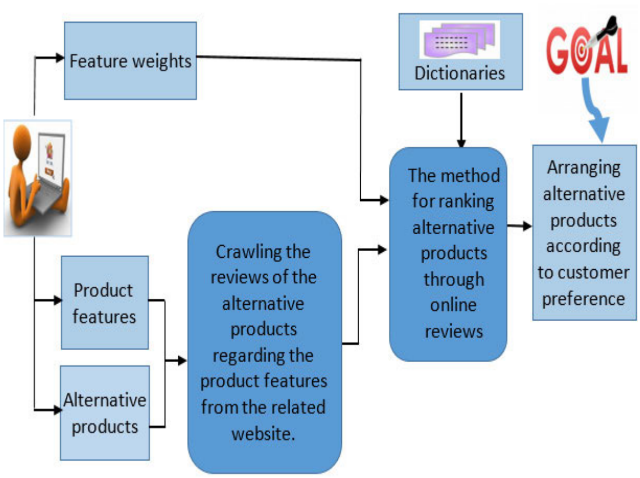


FIGURE 1. The problem of ranking alternative products through online reviews.

The following notations will be used to denote the sets and variables throughout the paper.

$w = (w_1, w_2, \ldots, w_m)$: denotes the vector of weights of m features, provided by the consumer, where w_j denotes the weight of feature f_j such that $w_j \geq 0$ and $\sum_{i=1}^{n} w_j = 1$, $j = 1, 2, \ldots,$ m.

$A = \{A_1, A_2, \ldots A_n\}$: denotes the set of n alternative products, where A_i denotes the i th alternative product, $i = 1, 2, \ldots, n$. Usually, the set A can be determined by the consumer.

$F = \{F_1, F_2 \ldots F_m\}$: denotes the set of product features, $D_{ik} = \left\{D_{ik}^1, D_{ik}^2, \ldots, D_{ik}^m\right\}$: denotes the sentence concerning feature f_j in the review D_{ik}, $i = 1, \ldots, n, j = 1, \ldots, m$, $k = 1, \ldots, q_i$.

$S_{ik}^j \left(a_{ik}^j, b_{ik}^j, c_{ik}^j\right)$: denotes the indicator vector which represents the sentiment score of sentence or review D_{ik}^j, where $a_{ik}^j, b_{ik}^j, c_{ik}^j$ are indicator variables for positive, neutral, and negative sentiment score, respectively.

u_{ij}^{pos} , u_{ij}^{neu} and u_{ij}^{neg}: denote the weighted aggregation values of a_{ik}^j, b_{ik}^j and c_{ik}^j are used to all reviews on alternative product A_i concerning feature f_j respectively.

2) THE RESOLUTION PROCESS FOR THE PROBLEM OF RANKING PRODUCTS THROUGH ONLINE REVIEWS

The resolution process for the problem of ranking products through online reviews is clearly shown in Fig. 2. It can be seen from Fig. 2 that the resolution process encompasses three stages, i.e., 1) crawling and preprocessing the online reviews, 2) computing sentiment scores of the online reviews based on the SA technique, and 3) ranking alternative products via NS theory.

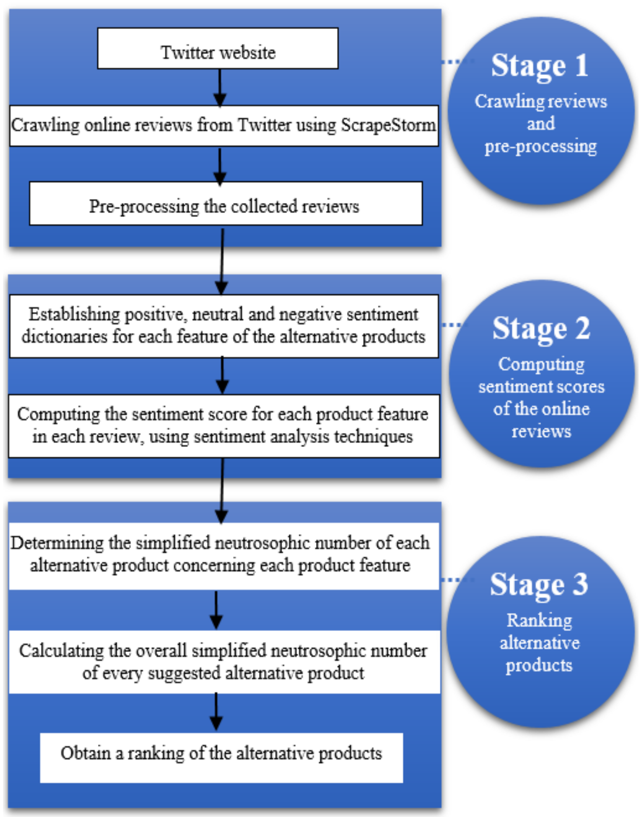


FIGURE 2. The resolution process for the problem of ranking products through online reviews.

In the first stage, the online reviews of the alternative products concerning multiple features concerned by the consumer are crawled and pre-processed.

The second stage aims to point out the positive, neutral, and negative sentiment scores as the truth, indeterminacy, and falsity memberships degrees of the neutrosophic number. Meanwhile, a neutral lexicon consists of 228 neutral words and phrases is compiled and the VADER for sentiment reasoning is adapted to handle the neutral data. The compiled neutral lexicon is used to establish a neutral dictionary in the light of the obtained online reviews. Further, the positive and negative dictionaries are established in the light of the obtained online reviews and Vader_lexicon. The positive, negative, and neutral SDs are established for each feature of the alternative products. These dictionaries as well as the adapted VADER are utilized to build a novel adaptation called Neutro-VADER which assigns positive, neutral, and negative sentiment scores to each review concerning the product feature.

The last stage aims to rank alternative products via NS theory. In this stage, the overall performance of each alternative concerning each feature based on a neutrosophic number is measured. The ranking of alternatives is also being evaluated through the simplified neutrosophic number weighted averaging (SNNWA) operator and cosine similarity measure methods.

B. THE METHOD FOR RANKING PRODUCTS THROUGH ONLINE REVIEWS

The detailed explanation of the resolution process shown in Figure 2, is presented in this section. We begin by first presenting the crawling and pre-processing stage. Then we move to the stage of computing sentiment scores. Finally, the ranking process is explained in more detail.

1) CRAWLING AND PRE-PROCESSING THE ONLINE REVIEWS

In this stage, the online reviews for multiple features of the alternative products are collected using the crawler software. Then, the obtained data are pre-processed. The details of this stage will be described as follows:

Crawling the Online Reviews: This stage includes collecting online textual reviews to do analysis, whether they are obtained from a specific online site for purchase through the internet or from social networking sites. Recently, the crawler software is used to obtain online reviews, which crawls reviews of related products from online shopping websites. In this paper, the crawler software ScrapeStorm is used to obtain Twitter online reviews about the mobile phone.

To choose a certain product from a various set of suggested alternative products, the features of these products are considered in the light of the consumer's personal choices. Moreover, the assignment of the consumer sets the weights of these features directly.

After crawling the data, there is a need to pre-processing the collected reviews.

Pre-Processing the Collected Reviews: Python has been used in this paper for the pre-processing stage. To begin with, Natural Language Toolkit (NLTK) has been imported, which is used in most of the processes of the pre-processing stage. The toolkit is one of the most powerful NLP libraries. NLTK consists of the most common algorithms, which makes machines understand human language and reply to that with an appropriate response. Therefore, NLTK is a powerful Python package that helps the computer to analyze, pre-process, and to understand the written text. In this paper, the pre-processing stage includes three main steps.

a: EXCLUDING REVIEWS BEFORE THE PRODUCT'S RELEASE DATE

The reviews before the official release date of the product have been excluded. This is because the opinions or the evaluations will not be based on the actual use or the real experience of the product by consumers.

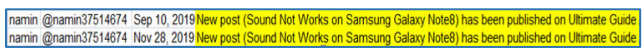
Fig. 3 shows an example of the reviews on iPhone 8 before the release date of the product on Sep 22, 2017. These reviews have been crawled from Twitter on Nov 10, 2016.

Ben Geskin	@BenGeskin	Nov 10, 2016	#iPhone8 design improvements: Thinner. Less round. No camera bump. No home button. Bigger display. Etc

FIGURE 3. An example of the review before the product's release date crawled from Twitter.

b: REMOVING DUPLICATE REVIEWS

In this step, the repeated reviews were removed. Fig. 4 shows an example of the repeated reviews being posted by the same reviewer, on the Samsung Galaxy Note 8 concerning sound feature.



namin	@namin37514674	Sep 10, 2019	New post (Sound Not Works on Samsung Galaxy Note8) has been published on Ultimate Guide
namin	@namin37514674	Nov 28, 2019	New post (Sound Not Works on Samsung Galaxy Note8) has been published on Ultimate Guide

FIGURE 4. An example of repeated review obtained from Twitter.

c: WORD SEGMENTATION AND PART-OF-SPEECH (POS) TAGGING

This method aims to tokenize the sentences into individual words and categorize these words as nouns, verbs, adjectives, etc.

d: FILTERING

This step encompasses the following

- Removing URL link, since URL link does not have significant information regarding the sentiment of the tweet [64].
- Removing numbers as they are not used when measuring sentiment.
- Removing Hashtags.
- Removing mention in Twitter.
- Removing stop words. Special stop words were added for this research. Stop word removal eliminates the common and frequent words, which do not have a significant influence in the sentence such as "a", "an", "the", "to", "of", "is", "are" and "for". Removing stop words saves time and improves the effectiveness and efficiency of SA.

After collecting data and pre-processing the collected review, then the sentiment scores will be computed.

2) COMPUTING SENTIMENT SCORES OF THE ONLINE REVIEWS

In this stage, the positive, neutral, and negative sentiment scores are calculated as the truth, indeterminacy, and falsity memberships degrees of the neutrosophic number. Two sections will be included in the stage, where the section of establishing the dictionaries will be presented in item (a) and the section of computing sentiment scores will be given in item (b).

a: ESTABLISHING THE DICTIONARIES

Establishing a neutral SD as well as positive and negative SDs are required to identify positive, neutral, and negative sentiment orientations for each review. This section is divided into two parts. The first part discussed establishing the neutral SD, where the second part on the positive and negative SDs.

Establishing Neutral Sentiment Dictionary: Conventional SA does not deal with a neutral or an indeterminate opinion since conventional SA merely gives an overall opinion as positive or negative.

VADER - as a popular SA tool - is sensitive to both polarity (positive/negative) and intensity (strength) of emotion but not to neutrality. Motivated by this, this paper is dedicated to handling the neutrality of data through SA. For this, VADER has to be adapted to handle the neutrality and due to its lexicon-based approach with a design focus on social media texts, this heightens the need for creating a neutral lexicon. Consequently, in this paper, a newly neutral lexicon has been compiled. This new lexicon consists of 228 neutral words and phrases, which is mainly compiled by collecting data from many resources. These resources are divided as follows.

- Encyclopedia (e.g. Encyclopedia Britannica, Encyclopedia Americana).
- Websites (e.g. Learnenglish.britishcouncil.org, Dictionary.com, and Thesaurus.com).
- Dictionaries (e.g. Oxford Dictionary, Cambridge Dictionary).
- Online reviews.

In this paper, five neutral dictionaries were compiled. A single dictionary for each of the different five features of the alternative products. The researcher established these dictionaries by making use of the common lexicon among the obtained online reviews and the new neutral lexicon.

Establishing Positive and Negative Sentiment Dictionaries: Both Positive and negative SDs differ in their features. Consequently, a word can belong to both the positive SD of a feature as well as the negative SD of another feature. For instance, the word "slow" belongs to both the positive SD of a feature "draining in battery" and the negative SD of a feature "response in screen". This leads to the necessity of establishing positive and negative dictionaries for every feature of the alternative products.

In this paper, five positive and negative SDs were compiled. A single dictionary for each of the different five features of the alternative products. The researchers established these dictionaries by making use of the common words among the obtained online reviews and the Vader_lexicon.

It is to be noted that there are some situations where identifying the sentimental orientations of some sentimental words using the Vader_lexicon SD is impossible. In this case, the SenticNet-5.0 SD has been used to determine the sentiment orientation and to score the sentiment words.

As mentioned above, Vader_lexicon has been adopted by establishing positive and negative dictionaries as Vader_lexicon contains many valuable advantages. To begin with, the VADER sentiment lexicon is especially attuned to social media contexts. Since the VADER sentiment lexicon is sensitive to both the polarity and the intensity of sentiments expressed in these microblogs. Also, an inspired list had been constructed by examining existing well-established sentiment word-banks (LIWC, ANEW, and GI). Where, numerous common lexical features had been incorporated into sentiment expression in microblogs, including:

A full list of Western-style emoticons, for example,:-) denotes a smiley face and generally indicates positive sentiment.

Sentiment-related acronyms and initials (e.g., LOL and WTF are both examples of sentiment-laden initials).

Commonly used slang with sentiment value (e.g., nah, meh and giggly).

b: COMPUTING THE SENTIMENT SCORES

This section aims to demonstrate the process of calculating the positive, negative as well as neutral sentiment scores of each review concerning each product feature. The sentiment score will be expressed by a neutrosophic number. So, this research uses the simplified NS whose membership values are singletons in the real standard interval. Thus, each simplified NS can be described by three real numbers in the real unit interval [0, 1].

Neutro-VADER is used to compute the positive, negative and, neutral sentiment scores of each review. Neutro-VADER is a new version of VADER which is sensitive to positivity, negativity, neutrality, and intensity (strength) of emotion.

The construction of Neutro- VADER has been achieved by integrating the positive, negative and neutral SDs that have been established in Section 4.3 with an algorithm that has been modified to compute the positive, negative and neutral sentiment scores as follows.

Positive sentiment score is calculated by summing the valence scores of each positive word in the lexicon, adjusted according to the rules, and then normalized to be between 0 (less extreme positive) and 1 (most extreme positive). Whereas the negative sentiment score is calculated by summing the valence scores of each negative word in the lexicon, adjusted according to the rules to be processed to achieve the absolute value. Then the absolute value is normalized to be between 0 (less extreme negative) and 1 (most extreme negative).

The compound score is computed by summing the valence scores of each word in the lexicon, adjusted according to the rules, and then normalized to be between −1 (most extreme negative), and +1 (most extreme positive).

To compute a neutral score, there are two cases:

1. In case there are neutral sentiment words or phrases in the sentence, then the neutral sentiment score will be 1.
2. In case there are no neutral sentiment words or phrases in the sentence, then we have the following cases.
   a) If there are both positive and negative sentiment words in the sentence, the neutral sentiment score will be (1 − absolute value of compound score).
   b) If there are only positive sentiment words in the sentence or only negative sentiment words in the sentence, then the neutral sentiment score will be 0.
   c) If there are neither positive nor negative sentiment words in the sentence, then the neutral sentiment score will be 0.

The last stage of the proposed method is ranking the products via neutrosophic set theory.

3) RANKING THE PRODUCTS IN THE LIGHT OF THE SIMPLIFIED NEUTROSOPHIC SET THEORY

The following is an approach used to rank the alternative products in the light of the simplified NS theory. The approach comprises three steps:

a: DETERMINING THE SIMPLIFIED NEUTROSOPHIC NUMBER OF EVERY ALTERNATIVE PRODUCT TO EACH PRODUCT FEATURE

The simplified neutrosophic number can be expressed as three components that represent simultaneously the degrees of support, hesitation, and opposition of the evaluations about some specific event. In this study according to the set of alternative products $A = \{A_1, A_2, \ldots A_n\}$ and the set of product features $F = \{F_1, F_2 \ldots F_m\}$ it is supposed that the sentences of the product features are extracted from the different online reviews. Therefore, the K_{th} online review of an alternative product A_i can be presented by $D_{ik} = \left\{D_{ik}^1, D_{ik}^2, \ldots, D_{ik}^m\right\}$. Where D_{ik}^j denotes the sentence concerning feature f_j in the review D_{ik},

$$i = 1, \ldots, n, \quad j = 1, \ldots, m, \quad k = 1, \ldots, q_i.$$

Suppose $S_{ik}^j\left(a_{ik}^j, b_{ik}^j, c_{ik}^j\right)$ denotes the indicator vector which represents the sentiment score of sentence or review D_{ik}^j, where $a_{ik}^j, b_{ik}^j, c_{ik}^j$ are indicator variables for positive, neutral, and negative sentiment score, respectively. According to the Neutro-VADER which assigns scores between 0 and 1 to these independent indicator variables, and if a_{ik}^j is considered a vote in support, b_{ik}^j as hesitation and c_{ik}^j as a vote in opposition, then $a_{ik}^j, b_{ik}^j, c_{ik}^j$ can be considered as simplified neutrosophic components, and a simplified neutrosophic number can be constructed to represent the performance of an alternative product concerning a product feature. Besides, in this paper, the greater importance degree of posted time will be assigned to the newer online reviews as they are more consistent and valid compared with earlier reviews. Let w_{ik}^j denotes the importance degree of the posted time of review D_{ik}^j, $i = 1, \ldots, n, j = 1, \ldots, m$ and $K = 1, 2, \ldots q_i$.

Najmi et al. [4] calculate w_{ik}^j as

$$w_{ik}^j = e^{\frac{t_{ik}^j - t_i}{t_c - t_i}}, \quad i = 1, \ldots, n, \ j = 1, \ldots, m \text{ and } K = 1, 2, \ldots q_i. \tag{1}$$

where t_{ik}^j denotes the posted time of review D_{ik}^j. t_i denotes the release time of the product A_i, t_c denotes the current time, $i = 1, \ldots, n, j = 1, \ldots, m, K = 1, 2, \ldots qi$.

Let u_{ij}^{pos}, u_{ij}^{neu} and u_{ij}^{neg} denote the weighted aggregation values of a_{ik}^j, b_{ik}^j and c_{ik}^j to all reviews on alternative product A_i concerning feature f_j respectively. The values of u_{ij}^{pos}, u_{ij}^{neu}, u_{ij}^{neg} can be respectively calculated by.

$$u_{ij}^{pos} = \sum_{k=1}^{q_i} w_{ik}^j a_{ik}^j \quad i = 1, \ldots, n, \ j = 1, \ldots, m, \quad K = 1, 2, \ldots, q_i, \tag{2}$$

$$u_{ij}^{neu} = \sum_{k=1}^{qi} w_{ik}^j b_{ik}^j \quad i = 1, \ldots, n, \ j = 1, \ldots, m, \quad K = 1, 2, \ldots, q_i, \tag{3}$$

$$u_{ij}^{neg} = \sum_{k=1}^{qi} w_{ik}^j c_{ik}^j \quad i = 1, \ldots, n, \ j = 1, \ldots, m, \quad K = 1, 2, \ldots, q_i. \tag{4}$$

Let V_{ij}^{pos}, V_{ij}^{neu} and V_{ij}^{neg} be the normalized values of u_{ij}^{pos}, u_{ij}^{neu} and u_{ij}^{neg}, respectively. Then V_{ij}^{pos}, V_{ij}^{neu} and V_{ij}^{neg} can be respectively calculated as:

$$V_{ij}^{pos} = \frac{\sum_{k=1}^{q_i} w_{ik}^j a_{ik}^j - \sum_{k=1}^{q_i} a_{ik}^j}{(e-1)\sum_{k=1}^{q_i} a_{ik}^j} \tag{5}$$

$$V_{ij}^{neu} = \frac{\sum_{k=1}^{q_i} w_{ik}^j b_{ik}^j - \sum_{k=1}^{q_i} b_{ik}^j}{(e-1)\sum_{k=1}^{q_i} b_{ik}^j} \tag{6}$$

$$V_{ij}^{neg} = \frac{\sum_{k=1}^{q_i} w_{ik}^j c_{ik}^j - \sum_{k=1}^{q_i} c_{ik}^j}{(e-1)\sum_{k=1}^{q_i} c_{ik}^j} \tag{7}$$

$$i = 1, \ldots, n, \ j = 1, \ldots, m \text{ and } k = 1, \ldots, q_i.$$

These values have been normalized, since they represent degrees of the influence of the reviewers' opinions in all reviews concerning the alternative product A_i, concerning each feature f_j.

It is clear from equations (5-7), that when $w_{ik}^{j} = e$, for all i, j, k. i.e. (the posted time of review is the current time), then V_{ij}^{pos}, V_{ij}^{neu} and V_{ij}^{neg} take their maximum values at 1, while its minimum takes the value 0 when $w_{ik}^{j} = 1$ for all $i,\ j,\ k$. i.e. (when the posted time of review D_{ik}^{j} is the release time of the product A_i).

From the previous discussion, the values $V_{ij}^{pos}, V_{ij}^{neu}, V_{ij}^{neg} \in [0, 1]$ and $0 \leq V_{ij}^{pos} + V_{ij}^{neu} + V_{ij}^{neg} \leq 3$, which are also independent.

According to the physical interpretation of simplified neutrosophic number [80], a neutrosophic number $M_{ij} = (T_{ij},\ I_{ij},\ F_{ij})$ is constructed to represent the performance of the alternative product A_i concerning the feature f_j, where

$T_{ij} = V_{ij}^{pos},\ I_{ij} = V_{ij}^{neu},\ F_{ij} = V_{ij}^{neg}$, denote respectively the support, hesitation as well as opposition degree of alternative product A_i concerning feature f_j.

b: CALCULATING THE OVERALL SIMPLIFIED NEUTROSOPHIC NUMBER OF EVERY SUGGESTED ALTERNATIVE PRODUCT

According to simplified neutrosophic numbers

$M_{i1} = (T_{i1}, I_{i1}, F_{i1}),\ M_{i2} = (T_{i2}, I_{i2}, F_{i2}), \ldots, M_{im} = (T_{im}, I_{im}, F_{im})$ and feature weights $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_m$, given by the customer, an overall simplified neutrosophic number of suggested alternative products A_i can be calculated.

Several operators are proposed to aggregate the different multiple simplified neutrosophic numbers such as the simplified neutrosophic number weighted averaging operator (SNNWA) and simplified neutrosophic number ordered weighted averaging operator (SNNOWA), etc [80]. Different aggregation operators are based on different assumptions and are suitable to be used in several types of problems. For example, the SNNWA operator is appropriate to be adopted to the problems in which the weights are pointed out to the product feature, while the SNNOWA operator is used in the problems where the weights are pointed out to the ranking positions of feature values. This research paper assigns the weights to the product features. Consequently, the SNNWA operator can be used to calculate the overall simplified neutrosophic number of every alternative product. For example,

$$SNNWA_{\alpha}(A_1, \ldots\ldots A_n) = \langle 1 - \prod_{j=1}^{m} (1 - T_{ij})^{\alpha_j},\ \prod_{j=1}^{m} I_{ij}^{\alpha_j},\ \prod_{j=1}^{m} F_{ij}^{\alpha_j} \rangle \quad (8)$$

where $t_i = 1 - \prod_{j=1}^{m} (1 - T_{ij})^{\alpha_j}$, $i_i = \prod_{j=1}^{m} I_{ij}^{\alpha_j}$ and $f_i = \prod_{j=1}^{m} F_{ij}^{\alpha_j}$ present in order the overall support, hesitation, and opposition degrees of product A_i, respectively.

c: DETERMINING RANKING OF THE ALTERNATIVE PRODUCTS

After calculating an overall simplified neutrosophic number of every alternative product, we proceed to rank the alternatives using two methods: (1) Cosine similarity measure method, and (2) Score function method.

1) *Cosine Similarity Measure:* The ranking order of alternatives is performed through the cosine similarity measure between an alternative A_i and the ideal alternative A^* and the best choice can be obtained according to the measure values.

The cosine similarity measure between A^* and the alternative $A_i : i = 1, \ldots,\ n$ can be defined as follows.

$$S_i(A_i, A^*) = \frac{t_i t_i^* + i_i i_i^* + f_i f_i^*}{\sqrt{t_i^2 + i_i^2 + f_i^2}\sqrt{(t_i^*)^2 + (i_i^*)^2 + (f_i^*)^2}} \quad (9)$$

where $A_i = (t_i, i_i, f_i)$ and $A^* = (t^*, i^*, f^*)$. Since the ideal simplified neutrosophic value is $A^* = (1, 0, 0)$, then

$$S_i(A_i, A^*) = \frac{t_i}{\sqrt{t_i^2 + i_i^2 + f_i^2}}. \quad (10)$$

and the bigger the measured value is the better the alternative A_i is because the alternative A_i is close to the ideal alternative A^*. We can also determine the better alternative using the score function method as follows.

2) *Score Function Method:* Peng et al [80] defined the score function $\lambda(A)$, accuracy function $\pi(A)$, and certainty function $\mu(A)$ of a *simplified neutrosophic number* as follows.

Definition (3) [81]: If $A = \langle T_A, I_A, F_A \rangle$ is a simplified neutrosophic number. Then the score function $\lambda(A)$, accuracy function $\pi(A)$, and certainty function $\mu(A)$ of a simplified neutrosophic number are defined as follows:

(a) $\lambda(A) = \dfrac{T_A + 1 - I_A + 1 - F_A}{3}$,

(b) $\pi(A) = T_A - F_A$,

(c) $\mu(A) = T_A$.

Based on the above definition, the method for comparing simplified neutrosophic numbers can be defined as follows.

Definition (4) [81]: If A and B are two simplified neutrosophic numbers. The comparison method can be defined as follows.

a) If $\lambda(A) > \lambda(B)$, then A is greater than B, that is, A is superior to B, denoted by $A > B$.

b) If $\lambda(A) = \lambda(B)$ and $\pi(A) > \pi(B)$. Then A is greater than B, that is, A is superior to B denoted by $A > B$,

c) If $\lambda(A) = \lambda(B)$, $\pi(A) = \pi(B)$ and $\mu(A) > \mu(B)$. Then A is greater than B, that is, A is superior to B, denoted by $A > B$:

d) If $\lambda(A) = \lambda(B)$, $\pi(A) = \pi(B)$ and $\mu(A) = \mu(B)$. Then A is equal to B, that is, A is indifferent to B, denoted by $A \sim B$.

V. CASE STUDY

The arrival of the mobile phone and its rapid and widespread growth may well be seen as one of the most significant developments in the fields of communication and information technology over the past two decades. Because of its durability and high-value consumers are discreet while selecting a satisfied one among many mobile phones with different brands. To support the consumer purchase decision and to further illustrate the proposed method, we use a case study of ranking mobile phone products based on online reviews.

Four alternative mobile phones are selected for the experiment in this case study. The proposed method in section IV is applied to rank these four alternative mobile phones. Experimental data statistics are shown in Table 1.

TABLE 1. Experimental data statistics for four alternative mobile phones.

Feature	Alternative mobile phone			
	IPhone8	iPhone 8 plus	Samsung Galaxy Note8	Huawei Mate 10 Pro
sound (audio)	56	25	35	3
Screen	36	27	14	15
Camera	50	234	92	35
Battery draining	163	81	32	12
Battery life	66	60	35	24
Total number of reviews		1095		

Table 1 shows that there are four alternative mobile phones which are listed as follows.

A_1 : iPhone 8, A_2: iPhone 8 plus, A_3: Samsung Galaxy Note8, A_4: Huawei Mate 10 Pro

These four phones have been chosen for the following reasons. To begin with, they have roughly the same official release date. Also, they have a large number of online reviews which is considered a great deal of information-rich content. Moreover, these phones are also very popular. Also, in terms of features, they can be considered flagship phones.

The following five features associated with alternative mobile phones are considered:

f_1 : Battery life. f_2: Camera, f_3: Screen, f_4: Sound (Audio), f_5: Battery draining.

Meanwhile, the vector of weights of the five features is provided by the consumer, i.e. $w = (0.3, 0.3, 0.2, 0.1, 0.1)$ In this case study, the dataset has been collected on April 4, 2020. As mentioned in Section (4), crawler software (Scrapestorm) is used to crawl data from Twitter. The original plan for this study was based on more than 1500 reviews which were later reduced to 1095 after removing the duplicate reviews and excluding the reviews that are published before the release date of the products.

Pre-processing is then performed using python for these 1095 reviews. The pre-processing stage includes Segmentation, POS tagging, and filtering (by removing unimportant URLs, mention, numbers, and hashtags).

The numbers of obtained reviews are 371,427,208 and 89, for the alternatives A_1, A_2, A_3 and A_4 respectively.

The obtained reviews can be expressed by

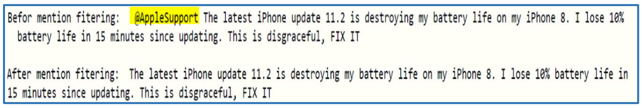
$$D_{ik} = \left(D_{ik}^1, D_{ik}^2, D_{ik}^3, D_{ik}^4, D_{ik}^5\right);$$

$i = 1, 2, 3, 4,\ k = 1, 2, \ldots, q_i,\ q_1 = 371,\ q_2 = 427,\ q_3 = 208$ and $q_4 = 89$.

Now, using the Neutro-VADER the sentiment score of D_{ik}^j is identified; i.e. indicator vector $S_{ik}^j = \left(d_{ik}^j, b_{ik}^j, c_{ik}^j\right)$ is set; $i = 1, 2, 3, 4, j = 1, 2, 3, 4, 5, k = 1, 2, \ldots, q_i, q_1 = 371, q_2 = 427, q_3 = 208$ and $q_4 = 89$.

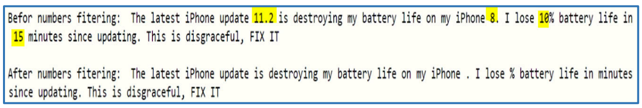
To illustrate this process, we take the following review D_{13}^1 = (@AppleSupport The latest iPhone update 11.2 is destroying my battery life on my iPhone 8. I lose 10% battery life in 15 minutes since updating. This is disgraceful, FIX IT).

For pre-processing stage, the process begins with applying segmentation and POS tagging. Filter process (includes removing URL, mention, numbers, and hashtags) is then applied. The review D_{13}^1 = (@AppleSupport The latest iPhone update 11.2 is destroying my battery life on my iPhone 8. I lose 10% battery life in 15 minutes since updating. This is disgraceful, FIX IT) includes mention which is removed as shown in Fig. 5, this also includes numbers which are removed as shown in Fig. 6.


Befor mention fitering: @AppleSupport The latest iPhone update 11.2 is destroying my battery life on my iPhone 8. I lose 10% battery life in 15 minutes since updating. This is disgraceful, FIX IT

After mention fitering: The latest iPhone update 11.2 is destroying my battery life on my iPhone 8. I lose 10% battery life in 15 minutes since updating. This is disgraceful, FIX IT

FIGURE 5. Removing mentions.


Befor numbers fitering: The latest iPhone update 11.2 is destroying my battery life on my iPhone 8. I lose 10% battery life in 15 minutes since updating. This is disgraceful, FIX IT

After numbers fitering: The latest iPhone update is destroying my battery life on my iPhone . I lose % battery life in minutes since updating. This is disgraceful, FIX IT

FIGURE 6. Removing numbers.

The final step in the pre-processing stage is removing stop words as shown in Fig. 7.

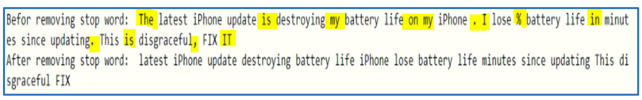

Befor removing stop word: The latest iPhone update is destroying my battery life on my iPhone . I lose % battery life in minutes since updating. This is disgraceful, FIX IT
After removing stop word: latest iPhone update destroying battery life iPhone lose battery life minutes since updating This disgraceful FIX

FIGURE 7. Applying stop word removing.

The next stage is to compute the sentiment score to the battery life of iPhone8 as shown in Fig. 8.

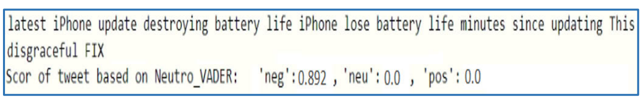

latest iPhone update destroying battery life iPhone lose battery life minutes since updating This disgraceful FIX
Scor of tweet based on Neutro_VADER: 'neg':0.892 , 'neu':0.0 , 'pos':0.0

FIGURE 8. Computing the sentiment score using Neutro-VADER.

The same steps are followed to find sentiment scores to other indicator vectors S_{ik}^j.

Then u_{ij}^{pos}, u_{ij}^{neu} and u_{ij}^{neg} are calculated using Eqs (2 - 4),

$i = 1, 2, 3, 4, \quad j = 1, 2, 3, 4, 5,$

$k = 1, \ldots, q_i,\ q_1 = 371, q_2 = 427,\ q_3 = 208$ and $q_4 = 89$.

The values of u_{ij}^{pos}, u_{ij}^{neu} and u_{ij}^{neg} are shown in Table 2.

Then, using Eqs. (5-7), the simplified neutrosophic number $M_{ij}\left(T_{ij}, I_{ij}, F_{ij}\right)$ of alternative A_i concerning feature F_j is determined, $i = 1, 2, 3, 4, j = 1, 2, 3, 4, 5$. The obtained simplified neutrosophic numbers are shown in Table 3.

TABLE 2. The values of u_{ij}^{pos}, u_{ij}^{neu} and u_{ij}^{neg}.

Feature	Alternative products			
	A1	A2	A3	A4
f1	37.5848, 15.2432, 14.0911	35.4402, 17.8331, 12.2813	19.1630, 11.1661, 5.73484	10.6963, 7.0184, 2.6531
f2	23.3017, 4.4291, 3.8083	91.9428, 36.9593, 19.8105	30.0173, 8.9054, 8.6746	7.5555, 4.5422, 1.9559
f3	23.8651, 17.3177, 8.8608	15.6723, 6.6668, 2.7663	5.0199, 0, 0.8388	6.9590, 0.8471, 1.1096
f4	40.0891, 27.9349, 21.8426	30.8528, 14.6354, 13.9158	10.6660, 4.5840, 5.8958	2.0183, 0, 0
f5	34.3228, 72.1261, 121.4497	17.3671, 30.5969, 48.6228	3.8922, 6.1529, 20.9535	4.3653, 0, 0

TABLE 3. The simplified neutrosophic number of A_i concerning feature f_j.

Feature	Alternative products			
	A1	A2	A3	A4
f1	0.33974, 0.38317, 0.28126	0.30261, 0.25644, 0.29573	0.22222, 0.19308, 0.13906	0.05041, 0.18239, 0.21438
f2	0.50735, 0.31552, 0.56491	0.14766, 0.24932, 0.20835	0.39447, 0.45217, 0.47949	0.17499, 0.34332, 0.19073
f3	0.24448, 0.28738, 0.23655	0.33304, 0.28568, 0.09458	0.24113, 0, 0.27164	0.29637, 0.11264, 0.15029
f4	0.33311, 0.30348, 0.36393	0.80440, 0.89432, 0.89467	0.30943, 0.26812, 0.30320	0.40966, 0, 0
f5	0.29546, 0.37424, 0.34664	0.28503, 0.32036, 0.33063	0.23930, 0.22671, 0.13023	0.03912, 0, 0

Using Equation (8), the overall simplified neutrosophic number R_i of alternatives A_i is calculated. i.e.

$R_1 = (0.3740\ ,\ 0.6674\ ,\ 0.6491)$,
$R_2 = (0.3519\ ,\ 0.3009\ ,\ 0.2398)$,
$R_3 = (0.2920\ ,\ 0\ ,\ 0.2474)$, and
$R_4 = (0.1813\ ,\ 0\ ,\ 0)$.

To determine the best alternative, the cosine similarity measure method is applied. Using equation (10), each cosine similarity can be measured $S_i\,(R_i,\ A^*)\quad (i = 1, \ldots, 4)$ as follows.

$$S_1\left(R_1,\ A^*\right) = 0.3727, S_2\left(R_2,\ A^*\right) = 0.6748,$$
$$S_3\left(R_3,\ A^*\right) = 0.7629,\ \text{and}\ S_4\left(R_4,\ A^*\right) = 1.$$

From the measured value $S_i\,(R_i,\ A^*)\quad (i = 1, 2, 3, 4)$ between an alternative and the ideal alternative, the ranking order of four alternative mobile phones is $A_4 > A_3 > A_2 > A_1$.

Therefore, Huawei mobile phone is the best choice among all mobile phones.

On the other hand, the scoring method to determine the best choice can be utilized as follows.

Using Definition (3), the score function of $R_i (i = 1, 2, 3, 4)$ can be obtained:

$$\lambda\,(R_1) = 0.3525,\quad \lambda\,(R_2) = 0.6037,$$
$$\lambda\,(R_3) = 0.6815,\quad \lambda\,(R_4) = 0.7271.$$

The score values are different; therefore, there is no need to compute both accuracy function value and certainty function value. Thus according to Definition (4), the final ranking is still $A_4 > A_3 > A_2 > A_1$ and the best alternative is Huawei mobile phone.

VI. COMPARATIVE ANALYSIS

This research paper proposes a novel method to the SA technique as well as neutrosophic fuzzy set theory for ranking alternative products. The most important contributions in each stage can be summarized. The contributions of this paper compared to the existing literature such as [1], [4], [5], [7], [73], [74], [81], [87], [90]–[98] can be divided into two parts, the sentiment analysis part and ranking part.

A. SENTIMENT ANALYSIS UNDER NEUTRALITY

In the SA part, the contributions can be presented as follows.

B. COLLECTION OF DATA

This paper collected a real dataset related to mobile phones. Where this dataset was obtained by crawling online mobile phone reviews on Twitter. These reviews were filtered by removing the repeated reviews being posted by the same reviewer. Also, the reviews before the official release date of the product have been excluded.

C. PRE-PROCESSING

The word segmentation is required to tokenize the sentences into individual words. In the light of this, reviews on Twitter may contain hashtags or emoticons that cannot be analyzed using word_tokenize. Fig. 9 shows the hashtag and emoticons that are not analyzed.

```
from nltk.tokenize import word_tokenize
tweet= "My #Iphone8 plus has a long battery life (-:< "
print(word_tokenize(tweet))

['>' ,':' ,'-' ,')' ,'My', '#', 'Iphone8', 'plus', 'has', 'a', 'long', 'battery', 'life']
```

FIGURE 9. An example of segmented review using word_tokenize.

TweetTokenizer (as a subset of word_tokenize) is mainly used to handle this issue. TweetTokenizer keeps the hashtag as shown in Fig. 10. However, there are still some shortcomings in analyzing some Emoticons as shown in Fig.11. To solve this issue, we adapt TweetTokenizer by adding some emoticon as shown in Fig. 12.

D. ESTABLISHING THE DICTIONARIES

Most studies and research on SA does not deal with a neutral or an indeterminate opinion, this merely gives an overall

```
from nltk.tokenize import TweetTokenizer
tweet= " My #Iphone8 plus has a long battery life (-:< "
q1 = TweetTokenizer()
print(q1.tokenize(tweet))

 ['>:-)' ,'My', '#Iphone8', 'plus', 'has', 'a', 'long', 'battery', 'life']
```

FIGURE 10. An example of a segmented review using TweetTokenizer.

```
t_impro = TweetTokenizer_our_imporvment()
tweet= " My #Iphone8 plus has a long battery life (^;o "
print( t_impro.tokenize(tweet)

['My', '#Iphone8', 'plus', 'has', 'long', 'battery', 'life', '(^;o']
```

FIGURE 11. An example of Emoticon was not analyzed using TweetTokenizer.

```
from nltk.tokenize import TweetTokenizer
tweet= " My #Iphone8 plus has a long battery life (^;o "
q1 = TweetTokenizer()
print(q1.tokenize(tweet))

 ['My', '#Iphone8', 'plus', 'has', 'a', 'long', 'battery', 'life', '(', '^', ';', 'o']
```

FIGURE 12. An example of emoticons that had been analyzed using adapted TweetTokenizer.

opinion as positive or negative [1], [5]. Therefore, this paper used VADER in SA. VADER is sensitive to both polarity (positive/negative) and intensity (strength) of emotion but not to neutrality. To deal the neutrality a newly neutral lexicon has been compiled in this paper. This new lexicon consists of 228 neutral words and phrases.

E. COMPUTING THE SENTIMENT SCORE CONCERNING EACH PRODUCT FEATURE IN EACH REVIEW

This paper adapted VADER to compute the positive, negative, and, neutral sentiment scores of each review, namely Neutro-VADER. Neutro-VADER is a new version of VADER which is sensitive to positivity, negativity, neutrality, and intensity (strength) of emotion. Neutro-VADER Highlights a novel method that helps to handle fuzzy, intuitionistic fuzzy, and neutrosophic data through SA.

The following example illustrates the difference between the Neutro-VADER and the traditional Vader. Fig. 13 shows how the traditional Vader computes the same neutral sentiment score (0.734) for two different sentences although, the consumer in the first sentence was sure of his opinion, while the consumer in the second sentence was completely hesitant in his opinion.

```
sentiment_analyzer_scores(" I used iphone8 yesterday , the sound is good")
sentiment_analyzer_scores(" I can't determine if the sound iphone8 is good")

 I used iphone8 yesterday , the sound is good--------- {'neg': 0.0, 'neu': 0.734, 'pos': 0.266
 I can't determine if the sound iphone8 is good------- {'neg': 0.0, 'neu': 0.734, 'pos': 0.266
```

FIGURE 13. Computing the neutral sentiment score using Vader.

On the other hand, Fig.14 shows how the Neutro-VADER can deal successfully with this case. In the first sentence,

```
NeutroVADER_scores(" I used iphone8 yesterday , the sound is good")
NeutroVADER_scores(" I can't determine if the sound iphone8 is good")

 I used iphone8 yesterday , the sound is good--------- {'neg': 0.0, 'neu':  0.0,  'pos': 0.4404
 I can't determine if the sound iphone8 is good------- {'neg': 0.0, 'neu':  1.0,  'pos': 0.4404
```

FIGURE 14. Computing the neutral sentiment score using Neutro-VADER.

when the consumer was sure of his opinion, the neutral score is (0). While in the second sentence, the consumer is not sure of his opinion. That is why the neutral score is 1.

F. RANKING PRODUCTS UNDER NEUTROSOPHIC ENVIRONMENT

To validate the feasibility of the proposed decision-making method based on the aggregation operators of simplified neutrosophic numbers and cosine similarity measure, a comparative study was conducted with the following methods: TODIM method for single-valued neutrosophic multiple attribute decision making extended by Xu et al. [87], PROMETHEE II along with the intuitionistic fuzzy weighted averaging (IFWA) operator presented by Liu et al. [1], TOPSIS method for multi-attribute group decision-making under single-valued neutrosophic environment presented by Biswas et al. [98] and the simplified neutrosophic number weighted geometric operator (SNNWG) operator introduced in Peng *et al.* [81].

Note that, to save space, in this section, only results are presented. The ranking results by these methods and the proposed method are shown in Table 4 by employing the case study in Section IV.

TABLE 4. Ranking results of five different methods.

Methods	Ranking	Optimal Alternative
Xu et al.[87]	$A_3 > A_3 > A_2 > A_1$	A_4
Liu et al.[1]	$A_3 > A_3 > A_2 > A_1$	A_4
Biswas et al.[98]	$A_3 > A_3 > A_2 > A_1$	A_4
Peng et al.[81]	$A_3 > A_1 > A_4 > A_2$	A_3
Proposed method	$A_3 > A_3 > A_2 > A_1$	A_4

From Table 4, it can be easily seen that the same or similar ranking results are obtained using the method proposed by Xu et al. [87], the method proposed by Liu et al. [1], the method proposed by Biswas et al. [98] and the method proposed in this paper. Also, It can be seen that the best alternative is A_4, followed by the alternative A_3, where the alternatives A_1 and A_2 lag behind these alternatives.

On the other hand, the final ranking is obtained by utilizing the method presented by Peng et al. [81] is $A_3 > A_1 > A_4 > A_2$ which is different from that obtained by using the proposed method and the other existing methods. There are different focal points between the weighted arithmetic average operator and the weighted geometric average operator. The weighted arithmetic average operator emphasizes the group's major points, whereas, the weighted geometric average operator emphasizes personal major points.

VII. CONCLUSION

This research paper presents a novel method, which fully uses the SA technique, NS theory, and MADM, to deal with the problem of ranking alternative products in the Twitter review. Firstly, the Twitter reviews of the alternative products have been crawled then pre-processed. Based on the SA technique, the sentiment scores are computed. Throughout this phase, a neutral lexicon of 228 neutral words and phrases is created. Neutro-VADER has been developed to handle the neutral data. The positive, negative, and neutral dictionaries are used in Neutro-VADER for the goal of assigning positive, neutral, and negative sentiment scores for each review of each product feature. A simplified neutrosophic number is constructed to represent the performance of each alternative product concerning each feature. Then the overall simplified neutrosophic number of each alternative product is calculated using the simplified neutrosophic number weighted average operator (SNNWA). Cosine similarity measure and score function methods are used to rank the alternative product. Further, a comparative analysis is provided to validate the proposed method. In the future, we will develop the Neutro-VADER by extending the lexicon to increase its classification accuracy. In particular, the neutral lexicon might need more words, especially neutral words that are commonly used in social media texts. As a generalization of neutrosophic logic, N-valued refined neutrosophic logic [99] can also be integrated into sentiment analysis and MADM methods to capture more room of uncertainty.

REFERENCES

[1] Y. Liu, J.-W. Bi, and Z.-P. Fan, "Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory," *Inf. Fusion*, vol. 36, pp. 149–161, Jul. 2017, doi: 10.1016/j.inffus.2016.11.012.

[2] Y. Liu, J.-W. Bi, and Z.-P. Fan, "A method for ranking products through online reviews based on sentiment classification and interval-valued intuitionistic fuzzy TOPSIS," *Int. J. Inf. Technol. Decis. Making*, vol. 16, no. 06, pp. 1497–1522, Nov. 2017, doi: 10.1142/S021962201750033X.

[3] X. Liang, P. Liu, and Z. Liu, "Selecting products considering the regret behavior of consumer: A decision support model based on online ratings," *Symmetry*, vol. 10, no. 5, p. 178, May 2018, doi: 10.3390/sym10050178.

[4] E. Najmi, K. Hashmi, Z. Malik, A. Rezgui, and H. U. Khan, "CAPRA: A comprehensive approach to product ranking using customer reviews," *Computing*, vol. 97, no. 8, pp. 843–867, Aug. 2015, doi: 10.1007/s00607-015-0439-8.

[5] C. Wu and D. Zhang, "Ranking products with IF-based sentiment word framework and TODIM method," *Kybernetes*, vol. 48, no. 5, pp. 990–1010, May 2019, doi: 10.1108/K-01-2018-0029.

[6] D. Zhang, C. Wu, and J. Liu, "Ranking products with online reviews: A novel method based on hesitant fuzzy set and sentiment word framework," *J. Oper. Res. Soc.*, vol. 71, no. 3, pp. 528–542, Mar. 2020, doi: 10.1080/01605682.2018.1557021.

[7] C. Guo, Z. Du, and X. Kou, "Products ranking through aspect-based sentiment analysis of online heterogeneous reviews," *J. Syst. Sci. Syst. Eng.*, vol. 27, no. 5, pp. 542–558, Oct. 2018, doi: 10.1007/s11518-018-5388-2.

[8] B. S. H. Karpurapu and L. Jololian, "A framework for social network sentiment analysis using big data analytics," in *Big Data and Visual Analytics*. Springer, 2017, pp. 203–217.

[9] C. W. Churchman, R. L. Ackoff, and E. L. Arnoff, "Introduction to operations research," Tech. Rep., 1957.

[10] F. Smarandache, M. Teodorescu, and D. Gifu, "Neutrosophy, a sentiment analysis model," in *Proc. RUMOUR Int. Conf.*, Toronto, ON, Canada, Jun. 2017, pp. 38–41.

[11] I. Kandasamy, W. B. Vasantha, J. M. Obbineni, and F. Smarandache, "Sentiment analysis of tweets using refined neutrosophic sets," *Comput. Ind.*, vol. 115, Feb. 2020, Art. no. 103180, doi: 10.1016/j.compind.2019.103180.

[12] K. Mishra, I. Kandasamy, W. B. Vasantha, and F. Smarandache, "A novel framework using neutrosophy for integrated speech and text sentiment analysis," *Symmetry*, vol. 12, no. 10, pp. 1–22, 2020, doi: 10.3390/sym12101715.

[13] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, nos. 1–2, pp. 1–35, 2008.

[14] B. Liu, "Sentiment analysis and opinion mining," *Morgan Claypool*, vol. 5, no. 1, pp. 1–167, 2012.

[15] T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in *Proc. Int. Conf. Knowl. Capture K-CAP*, 2003, pp. 70–77.

[16] I. Awajan and M. Mohamad, "A review on sentiment analysis in arabic using document level," *Int. J. Eng. Technol.*, vol. 7, no. 3.13, pp. 128–132, 2018.

[17] N. Boudad, R. Faizi, R. Oulad Haj Thami, and R. Chiheb, "Sentiment analysis in arabic: A review of the literature," *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 2479–2490, Dec. 2018.

[18] Y.-L. Yan, C.-L. Wang, and W.-L. Shi, "Survey of researches on Chinese sentiment analysis based on deep learning," *DEStech Trans. Comput. Sci. Eng.*, to be published.

[19] H. Peng, E. Cambria, and A. Hussain, "A review of sentiment analysis research in Chinese language," *Cognit. Comput.*, vol. 9, no. 4, pp. 423–435, Aug. 2017.

[20] B. G. Patra, D. Das, and A. Das, "Sentiment analysis of code-mixed Indian languages: An overview of SAIL_code-mixed shared task@ ICON-2017," 2018, *arXiv:1803.06745*. [Online]. Available: https://arxiv.org/abs/1803.06745

[21] S. Rani and P. Kumar, "A journey of indian languages over sentiment analysis: A systematic review," *Artif. Intell. Rev.*, vol. 52, no. 2, pp. 1415–1462, Aug. 2019.

[22] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *Stanford*, vol. 1, no. 12, p. 2009, 2009.

[23] F. Bütow, F. Schultze, and L. Strauch, "Semantic search: Sentiment analysis with machine learning algorithms on German news articles," Tech. Rep., 2017.

[24] M. Bouazizi and T. Ohtsuki, "Multi-class sentiment analysis on Twitter: Classification performance and challenges," *Big Data Mining Anal.*, vol. 2, no. 3, pp. 181–194, Sep. 2019.

[25] F. Hallsmar and J. Palm, "Multi-class sentiment classification on twitter using an emoji training heuristic," Tech. Rep., 2016, pp. 1–27.

[26] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," 2002, *arXiv:cs/0205070*. [Online]. Available: https://arxiv.org/abs/cs/0205070

[27] P. D. Turney, "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews," 2002, *arXiv:cs/0212032*. [Online]. Available: https://arxiv.org/abs/cs/0212032

[28] V. Vargas-Calderón, N. A. V. Sánchez, L. Calderón-Benavides, and J. E. Camargo, "Sentiment polarity classification of tweets using an extended dictionary," *Intel. Artif.*, vol. 21, no. 62, pp. 1–11, 2018, doi: 10.4114/intartif.vol21iss62pp1-12.

[29] F. H. Khan, U. Qamar, and S. Bashir, "ESAP: A decision support framework for enhanced sentiment analysis and polarity classification," *Inf. Sci.*, vols. 367–368, pp. 862–873, Nov. 2016.

[30] E. S. Tellez, S. Miranda-Jiménez, M. Graff, D. Moctezuma, R. R. Suárez, and O. S. Siordia, "A simple approach to multilingual polarity classification in Twitter," *Pattern Recognit. Lett.*, vol. 94, pp. 68–74, Jul. 2017.

[31] B. Pang and L. Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales," in *Proc. 43rd Annu. Meeting Assoc. Comput. Linguistics - ACL*, Jun. 2005, pp. 115–124, doi: 10.3115/1219840.1219855.

[32] S.-M. Kim and E. Hovy, "Determining the sentiment of opinions," in *Proc. 20th Int. Conf. Comput. Linguistics - COLING*, 2004, p. 1367, doi: 10.3115/1220355.1220555.

[33] X. Lei, X. Qian, and G. Zhao, "Rating prediction based on social sentiment from textual reviews," *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1910–1921, Sep. 2016.
[34] M. Rhanoui, M. Mikram, S. Yousfi, and S. Barzali, "A CNN-BiLSTM model for document-level sentiment analysis," *Mach. Learn. Knowl. Extraction*, vol. 1, no. 3, pp. 832–847, Jul. 2019.
[35] A. Tripathy, A. Anand, and S. K. Rath, "Document-level sentiment classification using hybrid machine learning approach," *Knowl. Inf. Syst.*, vol. 53, no. 3, pp. 805–831, Dec. 2017.
[36] M. E. Basiri and A. Kabiri, "Sentence-level sentiment analysis in persian," in *Proc. 3rd Int. Conf. Pattern Recognit. Image Anal. (IPRIA)*, Apr. 2017, pp. 84–89.
[37] O. Appel, F. Chiclana, J. Carter, and H. Fujita, "A hybrid approach to the sentiment analysis problem at the sentence level," *Knowl.-Based Syst.*, vol. 108, pp. 110–124, Sep. 2016.
[38] T. Hayashi and H. Fujita, "Word embeddings-based sentence-level sentiment analysis considering word importance," *Acta Polytech. Hungarica*, vol. 16, no. 7, p. 152, 2019.
[39] Y. Xia, E. Cambria, and A. Hussain, "AspNet: Aspect extraction by bootstrapping generalization and propagation using an aspect network," *Cognit. Comput.*, vol. 7, no. 2, pp. 241–253, Apr. 2015.
[40] Y. Wang, A. Sun, M. Huang, and X. Zhu, "Aspect-level sentiment analysis using AS-capsules," in *Proc. World Wide Web Conf. - WWW*, 2019, pp. 2033–2044.
[41] M. D. P. Salas-Zárate, J. Medina-Moreira, K. Lagos-Ortiz, H. Luna-Aveiga, M. Á. Rodríguez-García, and R. Valencia-García, "Sentiment analysis on tweets about diabetes: An aspect-level approach," *Comput. Math. Methods Med.*, vol. 2017, pp. 1–9, Feb. 2017.
[42] D. Bollegala, D. Weir, and J. Carroll, "Cross-domain sentiment classification using a sentiment sensitive thesaurus," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 8, pp. 1719–1731, Aug. 2013.
[43] J. Blitzer, M. Dredze, and F. Pereira, "Biographies, bollywood, boomboxes and blenders: Domain adaptation for sentiment classification," in *Proc. 45th Annu. Meeting Assoc. Comput. Linguistics*, Jun. 2007, pp. 440–447.
[44] N. Li, S. Zhai, Z. Zhang, and B. Liu, "Structural correspondence learning for cross-lingual sentiment classification with one-to-many mappings," 2016, *arXiv:1611.08737*. [Online]. Available: http://arxiv.org/abs/1611.08737
[45] R. González-Ibánez, S. Muresan, and N. Wacholder, "Identifying sarcasm in twitter: A closer look," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguistics, Hum. Lang. Technol.*, 2011, pp. 581–586.
[46] O. Tsur, D. Davidov, and A. Rappoport, "ICWSM-A great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews," in *Proc. ICWSM*, 2010, pp. 162–169.
[47] X. Fu, W. Liu, Y. Xu, C. Yu, and T. Wang, "Long short-term memory network over rhetorical structure theory for sentence-level sentiment analysis," in *Proc. Asian Conf. Mach. Learn.*, 2016, pp. 17–32.
[48] S. Moghaddam and M. Ester, "Opinion digger: An unsupervised opinion miner from unstructured product reviews," in *Proc. 19th ACM Int. Conf. Inf. Knowl. Manage. - CIKM*, 2010, pp. 1825–1828.
[49] M. Farhadloo, *Statistical Models for Aspect-Level Sentiment Analysis*. Merced, CA, USA: UC Merced, 2015.
[50] M. Farhadloo and E. Rolland, "Multi-class sentiment analysis with clustering and score representation," in *Proc. IEEE 13th Int. Conf. Data Mining Workshops*, Dec. 2013, pp. 904–912.
[51] F. Hemmatian and M. K. Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis," *Artif. Intell. Rev.*, vol. 52, no. 3, pp. 1495–1545, 2019.
[52] B. Verma and R. S. Thakur, "Sentiment analysis using lexicon and machine learning-based approaches: A survey," in *Proc. Int. Conf. Recent Advancement Comput. Commun.*, 2018, pp. 441–447.
[53] A. Abdul Aziz and A. Starkey, "Predicting supervise machine learning performances for sentiment analysis using contextual-based approaches," *IEEE Access*, vol. 8, pp. 17722–17733, 2020.
[54] E. Bastı, C. Kuzey, and D. Delen, "Analyzing initial public offerings' short-term performance using decision trees and SVMs," *Decis. Support Syst.*, vol. 73, pp. 15–27, May 2015.
[55] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986.
[56] F. S. Abdullah, N. S. A. Manan, A. Ahmad, S. W. Wafa, M. R. Shahril, N. Zulaily, R. M. Amin, and A. Ahmed, "Data mining techniques for classification of childhood obesity among year 6 school children," in *Proc. Int. Conf. Soft Comput. Data Mining*, 2016, pp. 465–474.
[57] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[58] B. Liu, *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. Springer, 2007.
[59] M. R. Saleh, M. T. Martín-Valdivia, A. Montejo-Ráez, and L. A. Ureña-López, "Experiments with SVM to classify opinions in different domains," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 14799–14804, Nov. 2011.
[60] M. Mohamad, W. N. S. W. Nik, Z. A. Zakaria, and A. C. Alhadi, "An analysis of large data classification using ensemble neural network," *Int. J. Eng. Technol.*, vol. 7, no. 2.14, pp. 53–56, 2018.
[61] A. El Abdouli, L. Hassouni, and H. Anoun, "Sentiment analysis of moroccan tweets using naive Bayes algorithm," *Int. J. Comput. Sci. Inf. Secur.*, vol. 15, no. 12, pp. 1–10, 2017.
[62] S. Taj, B. B. Shaikh, and A. Fatemah Meghji, "Sentiment analysis of news articles: A lexicon based approach," in *Proc. 2nd Int. Conf. Comput., Math. Eng. Technol. (iCoMET)*, Jan. 2019, pp. 1–5.
[63] M. P. Ashna and A. K. Sunny, "Lexicon based sentiment analysis system for malayalam language," in *Proc. Int. Conf. Comput. Methodologies Commun. (ICCMC)*, Jul. 2017, pp. 777–783.
[64] Y. M. Aye and S. S. Aung, "Senti-lexicon and analysis for restaurant reviews of myanmar text," *Int. J. Adv. Eng., Manage. Sci.*, vol. 4, no. 5, pp. 380–385, 2018.
[65] P. Ray and A. Chakrabarti, "Twitter sentiment analysis for product review using lexicon method," in *Proc. Int. Conf. Data Manage., Analytics Innov. (ICDMAI)*, Feb. 2017, pp. 211–216.
[66] L. Augustyniak, P. Szymánski, T. Kajdanowicz, and W. Tuliglowicz, "Comprehensive study on lexicon-based ensemble classification sentiment analysis," *Entropy*, vol. 18, no. 1, pp. 1–29, 2016, doi: 10.3390/e18010004.
[67] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining - KDD*, 2004, p. 168, doi: 10.1145/1014052.1014073.
[68] G. Qiu, B. Liu, J. Bu, and C. Chen, "Expanding domain sentiment lexicon through double propagation," in *Proc. IJCAI*, Jul. 2009, pp. 1199–1204.
[69] G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," *Comput. Linguistics*, vol. 37, no. 1, pp. 9–27, Mar. 2011.
[70] R. Pandarachalil, S. Sendhilkumar, and G. S. Mahalakshmi, "Twitter sentiment analysis for large-scale data: An unsupervised approach," *Cognit. Comput.*, vol. 7, no. 2, pp. 254–262, Apr. 2015.
[71] H. Saif, Y. He, M. Fernandez, and H. Alani, "Contextual semantics for sentiment analysis of Twitter," *Inf. Process. Manage.*, vol. 52, no. 1, pp. 5–19, Jan. 2016.
[72] P. V. Rajeev and V. S. Rekha, "Recommending products to customers using opinion mining of online product reviews and features," in *Proc. Int. Conf. Circuits, Power Comput. Technol. [ICCPCT-]*, Mar. 2015, pp. 1–5.
[73] W. Wang, H. Wang, and Y. Song, "Ranking product aspects through sentiment analysis of online reviews," *J. Experim. Theor. Artif. Intell.*, vol. 29, no. 2, pp. 227–246, Mar. 2017, doi: 10.1080/0952813X.2015.1132270.
[74] A. K. J and A. S, "Aspect-based opinion ranking framework for product reviews using a Spearman's rank correlation coefficient method," *Inf. Sci.*, vols. 460–461, pp. 23–41, Sep. 2018, doi: 10.1016/j.ins.2018.05.003.
[75] L. A. Zadeh, "Fuzzy sets," *Inf. Control.*, vol. 353, pp. 338–353, 1965.
[76] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets Syst.*, vol. 20, pp. 87–96, Aug. 1986.
[77] F. Smarandache, "Neutrosophic set- A generalization of the intuitionistic fuzzy set," *Int. J. pure Appl. Math.*, vol. 24, no. 3, p. 287, 2005.
[78] F. Smarandache, *Neutrosophy. Neutrosophic Probability, Set, And Logic*. Rehoboth, DE, USA: American Research Press, 1998.
[79] H. Wang, F. Smarandache, Y. Q. Zhang, and R. Sunderraman, "Single valued neutrosophic sets," *Multispace Multistruct.*, vol. 4, pp. 410–413, Oct. 2010.
[80] J. Ye, "A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets," *J. Intell. Fuzzy Syst.*, vol. 26, no. 5, pp. 2459–2466, 2014, doi: 10.3233/IFS-130916.
[81] J.-J. Peng, J.-Q. Wang, J. Wang, H.-Y. Zhang, and X.-H. Chen, "Simplified neutrosophic sets and their applications in multi-criteria group decision-making problems," *Int. J. Syst. Sci.*, vol. 47, no. 10, pp. 2342–2358, Jul. 2016, doi: 10.1080/00207721.2014.994050.
[82] F. Altun, R. Şahin, and C. Güler, "Multi-criteria decision making approach based on PROMETHEE with probabilistic simplified neutrosophic sets," *Soft Comput.*, vol. 24, no. 7, pp. 4899–4915, Apr. 2020, doi: 10.1007/s00500-019-04244-4.

[83] J.-J. Peng, J.-Q. Wang, H.-Y. Zhang, and X.-H. Chen, "An outranking approach for multi-criteria decision-making problems with simplified neutrosophic sets," *Appl. Soft Comput.*, vol. 25, pp. 336–346, Dec. 2014, doi: 10.1016/j.asoc.2014.08.070.

[84] J. Z. Tian, J. Wang, J. Q. Wang, and H. Y. Zhang, "Simplified neutrosophic linguistic multicriteria group decision-making approach to green product development," *Group Decis. Negotiation*, vol. 26, no. 3, pp. 597–627, 2017.

[85] R. Bausys, E. K. Zavadskas, and A. Kaklauskas, *Application of Neutrosophic Set to Multicriteria Decision Making by COPRAS*. Coimbatore, India: Infinite Study, 2015.

[86] X. Peng and C. Liu, "Algorithms for neutrosophic soft decision making based on EDAS, new similarity measure and level soft set," *J. Intell. Fuzzy Syst.*, vol. 32, no. 1, pp. 955–968, Jan. 2017.

[87] D. S. Xu, C. Wei, and G. W. Wei, "TODIM method for single-valued neutrosophic multiple attribute decision making," *Information.*, vol. 8, no. 4, pp. 1–18, 2017, doi: 10.3390/info8040125.

[88] S. Joshi and D. Deshpande, "Twitter sentiment analysis system," *Int. J. Comput. Appl.*, vol. 180, no. 47, pp. 35–39, Jun. 2018, doi: 10.5120/ijca2018917319.

[89] K. Chauhan, B. Ashish, and G. Amita, "Sentiment analysis using vader," *Int. J. Advance Res., Ideas Innov. Technol.*, vol. 4, no. 1, pp. 485–489, 2018.

[90] X. Chen, Y. Xue, H. Zhao, X. Lu, X. Hu, and Z. Ma, "A novel feature extraction methodology for sentiment analysis of product reviews," *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6625–6642, Oct. 2019, doi: 10.1007/s00521-018-3477-2.

[91] X. Fang and J. Zhan, "Sentiment analysis using product review data," *J. Big Data*, vol. 2, no. 1, pp. 1–4, Dec. 2015, doi: 10.1186/s40537-015-0015-2.

[92] L. Yang, X. Geng, and H. Liao, "A Web sentiment analysis method on fuzzy clustering for mobile social media users," *EURASIP J. Wireless Commun. Netw.*, vol. 2016, no. 1, pp. 1–13, Dec. 2016, doi: 10.1186/s13638-016-0626-0.

[93] S. Hossayni, M. Akbarzadeh-t, and D. Reforgiato, "Fuzzy synsets, and lexicon-based sentiment analysis," Tech. Rep., 2016.

[94] B. Anil, A. K. Singh, A. K. Singh, and S. Thota, "Sentiment analysis for product reviews," *Int. J. Adv. Res. Comput. Sci.*, vol. 9, no. 3, p. 23, 2018.

[95] H. Patil and P. M. Mane, "Survey on product review sentiment analysis with aspect ranking," *Int. J. Sci. Res.*, vol. 5, no. 9, pp. 749–752, 2016.

[96] A. Haque and T. Rahman, "Sentiment analysis by using fuzzy logic," *Int. J. Comput. Sci., Eng. Inf. Technol.*, vol. 4, no. 1, pp. 33–48, Feb. 2014, doi: 10.5121/ijcseit.2014.4104.

[97] J. Zhang, D. Chen, and M. Lu, "Combining sentiment analysis with a fuzzy kano model for product aspect preference recommendation," *IEEE Access*, vol. 6, pp. 59163–59172, 2018.

[98] P. Biswas, S. Pramanik, and B. C. Giri, "TOPSIS method for multi-attribute group decision-making under single-valued neutrosophic environment," *Neural Comput. Appl.*, vol. 27, no. 3, pp. 727–737, Apr. 2016.

[99] F. Smarandache, "N-valued refined neutrosophic logic and its applications to physics," *Infnite Study*, vol. 4, pp. 143–146, Oct. 2013.

IBRAHIM AWAJAN received the B.Sc. degree in computer information system from Al-Hussein Bin Talal University, Jordan, in 2006, and the M.Sc. degree in computer science from Al-Balqa' Applied University, Jordan. He is currently pursuing the Ph.D. degree in computer science with Universiti Sultan Zainal Abidin (UniSZA), Kuala Terengganu. He is working as a Lecturer with the Faculty of Computer Science and IT Deanship, Shaqra University, Saudi Arabia. His research interest includes sentiment analysis.

MUMTAZIMAH MOHAMAD received the B.Sc. degree in information technology from Universiti Kebangsaan Malaysia, Malaysia, in 2000, the M.Sc. degree in software engineering from University Putra Malaysia, in 2005, and the Ph.D. degree in computer science from Universiti Malaysia Terengganu, in 2014. She was appointed as a Lecturer with Kolej Ugama Sultan Zainal Abidin that currently known as Universiti Sultan Zainal Abidin (UniSZA), Kuala Terengganu, in 2000. She has been working as an Associate Professor and holds the responsibility as the Deputy Dean of Academic and Postgraduate for the Faculty of Informatics and Computing, since 2017. Her research interests include computer science especially in pattern recognition, machine learning, neural networks, and parallel processing.

ASHRAF AL-QURAN received the B.Sc. and M.Sc. degrees in mathematics from the Jordan University of Science and Technology, Jordan, and the Ph.D. degree in mathematics from Universiti Kebangsaan Malaysia, Malaysia. He is currently an Assistant Professor with the Department of Basic Sciences, Preparatory Year Deanship, King Faisal University, Saudi Arabia. His research interests include decision making under uncertainty, fuzzy set theory, neutrosophic set theory, and neutrosophic logic.