This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3183108

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.
Digital Object Identifier 10.1109/ACCESS. 2017.Doi Number

An Opinion Mining Approach to Handle Perspectivism and Ambiguity: Moving towards Neutrosophic Logic

REEM ESSAMELDIN[1], AHMED A. ISMAIL[2], AND SAAD M. DARWISH[1,*]
[1] Department of Information Technology, Institute of Graduate Studies and Research, Alexandria University, 163 Horreya Avenue, El–Shatby 21526, Alexandria, Egypt.
[2] The Higher Institute of Computers and Information Systems, Abo Qir, Alexandria 21913, Egypt.

Corresponding author: Saad M. Darwish (e-mail: saad.darwish@alexu.edu.eg).

ABSTRACT One main incentive for the utilization of opinion mining (OM) in social media is the impossibility of manually analyzing millions of opinions. However, applying OM for decision-making requires imitating human brain reasoning for more pragmatic results. Analyzing opinions in human-like intelligence is vital for avoiding misleading results. This occurs when the process can properly excogitate social influence and environmental uncertainty. In this study, an OM model for Twitter is proposed to handle perspectivism and opinion ambiguity. For perspectivism, social network analysis (SNA) is conducted, where users are ranked and then weighted using the UCINET tool and neural networks. An uncertainty classifier is used to integrate users' influence levels with the polarity scores of their texts, providing a new polarity score that can reflect the real-world reasoning of opinions. Polarity scores needed for integration is done using the lexicon resource TextBlob. In the proposed model, three uncertainty classifiers are tested: type1 fuzzy logic (T1- FL), type2 fuzzy logic (T2- FL), and neutrosophic logic (NL). A comparative analysis of the methods shows the ability of NL to deal with the uncertainty existing in the data more accurately, proving the benefit of NL in improving the power of OM in social media.

INDEX TERMS Social network analysis, opinion mining, fuzzy logic, neutrosophic logic.

I. INTRODUCTION

With billions of active users on Facebook and thousands of tweets every second on Twitter, opinions move on social media at a full speed [1]. Social media can provide businesses with a wealth of information about customers' thoughts and feelings. Moreover, social media becomes the primary source of information that expresses online public opinion in emergency or hot events (e.g. COVID-19) [2]. However, with data growing by the day, it is impossible to manually analyze this mass of information. This is where opinion mining (OM) came in [1],[3]. This information, if properly processed, allows businesses to understand opinions that are more important in supporting decisions to improve customer experience and public responsiveness. Thus, it is invaluable to businesses, governments, and individuals [2],[4],[5].

Whereas media is increasingly social, audiences are no longer disconnected from each other, can talk back, and influence each other [6]. This caused the practice of perspectivism to spread, where people can perceived opinions from different point of views depends on the influence degree of the opinion holders on their followers [6],[7]. Followers are naturally biased towards the more trusted opinion holders. Consciously recognizing this case, social media beneficiaries and researchers have tried to find a way to identify influencers and spread positive opinions with their help [7]. Reasoning the human brain in how people perceive texts is essential in the OM process for more rational results. Accordingly, Social Network Analysis (SNA) becomes part of the OM battle [4],[6],[7]. Using different methods, SNA succeeded in identifying influencers for business to gain their advantage for achieving predefined goals; business owners become allowed to measure their audience power to change their market image [6],[7].

In a deep look at the nature of opinion on social media, opinions, whether reflect emergency or nonemergency events, are usually contains uncertain information despite the uncertainty of the electronic opinion resulting from distance communication with strangers that challenges the

OM process in drawing conclusions, as would a human do [4],[8]. Moreover, emergencies increase the uncertainty [2]. Opinions, such as text, can face different types of uncertainties, including vagueness, indeterminacy, and ambiguity. This is when information has borderline cases, imprecise with conflicts and paradoxes, and leads to several feasible explications. These uncertainty types must be handled properly [9-11].

Fuzzy set theory, especially Type1-FL (T1-FL), has been considered the first to deal with the uncertainty problem (e.g., to handle vagueness in OM) and literature-wise, T1-FL is reviewed as the most adopted FL in the literature, especially in OM, where T1-FL deals with vagueness by following a fine-grained classification of sentiments[7],[12],[13]. However, the growing complexity of uncertainty and T1-FL limitations have driven the development of many extensions (e.g., intuitionistic fuzzy) [13],[14]. Type2- FL (T2-FL) has been proven to handle more complicated uncertainties (e.g., ambiguity or uncertainty resulting from different knowledge sources) [13],[15]. The main reason is in the T2-FL basic feature of determining the uncertainty range using a footprint of uncertainty (FOU). The FOU in T2-FL is constructed using two T1-FL membership functions (MF): upper MF and lower MF. Using FOU, uncertainty is captured and controlled, thus minimizing its adverse effects on systems [15].

Fuzzy set theory and its extensions can capture the positive and/or negative membership degrees to a set but fail to address the indeterminacy component [14]. To address indeterminacy, the concept of neutrosophy has been introduced [12]. The neutrosophic classifier is an extension of the fuzzy classifier that utilizes neutrosophic logic (NL). It is a more generalized logic that can effectively detect and reduce uncertain instances by providing undependable degrees of truth, uncertainty (indeterminacy), and falsity. Thus, NL can deal with the uncertainty nature of social media opinions and improve the process of decision-making [16].

This study attempts to overcome the problems of perspectivism and uncertainty to provide more accurate OM results from social media data that are suitable for any decision-making process. It proposes an integrated model that adopts SNA and OM processes in a unified framework using an uncertainty classifier to handle different types of environmental uncertainty. T1-FL, T2-FL, and NL were used to investigate the performance of the proposed model.

In line with the work done in [7], the sustained contributions of the proposed model are summarized and listed as follows:

-- Propose a modified opinion mining classifier that represents the polarity score of texts as a function of their users' influence level through a hybrid OM classifier that considering perspectivism.
-- Propose the integration of social network analysis with the regular opinion mining process through a neutrosophic classifier.
-- Quantify uncertainty for three machine-learning classifiers, including the proposed neutrosophic classifier.
-- The proposed model attempts to imitate the human brain reasoning in analyzing sentiments through reflecting perspectivism using social network analysis and detecting ambiguity using neutrosophic logic which is known for being highly integrated with the way humans think.
-- Implement the proposed model on a customized dataset collected for this purpose and benchmark datasets to investigate the advantage of the proposed model and its efficiency in dealing with the ambiguity compared to the other mentioned machine learning classifiers.

The remainder of this paper is organized as follows. In Section II, existing FL and NL research works on OM are considered. Section III illustrates the research problem, and in Section IV, the proposed model is defined. Experiments and analysis of results are presented and clarified in Section V. Conclusions and impending breadth for development are presented in Section VI.

II. RELATED WORK

In the literature, besides classifying its techniques into lexicon-based and machine learning (ML), OM research has chiefly pursued two target areas [1],[3]. Study the impact of various techniques on improving the main process of OM (i.e., accuracy and effectiveness) or applying different OM tasks to achieve predefined goals (e.g., stock market prediction and customer satisfaction) [13],[17]. Most studies have combined both targets; researchers are continuing to investigate the best OM techniques for better application accuracy [18],[19]. The same is the case when dealing with FL in the OM literature; FL has been studied from the perspective of finding a way to improve the idea of FL itself. In addition, the adoption of FL to address low accuracy has occurred in OM applications [13].

T1-FL is essentially adopted in the OM literature for minimizing vagueness to improve efficiency through fine-grained classification of opinions instead of the default classification (i.e., positive, neutral, negative) [7],[20]. One recent study in this area was conducted in 2021, where Batista *et al*. [20] attempted to facilitate decision-making based on a huge amount of texts extracted from social media on different topics at different times by creating a FL sentiment analysis (SA) dimension. The authors used OLAP for text storage and extraction in their multidimensional model. The idea behind FL usage is the ability to fine-grain sentiments, and thus, a more detailed analysis. The hierarchy levels of this fuzzy dimension consisted of five levels organized from the most general to the most specific (i.e., three, five, six, seven, and nine classes). Texts were clustered, sentiment scores were assigned, and processed using FL. TextBlob and VADER were used as they perform well with social texts and for being unsupervised tools so more generic. A comparative study was conducted on real tweets and movie reviews using six ML algorithms, showing the high performance and accuracy of this method. The advantage of this method is that it is both fully automated and unsupervised.

The idea of refining outputs using FL was adopted in the same year as Sharma *et al*. [21] used FL to obtain the final polarity score from the lexicon resource VADER. VADER gives positive and negative scores for the same text, whereas FL can classify texts into positive, negative, and neutral, considering positive and negative lexicon score intersections. The contribution of this work lies in the proposed fuzzy rules for assigning polarity to texts. Experiments were conducted on Twitter data and blogs, showing significant results for Twitter data but were not scalable for blogs. Recommendations were made to improve the adaptability of this method to any dataset.

Another fuzzy-based sentiment analyzer was proposed in 2020. Khattak *et al*. [22] attempted to classify customer sentiments at a fine-grained level. The authors attempted to improve the efficiency of the SA applied to measure customer satisfaction. They developed a fine-grained SA system by extending fuzzy linguistic hedges and rule sets to seven levels of granularity. In their work, 20 linguistic hedges were used to modify the shape of the fuzzy linguistic variable (e.g., extremely, too, or few). The opinion scores of words were obtained using SentiWordNet. The sentiment scores of linguistic hedges, along with the lexicon score of opinion words at a fine-grained level, provided the fuzzy input linguistic variable. Customer feedbacks from different products were collected and tested. Experiments and comparisons with different methods showed improved performance and efficiency of this method. One disadvantage of this method, as mentioned by the authors, is the inability to handle ambiguous words (e.g., the word "high" when used to express negative opinions). In 2017, Krishna *et al*. [23] discussed the analysis of the unstructured formatted data. They proposed a seven level classification of opinions and reviews' features using T1-FL. They believed that fine graining opinions and extracting opinions' features could improve the decision making process. Opinion words are classified using support vector machines (SVM) and Maximum entropy (MAXENT) then fed as input to a fuzzy inference system that provided a fine-grained sentiments. A comparative study was conducted between the two supervised classifiers in terms of accuracy and learning rate. The results showed a higher accuracy and learning rate for SVM.

Due to different types and reasons for uncertainty, T1-FL was found to be unable to deal with them all [24], especially when modeling ambiguous words as previously mentioned in [22]. Many extensions have been developed in the literature to improve FL performance when dealing with uncertainties, such as T2- FL [17],[24]. In 2020, Abbasimehr et al. [25] believed in the power of users to improve the effectiveness of the OM process by identifying reviewer credibility. They assumed that highly credible

TABLE I
SUMMARY OF DIFFERENT FL AND NL TECHNIQUES IN OM LITERATURE

Algorithm	Sentiment Analysis / Opinion Mining				SNA	Research Concern		Aim of Work
	Uncertainty Classification		Other methods			Process	Application	
	Type	# of classes	Lexicon	ML				
[20]	FL-T1	5 levels (3:9) classes	TextBlob/ VADER	×	×	√	×	1
[21]	FL-T1	3	VADER	×	×	√	×	2
[22]	FL-T1	7	SentiWordNet	×	×	×	√	4, 6
[23]	FL-T1	7	×	SVM, MAXENT	×	√	×	2, 4
[25]	×	×	×	×	FL-T2 AHP	√	×	3, 4
[26]	FL-T2	3	×	√	×	×	√	7
[27]	FL-T1&2	5	SentiWordNet	×	×	×	√	7
[28]	×	×	×	×	×	×	classifying appendicitis	5 (FL and NL)
[11]	×	×	×	×	×	×	systems evaluation	5, 7 (FL, IFL, and NL)
[12]	NL	7, 5, 3	×	FL-T1	×	√	×	2, 4
[29]	×	×	×	×	×	handle ambiguity	×	5, 7 (FL and NL)
[30]	NL	3	VADER, neutral lexicon	×	×	√	Rank products	1, 7
[31]	NL	8 classes	VADER	k-NN, SVM, K-means	×	√	×	4, 7

1 Handle big data to facilitate decision making 2 Refine polarity results 3 Detecting credible users 4 Improve accuracy of SA process
5 Comparative study 6 Customer satisfaction based on SA of reviews 7 To deal with the uncertain nature of data
'×' stands for "not applicable" and ' √' stands for "applicable"

users write high-quality reviews; thus, the OM process would be more accurate and efficient. They depended on T2- FL analytical hierarchy process (AHP) to rank source creditability. The hierarchy process is widely used in weighting, but it is difficult to deal with the uncertain judgments of decision makers. T2- FL was used in this process to handle uncertainty. Experiments were conducted using real data collected from epinions. The results show the significance of this method in determining creditable sources, thus improving the performance of the OM process based on online reviews.

In 2019, Bi *et al*. [26] introduced the problem of the low accuracy of SA results and its effect on the quality of decision making. They proposed the usage of T2- FL while representing aggregated sentiment scores for certain products. In their work, they considered the low-accuracy sentiment results to be uncertain, and thus represented them using T2- FL. The method was applied to a decision-making system for ranking products based on their aggregated sentiment results and decision maker preferences. Preferences were the aspects of products in which online reviews were analyzed. This study was the first attempt to consider accuracy rates while representing SA results.

In 2015, Ali *et al*. [27] highlighted the classical ontology limitation in dealing with the uncertain nature of Internet data when searching for reviews. They proposed searching for reviews (i.e., hotel reviews) using an OM ontology system that adopted T2-FL. Using such a system, opinions about hotels were preprocessed and converted into proper search queries, and feature opinions were extracted using T1-FL ontology. The T2-FL ontology was used to integrate the extracted features with the available information about hotels and what was required by users. Moreover, they aimed to improve the performance of the system. Features were assigned polarity scores based on the lexicon resource SentiWordNet. After aggregating the polarity scores of each word in an opinion sentence, T1-FL was applied to fine-grain the final polarity into five polarity classes using predefined rules. Experiments verified the ability of the proposed system to accurately analyze reviews of uncertain data types.

Some researchers have investigated the differences between FL and NL when dealing with uncertainty. Bhutani and Aggarwal [28] applied both FL and NL to classify appendicitis. They first identified the difference between classifications using fuzzy probability and neutrosophic probability, and their determination and entropy criteria. Second, both methods were implemented using the appendicitis dataset. The main difference was that neutrosophic allowed a real-world sample that included indeterminacy. Results using neutrosophic probability and entropy could be complete (i.e., have zero indeterminacy), incomplete (i.e., indicate the existence of some indeterminacy degree), and paraconsistent with some attribute inconsistency. More realistic and accurate classification results were obtained using neutrosophic probability and entropy than fuzzy probability.

Radwan *et al*. [11] discussed the problems of different uncertainty types facing knowledge representation and analysis in expert systems. They proposed an uncertain expert system for the quality evaluation of different management systems specified for the learning process based on predefined factors (e.g., ability to learn, ability to remember, and process efficiency). They applied three uncertainty models (T1-FL, intuitionistic FL, and NL) for this purpose. The results showed the superiority of NL in the evaluation due to FL and intuitionistic FL failure in handling indeterminacy and false membership representations, respectively.

Kandasamy *et al*. [12] attempted to represent the FL classification results of SA by using refined neutrosophic sets. They substituted the normal classes of SA into positive, indeterminate, and negative. They applied multi-refined, triple-refined, and single-value neutrosophic sets. The FL results were converted into seven classes (i.e., two positive, three indeterminate, two negative), five (i.e., two positive, indeterminate, two negative), and three (i.e., positive, indeterminate, and negative) classes. They applied refined sets to the SA results of ten social subjects on Twitter. The results showed that the multi-refined neutrosophic sets succeeded in improving the accuracy and performance of SA and proved their superiority in representing the indeterminacy and neutrality of opinions.

Ansaria *et al*. [32] discussed the neutrosophic phenomenon, how it differs from fuzzy, and how to implement it using the FL toolbox of MATLAB. They built a neutrosophic inference system (NIS) using three fuzzy inference systems (FIS): truth, indeterminacy, and falsity components of NL. They highlighted the overlapping region of fuzzy memberships as the region of indeterminacy and falsity in NL. The region of no overlap in the membership functions was represented by the truth component of the NL. A comparison was conducted between the results from both the NIS and FIS using the iris dataset special for values that lie in the ambiguity zone. Ambiguity in the FIS was defined by the results that lie in the overlapping region between two adjacent membership functions. However, for NIS, was identified by nonzero indeterminate and falsity values. The ambiguity resulting from the FIS was 36%, compared to 10% ambiguity resulting from the NIS. Therefore, the proposed NIS succeeded in handling ambiguity with a better accuracy than the FIS.

Until recent, the application of the neutrosophic theory to the OM process in social media is unfamiliar in the literature despite the continuous necessity to mine public opinion on social media and reduce the generated uncertainty of the emergency events and the ability of neutrosophic to deal with this type of uncertainty

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3183108

[2],[12],[29]. The authors in [30] attempted to combine neutrosophic with the sentiment analysis process of online reviews to rank products. They proposed a Neutro-VADAR (valence aware dictionary and sentiment reasoner) approach that assigned the online reviews a neutral score based on the generated positive and negative scores from VADAR and the combined neutral lexicon. If neutral words were detected, then the neutrality score equaled to 1. If both negative and positive words were only detected, then the neutrality score equals to the resulted subtraction of 1 and the compound score. Afterwards, products were ranked using neutrosophic averaging and similarity measures. The performance of the proposed model was tested on real data from Twitter. The results showed the ability of the proposed Neutro-VADAR to handle uncertainties better than the traditional VADAR.

To deal with the problem of mining opinions about hot events, the authors of [31] highlighted the indeterminate nature of such opinions on Twitter. They attempted to represent the opinions related to #MeToo movement based on the neutrosophic theory by clustering and classifying tweets into positive, indeterminate, and negative. The tweets were assigned a polarity score using VADAR then clustered and classified into positive, indeterminate, and negative memberships to be represented as neutrosophic. The authors applied k-means algorithm to cluster the neutrosophic sets and k-nearest neighbors (k-NN), and support vector machine classifiers (SVM) to classify the neutrosophic sets, then a performance evaluation was conducted on the chosen clustering and classification algorithms. The results showed a better performance for k-NN than SVM. The dataset analysis indicated that the movement's trend was indeterminate not positive as demonstrated by other analysis methods. Table I summarizes all the essential points covered in the literature.

III. PROBLEM DEFINITION

Decision quality is directly correlated with the accuracy of the results required for the decision process. The same is the case for decisions taken based on the OM process applied to social media. The more real the data processing, the more accurate and reliable the results, and the more inerrant the decisions therefrom. Moreover, the more emergent is the analyzed situation, the higher accurate is the required decisions. On this account, those involved in decision-making based on OM seek realistic opinions in terms of being candid in nature and reflect real intention when being analyzed. For a realistic OM, we should consider mining opinions from social media platforms such as Facebook and Twitter, which are far more popular among everyday people. For accurate analysis, the following problems need to be handled:

-- Opinions uncertainty nature: in the most cases in social media, the OM process is applied to opinions that reflect emergency events, where decision makers are required to accurately response to the online public opinion at a given time. The emergencies and natural human language increase the OM uncertainty in social media. Uncertainty of opinions comes in many shapes starting from using words that are vague with different degrees of same polarity to ambiguous words with different meanings for the same words.

-- Social media influence: When dealing with social media, we should consider the influence of users on each other. This influence can shift the polarity of opinions when they are read by others (e.g., ratifiers).

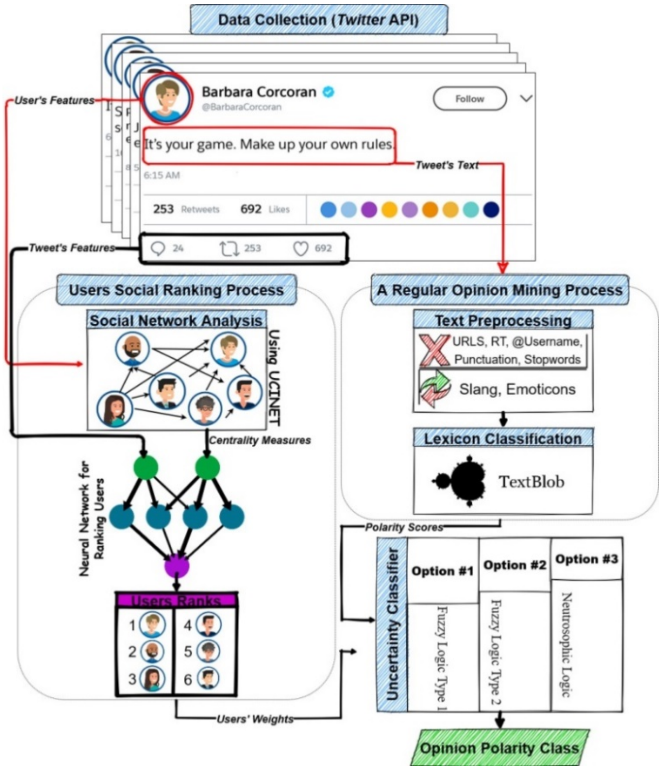


FIGURE 1. The proposed OM model.

IV. PROPOSED MODEL

To make the process of OM simulate human brain reasoning, we propose this model on Twitter data. The proposed model focuses primarily on the aforementioned problems. The output of this model represents a more accurate and realistic opinion polarity for effective decision-making. The contributions of the proposed model are as follows:

-- In line with the work done in [7], the proposed OM model considers social perspectivism concerning how people may perceive opinions on social media. It performs SNA by integrating two types of influence: not only the influence of users but also their opinion text by considering reactions on their tweets.

-- In terms of the main OM process, the proposed OM model attempts to improve the accuracy of the regular OM process using a hybrid technique. It combines a lexicon-based technique (i.e., TextBlob, effulgence in terms of accuracy) with ML that can handle uncertainty.

- -- The proposed OM model considers uncertainty in text classification through the implementation of an uncertainty classifier for opinions.
- -- A comparative analysis of different uncertainty classifiers was conducted to achieve better opinion classification accuracy.
- -- Finally, the proposed OM model is the first to integrate results from SNA with the text polarity score obtained from TextBlob using NL to handle ambiguity. Moreover, it can avoid vagueness by fine-graining tweets into seven classes (e.g., weak negative, negative, and strong negative). Fig. 1 presents the proposed model and its components. The phases of the proposed model are described as follows:

A. DATA COLLECTION (TWITTER API)

This phase is concerned with collecting the data essential for validating the proposed OM model. Twitter was chosen to apply the OM model. Twitter can provide a real-time trend list that reflects the hot topics of people's interest at a given time. Thus, Twitter, as an application domain, is an effective source to collect data about both the nonemergency (e.g., products reviews) and the emergency topics (e.g., COVID-19) that possess higher uncertainty. In our case, the data needed are of two types: main opinion texts along with reactions to them (e.g., likes, shares, comments). In addition, the social interaction of tweets' authors on social media includes the following and followers' lists of authors. There is no benchmark dataset of this type available online. Thus, the Twitter API is the chosen tool for collecting desired data. The required data were categorized as follows:

- -- Tweet's Text: includes text content in the English language – one tweet text per user,
- -- Tweet features: consist of different reactions to the collected opinion tweet, including favorite, reply, and retweet counts per each tweet, and
- -- User's features: include followers and following lists per user.

B. USERS SOCIAL RANKING PROCESS

This phase is still one of the main contributions of our work, although it was performed in [7]. It deals with social perspectivism and attempts to determine user influence on Twitter. The input to this phase was the tweet and user feature data collected in Phase 1. In this phase, two basic processes are performed: conducting the SNA and ranking users. The output of this phase is the users' weights ready to be integrated with the polarity scores obtained from upcoming phase #3.

1) SOCIAL NETWORK ANALYSIS (SNA)

In this analysis, a directed graph of users (G) is constructed to form a user network [7],[33]. Each user in the graph network is assigned topology measures based on their influence and importance in the entire constructed network. These topology measures indicate how central the user is (out-degree centrality $C_d^-(u)$), how effectively the user can control information diffusion (out-closeness centrality $C_c^-(u)$),), and how fast the user can spread information to the entire network (between-ness centrality $C_b(u)$). Besides being inspired by [33], the reason for choosing the centrality measures among the other well-known measures was the ability of the centrality measures to reflect the degree to which a user can support his published opinion to gain interest. Being central, close, and act as a bridge to different users' clusters can benefit opinions to spread in more diffusion paths with a higher diffusion rate [7],[33]. The UCINET software tool (University of California at Irvine Network) was used to perform SNA. It can automatically draw a direct graph and calculate the desired topology measures (see (1)–(4)). The input to UCINET is the follower/ following list of users. The output is the constructed direct graph (G) and each user topology measures (i.e., $C_d^-(u)$, $C_c^-(u)$), $C_b(u)$).

$$G = (V, E) \tag{1}$$

where, $V = \{u_1, u_2, u_3, \ldots, u_n\}$ is the set of nodes (users) in the micro-blog networks, and $E = \{e_1, e_2, e_3, \ldots, e_m\}$ is the set of edges (relationships between users) in the microblog network. If u_i is u_j's follower, then there exists a directed edge $e_{u_i u_j} \in E$ from u_j to u_i.

$$C_d^-(u) = |\overrightarrow{N_i}| \tag{2}$$

where, $N = \{ j \epsilon V: (i, j) \epsilon E\}$ is a set of neighbors of node i which i is connected to [33].

$$C_c^-(u) = \sum_j \frac{1}{d_{ij}} \tag{3}$$

where, d_{ij} is the harmonic mean of the length of the shortest paths between the i-th node and the rest of the network [33].

$$C_b(u) = \sum_{jk} \sigma_{jk}(i) \tag{4}$$

where, $\sigma_{jk}(i)$ is the number of shortest paths that connect j and k and contain i.

2) USER RANKING

This process is the final stage of the second phase. It is concerned with obtaining weights for users that describe how influent is the user in the social network constructed using the collected data in Phase 1. User weights are obtained using an artificial neural network (ANN). An ANN is an ML technique that can deal with complicated behaviors that have no mathematical basis. Social behavior on social media follows no model; a tweet of low influencer user in the network can cause him to be

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3183108

highly ranked owing to his valuable content that gained influence. Considering users and their tweets influence to classify users according to their combined influence level required ANN for being suitable in representing such complicated social behavior [34].

Based on [7], a single-layer feedforward ANN was constructed and trained on more than ten thousand samples. The trained data were well-prepared for this mission based on similar social media cases and top most followed users on Twitter. The input to this ANN is the output of the SNA processed in this phase (i.e., user topology measures) that represents the user's influence measures along with the text features obtained from the first phase (i.e., retweets, replies, and favorites counts) that represent the tweet's influence measures. The output of the constructed ANN is a user rank out of "100". Based on the given ranks, weights, ranged from [0 , 1], were calculated for each user. A weight of 0 is given for the lowest influencers, and a weight of 1 is given for the highest influencers. The resulted weights are integrated with the polarity score of the texts obtained for the upcoming phase.

C. A REGULAR OPINION MINING PROCESS

In this phase, the usual OM process documented in the literature is performed. As already mentioned in the literature, most adopted techniques for obtaining a polarity score from a text are lexicon-based or/and ML. In this phase, lexicon-based OM was performed, and the obtained polarity scores were processed using the ML technique in the next phase. Thus, a hybrid technique for classifying text polarity was applied. SentiWordNet, TextBlob, etc. are examples of lexicon-based techniques used in the literature [9],[12]. TextBlob is one efficient lexicon technique implemented in python. TextBlob was chosen for its highest provided performance among other lexicons when being hybridized with ML classifiers [35]. Moreover, TextBlob provides a subjective score that indicates the degree to which the analyzed text includes opinion or factual information [35]. In addition to text polarity classification, some important natural language processing (NLP) is applicable when dealing with texts [12].

Two processes are performed in this phase and are implemented using TextBlob: Text Preprocessing and Lexicon Classification. In text preprocessing, texts are cleared (e.g., remove stop words, URLs, numbers, usernames, punctuation) and edited (e.g., replace slangs and emoticons) to represent the main body of the text. In lexicon classification, the preprocessed text is assigned both a polarity score and a subjective score ranging from [-1, 1] and [0, 1], respectively. A subjective score of "0" indicates an objective text, whereas "1" means a highly subjective text [35]. The output from this phase and output from Phase 2 are the inputs for the last phase.

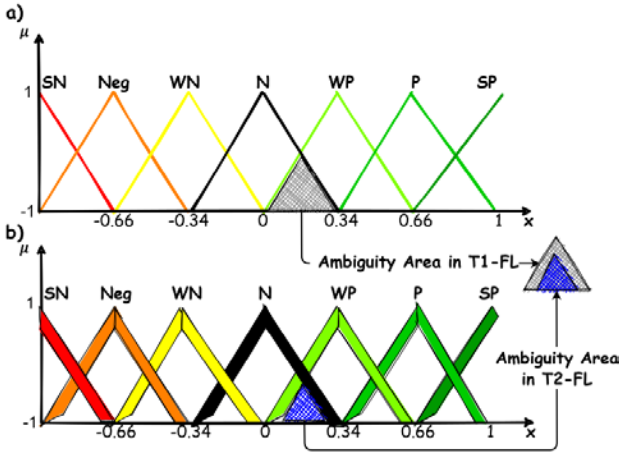


FIGURE 2. **Ambiguity areas: (a) T1-FL MF (b) T2-FL MF.**

D. OPINION CLASSIFICATION

The main purpose of this phase is to integrate the polarity scores of texts with their users influence to obtain a polarity score that is as close as possible to how people really perceived the texts. Because of the importance of such a step in our research, we attempted to obtain the best result from this phase. The best results are those with better classification accuracy, so they can improve the OM process in general and the decision quality in particular. Uncertainties classification is a real problem when dealing with social media data. Consequently, the core of this work focuses on choosing a better classifier that can handle environmental uncertainty and improve classification accuracy. A comparative analysis was conducted in this phase to choose a better classifier for the uncertainties from the chosen ones for comparison. The uncertainty classifiers used for comparison were FL-T1, FL-T2, and NL. The basic inference process of each classifier is shown in Fig. 3, where the main differences between the three classifiers are represented. Classification efficiency was tested for ambiguity. Ambiguity is the case when one instance has two classification possibilities (e.g., lies in an intersection area between two classes).

1) TYPE-1 FUZZY LOGIC CLASSIFICATION

The fuzzification, rule evaluation, and defuzzification steps were implemented based on the work done in [7]. For fuzzification, the chosen MFs for the inputs were inspired by [7],[36]. The user weight MF has a triangular shape ranging from [0 to 1] with three variables (i.e., low, moderate, and high influence). The same is true for the second input with a range of [-1 1] with seven variables (e.g., strong positive, positive, and weak positive). The 21 rules designed in [7] were used for rule evaluation. The designed rules, whether for T1-FL, T2-FL, or NL, played an important role in reflecting human-like intelligence of how humans could perceive texts of users of different influence levels. These designed rules, for example, should consider a lower polarity score, than estimated by machine, for texts that belong to low influencers. One example of rules is, IF user weight is ‘Low Influence’ and text polarity is ‘strong positive’, THEN Final tweet’s polarity is ‘positive’. The output from the FL-T1 system is a final polarity score of text ranging from [-1 to 1] with seven polarity classes. Fine-grained polarity classes can handle text vagueness.

2) TYPE-2 FUZZY LOGIC CLASSIFICATION

The main difference between T1-FL and T2-FL is that a bounded region exists in the membership functions of T2-FL, so called Footprint of Uncertainty (FOU). This FOU can handle ambiguity to an acceptable degree owing to its ability to minimize the intersection between two classification classes [15]. T2-FL in the implementation is the same as T1-FL; thus, the same steps and ranges are considered. The output of this T2- FL system was a polarity score based on the same seven classes in the T1-FL system.

3) NEUTROSOPHIC LOGIC CLASSIFICATION

T1-FL provides the degree membership of an object in a specific set, and T2-FL does the same with more precision (owing to the presence of FOU). NL believes in the existence of a gray area between the white (degree of membership, i.e., truth) and the black (degree of non-membership, i.e., false), so called indeterminacy membership [12]. The main problem that can cause ambiguity in the classification is the presence of objects in the intersection area between two adjacent classes. In this study, this problem is enlarged owing to the chosen design of the MFs whose variables (i.e., classes) are fully overlapped. As illustrated in Fig. 2, all classified objects are in the overlapping region between any two adjacent classes. NL can provide realistic results when applied to the overlapping region of classification. It can give an object a degree of truth, falsity, and indeterminacy belonging to a certain class rather than the adjacent one [28],[32]. Accordingly, based on the basic features of each uncertainty classifier summarized in Tables II, III, and IV, is expected to provide the best detailed classification results for our research case, with a minor presentation of ambiguity in classification.

In the implementation process, NL differs from FL with its two types in the concept of having not only one MF per parameter but three MF per parameter. For example, in FL, one MF is constructed for user influence (UI) and the other for the polarity score. In the case of NL, the user’s influence has three MFs instead of only one UI_{True}, $UI_{Indeterminate}$, and UI_{False}. Thus, three outputs are expected from the neutrosophic inference system (NIS) compares to only one from the fuzzy inference system (FIS) [32]. The FIS results show ambiguity if they lie in the MF overlap region as shown in Fig. 2. For NIS, MFs are designed with no overlap areas and results show ambiguity if they have nonzero falsity and indeterminacy values [32]. Based on [32] and what has been mentioned in Table II regarding how FL and NL can mathematically and graphically represent numbers, UI MFs for NL can be designed as shown in Fig. 4, where, for example, the membership value to the LI class (MV_{LI}) can be mathematically described as follows:

$$MV_{LI} = \begin{cases} \text{higher to LI}, \ 0 \le x < a \\ \text{decreased to LI}, \ a \le x < (a+b)/2 \\ \text{equal to LI and MI}, \ x = (a+b)/2 \\ \text{increased to MI}, \ (a+b)/2 \le x < b \\ \text{higher to MI}, \ b \le x < c \end{cases} \quad (5)$$

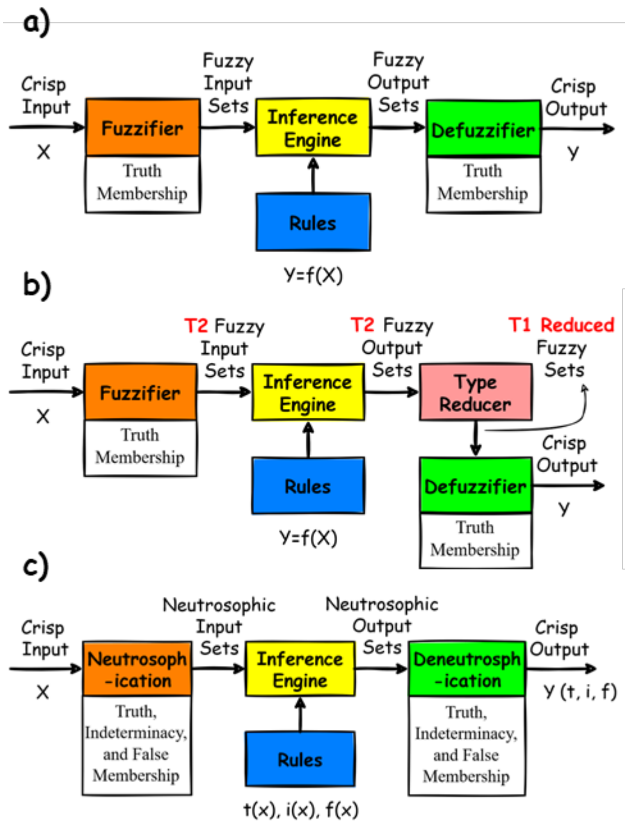


FIGURE 3. **Inference systems: (a) T1-FL (b) T2-FL (c) NL.**

The adopted neutrosophic steps are illustrated as follows:

Step1: Neutrosophication

The three MFs, representing truth, indeterminate, and false components for each input, are designed to be used in converting crisp inputs into neutrosophic input sets. The expected inputs for our NIS are as follows:

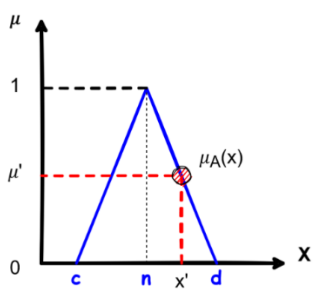
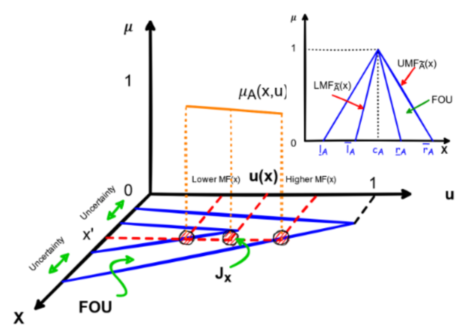
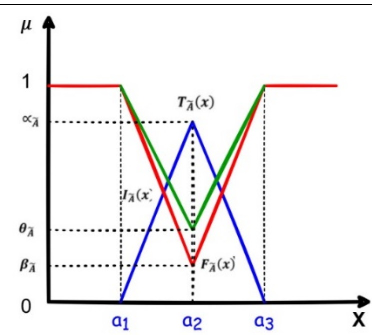
- User Influence (UI): The first input in the neutrosophication process is classified into three levels: low influence level (LI), moderate influence level (MI), and high influence level (HI). The designed MFs shown in Fig. 5 are trapezoidal inspired by the work done in [7], and range from 0 to 1. For example, UI=0.61→ UI_{True}=0.9, $UI_{Indeterminate}$ = 0.1, and UI_{False}= 0.1.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3183108

TABLE II
COMPARATIVES ANALYSIS BETWEEN T1-FL, T2-FL, AND NL [9],[11],[14],[15],[16],[21],[25],[26],[38-49]

Points of Comparison	**Fuzzy Logic**		**Neutrosophic Logic**
	Type – 1	**Type – 2**	
Logic's Type	1 – infinite	1 – infinite	3 – infinite
Components	t, f (dependent) where, t corresponds to the truth value component and f to the falsity value.	t, f (dependent) where, t corresponds to the truth value component and f to the falsity value.	t, i, f (independent) where, t corresponds to the truth value component and i and f for indeterminacy and falsity values.
Components relations	$t + f = 1$	$t + f = 1$	$t + i + f \rightarrow]^{-}0, 1^{+}[$
Output Processing	Called "defuzzification", maps a type-1 fuzzy set into a number. There are many ways for doing this, e.g., COG (Centre Of Gravity): compute the union of the fired-rule output fuzzy sets (the result is another type-1 fuzzy set) and then compute the center of gravity of the membership function for that set.	Conversion from an interval type-2 fuzzy set to a number (usually) requires two steps called "type-reduction"then "defuzzification". Results of rules is distilled then defuzzified. There are as many type-reduction methods as there are type-1 defuzzification methods.	Called deneutrosophication, maps neutrosophic value to crisp value using the three membership functions (T, I, F). Modulation techniques like COG (linear modulation) or mean of maxima (modulation by clipping), can be applied to the neutrosophic sets to generate the crisp true, indeterminacy and falsity values components.
Output	Only one output (Crisp numerical values)	There can be two outputs to an interval type-2 FLS—crisp numerical values and the type-reduced set (is where an interval type-2 fuzzy set is reduced to an interval-valued type-1 fuzzy set)	Final output generated is represented as single crisp value which has triplet format(t, i, f)
Uncertainty Data Type	Vaguness	Vaguness and ambiguty	Vagueness, ambiguity, inconsistent, and imprecision
Core Advantage	The first to manipulate human thinking and deal with vaguness	The presence of FOU to deal with uncertainties	The first to handle indeterminacy
Set A Representation	A fuzzy set A in a universe of discourse X is defined as the following set of pairs: $A = \{(x_i, \mu_A(x_i)) \mid x_i \in X\}$ where, $\mu_A: X \rightarrow [0,1]$is a mapping called the degree of membership function of the fuzzy set A, $\mu_A(x_i) \in [0,1]$is called the membership value of $x \in X$ in the fuzzy set A.	A type-2 fuzzy set $\check{A}$defined in the universe of discourse$X$, $\check{A} = \{((x,u), \mu_{\check{A}}(x,u)) \forall x \in X, \forall u \in J_x \subseteq [0,1]\}$ where, $\mu_{\check{A}}: X \times J_x \rightarrow [0,1]: 0 \leq \mu_{\check{A}}(x,u) \leq 1, x \in A$, $J_x \subset [0,1]$ is an interval for each $x \in X$.	A set A in X is characterized by three functions $t_A(x)$ truth, $i_A(x)$ indeterminacy, and $f_A(x)$falsity memberships, $A = \{< x: t_A(x), i_A(x), f_A(x) >, x \in X\}$ where, $t_A, i_A, f_A: X \rightarrow [0,1]: 0 \leq t_A + i_A + f_A \leq 3, x \in A$

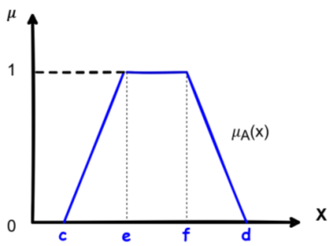
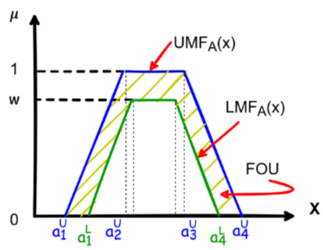
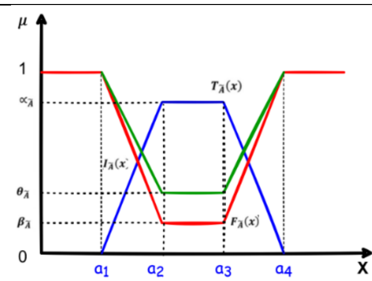
TABLE III
TRIANGULAR NUMBERS REPRESENTATIONS BETWEEN T1-FL, T2-FL, AND NL [9],[11],[14],[15],[16],[21],[25],[26],[38-49]

Points of Comparison	**Fuzzy Logic**		**Neutrosophic Logic**
	Type – 1	**Type – 2**	
Triangular Numbers Graphical Representation	 Graphical representation of triangular fuzzy number $(\boldsymbol{c}, \boldsymbol{n}, \boldsymbol{d};\ \mathbf{1})$	 Upper and lower membership functions are represented by a triangular fuzzy number $\tilde{A} = \langle[\underline{l_A}\ \overline{l_A}], c_A, [\underline{r_A}\ \overline{r_A}]\rangle$	 $\check{A} = \langle(a_1, a_2, a_3);\ \alpha_{\check{A}}, \theta_{\check{A}}, \beta_{\check{A}}\rangle$, where, $a_1 \leq a_2 \leq a_3$ and $\alpha_{\check{A}}, \theta_{\check{A}}, \beta_{\check{A}} \in [0,1]$ denote the maximum truth membership degree, minimum indeterminacy-membership degree, and minimum falsity-membership degree
Triangular Numbers Mathematical Representation	$\mu_A(X) = \begin{cases} 0, x \leq c \\ \dfrac{x-c}{n-c}, c < x \leq n \\ \dfrac{d-x}{d-n}, n < x < d \\ 0, x \geq d \end{cases}$ The membership function for a fuzzy set A on the universe X is represented as $\mu_A : X \rightarrow [0,1]$.	$LMF_{\tilde{A}}(x) = \begin{cases} \dfrac{x - \overline{l_A}}{c_A - \overline{l_A}}, & \overline{l_A} \leq x < c_A \\ 1, & x = c_A \\ \dfrac{x - \underline{r_A}}{c_A - \underline{r_A}}, & c_A \leq x < \underline{r_A} \\ 0, & otherwise \end{cases}$	$T_{\check{A}}(x) = \begin{cases} \alpha_{\check{A}}\left(\dfrac{x - a_1}{a_2 - a_1}\right) & (a_1 \leq x \leq a_2) \\ \alpha_{\check{A}} & (x = a_2) \\ \alpha_{\check{A}}\left(\dfrac{a_3 - x}{a_3 - a_2}\right) & (a_2 \leq x \leq a_3) \\ 0 & otherwise \end{cases}$

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3183108

		$UMF_{\tilde{A}}(x)=\begin{cases}\frac{x-\underline{l_A}}{c_A-\underline{l_A}}, & \underline{l_A}\le x< c_A\\ 1, & x=c_A\\ \frac{x-\overline{r_A}}{c_A-\overline{r_A}}, & c_A\le x<\overline{r_A}\\ 0, & otherwise\end{cases}$ where, $\underline{l_A},\overline{l_A},c_A,\underline{r_A},\overline{r_A}$ are the reference points on the triangular interval type-2 fuzzy set that satisfying the condition , $0\le\underline{l_A}\le\overline{l_A}\le c_A\le\underline{r_A}\le\overline{r_A}\le 1$	$I_{\breve{A}}(x)=\begin{cases}\frac{(a_2-x+\theta_{\breve{A}}(x-a_1))}{a_2-a_1} & (a_1\le x\le a_2)\\ \theta_{\breve{A}} & (x=a_2)\\ \frac{(x-a_2+\theta_{\breve{A}}(a_3-x))}{a_3-a_2} & (a_2\le x\le a_3)\\ 1 & otherwise\end{cases}$ $F_{\breve{A}}(x)=\begin{cases}\frac{(a_2-x+\beta_{\breve{A}}(x-a_1))}{a_2-a_1} & (a_1\le x\le a_2)\\ \beta_{\breve{A}} & (x=a_2)\\ \frac{(x-a_2+\beta_{\breve{A}}(a_3-x))}{a_3-a_2} & (a_2\le x\le a_3)\\ 1 & otherwise\end{cases}$ where, $\breve{A}$is a special neutrosophic set on the real line set R, $\alpha_{\breve{A}},\theta_{\breve{A}},\beta_{\breve{A}}$ are used to calculate $T_{\breve{A}}: R\in[0,\alpha_{\breve{A}}]$, $I_{\breve{A}}: R\in[\theta_{\breve{A}},1]$, $F_{\breve{A}}$

TABLE IV
TRAPEZOIDAL NUMBERS REPRESENTATIONS BETWEEN T1-FL, T2-FL, AND NL [9],[11],[14],[15],[16],[21],[25],[26],[38-49]

Points of Comparison	Fuzzy Logic		Neutrosophic Logic
	Type – 1	**Type – 2**	
Trapezoidal Numbers Graphical Representation	 Graphical representation of trapezoidal fuzzy number $(c,e,f,d;\ 1)$	 Upper and lower membership functions are represented by a trapezoidal fuzzy number $\tilde{A}=\langle(a_1^U,a_2^U,a_3^U,a_4^U;1),(a_1^L,a_2^L,a_3^L,a_4^L;w)\rangle$	 $\breve{A}=\langle(a_1,a_2,a_3,a_4);\ \alpha_{\breve{A}},\theta_{\breve{A}},\beta_{\breve{A}}\rangle$, where, $a_1\le a_2\le a_3\le a_4$and $\alpha_{\breve{A}},\theta_{\breve{A}},\beta_{\breve{A}}\in[0,1]$ denote the maximum truth membership degree, minimum indeterminacy-membership degree, and minimum falsity-membership degree
Trapezoidal Numbers Mathematical Representation	The membership function for a fuzzy set A on the universe X is represented as $\mu_A: X\rightarrow[0,1]$. $\mu_A(X)=\begin{cases}\frac{x-c}{e-c}, & c\le x\le e\\ 1, & e\le x\le f\\ \frac{d-x}{d-f}, & f\le x\le d\\ 0, & otherwise\end{cases}$ If $e=f$ then the trapezoidal T1Fuzzy number is converted into triangular one.	$LMF_{\tilde{A}}(x)=\begin{cases}w\frac{x-a_1^L}{a_2^L-a_1^L}, & a_1^L\le x\le a_2^L\\ w, & a_2^L\le x\le a_3^L\\ w\frac{a_4^L-x}{a_4^L-a_3^L}, & a_3^L\le x\le a_4^L\\ 0, & otherwise\end{cases}$ $UMF_{\tilde{A}}(x)=\begin{cases}\frac{x-a_1^U}{a_2^U-a_1^U}, & a_1^U\le x\le a_2^U\\ 1, & a_2^U\le x\le a_3^U\\ \frac{a_4^U-x}{a_4^U-a_3^U}, & a_3^U\le x\le a_4^U\\ 0, & otherwise\end{cases}$ where, $a_1^U,a_2^U,a_3^U,a_4^U,a_1^L,a_2^L,a_3^L,a_4^L$ are the reference points on the trapezoidal interval type-2 fuzzy set.	$T_{\breve{A}}(x)=\begin{cases}\alpha_{\breve{A}}\left(\frac{x-a_1}{a_2-a_1}\right), & a_1\le x\le a_2\\ \alpha_{\breve{A}}, & a_2\le x\le a_3\\ \alpha_{\breve{A}}\left(\frac{a_4-x}{a_4-a_3}\right), & a_3\le x\le a_4\\ 0, & otherwise\end{cases}$ $I_{\breve{A}}(x)=\begin{cases}\frac{(a_2-x+\theta_{\breve{A}}(x-a_1))}{a_2-a_1}, & a_1\le x\le a_2\\ \theta_{\breve{A}}, & a_2\le x\le a_3\\ \frac{(x-a_3+\theta_{\breve{A}}(a_4-x))}{a_4-a_3}, & a_3\le x\le a_4\\ 1, & otherwise\end{cases}$ $F_{\breve{A}}(x)=\begin{cases}\frac{(a_2-x+\beta_{\breve{A}}(x-a_1))}{a_2-a_1}, & a_1\le x\le a_2\\ \beta_{\breve{A}}, & a_2\le x\le a_3\\ \frac{(x-a_3+\beta_{\breve{A}}(a_4-x))}{a_4-a_3}, & a_3\le x\le a_4\\ 1, & otherwise\end{cases}$ where, $\breve{A}$is a special neutrosophic set on the real line set R, $\alpha_{\breve{A}},\theta_{\breve{A}},\beta_{\breve{A}}$ are used to calculate $T_{\breve{A}}: R\in[0,\alpha_{\breve{A}}]$, $I_{\breve{A}}: R\in[\theta_{\breve{A}},1]$

- Polarity Score (PS): The second input and is classified into seven polarity classes: strong positive (SP), positive (P), weak positive (WP), neutral (N), weak negative (WN), negative (Neg), and strong negative (SN). The designed MFs for the neutrosophication process of this input, as shown in Fig. 5, are in the range of 0–1 triangular shaped inspired by [7],[36],[39]. The output neutrosophic sets of these two inputs were then evaluated and aggregated in the next step.

STEP 2: Rule Evaluation (Inference Engine)

In this step, rules are designed and evaluated for strength using the neutrosophic inputs output from the previous step. This evaluation results in another neutrosophic set, namely, neutrosophic output sets. Rules are designed for all truth, indeterminacy, and falsity input-output MFs. The antecedences' measurements of the designed rules (i.e., UI and PS neutrosophic inputs) are used to fire the satisfied rules and then aggregate the rules' strength value using the chosen operator AND. A sample of the design rules is presented in Table V. For example, for the obtained neutrosophic inputs, Rule #15 and Rule #27 are fired for the true component of indeterminate. The neutrosophic outputs are: 0.76 for WP-t, and 0.1 & 0.1 for P-WP-i.

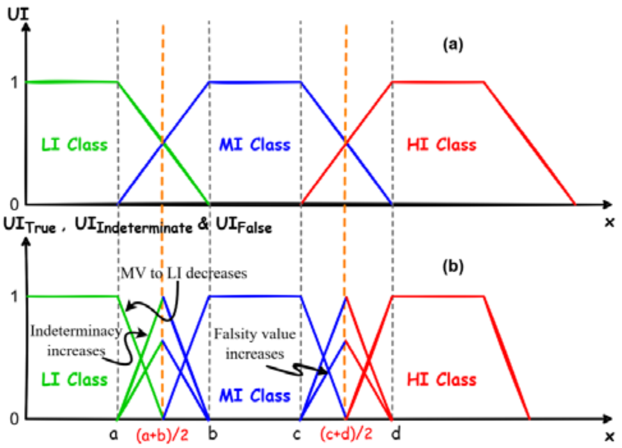


FIGURE 4. NL MF design for UI based on FL MF: (a) FL MF (b) NL true, indeterminacy, and falsity components.

STEP 3: De-neutrosophication

The output of this step is the proposed model output (PS). In this step, three MFs are designed for the output, and a modulation technique is chosen (i.e., COG in our case) to convert the neutrosophic output sets into crisp outputs. The final output is represented in a triple format$\mathrm{PS}(\mathrm{t},\mathrm{i},\mathrm{f})$, donating true, indeterminate, and false component values. Ambiguity can be generated in the results. In dealing with ambiguity, NL can set a confidence value for the truth component so that the falsity and indeterminacy components are not significant if this truth value exceeds the predefined confidence value. As a result of that, the final output for ambiguity results is the truth component. Ambiguity is significant if the truth component is less than a predefined confidence value.

TABLE V
SAMPLE OF THE DESIGNED NL RULES

#	Rules
1	**IF** UI is 'LI-t' and PS is 'SN-t', **THEN** PS is 'NEG-t'
2	**IF** UI is 'LI-t' and PS is 'SP-t', **THEN** PS is 'P-t'
3	**IF** UI is 'LI-t' and PS is 'NEG', **THEN** PS is 'WN-t'
4	**IF** UI is 'LI-t' and PS is 'P-t', **THEN** PS is 'WP-t'
5	**IF** UI is 'LI-t' and PS is 'WN-t', **THEN** PS is 'N-t'
6	**IF** UI is 'LI-t' and PS is 'WP-t', **THEN** PS is 'N-t'
7	**IF** UI is 'LI-t' and PS is 'N-t', **THEN** PS is 'N-t'
8	**IF** UI is 'MI-t' and PS is 'N-t', **THEN** PS is 'N-t'
9	**IF** UI is 'HI-t' and PS is 'N-t', **THEN** PS is 'N-t'
10	**IF** UI is 'MI-t' and PS is 'SN-t', **THEN** PS is 'SN-t'
11	**IF** UI is 'MI-t' and PS is 'NEG-t', **THEN** PS is 'NEG-t'
12	**IF** UI is 'MI-t' and PS is 'WN-t', **THEN** PS is 'WN-t'
13	**IF** UI is 'MI-t' and PS is 'WP-t', **THEN** PS is 'WP-t'
14	**IF** UI is 'MI-t' and PS is 'P-t', **THEN** PS is 'P-t'
15	**IF** UI is 'MI-t' and PS is 'SP-t', **THEN** PS is 'SP-t'
16	**IF** UI is 'HI-t' and PS is 'SN-t', **THEN** PS is 'SN-t'
17	**IF** UI is 'HI-t' and PS is 'SP-t', **THEN** PS is 'SP-t'
18	**IF** UI is 'HI-t' and PS is 'NEG-t', **THEN** PS is 'SN-t'
19	**IF** UI is 'HI-t' and PS is 'P-t', **THEN** PS is 'SP-t'
20	**IF** UI is 'HI-t' and PS is 'WN-t', **THEN** PS is 'NEG-t'
21	**IF** UI is 'HI-t' and PS is 'WP-t', **THEN** PS is 'P-t'
22	**IF** UI is 'MI-HI-i' and PS is 'SN-NEG-i', **THEN** PS is 'SN-NEG-i'
23	**IF** UI is 'MI-HI-i' and PS is 'SP-P-i', **THEN** PS is 'SP-P-i'
24	**IF** UI is 'MI-HI-i' and PS is 'NEG-WN-i', **THEN** PS is 'SN-Neg-i'
25	**IF** UI is 'MI-HI-i' and PS is 'P-WP-i', **THEN** PS is 'SP-P-i'
26	**IF** UI is 'MI-HI-i' and PS is 'N-WN-i', **THEN** PS is 'NEG-WN-i'
27	**IF** UI is 'MI-HI-i' and PS is 'N-WP-i', **THEN** PS is 'P-WP-i'
28	**IF** UI is 'LI-MI-i' and PS is 'NEG-WN-i', **THEN** PS is 'N-WN-i'
29	**IF** UI is 'LI-MI-i' and PS is 'P-WP-i', **THEN** PS is 'N-WP-i'

Accordingly, in this case, the indeterminacy and falsity components should be considered, and the final output of the NIS is manually determined. In line with the chosen confidence level value for the truth component in [32], this study considered the same value as follows:

$$\mathrm{i/f} = \begin{cases} \text{significant}, & \mathrm{t} < 50\% \\ \text{insignificant}, & \mathrm{t} \geq 50\% \end{cases} \tag{6}$$

where, t, i, f are the truth, indeterminacy, and falsity components, respectively. A full trace of the example is shown in Fig. 5. The final output is $PS(0.34, 0.53, 0.53)$ which indicates ambiguity. Based on (6), the generated ambiguity is insignificant because the true value is greater than the assigned confidence value. Accordingly, the final polarity class of this example was P.

E. OPINION POLARITY CLASS

The output from this OM proposed model is a polarity class that reflects the real perceived opinion polarity based on users' influence in their network, considering classification ambiguity problem. The polarity score obtained from the 5th phase is given a class label out of the polarity classes based on each class polarity range. Table VI represents the seven polarity ranges and their corresponding classes.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3183108

TABLE VI
SEVEN POLARITY LEVELS [37]

Score	Polarity Class
$n > 0.75$	Strong Positive
$0.25 < n \leq 0.75$	Positive
$0 < n \leq 0.25$	Weak Positive
0	Neutral
$-0.25 \leq n < 0$	Weak Negative
$-0.75 \leq n < -0.25$	Negative
$n < -0.75$	Strong Negative

V. EXPERIMENTAL RESULTS

To obtain performance-related insights into the proposed model, a desktop program was implemented using Python and MATLAB libraries. The desired dataset for application is the one used in [7], consists of tweets and their associated counting of various reactions along with their authors' social relationship lists.

TABLE VII
SAMPLE OF COLLECTED DATA PER USER

<table>
<tr><th>Author</th><th colspan="2">Tweet</th><th>Created at</th><th>Retweets Count</th><th>Replies Count</th><th>Favorite Count</th></tr>
<tr><td rowspan="5">#1</td><td colspan="2">His face after he accidentally scored that goal just broke my heart. The fact that he is going to get back today while blaming himself for it is just too hard for anyone to handle, I hope people go easy on him and remember that mistakes can happen. #WorldCup18</td><td>June 20</td><td>2</td><td>3</td><td>5</td></tr>
<tr><td>Following List</td><td colspan="2">Followers List</td><td colspan="3">Time</td></tr>
<tr><td>#1</td><td colspan="2">2</td><td colspan="3" rowspan="3">June 20</td></tr>
<tr><td>#1</td><td colspan="2">15</td></tr>
<tr><td>#134</td><td colspan="2">#1</td></tr>
</table>

This dataset was collected from Twitter, based on the provided trend-list at the time of data collection. Thus, the proposed model was presented in the context of the World's Football Cup, but can be generalized to any other emergency or nonemergency opinions. Table VII shows a sample example of what this dataset may contain. Users were given numerical symbols to avoid privacy violations. Based on what is reported in Table VII, user #1 has two followers (#2 and #15) and the following one (#134). In addition to that, to assess applicability of the proposed model, another two online available datasets were utilized. The first one was the "farm loan" dataset used in [12], the other one is publicly available at [50] and contained more than 200,000 tweets about COVID-19 vaccines. Authors in [12] followed a similar analysis to the proposed model; they performed SA on the "farm loan" dataset using both TextBlob and fuzzy technique then refined the output as neutrosophic sets. Although the "farm loan" dataset possessed no information about the opinion holders of the provided texts, applying this dataset on the proposed model can give some insights on the effectiveness of the adopted hybrid classification approach in the proposed model.

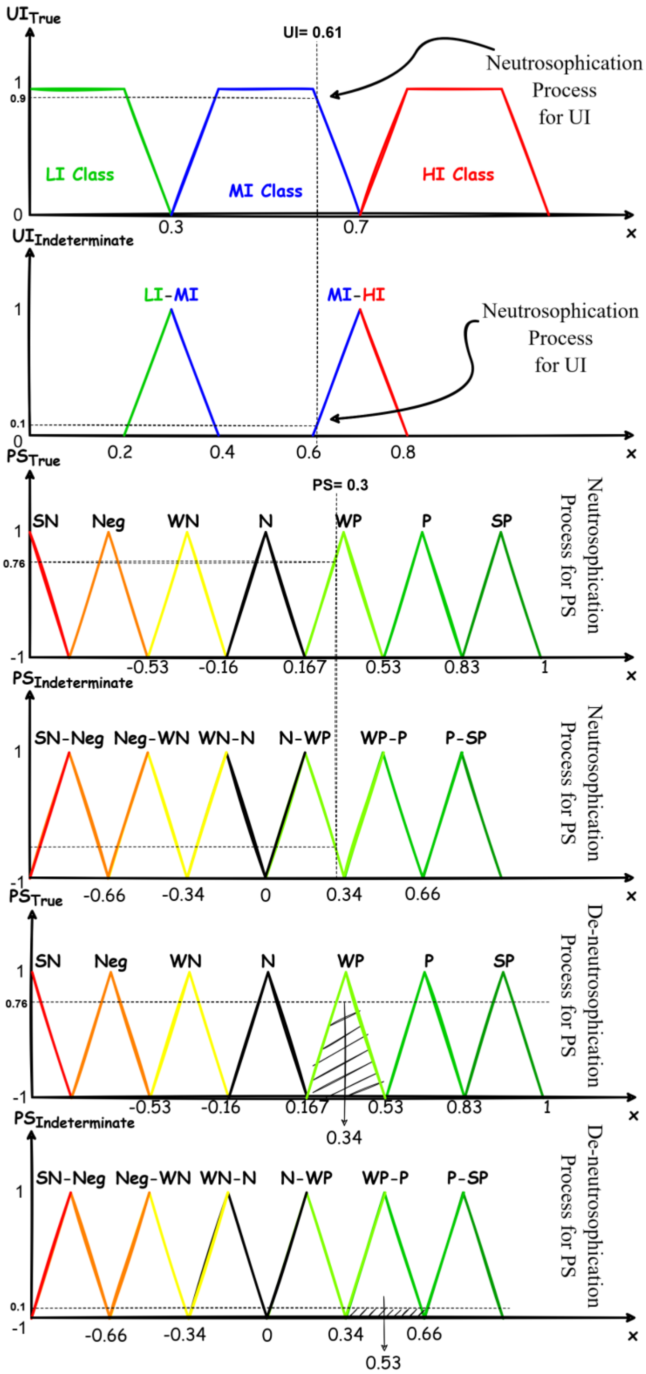


FIGURE 5. Neutrosophic inference system steps.

Whereas, the enthusiastic concern to deal with a larger dataset and to manage an emergency event of massive interest and heterogeneous responses on social media, are the reasons for choosing the second dataset. Moreover, the vaccination dataset contained some important data regarding perspectivism (e.g. the reactions on tweets counts, followers count). Thus, considering perspectivism can be achieved without measuring both the closeness and the between-ness centrality measures because the required data to perform the task are missed.

Machine usually deals with the texts as if their writers possess the same influence level (i.e. moderate influence level). The first set of experiments was conducted to investigate the significance of considering perspectivism in the OM process. To achieve the experimental objective, the constructed ANN was used to weight the users of both the customized and the vaccination datasets based on their importance in their social networks. The results reported in Table Ⅷ shows that not all the users were moderate influencers; they could be low influencers with the highest percentage. Thus, their texts should not be considered as valuable as the texts of the other types of influencers. To highlight how exactly perspectivism can change the final polarity of texts, the second set of experiments were run. Table Ⅸ and Table Ⅹ summarize the statistical representation of the polarity classes using the same datasets. The results showed how perspectivism could cause a makeable change in the polarity scores and in the decision-making process in return.

The next set of experiments was performed to verify the effectiveness of the proposed compared to the model proposed in [12]. The "farm loan" dataset was utilized for this experimental purpose. For this dataset to be implemented on the proposed model, the influence levels of users were considered "moderate" to assure a fair comparison with [12]. It was obvious that the two approaches shared close polarity distributions proving the advantage of the proposed model as a hybrid classifier. To highlight the differences in the polarity distribution that could be resulted when the work done in [12] consider the influence factor of the users, estimated influence weights were given to the users in the datasets. The weights distribution process was correlated with the results obtained in Table Ⅷ. The results showed an increase in the neutral opinions and a corresponding decrease in both the positive and the negative opinions. Some positive and negative opinions had a neutral effect on readers. Accordingly, it was concluded that ignoring perspectivism can doubt the effectiveness of the OM-based decisions. The related results were reported in Table XI.

TABLE Ⅷ
CLASSIFICATION OF USERS BASED ON ANN

Influence Level	Using the customized dataset	Using the vaccination dataset
Low (LI)	72.69%	54.46%
Moderate (MI)	16.85%	26.39%
High (HI)	10.46%	19.15%

TABLE Ⅸ
POLARITY CLASSES DISTRIBUTIONS FOR THE VACCINATION DATASET

Polarity Class	Before applying the proposed model (TextBlob)	After applying the proposed model (TextBlob+ NL+UI)
SP	1.74%	1.16%
P	16.97%	18.35%
WP	23.50%	17.19%
N	48.50%	60.70%
WN	5.66%	0.58%
Neg	3.48%	2.39%
SN	0.15%	0.15%

TABLE Ⅹ
POLARITY CLASSES DISTRIBUTIONS FOR THE CUSTOMIZED DATASET

Polarity Class	Before applying the proposed model (TextBlob)	After applying the proposed model (TextBlob+ NL+UI)
SP	13.89%	3.70%
P	21.48%	22.22%
WP	16.39%	14.91%
N	42.22%	57.04%
WN	2.96%	0.83%
Neg	2.96%	0.93%
SN	0.10%	0.37%

TABLE XI
POLARITY CLASSES DISTRIBUTIONS FOR THE FARM LOAN DATASET

Polarity Class	The work done in [12]	The proposed model (without users influence)	The proposed model (with the estimated users influence)
Positive	0.419	0.450	0.316
Neutral	0.438	0.438	0.639
Negative	0.143	0.112	0.045

To quantify classification ambiguity, which is the main contribution of the proposed model, experiments were conducted. The proposed model was implemented three times using the customized dataset and three classifiers (T1-FL, T2-FL, and NL). These experiments were designed to test the ability of the three adopted uncertainty classifiers to detect ambiguity in opinion sentiment classification. Fully overlapped output variables of MFs were applied to obtain the maximum ambiguity area that any classifier might face and to show how the three classifiers could deal with such a threat. The ambiguity percentage per each classifier is calculated to investigate the classifiers' performance. According to (5), Tables XII and XIII document the distribution of our data instances in the overlap areas of each polarity class using the TL-FL and T2-FL classifiers. The results showed higher ambiguity in T1 than in T2 images. The reason for this result is that the existence of an FOU in T2-FL by default decreases ambiguity.

Table XIV presents the differences between the three uncertainty classifiers in terms of the existence of significant ambiguity. The results were obtained based on equation (6) for NL and the percentage of data points belonging to the wrong class in Tables XII and XIII for T1-FL and T2-FL, respectively. The results demonstrate the superiority of NL in dealing with ambiguity in opinion classification. The reasons for such results are the nature of NL in giving each instance true, indeterminate, and false values that precisely describe each instance's classification possibilities. Based on a 400 test samples, a test is conducted to verify whether – for a chosen 400 samples – NL provided the best performance in terms of classification accuracy. The results recorded in Table XV confirm the superiority of NL not only in handling ambiguity, but also in providing better classification accuracy than its alternatives. To handle ambiguity, especially the significant cases, human expert should be involved for the final analysis. The proposed NL model decreased the significant ambiguity in dataset by about 10% compared to T1-FL and 5% compared to T2-FL.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3183108

Moreover, this proposed model tried to decrease the human involvement in determine the proper class for the ambiguity cases by an average of 20% for both significant and insignificant cases. Intelligence is generally achieved when machine performs the human task in an accurate way. Thus, the proposed model stepped forward for a human-like intelligence in dealing with ambiguity.

In an in-depth investigation of the classification accuracy case, another experiment was designed on the entire dataset to analyze the classification behavior of each uncertainty classifier. The results summarized in Table XVI highlight the ability of NL to correctly classify classes that have been incorrectly classified by either or only one of its alternative classifiers. Once again, according to the presented results, the proposed model considered a different polarity score than assigned by TextBlob (i.e. machine) based on the designed rules that reflect how people might perceived texts based on their authors' influence on them. The proposed model achieved the goal with a higher accuracy than the other classifiers.

TABLE XII
SUMMARY OF T1-FL LYING IN OVERLAPPING RANGE

Overlap Range	Higher to Correct Class	Equal to 2 Adjacent Classes	Higher to Wrong Class
SN-Neg	0	0	0
Neg-WN	5	0	7
WN-N	7	0	20
N-WP	68	0	101
WP-P	133	0	37
P-SP	55	0	8

TABLE XIII
SUMMARY OF T2-FL LYING IN OVERLAPPING RANGE

Overlap Range	Higher to Correct Class	Equal to 2 Adjacent Classes	Higher to Wrong Class
SN-Neg	0	0	0
Neg-WN	2	0	6
WN-N	5	0	20
N-WP	92	0	67
WP-P	108	0	27
P-SP	16	0	1

TABLE XIV
RESULTANT AMBIGUITY IN EACH UNCERTAINTY CLASSIFIER

Uncertainty Classifier	Ambiguity Recorded	
T1-FL	40.83%	**Sig: 16.02%** InSig: 24.81%
T2-FL	31.85%	**Sig: 11.20%** InSig: 20.65%
NL	16.02%	**Sig: 6.02%** InSig: 10%

TABLE XV
CLASSIFICATION ERROR OF 400 SAMPLE INSTANCES

The uncertainty classifier	Correct classified Instances	Classification Error (%)
FL-T1	400-34	8.5
FL-T2	400-36	9
NL	400-18	4.5

TABLE XVI
CLASSIFICATION DIFFERENCES AMONG CLASSIFIERS

User #	TextBlob	UI	Desired PS	T1-FL PS	T2-FL PS	NL PS
57	0.80	High	SP	0.74	0.69	0.93
62	0.13	Low	N	0.02	0.08	0.00
121	0.90	Mo	SP	0.71	0.71	0.94
128	0.80	Mo	SP	0.61	0.64	0.87
132	0.45	Low	WP	0.26	0.33	0.19
137	0.29	Mo	P	0.25	0.24	0.34
212	0.13	Low	Neu	0.14	0.13	0.00
214	0.77	Mo	SP	0.68	0.68	0.84
231	0.70	High	SP	0.74	0.64	0.93
281	0.38	High	SP	0.64	0.52	0.85
357	0.75	High	SP	0.70	0.68	0.86
602	0.08	Low	Neu	0.22	0.14	0.00
677	-0.40	High	SN	-0.51	-0.42	-0.83
686	0.65	High	SP	0.73	0.62	0.94
749	0.50	Low	WP	0.29	0.34	0.17
67	-0.22	High	Neg	-0.37	-0.25	-0.68
123	0.13	High	P	0.29	0.19	0.53
188	0.13	Low	Neu	0.00	0.05	0.00
204	0.29	Mo	P	0.3	0.24	0.51
591	0.08	Low	Neu	0.00	0.05	0.00
696	-0.21	Low	Neu	0.00	-0.05	0.00
775	0.20	Low	Neu	0.00	0.04	0.00
240	0.33	Low	WP	0.00	0.04	0.17
395	0.29	Low	WP	0.00	0.04	0.17

Note:
NL is the best T1-FL is the worst T2-FL is the worst

VI. CONCLUSIONS AND FUTURE WORK

In this study, we proposed an OM model suitable for social media that can deal with perspectivism and handle ambiguity in opinion classification. In the proposed model, ANN had recourse to classify social media users into three levels of influence owing to its human-like intelligence in dealing with representing behaviors that follow no rule. The advantage of the proposed model was investigated using a customized and benchmark datasets. The results emphasize the significance of considering perspectivism in the OM process. A comparative study was conducted among three uncertainty classifiers with credibility in handling ambiguity in classification – T1-FL, T2-FL, and NL. NL was adopted for the first time to integrate polarity scores of texts with their authors' influence level and was tested for significance in dealing with classification ambiguity in comparison with its classical alternatives T1 and T2 FL. The projected output of this model is a text's polarity score of seven classification classes that represent the audience's trust of authors with as high a degree of certainty as possible. Fully overlapped MF variables of the output polarity score were applied to amplify the ambiguity for a worthwhile test. The results show the ability of T2-FL, with its developed features, to decrease ambiguity compared to T1-FL. At the same time, NL reported the highest attitude in dealing with classification ambiguity, giving it primacy in such an application. Including images of texts, considering more factors that affect user influence, building a new benchmark dataset for such a type of research, and more multiple comparisons with state-of-art works will be our future work.

REFERENCES

[1] Ch. Messaoudi, Z. Guessoum, and L. Romdhane, "Opinion mining in online social media: a survey", *Social Network Analysis and Mining*, Vol. 12, No. 1, pp. 1-18, 2022.

[2] Y. Chen, Y. Li, Z. Wang, A. Quintero, Ch. Yang, and W. Ji, "Rapid perception of public opinion in emergency events through

social media", *Natural Hazards Review*, vol. 23, no. 2, 4021066, pp.1-12, 2022.

[3] R. Nandal, A. Chawla, and K. Joshi, "Opinion mining and analyzing real-time tweets using rapid miner", In *Proceedings of the 3rd International Conference on Sustainable Computing*, Singapore, pp. 213- 221, 2022.

[4] M. Ghiasi, and S. Gharehgozli, "Social sentiment analysis using fuzzy logic: addressing context effect on sentiment analysis by fuzzy lexicon dynamism", In *Proceedings of IEEE International Conference on Application of Information and Communication Technologies*, Azerbaijan, pp. 1- 6, 2019.

[5] N. Tsapatsoulis, and C. Djouvas, "Opinion mining from social media short texts: does collective intelligence beat deep learning?", *Frontiers in Robotics and AI*, vol. 5, pp. 138-151, 2019.

[6] D. Camacho, M. Luzón, and E. Cambria, "New research methods and algorithms in social network analysis", *Future Generation Computer Systems*, vol. 114, pp. 290–293, 2021.

[7] M. Madbouly S. Darwish, and R. Essameldin, "A modified fuzzy sentiment analysis approach based on user ranking suitable for online social networks", *IET Software*, vol. 14, no. 3, pp. 300-307, 2020.

[8] I. P. Jagiripu, P. K. Mishra, A. Saini, and A. Biswal, "Testing the Impact of Uncertainty Reducing Reviews in the Prediction of Cross Domain Social Media Pages Ratings", *Journal of Indian Business Research*, vol. 2021, pp. 1-17, 2021.

[9] F. Smarandache, M. Colhon, Ş. Vlăduţescu, and X. Negrea, "Word-level neutrosophic sentiment similarity", *Applied Soft Computing*, vol. 80, pp. 167–176, 2019.

[10] S. Chao, W. Xiao-Kang, C. Peng-fei, W. Jian-qiang, and L. Lin, "SACPC: a framework based on probabilistic linguistic terms for short text sentiment analysis", *Knowledge-Based Systems*, vol. 194, pp. 105572- 105583, 2020.

[11] N. Radwan, M. Senousy, and A. Riad, "Neutrosophic logic approach for evaluating learning management systems", *Neutrosophic Sets and Systems*, vol. 11, pp. 3- 7, 2016.

[12] I. Kandasamya, W. Vasanthaa, J. Obbinenib, and F. Smarandache, "Sentiment analysis of tweets using refined neutrosophic sets", *Computers in Industry*, vol. 115, pp. 103180-103190, 2020.

[13] J. Serrano-Guerrero, F. Romero, and J. Olivas, "Fuzzy logic applied to opinion mining: a review", *Knowledge-Based Systems*, vol. 222, pp. 107018-107038, 2021.

[14] E. AboElHamd, H. Shamma, M. Saleh, and I. El-Khodary, "Neutrosophic logic theory and applications", *Neutrosophic Sets and Systems*, vol. 41, no. 1, pp. 4- 25, 2021.

[15] A. Naik, and C. Gupta, " Performance comparison of type-1 and type-2 fuzzy logic systems", In P*roceedings of the IEEE International Conference on Signal Processing Computing and Control*, India, pp. 72- 76 , 2017.

[16] S. Basha, A. Abdalla1, and A. Hassanien, "NRCS: neutrosophic rule-based classification system", In *Proceedings of SAI Intelligent Systems Conference*, pp. 627-639, 2016.

[17] D. Mai, L. Ngo, and L. Trinh, "Approach the interval type-2 fuzzy system and PSO technique in land cover classification", In *Proceedings of the Asian Conference on Intelligent Information and Database Systems*, Thailand, pp. 402- 414, 2020.

[18] S. Verma, and J. Kumar, *A Survey on Sentiment Analysis Techniques for Twitter*, Data Mining Approaches for Big Data and Sentiment Analysis in Social Media, IGI Global, pp. 57- 90, 2022.

[19] M. Wankhade, A. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges", A*rtificial Intelligence Review*, vol. 55, pp. 1- 50, 2022.

[20] K. Gutiérrez-Batista, M. Vila, and M. Martin-Baut, "Building a fuzzy sentiment dimension for multidimensional analysis in social networks", *Applied Soft Computing*, vol. 108, pp. 107390- 107401, 2021.

[21] S. Sharma, and V. Kalra, "Fuzzy lexicon-based approach for sentiment analysis of blog and microblog text", In *Proceedings of Research in Intelligent and Computing in Engineering*, Singapore, pp. 705- 714, 2021.

[22] A. Khattak, W. Paracha, M. Asghar, N. Jillani, U. Younis, F. Saddozai, and I. Hameed, "Fine grained sentiment analysis for measuring customer satisfaction using an extended set of fuzzy linguistic hedges", *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 744- 756, 2020.

[23] B. V. Krishna, A. K. Pandey, and A. S. Kumar, "Feature based opinion mining and sentiment analysis using fuzzy logic", *Cognitive Science and Artificial Intelligence*, Springer, Singapore, pp. 79-89, 2018.

[24] J. M. Mendel, "Type-2 fuzzy sets as well as computing with words", *IEEE Computational Intelligence Magazine*, vol.14, no.1, pp. 82- 95, 2019.

[25] H. Abbasimehr, E. Nourani, and M. Shabani, "A hybrid framework for ranking reviewers based on interval type-2 fuzzy AHP and VIKOR", *International Journal of Intelligent Engineering Informatics*, vol. 8, no. 2, pp. 95- 116, 2020.

[26] J. Bi, Y. Liu, and Z. Fan, "Representing Sentiment analysis results of online reviews using interval type-2 fuzzy numbers and its application to product ranking", *Information Sciences*, vol. 504, pp. 293–307, 2019.

[27] F. Ali, E. Kim, and Y. Kim, "Type-2 fuzzy ontology-based opinion mining and information extraction: a proposal to automate the hotel reservation system", *Applied Intelligence*, vol. 42, no. 3, pp. 481- 500, 2015.

[28] K. Bhutani, and S. Aggarwal, "A novel approach for data classification using neutrosophic entropy", In *Proceedings of the International Conference on System and Architecture*, Singapore, pp. 305- 317, 2018.

[29] F. Smarandache, and S. Vladutescu, *Neutrosophic emergencies and incidences*, LAP LAMBERT Academic Publishing, 2013.

[30] I. Awajan, M. Mohamad, and A. Al-Quran, "sentiment analysis technique and neutrosophic set theory for mining and ranking big data from online reviews", *IEEE Access*, vol. 9, pp. 47338-47353, 2021.

[31] I. Kandasamy, W. Vasantha, N. Mathur, M. Bisht, and F. Smarandache, "Sentiment analysis of the #MeToo movement using neutrosophy: application of single-valued neutrosophic sets", *Optimization Theory Based on Neutrosophic and* Plithogenic Sets, Academic Press, pp. 117-135, 2020.

[32] A. Ansaria, R. Biswasb, and S. Aggarwal, "Neutrosophic classifier: an extension of fuzzy classifier", *Applied Soft Computing*, vol. 13, pp. 563–573, 2012.

[33] Z. Jianqiang, G. Xiaolin, and T. Feng, "A New method of identifying influential users in the micro- blog networks", *IEEE Access*, vol. 5, pp. 3008-3015, 2017.

[34] A. Ivaschenko, A. Krivosheev, A. Stolbova, and P. Sitnikov, "Approximate analysis of deviant behavior on social media", *Intelligent Computing*, pp. 539- 547, Springer, 2022.

[35] N. M. Sham, A. Mohamed, "Climate change sentiment analysis using lexicon, machine learning and hybrid approaches", *Sustainability*, vol. 14, no. 8, pp. 4723- 4750, 2022.

[36] F. Ali, D. Kwak, P. Khan, S. R. Islam, K. Kim, and K. Kwak, "Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling", *Transportation Research Part C: Emerging Technologies*, vol. 77, pp. 33- 48, 2017.

[37] F. Benedetto, and A. Tedeschi, "Big data sentiment analysis for brand monitoring in social media streams by cloud computing", *Sentiment Analysis and Ontology Engineering Studies in Computational Intelligence*, vol. 639, pp. 341-377, 2016.

[38] R. Jafelice, A. Bertone, and R. Bassanezi, "A study on subjectivities of type 1 and 2 in parameters of differential equations", *TEMA* (São Carlos), vol. 16, no.1, pp. 51- 60, 2015.

[39] S. Al-subhi, E. Papageorgiou, P. Pe´rez, G. Mahdi, and L. Acun˜a, "Triangular neutrosophic cognitive map for multistage sequential decision-making problems", *International Journal of Fuzzy Systems*, vol. 23, no. 3, pp. 657- 679, 2021.

[40] D. Sagayakavitha, and A.Sudha, "New approach to solve symmetric triangular neutrosophic linear programming problem by score function", *Journal of Information and Computational Science*, vol. 10, no. 3, pp. 908- 915, 2020.

[41] S. Broumi, D. Nagarajan, A. Bakali, M. Talea, F. Smarandache, M. Lathamaheswari, and J. Kavikumar, "Implementation of

neutrosophic function memberships using MATLAB program", *Neutrosophic Sets and Systems*, vol. 27, pp. 44- 52, 2019.

[42] M. Abdel-Basset, G. Manogaran, A. Gamal, and F. Smarandache, "A hybrid approach of neutrosophic sets and DEMATEL method for developing supplier selection criteria", *Design Automation for Embedded Systems*, vol. 22, No. 3, pp. 257- 278, 2018.

[43] A. Minj, and T. Pathinathan, "Totally symmetric type-2 triangular fuzzy numbers", *Journal of Computer and Mathematical Sciences*, Vol. 9, No. 11, pp. 1717- 1727, 2018.

[44] D. Dinagar, and K. Latha, "Some types of type-2 triangular fuzzy matrices", *International Journal of Pure and Applied Mathematics*, vol. 82, no. 1, pp. 21- 32, 2013.

[45] M. Lathamaheswari, D. Nagarajan, J. Kavikumar, and S. Broumi, "Triangular interval type-2 fuzzy soft set and its application", *Complex and Intelligent Systems*, vol. 6, pp. 531– 544, 2020.

[46] J.M. Mendel, *Type-2 Fuzzy Sets*, Uncertain Rule-Based Fuzzy Systems, Springer, pp. 259- 306, 2017.

[47] A. Ansari, R. Biswas, and S. Aggarwal, "Extension to fuzzy logic representation: moving towards neutrosophic logic-a new laboratory rat", In *Proceedings of the IEEE International Conference on Fuzzy Systems*, India, pp. 1- 8, 2013.

[48] A. Abu Bakar, K. Khalif, and A. Gegov, "Ranking of interval type-2 fuzzy numbers based on centroid point and spread", In *Proceedings of the International Joint Conference on Computational Intelligence*, Portugal, vol.2, pp. 131-140, 2015.

[49] M. Anand, and M. Lathamaheswari, "A new generalized normal type-2 trapezoidal fuzzy number", *Global Journal of Pure and Applied Mathematics*, vol. 12, no. 3, pp. 565-574, 2016.

[50] G. Preda, "COVID-19 all vaccines tweets", *Kaggle data*, 2021.https://doi.org/10.34740/KAGGLE/DSV/2845240

REEM ESSAMELDIN received the B.Sc. degree in Electrical Engineering (Communications and Electronics Section) from the Faculty of Engineering, Alexandria University, Egypt in 2012. She held the M.Sc. degree in information technology from the Institute of Graduate Studies and Research (IGSR), Department of Information Technology, University of Alexandria in 2019. She held another M.Sc. degree in Masters of Business Administration (MBA) from Zhejiang Normal University, China in 2016 for a thesis in electronic human resource management (e-HRM). She is researcher and assistant teacher at the Institute of Graduate Studies and Research (IGSR). Her research interests include machine learning, social network analysis, opinion mining, e-marketing, business intelligence, e-commerce, and communication networks.

AHMED A. ISMAIL received the B.Sc. degree in Management Information System (MIS) from the College of Business Administration, Arab Academy for Science & Technology and Maritime Transport in 1999. He held the M.Sc. degree in information technology from the Institute of Graduate Studies and Research (IGSR), Department of Information Technology, University of Alexandria in 2007. He received his Ph.D. degree from Helwan University for a thesis in tracing and detecting explosives. He has more than 15 year experience in the field of Information Technology. His research and professional interests include e-marketing, business intelligence, business information systems, management information systems, database management systems, web programming, data communication and computer networks.

SAAD M. DARWISH received the B.Sc. degree in Statistics and Computer Science from the Faculty of Science, Alexandria University, Egypt in 1995. He held the M.Sc. degree in information technology from the Institute of Graduate Studies and Research (IGSR), Department of Information Technology, University of Alexandria in 2002. He received his Ph.D. degree from Alexandria University for a thesis in image mining and image description technologies. He is the author or coauthor of 50+ papers publications in prestigious journals and top international conferences and also has received several citations. He has served as a reviewer for several international journals and conferences. He has supervised around 80 M.Sc. and Ph.D. students. His research and professional interests include image processing, optimization techniques, security technologies, database management, machine learning, biometrics, digital forensics, and bioinformatics. Since June 2017, he has been a professor in the department of information technology at IGSR.