Computers in Industry 115 (2020) 103180

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

Sentiment analysis of tweets using refined neutrosophic sets

Ilanthenral Kandasamy[a,*], W.B. Vasantha[a], Jagan M. Obbineni[b], F. Smarandache[c]

[a] School of Computer Science and Engineering, VIT, Vellore, Tamil Nadu 632014, India
[b] VIT School for Agricultural Innovations and Advanced, Learning (VAIAL), VIT, Vellore, India
[c] Department of Mathematics, University of New Mexico, 705 Gurley Avenue, Gallup, NM 87301, USA

ARTICLE INFO

Article history:
Received 30 January 2019
Received in revised form
24 November 2019
Accepted 3 December 2019

Keywords:
Opinion mining
Sentiment analysis
Text mining
Social media
Tweets
Neutrosophy
Neutrosophic logic
Indeterminacy
Single valued neutrosophic set (SVNS)
Refined neutrosophic sets
Multi refined neutrosophic set (MRNS)
Triple refined indeterminate neutrosophic set (TRINS)

ABSTRACT

In the last decade, opinion mining and sentiment analysis have been the subject of fascinating interdisciplinary research. Alongside the evolution of social media networks, the sheer volume of social media text available for sentiment analysis has increased multi-fold, leading to a formidable corpus. Sentiment analysis of tweets have been carried out to gauge public opinion on breaking news, various policies, legislations, personalities and social movements. Fuzzy logic has been used in the sentiment analysis of twitter data, whereas neutrosophy which factors in the concept of indeterminacy has not been used to analyse tweets. In this paper, the concept of multi refined neutrosophic set (MRNS) with two positive, three indeterminate and two negative memberships is proposed. Single valued neutrosophic set (SVNS), triple refined indeterminate neutrosophic set (TRINS) and MRNS have been used in the sentiment analysis of tweets on ten different topics. Eight of these topics chosen for sentiment analysis are related to Indian scenario and two topics to international scenario. A comparative analysis of the methods show that the approach with MRNS provides better refinement to the indeterminacy present in the data.

1. Introduction

The evolution of public opinion as an influential force in the political sphere can be dated back to French revolution in 17^{th} century [1]. The analysis of public opinion began in the 20^{th} century. Opinion mining is a research domain that has seen speedy evolution in the preceding decade. In a contemporary trend mostly sentiment analysis is carried out on social media texts from Twitter and Facebook. A recent review paper on sentiment analysis states that more than 7000 research papers have been published on this topic and that modern sentiment analysis has established a 50-fold growth in over ten years (2005 to 2016) [2]. Conventional sentiment analysis does not deal with a neutral or an indeterminate opinion, it merely gives an overall opinion as positive or negative.

Fuzzy theory has been helpful in improving sentiment analysis techniques on twitter data [3]. Fuzzy set theory [4] that permits soft partition of sets, is stretched to Intuitionistic Fuzzy Set (A-IFS), in which a membership and a non-membership degree is allotted to every single constituent element [5], whereas in neutrosophic set, an indeterminacy membership is represented independently, together with truth membership and falsity membership to separately represent indeterminate, unpredictable, vague and uncertain information from the real world [6]. It simplifies from a philosophical point of view the idea of several sets, and its functions; $T_A(x)$, $I_A(x)$, and $F_A(x)$ and the functions are real standard or non-standard subsets for any object x in the universal space of points or objects.

Wang et al. [7] presented a single valued neutrosophic set (SVNS), to achieve an improved solution to the problem of applying neutrosophy in real world scientific and engineering problems. Neutrosophy and neutrosophic logic have found manifold applications in real world practical problems like image processing [8–10], decision-making [11–18], social network analysis [19] and social issues [20,21] etc.

In double valued neutrosophic set (DVNS) [22,23], an indeterminacy membership of the neutrosophic set has been characterised into two memberships to enable more accuracy in the indeterminacy present. Distance measure, cross entropy measure, dice

* Corresponding author.
E-mail addresses: ilanthenral.k@vit.ac.in (I. Kandasamy), vasantha.wb@vit.ac.in (W.B. Vasantha), jagan.obbineni@vit.ac.in (J.M. Obbineni), smarand@unm.edu (F. Smarandache).

https://doi.org/10.1016/j.compind.2019.103180
0166-3615/

measure, and clustering algorithm of DVNS was introduced and studied in [22,24]. The indeterminacy notion was separated into three, as indeterminacy inclined towards truth, indeterminacy and indeterminacy inclined towards false memberships in triple refined indeterminate neutrosophic set (TRINS), to improve the accuracy and precision of the uncertain data and to adapt it to the Likert's scale which is a habitually used psychometric scale. It was utilized for personality testing and classification [25]. TRINS was refined recently with positive, positive indeterminate, indeterminate, negative indeterminate and negative memberships, to give the finest conceivable mapping of the Likert scaling. This was defined as indeterminate likert scaling [26,27].

To capture the indeterminacy present in sentiment analysis of tweets, neutrosophy is used. Multi refined neutrosophic set (MRNS) with 2 positive, 3 indeterminate and 2 negative memberships is introduced and utilized for sentiment analysis. These seven memberships aid in capturing the polarity with better accuracy.

Section one is preliminary in nature. The rest of the paper is planned as follows: Section two presents some elementary concepts about sentiment analysis and different neutrosophic sets like SVNS, DVNS, TRINS and refined neutrosophic sets. In section three MRNS with 3 indeterminate memberships is defined and its properties are discussed. Section four discusses the limitation and problems with normal sentiment analysis and provides justification for using indeterminacy in sentiment analysis. In section five, sentiment analysis using neutrosophy is proposed. In the next section sentiment analysis of tweets of eight domestic and two international issues using three neutrosophic sets namely SVNS, TRINS and MRNS is carried out. Evaluations and discussions of these different models are analyzed in section seven. A sample topic from SemEval 2017 is taken for comparative analysis of our models. Results and further probable studies in this direction are provided in the last section.

2. Basic concepts

2.1. Sentiment Analysis

News articles, blogs, film reviews and social media information have been investigated extensively to comprehend public opinion. Typically, tweets are scrutinized and classified as positive, neutral or negative; this methodology is carried out to discover how society is feeling about a specific trending topic. Usually keyword-based tools are used to classify data (mostly social media posts, news, reviews, etc.) as positive or neutral or negative.

With the increase in data available online from early 2000, modern sentiment analysis started to take shape in mid-2000s. It has resulted in various concepts like web product reviews [28], prediction of financial markets [29], reactions to terrorist attacks [30] and multi-lingual support [31].

Sentiment analysis overlaps or relies heavily on information and knowledge management, data mining, text mining, web mining, natural language processing (NLP) and computational linguistics. Recently work is being carried out in evolving from humble polarity detection to complex gradations of emotions and distinguishing negative emotions such as anger from grief [32] and figurative language [33].

Irony detection [34–36] has a huge impact on sentiment analysis, since they write the opposite of what they feel. Recent deep learning based approaches like transfer learning have been applied in irony detection [37], aspect-based sentiment analysis using attentive long short-term memory (LSTM) [38], word vectors representations for sentiment analysis [39] and capsule networks for sentiment classification [40]. Other deep learning techniques used in sentiment analysis are reviewed in [41,42].

Fuzzy theory has been used for sentiment analysis in [3,43]. Fuzzy theory captures the positive and negative of the topic but fails to address the indeterminacy present. To address the indeterminacy present, the concept of neutrosophy in the form of SVNS, TRINS and MRNS are used to analyse the twitter data. Recently in [44] sentiment analysis of tweets about #MeToo movement, each tweet was analysed and represented as SVNS.

The discussion about the tools and methodology used for sentiment analysis in this paper are carried out in later sections. We briefly describe the notion of neutrosophy, SVNS and TRINS in the following subsection.

2.2. Neutrosophy, SVNS and TRINS

Neutrosophy, studies an opinion or sentiment, "*A*" and its relation to its opposite sentiment, "anti-*A*" and not *A*, "non-*A*", and as neither "*A*" nor "anti-*A*".

Definition 1. Let X be a space of points (objects) with generic elements x in X. The set A in X is characterized by three functions $T_A(x)$ truth membership, $I_A(x)$ indeterminacy membership, and $F_A(x)$ falsity membership. For each $x \in X$, $T_A(x), I_A(x), F_A(x) \in [0, 1]$ and $0 \le T_A(x) + I_A(x) + F_A(x) \le 3$. Single valued neutrosophic set (SVNS) A is represented by $A = \{\langle x, T_A(x), I_A(x), F_A(x)\rangle \mid x \in X\}$.

The refined neutrosophic set [45] is defined as follows:

Definition 2. The truth T is divided into several types of truths: $T_1, T_2, \ldots, T_p$, and I into various indeterminacies: $I_1, I_2, \ldots, I_r$, and F into various falsities: $F_1, F_2, \ldots, F_s$, where all $p, r, s \ge 1$ are integers, and $p + r + s = n$.

Definition 3. A triple refined indeterminate neutrosophic set (TRINS) A in X, as given above is characterized by positive $P_A(x)$, positive indeterminate $I_{PA}(x)$, indeterminate $I_A(x)$, negative indeterminate $I_{NA}(x)$ and negative $N_A(x)$ membership functions. Each has a weight $w_m \in [0, 5]$ associated with it. For each $x \in X$, there are

$P_A(x), I_{PA}(x), I_A(x), I_{NA}(x), N_A(x) \in [0, 1]$, with weights $w_P^m(P_A(x))$, $w_{I_P}^m(I_{PA}(x))$, $w_I^m(I_A(x))$, $w_{I_N}^m(I_{NA}(x))$, $w_N^m(N_A(x)) \in [0, 5]$,

and $0 \le P_A(x) + I_{PA}(x) + I_A(x) + I_{NA}(x) + N_A(x) \le 5$.

Therefore, TRINS A can be represented by

$A = \{\langle x, P_A(x), I_{PA}(x), I_A(x), I_{NA}(x), N_A(x)\rangle \mid x \in X\}$.

The different properties and set theoretic operators like commutativity, idempotency, distributivity, associativity, absorption and the DeMorgan's Laws have been defined over TRINS [25]. As future research it is proposed to map the middle 3 terms of TRINS to neutrosophic triplets [46] and then they can be automatically mapped to neutrosophic duplets [47,48] in case of the indeterminacy leaning towards false is zero.

Neutrosophy has been applied to several different fields ranging from medical diagnosis [49,50] image processing [51], decision making [52,53], personnel selection [54], supply chain management [55,56], internet of things [57], psychology [25,58] and social science [21,59], but has not been used in sentiment analysis, until recently in [44].

3. Multi refined neutrosophic set (MRNS)

In the newly proposed Multi refined neutrosophic set (MRNS), the concept of positive (truth) is divided into two memberships, as strong positive and positive, similarly the concept of negative (false) is divided into two memberships as strong negative and negative. Also the indeterminate membership is divided into three memberships as positive indeterminate, indeterminate and negative indeterminate. This division helps in increasing the accuracy and precision in data analysis and fits the multipoint likert scale kind of structure where there are different degrees of acceptance. This refined neutrosophic set is defined as MRNS.

Definition 4. A multi refined neutrosophic set (MRNS) A in X is characterized by strong positive $SP_A(x)$, positive $P_A(x)$, positive indeterminate $PI_A(x)$, indeterminate $I_A(x)$, negative indeterminate $NI_A(x)$, negative $N_A(x)$ and strong negative $SN_A(x)$ membership functions. Each has a weight $w_m \in [0, 7]$ associated with it. For each $x \in X$, there are

$SP_A(x), P_A(x), PI_A(x), I_A(x), NI_A(x), N_A(x), SN_A(x) \in [0, 1]$,

and $0 \leq SP_A(x) + P_A(x) + PI_A(x) + I_A(x) + NI_A(x) + N_A(x) + SN_A(x) \leq 7$.

Therefore, a MRNS A can be represented by

$A = \{\langle x, SP_A(x), P_A(x), PI_A(x), I_A(x), NI_A(x), N_A(x), SN_A(x)\rangle \mid x \in X\}$.

To illustrate the applications of MRNS in a real world problem, consider parameters like work satisfaction, occupational stress and role of technology that are commonly used to measure work-life balance. The evaluation of work-life balance is used to illustrate set-theoretic operations on MRNSs.

Example 1. Let $WL = [w_1, w_2, w_3]$ where w_1 is work satisfaction, w_2 is occupational stress and w_3 is role of technology. The values of w_1, w_2 and w_3 are in $[0, 1]$. They are obtained using a questionnaire answered by an anonymous working women. A is a MRNS of WL defined by

$A = \langle 0.3, 0.2, 0.1, 0.2, 0.1, 0.2, 0.5\rangle/w_1 + \langle 0.4, 0.1, 0.1, 0.1, 0.2, 0.2, 0.3\rangle/w_2 + \langle 0.5, 0.2, 0.2, 0.1, 0, 0.1, 0.1\rangle/w_3$.

B is a MRNS of WL defined by

$B = \langle 0.4, 0.2, 0.1, 0.2, 0.2, 0.1, 0.2\rangle/w_1 + \langle 0.2, 0.2, 0, 0.1, 0.1, 0.2, 0.4\rangle/w_2 + \langle 0.4, 0.2, 0.1, 0.1, 0.1, 0.2, 0.3\rangle/w_3$

Definition 5. The complement $c(A)$ of A is defined as

1 $SP_{c(A)}(x) = SN_A(x)$
2 $P_{c(A)}(x) = N_A(x)$
3 $PI_{c(A)}(x) = 1 - PI_A(x)$
4 $I_{c(A)}(x) = 1 - I_A(x)$
5 $NI_{c(A)}(x) = 1 - NI_A(x)$
6 $N_{c(A)}(x) = P_A(x)$
7 $SN_{c(A)}(x) = SP_A(x)$ for all x in X.

Definition 6. A is contained in B, that is $A \subseteq B$, if and only if

1 $SP_A(x) \leq SP_B(x)$
2 $P_A(x) \leq P_B(x)$
3 $PI_A(x) \leq PI_B(x)$
4 $I_A(x) \leq I_B(x)$
5 $NI_A(x) \leq NI_B(x)$
6 $N_A(x) \geq N_B(x)$
7 $SN_A(x) \geq SN_B(x)$ for all x in X.

Note that by the definition of containment relation, X is a partially ordered set and not a totally ordered set.

For example, consider the MRNS A and B as mentioned in Example 1, then A is not contained in B and vice versa.

Definition 7. Two MRNS A and B are equal, that is $A = B$, if and only if $A \subseteq B$ and $B \subseteq A$.

Definition 8. The union of A and B is a MRNS G, denoted as $G = A \cup B$, whose seven membership functions are related to those of A and B by the following

1 $SP_G(x) = max(SP_A(x), SP_B(x))$
2 $P_G(x) = max(P_A(x), P_B(x))$
3 $PI_G(x) = max(PI_A(x), PI_B(x))$
4 $I_G(x) = max(I_A(x), I_B(x))$
5 $NI_G(x) = max(NI_A(x), NI_B(x))$
6 $N_G(x) = min(N_A(x), N_B(x))$
7 $SN_G(x) = min(SN_A(x), SN_B(x))$ for all x in X.

Definition 9. The intersection of A and B is a MRNS F, denoted as $F = A \cap B$, whose seven membership functions are related to those of A and B by the following

1 $SP_F(x) = min(SP_A(x), SP_B(x))$
2 $P_F(x) = min(P_A(x), P_B(x))$
3 $PI_F(x) = min(PI_A(x), PI_B(x))$
4 $I_F(x) = min(I_A(x), I_B(x))$
5 $NI_F(x) = min(NI_A(x), NI_B(x))$
6 $N_F(x) = max(N_A(x), N_B(x))$
7 $SN_F(x) = max(SN_A(x), SN_B(x))$ for all $x \in X$.

Theorem 1. *$A \cap B$ is the largest MRNS contained in both A and B.*

Proof. It is straightforward from the definition of intersection operator. □

Definition 10. The difference D, written as $D = A\ B$, whose seven membership functions are related to those of A and B by

1 $SP_D(x) = min(SP_A(x), SN_B(x))$
2 $P_D(x) = min(P_A(x), N_B(x))$
3 $PI_D(x) = min(PI_A(x), 1 - PI_B(x))$
4 $I_D(x) = min(I_A(x), 1 - I_B(x))$
5 $NI_D(x) = min(NI_A(x), 1 - NI_B(x))$
6 $N_D(x) = max(N_A(x), P_B(x))$
7 $SN_D(x) = max(SN_A(x), SP_B(x))$ for all x in X.

Two operators positive favourite ($\triangle$) and negative favourite ($\triangledown$) are defined to remove the indeterminacy in the MRNSs and transform it into intuitionistic fuzzy sets or paraconsistent sets. Similarly a MRNS can be transformed into a SVNS by applying the indeterminacy neutral (∇) operator that combines the indeterminacy values of the MRNS.

Definition 11. The positive favourite of A represented as $B = \triangle A$, whose membership functions are related to those of A by

1 $T_B(x) = min(SP_A(x) + P_A(x) + PI_A(x), 1)$
2 $F_B(x) = N_A(x)$

Definition 12. The negative favourite of A, represented as $B = \triangledown A$, whose membership functions are related to those of A by

1 $T_B(x) = P_A(x)$
2 $F_B(x) = min(SN_A(x) + N_A(x) + NI_A(x), 1)$

Definition 13. The indeterminacy neutral of a MRNS A, written as $B = \nabla A$, whose membership functions are related to those of A by

1 $T_B(x) = min(SP_A(x) + P_A(x), 1)$
2 $I_{TB}(x) = min(PI_A(x) + I_A(x) + NI_A(x), 1)$
3 $F_B(x) = min(SN_A(x) + N_A(x), 1)$

The set theoretic operators like commutativity, associativity, distributivity, idempotency, absorption, involution and De Morgan's Laws are similar to the ones defined on SVNS. MRNS satisfies most of the properties of classical set, fuzzy set, intuitionistic fuzzy set and SVNS. Similar to fuzzy set, IFS set, SVNS, DVNS and TRINS, MRNS does not satisfy the principle of middle exclude.

4. Justification for applying indeterminacy

In opinion mining and sentiment analysis, the major division of opinion is done in terms of positive, neutral and negative opinion. Mostly Likert scaling based questionnaires are used for opinion mining. Even if Likert scaling is used in gauging the opinion of the user, the user is forced to select the most dominant choice. Generally, a person has feelings which actually vary from "strongly agree" to "strongly disagree", and which are indefinite in nature, mostly a mixture of feelings. A little disagreement might force the opinion from "strongly agree" to "agree"; whereas a different person might still choose to go ahead with the dominant opinion of "strongly agree" ignoring the little disagreement. A different respondent might mark the option "neither agree nor disagree" due to a little disagreement.

Evidently people respond differently to experiences and issues (political, economic or social in nature) while answering the questions. The questionnaire based on Likert scale will fail to capture the feelings accurately. The respondent generally goes with the dominant choice or the choice which he feels at that time or the choice which may be only a shade dominant than the other choice, thereby the degree of the memberships with other choices is completely lost. Only a measure of coarse ordinal scale with closed format is used by Likert method [60].

Similarly, when opinion mining is carried out on a specific topic to gauge the public reaction; only positive, neutral or negative categorization is done. Every person will have opinion that has various degrees of different memberships and the analysis needs to go with the dominant choice. The innumerable degrees and choices has to be captured accurately with greater precision; in fact, in a sensitive, accurate and realistic way and not in an approximate way. This will eventually aid in better understanding of people opinion or public or customers.

There is actually a lot of difference between someone who is undecided and someone who is taking a neutral stance, in a MRNS, there can be a separate option for undecided, since equal amount of agreement and disagreement can be represented in degree of weak agreement and degree of weak disagreement, individually.

Consider the seven point Likert scale, the various options given will be strongly disagree, disagree, weakly disagree, neither agree or disagree, weakly agree, agree and strongly agree. They will get mapped in MRNS; independently and appropriately under the seven heads.

Take a typical situation were the researchers must ascertain the public opinion about a political party. Usually they will project as positive or neutral or negative opinion. If the same is ascertained by making use of MRNS, the results obtained will be very accurate and clearly show the different degrees of strong negative, negative, indeterminate negative, indeterminate, indeterminate positive, positive and strong positive memberships.

5. Sentiment analysis using neutrosophy

Commonly sentiment analysis is done on tweets to classify the tweet as positive or neutral or negative. The typical scenario in which sentiment analysis of tweets is carried out is to discover how people are feeling about a specific trending topic. It is well known that there are many shades of agreement, disagreement and neutrality and there is indeterminacy involved in that neutrality. Two tweets which are classified as positive need not exactly have the same amount of positivity in them. One of them might be very strongly positive, whereas the other might be a little positive and have a lot of uncertainty or indeterminacy in it. This newly proposed method of analyzing using neutrosophic sets will give greater complexity but the accuracy in prediction of the tweet's polarity is better.

Table 1
Case study and data collection time line

Case study	Date of data collection
Farm loan wavier (*FL*)	28-12-2018
Onion price (*OP*)	29-12-2018
Foreign trips of PM (*FT*)	29-12-2018
Women reservation bill (*WB*)	29-12-2018
Triple talaq bill (*TT*)	29-12-2018
POCSO act (*PA*)	29-12-2018
UP mob violence (*UP*)	30-12-2018
Trump wall (*TW*)	30-12-2018
Yellow vests protest (*YV*)	30-12-2018
#Metoo movement (*MM*)	30-12-2018

5.1. Tools and methodology

A Twitter application programming interface (API) was created and consumer key, consumer secret, access token and access token secret were generated for obtaining tweets. Python language was used for data analysis of the collected tweets. Tweets were fetched from Twitter using tweepy python client. TextBlob is a Python library that helps in processing textual data. It provides a simple API for NLP. In our study, TextBlob was utilized for sentiment analysis. The sentiment property of TextBlob returns a tuple named Sentiment. It is of the form Sentiment(polarity, subjectivity), it returns two properties, polarity, and subjectivity. The polarity score is a float value in the range $[-1.0, 1.0]$, where 1 means positive statement and -1 means a negative statement. Let $p(x)$ denote the polarity score of the tweet. A part of the tweets extracted and used for analysis is available at [61].

5.2. Pre-processing of the Twitter data

The presence of mention, numbers, special characters, stopwords, hashtags, links and other jargon decreases the efficiency of the model and hence the tweets were cleaned. Several python libraries like pandas, numpy, matplotlib, BeautifulSoup and WordPunctTokenizer were used for cleaning. The cleaned tweet was saved as.CSV file and later used for analysis, using TextBlob.

The methodology of using different neutrosophic sets are discussed below:

5.3. Using SVNS

The normal classification of tweets as positive or negative or neutral is generally carried out. This classification is represented as SVNS. If the polarity calculated is greater than 0, i.e., $p(x) \in (0, 1]$, it is mapped to positive membership, if polarity is less than 0; $p(x) \in [-1, 0)$, it is mapped as negative membership, and if polarity is 0 it is mapped to indeterminate membership.

In the considered scenario of farm loan (denoted by *FL*), from the analysis it was obtained that, 41.90% was positive, 43.8% was indeterminate and 14.30% was negative as given in Table 2. Tweets are classified as neutral, when the classifier is not able to decide on the polarity of the tweet, that is when it is indeterminate. The aggregated result is normalized before it is converted to a neutrosophic set representation. The result of the analysis carried out is represented as SVNS is $FL_{SVNS} = \langle 0.419, 0.438, 0.143\rangle$. It is clearly seen that positive membership and indeterminate membership have a very small difference.

Table 2
Case 1: Farm loan waiver (*FL*) and Case 2: Onion prices (*OP*)

	SVNS	TRINS	MRNS		SVNS	TRINS	MRNS
SP		0.048	0.026	SP		0.025	0.012
P	0.419	0.311	0.164	P	0.253	0.228	0.036
PI			0.229	PI			0.205
I	0.438	0.438	0.438	I	0.664	0.664	0.664
NI			0.111	NI			0.059
N	0.143	0.136	0.025	N	0.082	0.076	0.017
SN		0.007	0.007	SN		0.006	0.006

5.4. Using TRINS

The classification can be made more precise by dividing the positive polarity tweets into two different classifications and the negative polarity into two different classifications. The classification is strong positive, positive, indeterminate, negative and strong negative. If polarity of the tweet is from 1 to 0.5, i.e., $p(x) \in (0.5, 1]$, it is classified as strong positive, if $p(x) \in (0, 0.5]$ is from 0.5 to greater than 0, it is classified as positive, if $p(x) \in (0)$ is mapped to indeterminate, less than 0 to −0.5 i.e., $p(x) \in [-0.5, 0)$, it is classified as negative, greater than −0.5 to −1 i.e., $p(x) \in [-1, -0.5)$ is classified as strong negative. The data given is normalised and it is represented as TRINS is $FL_{TRINS} = \langle 0.048, 0.371, 0.438, 0.136, 0.007\rangle$, in the case study of farm loan.

5.5. Using MRNS

The classification is made even more precise by dividing the positive polarity tweets and the negative polarity tweets into three different classification. The classification scheme that is introduced is strong positive, positive, positive indeterminate, indeterminate, negative indeterminate, negative, strong negative. If polarity of the tweet is from +1 to greater than 0.6; $p(x) \in (0.6, 1]$, it is classified as strong positive, from 0.6 to greater than 0.3; $p(x) \in (0.3, 0.6]$ it is classified as positive, from 0.3 to greater than 0; $p(x) \in (0, 0.3]$ it is mapped as positive indeterminate, if $p(x) \in (0)$; it is mapped to indeterminate, less than 0 to −0.3, $p(x) \in [-0.3, 0)$ it is classified as negative indeterminate, lesser than −0.3 to −0.6; $p(x) \in [-0.6, -0.3)$ it is mapped as negative, less than −0.6 to −1; $p(x) \in [-1, -0.6)$ it is classified as strong negative. According to this classification the results obtained is normalized and represented as MRNS is $FL_{MRNS} = \langle 0.026, 0.164, 0.229, 0.438, 0.111, 0.025, 0.007\rangle$.

The analysis of each individual case scenario is carried out in next section.

6. Sample case scenarios

Utilizing the twitter API created for this purpose, 1000 tweets were obtained for each case under consideration. Preprocessing of the tweets were carried out to remove links and emojis, after which sentiment analysis was done. For each case study a background is provided and then analysis of tweets is discussed. The topic with abbreviation and the period of data collection is tabulated in Table 1. All tweets collected were in English language and from across the world, despite some topics being related only to some geographic location. The 10 topics selected were trending topics at the time of data collection, since we needed at least 1000 tweets about the topic.

The analysis result of each case study is given in the form of tables were the following abbreviations are used. In the first column, SP refers to strong positive membership value, P is positive membership value, PI is positive indeterminate membership value, I is indeterminate membership value, NI is negative indeterminate membership value, N is negative membership value and SN is strong negative membership value. Along the header row, SVNS stands for single valued neutrosophic set, TRINS is triple refined indeterminate neutrosophic set and MRNS is used for multi refined neutrosophic set.

6.1. Case 1: Farm loan wavier by government

Introduction: Agriculture remains to be the primary source of livelihood for nearly more than half of India's population. The country is dependent on farmers, the systematic failures of the state and center government to address their issues, pushes farmers to protest regularly and in recent years many farmers have committed suicide. Farmers have marched in Delhi and Mumbai cities to highlight the reality of their deprivation and anger. This is due to lack of compensation from drought and natural disasters like cyclone etc., crop insurance scheme failures, and the deficit created due to prices decreasing below the minimum support prices and so on. These losses are estimated to be around thousands of crores every year.

Most leaders of major Indian political parties have pledged their support to the farmers issue. In late 2018, the first step taken by three newly formed state governments (Rajasthan, Madhya Pradesh and Chhattisgarh) was a farm loan waiver. This has understandably started a debate about the usefulness of loan waiver since it is only an element of immediate relief. Farm loans and wavier have been a topic that has invoked mixed responses from people. Here tweets were collected about farm loan for analysis using the search term "farm loan wavier".

Analysis: While applying SVNS for analysis, the result obtained is $FL_{SVNS} = \langle 0.419, 0.438, 0.143\rangle$. The indeterminate and positive membership values have little difference; hence the opinion is indeterminate and positive opinion. Next TRINS was used for analysing the same set of tweets, the result obtained is $FL_{TRINS} = \langle 0.048, 0.311, 0.438, 0.136, 0.007\rangle$, which also implies an indeterminate and positive opinion. When MRNS was applied a change in the scenario is seen. The resultant obtained is $FL_{MRNS} = \langle 0.026, 0.164, 0.229, 0.438, 0.111, 0.025, 0.007\rangle$; most of the positive is indeterminate positive, even the negative opinion is mostly indeterminate negative. It is seen that the public are undecided about farm loan wavier. The resultant of each neutrosophic representation is given in Table 2.

6.2. Case 2: Decrease in onion price

Introduction: The last weeks of December 2018 saw steep drops in the prices of onions and potatoes in India, it crashed as much as 86 percent. Both are staple food for the India's huge population, such a steep decrease has badly hit the rural economy in large states. Onion price hit a low of Re.1 per kilogram, while it cost nearly Rs.8 to produce one kilogram. These kind of unsteady market prices cause more distress to farmers. This also does not benefit the urban population because there are too many middlemen between the farmer and customer. The search term "onion price" was used to collect 1000 tweets for analysis.

Analysis: In SVNS representation the result obtained is $OP_{SVNS} = \langle 0.253, 0.664, 0.082\rangle$. It is indeterminate, even though the steep decrease in onion price, has affected the farmers adversely. While using TRINS it is observed that $OP_{TRINS} = \langle 0.025, 0.228, 0.664, 0.076, 0.006\rangle$; it indicates that people neither have a strong negative or strong positive opinion about the price decrease, but it is in general more positive than negative. MRNS was used to analyse the same dataset of tweets. The resultant is $OP_{MRNS} = \langle 0.012, 0.036, 0.205, 0.664, 0.059, 0.017, 0.006\rangle$ as given in Table 2. It shows that even most of the positive opinion was tending towards indeterminacy. Hence, the indeterminate positive has the second highest value. People who are twitter users are unaffected by the steep drop in price which affects farmers unfavorably. More so the affected

Table 3
Case 3: Women reservation bill and Case 4: Triple talaq bill

	SVNS	TRINS	MRNS		SVNS	TRINS	MRNS
SP		0.005	0.004	SP		0.152	0.149
P	0.084	0.079	0.038	P	0.315	0.163	0.109
PI			0.042	PI			0.057
I	0.165	0.165	0.165	I	0.538	0.538	0.538
NI			0.019	NI			0.110
N	0.751	0.750	0.731	N	0.146	0.145	0.035
SN		0.001	0.001	SN		0.001	0.001

Table 4
Case 5: POCSO act death penalty and Case 6: MeToo movement

	SVNS	TRINS	MRNS		SVNS	TRINS	MRNS
SP		0.001	0.001	SP		0.04	0.014
P	0.909	0.899	0.863	P	0.562	0.522	0.375
PI			0.036	PI			0.173
I	0.038	0.038	0.038	I	0.291	0.291	0.291
NI			0.052	NI			0.086
N	0.053	0.053	0.001	N	0.147	0.132	0.046
SN		0	0	SN		0.015	0.015

farmers are not the tweeters; so only this opinion of tweeter clearly reflects the situation is indeterminable.

6.3. Case 3: Women reservation bill

Introduction: The women reservation bill (108^{th} amendment to the constitution of India) is a lapsed bill in the parliament of India that was proposed in 2008. It proposed to amend the constitution of India to reserve for women 33% of all seats in the lok sabha (lower house of parliament of India), and in all state legislative assemblies, in rotational basis. With the 2019 general elections in a few months' time, the demand for the bill in the parliament has been gathering support. The bill has been around for nearly ten years and the people have mixed opinion. The term "women reservation bill" was used to query and collect 1000 tweets for analysis.

Analysis: Sentiment analysis was carried out on women reservation bill using SVNS. The resultant is $WB_{SVNS} = \langle 0.084, 0.165, 0.751\rangle$; it clearly shows that the general public (that are on twitter) are against the bill. Even when TRINS is applied, the resultant $WB_{TRINS} = \langle 0.005, 0.079, 0.165, 0.750, 0.001\rangle$, shows the same sentiment with the value of 0.750 for negative and 0.001 for strong negative. When MRNS is applied for analysis, the resultant obtained is $WB_{MRNS} = \langle 0.004, 0.038, 0.042, 0.165, 0.019, 0.731, 0.001\rangle$, it shows a meagre amount of negative indeterminate (0.019), decreasing the value of negative membership. It is clearly seen that most people are openly against the bill. The results are tabulated in Table 3.

6.4. Case 4: Triple talaq bill

Introduction: From 2011 census, it is known that 2.37 million women across India have identified themselves as "separated", though it is not known if these women voluntarily separated from their husbands or were abandoned or worse sent away. The vast majority (1.9 million) are hindu women, and nearly 0.28 million were "separated" muslim women. It is known that India‘s family laws permit for divorce, but they also allow husbands to leave a marriage without the divorce formalities. The salient features of the triple talaq bill states that any declaration of talaq by a muslim man upon his wife shall be void and illegal, shall be punished with an imprisonment term and are liable to fine. The custom of triple talaq is both harsh and unjust, muslim women have crusaded long to get free of it. Despite the law, men can choose to walk out of the marriage without saying talaq and they go free without punishment. The bill does not address the issue of non-muslim women who are abandoned by their husbands and provide punishment for those people who are equally guilty of abandonment. The keyword "triple talaq bill" was used for collecting 1000 tweets for analysis.

Analysis: Using SVNS, the resultant obtained is $TT_{SVNS} = \langle 0.315, 0.538, 0.146\rangle$, it clearly shows that the general public have not made up their mind, they are undecided, and the second leading opinion was positive. Even when TRINS is applied, the resultant $TT_{TRINS} = \langle 0.152, 0.163, 0.538, 0.145, 0.001\rangle$, shows the absence of a strong negative, whereas the positive sentiment is divided with the value of 0.152 for positive and 0.163 for strong positive. When the newly constructed MRNS is applied, the resultant is $TT_{MRNS} = \langle 0.149, 0.109, 0.057, 0.538, 0.110, 0.035, 0.001\rangle$, a meagre amount of positive indeterminate (0.019) and negative indeterminate (0.110) comes into picture, increasing the indeterminacy of the opinion. It is clearly seen that mostly people are undecided about the bill. The resultants are given in Table 3.

6.5. Case 5: Protection of children from sexual offences (POCSO) act death penalty amendment

Introduction: The Indian cabinet has approved amendments to the protection of children from sexual offences (POCSO) act in December 2018, to give more stringent punishment for committing sexual crimes against children. To discourage the current trend of child sexual abuse to act as a warning, it has been amended to provide death penalty in case of aggravated penetrative sexual assault on a child as option of stringent punishment. The search term "POCSO Act" was used for collecting the tweets for analysis.

Analysis: When sentimental analysis was carried out using SVNS, we obtained $PA_{SVNS} = \langle 0.909, 0.038, 0.053\rangle$, where a 90% majority felt it be a good move as given in Table 4. When the analysis was carried out with TRINS, we obtained $PA_{TRINS} = \langle 0.001, 0.899, 0.038, 0.053, 0\rangle$, it was clearly seen that no one felt strongly negative about it. People have a positive opinion about the law. Only a meagre amount of people had a negative opinion. Lastly, MRNS was used for analysing the same set of tweets. The result obtained was $PA_{MRNS} = \langle 0.001, 0.863, 0.036, 0.038, 0.052, 0.001, 0\rangle$. We are clearly able to capture that even the meagre negative is negative indeterminate; implying that people who are not sure about the amendments in the act, but they have some indeterminate negative opinion. Hence, we conclude that the negative opinion is also an indeterminate negative making up the overall opinion to be a positive opinion.

6.6. Case 6: #Metoo movement

Introduction: In past year the "#MeToo" movement was started against sexual harassment and assault and has gather a lot of attention and created several controversies. The term "MeToo" was used to extract related tweets for analysis, the term is ambiguous and can also lead to tweets unconnected to the movement.

Analysis: The result of the analysis is given in Table 4. Using SVNS, we obtained $MM_{SVNS} = \langle 0.559, 0.291, 0.144\rangle$, where a majority have a positive opinion. While using TRINS, we obtained $MM_{TRINS} = \langle 0.04, 0.519, 0.291, 0.13, 0.014\rangle$, it was clearly seen that the opinion is positive. Only a meagre amount of people had a negative opinion. Lastly, MRNS was used for analysis. The result obtained was $MM_{MRNS} = \langle 0.013, 0.374, 0.172, 0.291, 0.085, 0.045, 0.014\rangle$. We could clearly capture that even the negative is mostly negative indeterminate opinion. The overall opinion happens to be a positive opinion.

6.7. Case 7: Foreign trips of prime minister

Introduction: Over Rs. 2,021 crores (from June 2014) was spent on chartered flights, aircraft maintenance and hot-line facilities for the Indian Prime Minister Narendra Modi's visits to foreign coun-

Table 5
Case 7: Foreign trip and Case 8: UP mob violence

	SVNS	TRINS	MRNS		SVNS	TRINS	MRNS
SP		0.001	0.001	SP		0	0
P	0.474	0.473	0.003	P	0.104	0.104	0.013
PI			0.470	PI			0.091
I	0.123	0.123	0.123	I	0.223	0.223	0.223
NI			0.180	NI			0.434
N	0.403	0.399	0.219	N	0.673	0.670	0.236
SN		0.004	0.004	SN		0.003	0.003

Table 6
Case 9: Trump wall shutdown and Case 10: Yellow vests protest

	SVNS	TRINS	MRNS		SVNS	TRINS	MRNS
SP		0.030	0.020	SP		0	0
P	0.188	0.158	0.046	P	0.018	0.018	0
PI			0.122	PI			0.018
I	0.581	0.581	0.581	I	0.319	0319	0.319
NI			0.199	NI			0.654
N	0.231	0.224	0.025	N	0.663	0.662	0.008
SN		0.007	0.007	SN		0.001	0.001

tries. The numerous visits of the PM have been a topic of debate. The search term used for extracting tweets is “Foreign trip”.

Analysis: Here 1000 tweets were analysed using SVNS, the resultant obtained is $FT_{SVNS} = \langle 0.474, 0.123, 0.403\rangle$. It shows an overall positive opinion about the amount spent on foreign visits by the PM, the negative opinion is also prevalent among the public as the positive membership and negative membership values have very little difference. When TRINS is used for analysis, $FT_{TRINS} = \langle 0.001, 0.473, 0.123, 0.399, 0.004\rangle$ is the resultant. Despite an overall positive opinion; it is seen that very few had a strong positive opinion, but more people had a strong negative opinion about the trips. While using MRNS we arrive at a clearer picture where the resultant is $FT_{MRNS} = \langle 0.001, 0.003, 0.470, 0.180, 0.219, 0.004\rangle$. Most of the positive opinion is indeterminate positive and not actually positive; whereas most of the negative is negative. People are undecided about PM’s foreign visits, and more people have a decisive negative opinion than a positive opinion about the visit. It can be clearly seen that MRNS provides a better realistic picture of the actual sentiment of public opinion.

6.8. Case 8: Uttar Pradesh mob violence

Introduction: In Uttar Pradesh (India), a police constable was stoned to death in Ghazipur district, the head constable’s death is the second such incident in a month. A police inspector was killed in Bulandshahr when he tried to stop a mob from keeping cattle carcasses to block traffic. While reacting to such horrific incidents, BJP MP Udit Raj called it an isolated incident and refused to admit law and order lapse saying that such incidents can happen in a huge state like UP. Nearly 1000 tweets were collected using the term “UP mob violence”, they were used for analysis.

Analysis: While using SVNS, we obtained $UP_{SVNS} = \langle 0.104, 0.223, 0.673\rangle$; it shows that the public opinion was a negative one, whereas while using TRINS; $UP_{TRINS} = \langle 0, 0.104, 0.223, 0.67, 0.003\rangle$; it was found that no one had a strong positive opinion and very little population had a positive opinion, the majority was negative. When analysis was further carried out using MRNS, the result was $UP_{MRNS} = \langle 0, 0.013, 0.091, 0.223, 0.434, 0.236, 0.003\rangle$; most of the positive also turns out to be indeterminate positive. It is clearly seen that much of the negative opinion is negative indeterminate. The overall major opinion was indeterminate in nature. MRNS highlights the indeterminacy involved in this case accurately, whereas SVNS failed to capture the data with this amount of accuracy as shown in Table 5.

6.9. Case 9: Trump wall shutdown

Introduction: A critical feature of the standoff that has led US President Trump in 2018 to partially shut down the government is for funding of billions of dollars to build wall on the US-Mexico border.

Analysis: The result of the analysis is given in Table 6. While using SVNS for analysis it is seen that majority have an indeterminate opinion about the wall and related forced shut down of government. The resultant obtained is $TW_{SVNS} = \langle 0.186, 0.581, 0.229\rangle$; it is also noted that the difference between positive and negative is minimal. While TRINS was used for analysis, the resultant is $TW_{TRINS} = \langle 0.03, 0.156, 0.581, 0.223, 0.006\rangle$; it is seen from the membership values that not many people felt strongly negative about it, but people felt strongly positive about it. When MRNS was used to analyze the same set of tweets, the resultant obtained is $TW_{MRNS} = \langle 0.02, 0.045, 0.121, 0.581, 0.199, 0.024, 0.006\rangle$; it is clearly seen that majority fell into indeterminate, positive indeterminate and negative indeterminate membership intervals. Meagre value is associated with positive or negative membership. This clearly proves that people neither hold a positive nor negative opinion that is they are not able to make up their mind.

6.10. Case 10: Yellow vests protest

Introduction: The yellow vests movement is a populist political movement from grass roots for economic justice which started in France in 2018. Regular mass demonstrations against the French government began on 17 November 2018. The movement is motivated by high fuel prices and high cost of living and together with the burden on the working and middle classes due to government’s tax reforms. The protests had turned violent and tear gas was used on the protesters. The tweets taken for analysis are right after this news broke out.

Analysis: The majority opinion about yellow vests protest and the police action was negative in nature when analysis with SVNS was carried out, the resultant SVNS was $YV_{SVNS} = \langle 0.018, 0.319, 0.663\rangle$. When TRINS is used, the resultant TRINS is $YV_{TRINS} = \langle 0, 0.018, 0.319, 0.662, 0.001\rangle$; it is seen that there is a meagre strong negative and no strong positive. When MRNS is used the resultant obtained is $YV_{MRNS} = \langle 0, 0, 0.018, 0.319, 0.654, 0.008, 0.001\rangle$; it is clearly seen that most people had a negative indeterminate opinion. MRNS provides a clear result about the indeterminacy involved and is tabulated in Table 6.

7. Comparison and discussions

Till date neutrosophy has not been used for sentiment analysis of twitter data. A brief comparison of conventional, fuzzy and neutrosophic sentiment analysis is illustrated by Table 7 to highlight the results of our study.

If fuzzy set theory based sentiment analysis given in [3] is carried on the same set of tweets, only positive and negative memberships will get mapped, indeterminacy concept will not be dealt with, hence there will be information loss, and results will not be accurate. Henceforth it is clear that when splitting of these three memberships is done, better accuracy is achieved. There is an increased complexity in handling seven memberships when MRNS is taken, but better accuracy is achieved.

A comprehensive tabulation of all the ten topics is given in Table 8 with respect to fuzzy sentiment analysis, analysis using SVNS, TRINS, and MRNS. It is evident from Table 8 each of the case studies given above, that MRNS gives more accurate results than TRINS or SVNS or fuzzy.

Table 7
Comparison of different sentiment analysis approaches

Normal sentiment analysis	Fuzzy sentiment analysis	Neutrosophic sentiment analysis
Major overall sentiment	Percentage of positive sentiment Percentage of negative sentiment	Proportion of positive sentiment Proportion of indeterminate sentiment Proportion of negative sentiment

Table 8
Neutrosophic representation of each case study

Type of NS	Positive			Neutral	Negative		
			Case 1: Farm loan Wavier				
Fuzzy		0.857				0.143	
SVNS		0.419		0.438		0.143	
TRINS	0.048	0.371		0.438		0.136	0.007
MRNS	0.026	0.164	0.229	0.438	0.111	0.255	0.007
			Case 2: Onion Price				
Fuzzy		0.917				0.083	
SVNS		0.253		0.664		0.083	
TRINS	0.026	0.228		0.664		0.076	0.006
MRNS	0.012	0.037	0.205	0.664	0.059	0.017	0.006
			Case 3: Women Reservation Bill				
Fuzzy		0.249				0.751	
SVNS		0.084		0.165		0.751	
TRINS	0.005	0.079		0.165		0.750	0.001
MRNS	0.004	0.038	0.042	0.165	0.019	0.731	0.001
			Case 4: Triple Talaq Bill				
Fuzzy		0.854				0.146	
SVNS		0.316		0.538		0.146	
TRINS	0.153	0.163		0.538		0.145	0.001
MRNS	0.149	0.11	0.057	0.538	0.11	0.035	0.001
			Case 5: POCSO Act Death Penalty				
Fuzzy		0.947				0.053	
SVNS		0.909		0.038		0.053	
TRINS	0.01	0.899		0.038		0.053	0
MRNS	0.01	0.863	0.036	0.038	0.052	0.001	0
			Case 6: #Me too Movement				
Fuzzy		0.853				0.147	
SVNS		0.562		0.291		0.147	
TRINS	0.04	0.522		0.291		0.132	0.015
MRNS	0.014	0.375	0.173	0.291	0.086	0.046	0.015
			Case 7: Foreign Trip PM				
Fuzzy		0.597				0.403	
SVNS		0.474		0.123		0.403	
TRINS	0.001	0.473		0.123		0.399	0.004
MRNS	0.001	0.003	0.470	0.123	0.180	0.219	0.004
			Case 8: UP Mob Violence				
Fuzzy		0.327				0.673	
SVNS		0.104		0.223		0.673	
TRINS	0	0.104		0.223		0.670	0.003
MRNS	0	0.013	0.091	0.223	0.434	0.236	0.003
			Case 9: Trump Wall Shutdown				
Fuzzy		0.769				0.231	
SVNS		0.188		0.581		0.231	
TRINS	0.03	0.158		0.581		0.224	0.007
MRNS	0.02	0.046	0.122	0.581	0.199	0.025	0.007
			Case 10: Yellow Vest Protest				
Fuzzy		0.337				0.663	
SVNS		0.018		0.319		0.663	
TRINS	0	0.018		0.319		0.662	0.001
MRNS	0	0	0.018	0.319	0.654	0.008	0.001

To further enhance the results and enable comparison, indeterminacy neutral operator (Definition 13) was used on MRNS to convert it to SVNS kind of representation. The comparison of fuzzy, SVNS and indeterminate neutral of MRNS is given in Table 9, it is seen in several cases that a positive opinion or a negative opinion is actually indeterminate in nature. Also in some cases a positive opinion turns out to indeterminate positive and a negative opinion has an edge over positive opinion.

It is clearly seen in case of “foreign trips of PM” and “Me too”, the overall positive opinion turns out to be indeterminate in nature when analysed with MRNS. Similarly the overall negative opinion of “yellow vest protest” and “UP mob violence” turns out to be indeterminate opinion with analysed with MRNS. In case of “farm loan wavier” the opinion changes from positive to negative, and in case of ”Trump wall”, the opinion changes from more negative to more positive when analysed with MRNS. In the other cases, like “Women reservation bill”, ”triple talaq bill”, “POSCO act” and “onion price”, the opinion obtained with SVNS and TRINS is confirmed using MRNS for analysis. The trendy that is predicted by fuzzy is sometimes not exactly accurate, mostly when indeterminacy happens to be the actual trend.

Table 9
Comparison between Fuzzy, SVNS and Indeterminate neutral of MRNS

Case no	Fuzzy	SVNS	Indeterminate neutral MRNS
1	Positive	Indeterminate; More positive than negative	⟨0.19, 0.778, 0.263⟩ Indeterminate more negative than positive
2	Positive	Indeterminate more positive than negative	⟨0.048, 0.928, 0.023⟩ Indeterminate; meagre negative and positive
3	Negative	Negative	⟨0.042, 0.226, 0.732⟩ Negative; negligible positive
4	Positive	Indeterminate more positive than negative	⟨0.258, 0.705, 0.03⟩ Indeterminate; more positive than negative
5	Positive	Positive	⟨0.873, 0.126, 0.001⟩ Positive; no negative
6	Positive	Positive	⟨0.387, 0.548, 0.059⟩ Indeterminate more positive than negative
7	Positive	Positive	⟨0.004, 0.773, 0.223⟩ Indeterminate more negative than positive
8	Negative	Negative	⟨0.014, 0.743, 0.239⟩ Indeterminate more negative than positive
9	Positive	Indeterminate more negative than positive	⟨0.065, 0.802, 0.03⟩ Indeterminate; more positive than negative
10	Negative	Negative	⟨0, 0.991, 0.009⟩ Indeterminate; negligible negative and no positive

Table 10
Case study of Amazon with SemEval 2017

Values	Positive			Neutral	Negative		
Original Values	0.05	0.52		0.27		0.15	0.01
TRINS	0.06	0.47		0.28		0.16	0.03
MRNS	0.06	0.03	0.44	0.28	0.13	0.06	0

7.1. Working on SemEval2017 Task 4 dataset

Recently, SemEval-2017 Task 4 [62], the sentiment analysis in Twitter task was conducted for the fifth year. Several tasks are given to the participating teams. It includes identifying the overall sentiment of the tweet, sentiment towards a topic with classification on a two point and on a five-point ordinal scale. The five-point scale used was strongly positive, weakly positive, neutral, weakly negative, and strongly negative. This five-point scale is similar to the TRINS concept in neutrosophy.

SemEval-2017 Task 4: Sentiment analysis in twitter is an important benchmark in sentiment analysis of twitter data. The various subtasks associated with it are subtasks A, B, C, D and E. Subtask A is about message polarity classification, subtasks B-C are about topic based message polarity classification into two-point and five-point scale and subtasks D-E are tweet quantification into two-point and five-point scale.

Our model using neutrosophy is closely related with polarity classification and tweet quantification (Subtask E), it works on three-point scale (SVNS), five-point scale (TRINS) and seven point scale (MRNS). An in-depth analysis of using neutrosophy on SemEval 2017 dataset will be done later for all the 5 subtasks of task 4, since almost all participating teams had made use of several machine learning algorithms. Our existing model is not making use of any machine learning algorithm and hence is at disadvantage for complete comparison with the baselines and benchmarks set by them. The possible comparative analysis can be carried out on five-point scale (TRINS) as subtask E of tweet quantification is on a five-point scale. For illustrative purpose we have worked with just one topic from SemEval-2017 dataset. We have dealt with 100 tweets related to the topic "Amazon" from the SemEval-2017 training dataset. The original values are taken from the polarity value given in the dataset, it is mapped to the percentage of -2, -1, 0, 1, 2 present in the polarity column. The values obtained for subtask E in case of topic "Amazon" using TRINS is tabulated below in Table 10.

The TRINS values are obtained by polarity classification and tweet quantification of the tweets in the dataset. The values obtained are ⟨0.06, 0.47, 0.28, 0.16, 0.03⟩. There is a slight variation when compared to the expected values given by the experts. The same data set was used for analysis using MRNS and the following was obtained ⟨0.06, 0.03, 0.44, 0.28, 0.13, 0.06, 0⟩ on a 7-point scale. This cannot be compared with the expected values of the SemEval dataset, since they are in different point scale, but it is pertinent to mention that the 7-point scale gives more detailed accuracy to the analysis. For example, it can be seen that none had a strongly negative opinion about Amazon, this is not captured with TRINS.

Since SemEval-2017 dataset is an interesting, important and bench-marked dataset to work on, a detailed analysis of the SemEval 2017 dataset using neutrosophy will be carried out soon. It shall be a separate analysis on its own along with creation of appropriate machine learning and natural language processing algorithms which make use of neutrosophy. This exciting and innovative study will be taken up in future. For further study, we are planning to work on datasets from [33–35], combined together with appropriate NLP and machine learning algorithms that are based on neutrosophy.

8. Results and future study

Both conventional and fuzzy sentiment analysis fail to capture the indeterminacy and neutrality that is present in the content. To handle the indeterminacy, neutrosophy is used for analysis of tweets. In this paper, a new concept called MRNS with two positive memberships, three indeterminate memberships and two negative memberships is defined. Its properties and various operators are discussed. For the purposes of this research work, ten subjects (data sets) which are political or social in nature were taken for sentiment analysis and for each case study, 1000 tweets were collected and used for analysis. The collection was carried out through a specifically created twitter API, through which tweets were extracted, programming was carried out using Python and necessary libraries for NLP. It is the first time that refined neutrosophic sets have been used for sentiment analysis of tweets. Three neutrosophic sets namely SVNS, TRINS and MRNS were used to analyse the tweets. In each case, it was clearly seen that MRNS gives a better and accurate result when compared to SVNS or TRINS. An illustrative comparison with one topic from SemEval-2017 was also done with TRINS and MRNS. From the discussions and comparisons carried out, it is seen that when MRNS is used the best accuracy was obtained in sentiment analysis.

References

Speier, H., 1950. Historical development of public opinion. American Journal of Sociology 55 (4), 376–388.
Mäntylä, M.V., Graziotin, D., Kuutila, M., 2018. The evolution of sentiment analysis-a review of research topics, venues, and top cited papers. Computer Science Review 27, 16–32, doi:https://doi.org/10.1016/j.cosrev.2017.10.002. URL http://www.sciencedirect.com/science/article/pii/S1574013717300606.
Haque, A., et al., 2014. Sentiment analysis by using fuzzy logic. International Journal of Computer Science, Engineering and Information Technology 4 (1), 33–48.

Zadeh, L.A., 1965. Fuzzy sets. Information and control 8 (3), 338–353.
Atanassov, K.T., 1986. Intuitionistic fuzzy sets. Fuzzy sets and Systems 20 (1), 87–96, doi:https://doi.org/10.1016/S0165-0114(86)80034-3.
Smarandache, F., 2000. A Unifying Field in Logics: Neutrosophic Logic. Neutrosophy, Neutrosophic Set, Probability, and Statistics. American Research Press, Rehoboth, URL https://arxiv.org/pdf/math/0101228.
Wang, H., Smarandache, F., Zhang, Y., Sunderraman, R., 2010. Single valued neutrosophic sets. Review, 10, URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.309.9470&rep=rep1&type=pdf.
Cheng, H.-D., Guo, Y., 2008. A new neutrosophic approach to image thresholding. New Mathematics and Natural Computation 4 (03), 291–308, doi:10.1142/S1793005708001082.
Sengur, A., Guo, Y., 2011. Color texture image segmentation based on neutrosophic set and wavelet transformation. Computer Vision and Image Understanding 115 (8), 1134–1144, doi:10.1016/j.cviu.2011.04.001.
Zhang, M., Zhang, L., Cheng, H., 2010. A neutrosophic approach to image segmentation based on watershed method. Signal Processing 90 (5), 1510–1517, doi:10.1016/j.sigpro.2009.10.021.
Liu, P., Wang, Y., 2014. Multiple attribute decision-making method based on single-valued neutrosophic normalized weighted bonferroni mean. Neural Computing and Applications 25 (7-8), 2001–2010, doi:10.1007/s00521-014-1688-8.
Liu, P., Shi, L., 2015. The generalized hybrid weighted average operator based on interval neutrosophic hesitant set and its application to multiple attribute decision making. Neural Computing and Applications 26 (2), 457–471, doi:10.1007/s00521-014-1736-4.
Liu, P., Teng, F., 2017. Multiple attribute group decision making methods based on some normal neutrosophic number heronian mean operators. Journal of Intelligent & Fuzzy Systems 32 (3), 2375–2391, doi:10.3233/JIFS-16345.
Liu, P., Li, H., 2017. Multiple attribute decision-making method based on some normal neutrosophic bonferroni mean operators. Neural Computing and Applications, 179–194, doi:10.1007/s00521-015-2048-z.
Ye, J., 2013. Multicriteria decision-making method using the correlation coefficient under single-valued neutrosophic environment. International Journal of General Systems 42 (4), 386–394, doi:10.1080/03081079.2012.761609.
Ye, J., 2014. A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets. Journal of Intelligent & Fuzzy Systems 26 (5), 2459–2466, doi:10.3233/IFS-130916.
Ye, J., 2014. Single valued neutrosophic cross-entropy for multicriteria decision making problems. Applied Mathematical Modelling 38 (3), 1170–1175, doi:10.1016/j.apm.2013.07.020.
Ye, J., 2014. Similarity measures between interval neutrosophic sets and their applications in multicriteria decision-making. Journal of Intelligent & Fuzzy Systems 26 (1), 165–172, doi:10.3233/IFS-120724.
A. Salama, A. Haitham, A. Manie, M. Lotfy, Utilizing neutrosophic set in social network analysis e-learning systems, International Journal of Information Science and Intelligent System 3 (2), 2014, 1-12. URL http://fs.gallup.unm.edu/SN/Neutro-UtilizingNeutrosophicSet.pdf.
Vasantha, W., Smarandache, F., 2003. Fuzzy cognitive maps and neutrosophic cognitive maps. Xiquan, URL https://arxiv.org/pdf/math/0311063.
W. Vasantha, F. Smarandache, Analysis of social aspects of migrant labourers living with hiv/aids using fuzzy theory and neutrosophic cognitive maps: With special reference to rural tamil nadu in india, arXiv preprint math/0406304.
Kandasamy, I., 2018. Double-valued neutrosophic sets, their minimum spanning trees, and clustering algorithm. Journal of Intelligent Systems 27 (2), 163–182, doi:10.1515/jisys-2016-0088.
Kandasamy, I., Smarandache, 2016. Multicriteria decision making using double refined indeterminacy neutrosophic cross entropy and indeterminacy based cross entropy. Applied Mechanics and Materials 859, 129–143, doi:10.4028/www.scientific.net/AMM.859.129.
Q. Khan, P. Liu, T. Mahmood, Some generalized dice measures for double-valued neutrosophic sets and their applications, Mathematics 6 (7). doi:10.3390/math6070121. URL http://www.mdpi.com/2227-7390/6/7/121.
Kandasamy, I., Smarandache, F., 2016. Triple refined indeterminate neutrosophic sets for personality classification. In: Computational Intelligence (SSCI), 2016 IEEE Symposium Series on, IEEE. doi:10.1109/SSCI.2016.7850153, pp. 1–8.
Kandasamy, I., Vasantha, W.B., Obbineni, J., Smarandache, F., 2019. Indeterminate likert scaling. Soft Computing, http://dx.doi.org/10.1007/s00500-019-04372-x.
I. Kandasamy, "Indeterminate likert scale - sample dataset - customer feedback of restaurant", Mendeley Data, v1 doi:https://doi.org/10.17632/ywjxpyw95w.1.
Dave, K., Lawrence, S., Pennock, D.M., 2003. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. Proceedings of the 12th international conference on World Wide Web, ACM, 519–528.
Nassirtoussi, A.K., Aghabozorgi, S., Wah, T.Y., Ngo, D.C.L., 2014. Text mining for market prediction: A systematic review. Expert Systems with Applications 41 (16), 7653–7670.
Burnap, P., Williams, M.L., Sloan, L., Rana, O., Housley, W., Edwards, A., Knight, V., Procter, R., Voss, A., 2014. Tweeting the terror: modelling the social media reaction to the woolwich terrorist attack. Social Network Analysis and Mining 4 (1), 206.
Hogenboom, A., Heerschop, B., Frasincar, F., Kaymak, U., de Jong, F., 2014. Multilingual support for lexicon-based sentiment analysis guided by semantics. Decision support systems 62, 43–53.
Munezero, M.D., Montero, C.S., Sutinen, E., Pajunen, J., 2014. Are they different? affect, feeling, emotion, sentiment, and opinion detection in text. IEEE transactions on affective computing 5 (2), 101–111.
Ghosh, A., Li, G., Veale, T., Rosso, P., Shutova, E., Barnden, J., Reyes, A., 2015. Semeval-2015 task 11: Sentiment analysis of figurative language in twitter. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), 470–478.
Reyes, A., Rosso, P., 2014. On the difficulty of automatically detecting irony: beyond a simple case of negation. Knowledge and Information Systems 40 (3), 595–614.
Reyes, A., Rosso, P., 2011. Mining subjective knowledge from customer reviews: A specific case of irony detection. In: Proceedings of the 2nd workshop on computational approaches to subjectivity and sentiment analysis, Association for Computational Linguistics, pp. 118–124.
Farias, D.H., Rosso, P., 2017. Chapter 7 - irony, sarcasm, and sentiment analysis. In: Pozzi, F.A., Fersini, E., Messina, E., Liu, B. (Eds.), Sentiment Analysis in Social Networks. Morgan Kaufmann, Boston, pp. 113–128, doi:https://doi.org/10.1016/B978-0-12-804412-4. 00007-3. URL http://www.sciencedirect.com/science/article/pii/B9780128044124000073.
Zhang, S., Zhang, X., Chan, J., Rosso, P., 2019. Irony detection via sentiment-based transfer learning. Information Processing & Management 56 (5), 1633–1644.
Ma, Y., Peng, H., Cambria, E., 2018. Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive lstm. Thirty-Second AAAI Conference on Artificial Intelligence.
Maas, A.L., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y., Potts, C., 2011. Learning word vectors for sentiment analysis. In: Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1, Association for Computational Linguistics, pp. 142–150.
Yin, H., Liu, P., Zhu, Z., Li, W., Wang, Q., 2019. Capsule network with identifying transferable knowledge for cross-domain sentiment classification. IEEE Access 7, 153171–153182.
Zhang, L., Wang, S., Liu, B., 2018. Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 8 (4), e1253.
Young, T., Hazarika, D., Poria, S., Cambria, E., 2018. Recent trends in deep learning based natural language processing. Ieee Computational Intelligence Magazine 13 (3), 55–75.
Jefferson, C., Liu, H., Cocea, M., 2017. Fuzzy approach for sentiment analysis, doi:10.1109/FUZZ-IEEE.2017.8015577.
I. Kandasamy, W.B. Vasantha, N. Mathur, M. Bisht, F. Smarandache, Chapter 6 sentiment analysis of the metoo movement using neutrosophy: Application of single-valued neutrosophic sets, In: F. A. Pozzi, E. Fersini, E. Messina, B. Liu (Eds.), Optimization Theory Based on Neutrosophic and Plithogenic Sets, Elsevier, 2020. doi:s. https://doi.org/10.1016/B978-0-12-819670-0. 00006-8.
Smarandache, F., 2013. n-valued refined neutrosophic logic and its applications in physics. Progress in Physics 4, 143–146, URL https://arxiv.org/pdf/1407.1041.
W. B. Vasantha, I. Kandasamy, F. Smarandache, A classical group of neutrosophic triplet groups using $Z_{2p}, \times$, Symmetry 10 (6). doi:10.3390/sym10060194. URL http://www.mdpi.com/2073-8994/10/6/194.
W. B. Vasantha, I. Kandasamy, F. Smarandache, Neutrosophic duplets of $Z_{pn}, \times$ and $Z_{pq}, \times$ and their properties, Symmetry 10 (8). doi:10.3390/sym10080345. URL http://www.mdpi.com/2073-8994/10/8/345.
Vasantha, W., Kandasamy, I., Smarandache, F., 2018. Algebraic structure of neutrosophic duplets in neutrosophic rings. Neutrsophic Sets and Systems 23, 85–95.
Ali, M., Thanh, N.D., Van Minh, N., et al., 2018. A neutrosophic recommender system for medical diagnosis based on algebraic neutrosophic measures. Applied Soft Computing 71, 1054–1071.
Nguyen, G.N., Ashour, A.S., Dey, N., et al., 2019. A survey of the state-of-the-arts on neutrosophic sets in biomedical diagnoses. International Journal of Machine Learning and Cybernetics 10 (1), 1–13.
Ali, M., Khan, M., Tung, N.T., et al., 2018. Segmentation of dental x-ray images in medical imaging using neutrosophic orthogonal matrices. Expert Systems with Applications 91, 434–441.
Abdel-Basset, M., Manogaran, G., Gamal, A., Smarandache, F., 2019. A group decision making framework based on neutrosophic topsis approach for smart medical device selection. Journal of medical systems 43 (2), 38.
Broumi, S., Bakali, A., Talea, M., Smarandache, F., Singh, P.K., Uluçay, V., Khan, M., 2019. Bipolar complex neutrosophic sets and its application in decision making problem. In: Fuzzy Multi-criteria Decision-Making Using Neutrosophic Sets, Springer, pp. 677–710.
Ji, P., Zhang, H.-y., Wang, J.-q., 2018. A projection-based todim method under multi-valued neutrosophic environments and its application in personnel selection. Neural Computing and Applications 29 (1), 221–234.
Abdel-Baset, M., Chang, V., Gamal, A., 2019. Evaluation of the green supply chain management practices: A novel neutrosophic approach. Computers in Industry 108, 210–220.
Nirmal, N., Bhatt, M., 2019. Development of fuzzy-single valued neutrosophic madm technique to improve performance in manufacturing and supply chain functions. In: Fuzzy Multi-criteria Decision-Making Using Neutrosophic Sets, Springer, pp. 711–729.
N. A. Nabeeh, M. Abdel-Basset, H. A. El-Ghareeb, A. Aboelfetouh, Neutrosophic multicriteria decision making approach for iot-based enterprises, IEEE Access 7 (2019) 59559-59574.

Smarandache, F., 2018. Neutropsychic Personality: A mathematical approach to psychology. Infinite Study.

Vasantha, W., Smarandache, F., 2014. Fuzzy Neutrosophic Models for Social Scientists. Infinite Study.

Russell, C.J., Bobko, P., 1992. Moderated regression analysis and likert scales: Too coarse for comfort. Journal of Applied Psychology 77 (3), 336, URL https://www.ncbi.nlm.nih.gov/pubmed/1601825.

I. Kandasamy, "Tweets on political and social issues for analysis using neutrosophic sets", Mendeley Data, v1doi:https://doi.org/10.17632/fnzmfgy2bd.1.

Rosenthal, S., Farra, N., Nakov, P., 2017. SemEval-2017 task 4: Sentiment analysis in Twitter. In: Proceedings of the 11th International Workshop on Semantic Evaluation, SemEval '17, Association for Computational Linguistics, Vancouver, Canada.