Senti-NSetPSO: large-sized document-level sentiment analysis using Neutrosophic Set and particle swarm optimization

Article *in* Soft Computing · July 2019

DOI: 10.1007/s00500-019-04209-7

4 authors, including:

Amita Jain
Ambedkar Institute of Advanced Communication Technologies and Research

Charu Gupta
Bhagwan Parshuram Institute of Technology

Soft Computing
https://doi.org/10.1007/s00500-019-04209-7

FOCUS

Senti-NSetPSO: large-sized document-level sentiment analysis using Neutrosophic Set and particle swarm optimization

Amita Jain[1] · Basanti Pal Nandi[2] · Charu Gupta[3] · Devendra Kumar Tayal[4]

Abstract

In the last decade, opinion mining has been explored by using various machine learning methods. In the literature, document-level sentiment analysis has been majorly dealt with short-sized text only. For large-sized text, document-level sentiment analysis has never been dealt. In this paper, a hybrid framework named as "*Senti-NSetPSO*" is proposed to analyse large-sized text. *Senti-NSetPSO* comprises of two classifiers: binary and ternary based on hybridization of particle swarm optimization (PSO) with Neutrosophic Set. This method is suitable to classify large-sized text having more than 25 kb of size. Swarm size generated from large text can give a suitable measurement for implementation of PSO convergence. The proposed approach is trained and tested for large-sized text collected from Blitzer, aclIMDb, Polarity and Subjective Dataset. The proposed method establishes a co-relation between sentiment analysis and Neutrosophic Set. On Blitzer, aclIMDb and Polarity dataset, the model acquires satisfactory accuracy by ternary classifier. The accuracy of ternary classifier of the proposed framework shows significant improvement than review paper classifier present in the literature.

Keywords Evolutionary algorithm · Neutrosophic Set · Opinion mining · Particle swarm optimization · Sentiment analysis

1 Introduction

The field of opinion mining has grown and matured enough to understand the relationship between buyers and sellers (manufacturers) in e-commerce. It has been demonstrated that the online user reviews play a vital role in influencing the buyers to buy a desired product. For this reason, the polarity detection has become an overly blooming field of research. In this regard, researchers have applied different techniques for polarity detection (classification) of a given text. The techniques applied depend upon the text nature, which can be divided into two categories: *Subjective text and Objective text*. Subjective text displays emotion of the author, while objective text shares a fact. Personal views, feelings or opinion about some facts lead to a polarity of mind, may it be positive, negative or neutral or it can express emotions. Authors of previous research have captured such classification and sentiment score with varied algorithm which includes uni-gram, bi-gram, support vector machine, Naive Bayes, maximum entropy, neural network, vector kernel, etc. Apart from this classification of text, the underlying sentiment can be analysed from three basic levels—*Document-level, Sentence level* and *Aspect level* (Agarwal et al. 2015; Anne et al. 2017; Joshi and Nigam 2011). The polarity of emotions can be determined on document-level or on sentence level as per the text present in whole document or in a single sentence. At *aspect level*, the text provides positive polarity for a definite aspect (feature) where as it provides negative polarity for other aspects. Each of the aspects of analysis or classification has its own direction for merits. Document-level analysis gives polarity of a document having positive or negative sentiment over a text consisting more than one sentence (Pu et al. 2019; Sharma et al. 2014; Yessenalina et al. 2010). Generally any opinion about products, movies

Communicated by O. Castillo, D. K. Jana.

✉ Amita Jain
amita_jain_17@yahoo.com

1 Ambedkar Institute of Advanced Communication Technologies and Research, Delhi, India

2 Guru Tegh Bahadur Institute of Technology, Delhi, India

3 Bhagwan Parshuram Institute of Technology, Delhi, India

4 Indira Gandhi Delhi Technical University for Women, Delhi, India

Published online: 16 July 2019

or other facts mentioned by people contains a text with at least 2–3 sentences. For opinion mining, it is significant to generate a polarity detection system for such documents. In this paper, the proposed method focuses on large-sized document-level analysis.

Recently, researchers have shown comparison of machine learning algorithms on single instance of a review to detect the opinion of the text (Tripathy et al. 2017). They have reached to greater than 90% accuracy on IMDb movie review dataset by applying SVM with ANN as binary classifier. This method outperformed with other methods applied earlier such as application of only SVM or Naive Bayes (Pang et al. 2002). It is observed that the methods in the literature work are on the approach of a single instance of review which is merely 2–3 lines of opinion. State-of-the-art method for short text has been dealt in the work of Bouazizi and Ohtsuki (2017) using a multiclass ternary classifier for analysing twitter dataset. This approach again suffers from the same limitation of being used only on very short text twitter comment as opposed to large-sized text which is dealt in Senti-NSetPSO.

This gives a forceful motivation to work with machine learning algorithms which can work on large-sized text. The aim of the proposed work is *twofolds:* to understand the intricacies with large-sized text and to determine the polarity of the sentiment with high accuracy. The accuracy of the proposed approach is measured using standard methods like precision, recall and F-measure. It has been analysed that the use of soft-computing techniques is widely dealt in the literature primarily the fuzzy logic approach. It is seen that the component membership and non-membership are dependent. Sentiment analysis is among some of the problems where membership and non-membership are not dependent.

Neutrosophic logic helps in providing the model for such kind of problems. Neutrosophic approach defines a tuple with a three-valued logic namely Truth, Indeterminacy and Falsehood. It is worth understanding that these three measures are independent of each other unlike fuzzy logic in classification. Neutrosophic classifier has the power to capture the indeterminacy beyond the crisp set of truth or false. Neutrosophic Set for further development is used for neutral classes of the document. Most recent development shows that PSO hybridization with Neutrosophic logic has been tried in two different cases: *one*, for CT image segmentation and *two*, for information theory measures for forest fire (Anter and Hassenian 2018; Gafar et al. 2018). In this paper, the variation of this algorithm (PSO with Neutrosophic Set) is analysed for sentiment analysis. PSO as evolutionary algorithm is chosen (correlated with the literature: Sect. 2) as it is simpler in computation as well as efficient in terms of time and accuracy than other evolutionary algorithms.

Senti-NSetPSO is the first approach to analyse large-sized text as no other work in available in the literature. In this paper, the objective is to get the overall polarity of a document with the application of PSO on Neutrosophic values of the words of the review texts. Application of nature-inspired algorithms in sentiment analysis is in their infancy still now, as well as fuzzy logic or the extension of fuzzy logic such as Neutrosophic logic, needs more exposure on this area.

In this paper, Senti-NSetPSO uses Neutrosophic concept for values of seeded words to optimize through PSO to get an overall opinion. The polarity score of words from the text is generated using SentiWordNet. The three Neutrosophic values (positive, negative and indeterminacy) for each meaningful word are then fed to the PSO optimizer. Finally, optimized values of the document are classified for polarity determination.

In sentiment analysis, Neutrosophic concept has not been applied earlier. In this paper, PSO is applied directly on Neutrosophic values of the document, whereas in earlier approaches, PSO has been used for feature selection. The main contribution of the proposed work is to classify large-sized document where feature selection is a cumbersome and time-consuming task. Here, Neutrosophic values give a direct approach to apply PSO to categorize the document easily. The approach to classify large-sized documents has been modelled in *two parts*: ternary classifier and binary classifier. With the proposed approach, the ternary classifier showed improved results than machine learning algorithms applied earlier in the literature for review sentiment classifier.

The work is organized as follows: Sect. 2 gives motivation of the proposed work. Section 3 discusses the related work available in literature. Section 4 gives an insight to Neutrosophic approach. Section 5 focuses on the proposed methodology of Senti-NSetPSO. Section 6 analyses the experimental results. Conclusion and future work are presented in Sect. 7.

2 Motivation

This section describes the motivation of the proposed approach and the importance of the Neutrosophic logic. This section gives a clear understanding of the concept of Neutrosophic logic in developing the proposed framework.

In crisp logic, there are two extreme values as True and False where the data are assigned 0 or 1 as membership of a system. In real-world applications, the parameters defining the domain may not be captured within these two values. The conditions like usually, few, likely, not likely, mostly, unlikely, probably, nearly cannot be completely understood by truth or falsehood. Fuzzy set proposed by Zadeh gives a natural implementation for such situation.

Fuzzy logic applied to a variable allowed values with a *degree of membership* (Cheung et al. 2005; Kia et al. 2009). The membership value is between 0.0 and 1.0 and any intermediate value of a variable "x" say 0.2 implies that "x" is true for 20% of the time.

Neutrosophic logic (Smarandache 2016) is an extension of fuzzy logic which gives a value of indeterminacy besides truth and falsehood. Neutrosophic implies "The knowledge of neutral thought" (Ashbacher 2002). Neutrosophic can be understood by considering some real-world examples like for games: win/loss/draw, for voting: pro/cont/blank, for numbers: positive/negative/zero, etc. The inherent neutral or indeterminate value due to system imbalance can be measured and decision needs to be taken by human expert or rule-based systems in such situation. In Neutrosophic logic, sum of the components need not necessarily to be 1 but any value between -0 to $3+$. It can define paradoxes where variables can be truth or false at the same time which is not possible in classic fuzzy system.

In this paper, large documents (more than 25 KB size) are selected for classification. On large-sized documents ternary classifiers has not been applied in the state-of-the-art. For the first time, Neutrosophic values are used for sentiment analysis as a ternary and binary classifier application.

The inherent motivation is to work on large-sized documents like news articles, reports of any investigation, review comments, etc., which have not been taken for classification purposes. As discussed in Sect. 1, for such large documents, feature selection is a time-consuming process. Moreover, it is applicable in cross-domain texts where only positive, negative or neutral values will be relevant. Cross-domain texts are still not used for sentiment analysis purpose. The proposed model can be used for classification purpose of cross-domain texts as well.

3 Related work

Opinion mining to get the decision values for movie reviews, electronic product reviews, election opinion has become an integrated part of our life. Fuzzy logic has been applied to evaluate the review tendency towards a score which will give the mass opinion (Ali et al. 2017; Bing and Chan 2014; Dragoni et al. 2015). The fuzzy behaviour of objective sentences can be categorized in positive, negative and indeterminacy values. Li and Tsai (2013) used fuzzy conceptual model for text mining. They used formal concepts which are relevant to new document to determine the cohesion of the document to the specified classes. The research was conducted for single class problem. Fuzzy formal concept analysis construct lattice for the document and term features which are used for further determination of classes of the new document.

Document-level sentiment classification has been explored using various machine learning algorithms. Tripathy et al. (2017) tested these models on IMDb and polarity dataset. SVM is used for selecting the best feature, and then these selected features are used as the input to ANN. ANN classifies the review as positive or negative. Many researchers have used hybridization of machine learning techniques such as Naive Bayes, decision tree, SVM with n-gram model and genetic algorithms for the same purpose.

Feature selection based on metaheuristic algorithms has increased the accuracy for document classification. Swarm intelligent algorithms have been used to perform *two* basic tasks: first, select prominent feature, second, to remove noisy and overlapping features. A survey on application of swarm intelligence algorithm for sentiment analysis has been done which explores the advantages of these algorithms for feature selection (Kumar et al. 2016). ABC with SVM, ABC with Naive Bayes, FURIA, RIDOR, Hybrid PSO, ACO with decision tree, PSO with SVM, PSO with CRF are various other feature optimization algorithms which show considerable amount of increase in accuracy for sentiment classification. PSO and ACO have the proven power to get state-of-the-art accuracy with about 36% less selected feature set. In this paper, the convergence modality of PSO on Neutrosophic score of document concepts is studied.

Neutrosophic means the "study of neutrality" and it is in its very early stage of development where very few research articles have used it for real-time applications. Following this concept, Colhon et al. (2017) have given a neutral score of a word to determine *how much* neutral it is. A word has *two* possibilities of neutrality—*first,* its score of Truth and Falsehood is zero with Indeterminacy value as 1 and second, to have a balanced Truth and Falsehood score which leads to neutrality. In the literature, there are only few applications of Neutrosophic value. For instance, Ansari et al. (2013) applied this logic to classify Iris dataset based on some predetermined rules. Some multiclass classifiers have been developed. For example, Bouazizi and Ohtsuki (2017) gave a pattern-based multiclass classifier on Twitter dataset. Their ternary classifier gives 70.1% of accuracy, whereas multiclass classifier gives 60.2% result. Nagarajan and Gandhi (2019) have reached 90.4% accuracy on ternary classifier over Twitter dataset. Their hybrid method PSO and Genetic Algorithm for feature selection and Decision Tree for classification have outperformed all earlier methods including machine learning and metaheuristic. The usefulness of Neutrosophic concept on ternary classifier is studied in Jain et al. 2019.

In earlier approaches, Valdez (2015) has defined an optimization approach using PSO with Genetic Algorithms using fuzzy logic. Also in Valdez et al. (2017), a comparative study of fuzzy logic with PSO has also been studied in great detail. The comparative study also proposes a fuzzy particle swarm optimization based on the results of inertia weight adaptation strategies.

Bai et al. (2018) proposed a two-stage PSO-based algorithm for feature selection. In the first stage, four filters—Correlation, Information gain, Gain ratio and symmetrical uncertainty—are used for feature selection, PSO is then used for further reducing the size of selected feature subset for better performance. In the second stage Pbest, GBest model is used for next selection of feature subset. After the iteration goes on for selection of global particle, the gBest particle of swarm gives the final selection of the features. Maiyar et al. (2018) have calculated a comparison of PSO and its variant self-learning PSO combined with Bayesian network structure learning K2 score maximizing model. They have tested their module on components of fashion blogs which influences the customer and studies their underlying structure.

The most recent development in sentiment analysis is the use of deep learning techniques (Samui et al. 2017). Hassan and Mahmood (2018) used a framework that comprises of deep learning network for feature map learned by a convolutional network. CNN gives high-level features which are invariant to local translation. This framework is used for sentence classification and gives high accuracy in two benchmark dataset. Lee et al. (2018) also used convolutional neural network for sentence classification of positive and negative sentences using a weakly supervised learning method. This model not only classifies sentences but also identifies high polarity words present in the text.

In this paper, the proposed method emphasizes on ternary classification which classifies positive, negative and neutral document. The same concept is used for binary classifier and it shows higher accuracy percentage. The result for binary and ternary classifications is discussed in the subsequent sections.

4 Developing insight into Neutrosophic Set and PSO

In this section, an insight into Neutrosophic Set and PSO is explained. These approaches form the basic architecture of Senti-NSetPSO. Classical set gives a crisp value for an element either 0 or 1. In advancement to this, fuzzy set proposed by Zadeh (1965) considers that every element is mapped to area value between 0 and 1, which shows its belongingness to the set. Zadeh's classical fuzzy set assigns a membership value which is determined by some membership function which can be triangular function, trapezoidal function, etc.

In contrast to classical fuzzy set, Smarandache (2014) introduces Neutrosophic Set for an element in the universe, where each element has a three-way decision values. With the degree of membership, it generates a degree of non-membership and a degree of indeterminacy. In Neutrosophic Set, three standard and non-standard subsets—T (Truth), I (Indeterminacy), F (False)—are defined for each element of the Universe. Each element of the set have ranges which lies in]$^{-}$0,1^{+}[. An element "*x*" in universe "*U*" within a set "*M*" has *three* components namely (*t, i, f*) which interprets as the belongingness of "*x*" in "*M*" represents with t % truth, i % indeterminacy, and f % falsehood. The components vary from 0 to 1 and can even be less than 0 or greater than 1. The components are not necessarily a number but can be a subset of type discrete or continuous set. The other categories can be open, closed, half open, half closed sets and it can also be union or intersection of previous sets (Smarandache 2014).

Neutrosophic fuzzy set again have some categorization like Interval Neutrosophic Set, Single valued Neutrosophic Set or Refined Neutrosophic Set (Wang et al. 2005a, b; Deli et al. 2015). All such Neutrosophic variations are defined over discrete or continuous space and also vary in terms of the range of the three components of the elements. The idea proposed by Smarandache has a great impact in representation of real-world problems. Neutrosophic fuzzy logic has already been applied in physics, robotics as well as in NLP to generate word's neutral value calculation (Zhang et al. 2010; Smarandache and Vlădăreanu 2011; Colhon et al. 2017). In Neutrosophic logic, every logical variable "x" is a three-value tuple as given in Eq. 1.

$$X = \{t, i, f\} \tag{1}$$

where t = degree of truth, f = degree of false, i = level of indeterminacy

The characteristics of Neutrosophic logic help in maintaining a powerful construct for the logic. These characteristics are as follows:

- In order maintain the compatibility with fuzzy logic, the values of Neutrosophic logic made a special construction where $t + i + f = 1.0$
- To indicate incomplete information, the Neutrosophic value $t + i + f < 1.0$
- For contradictory information on a variable gives $t + i + f > 1.0$

A Neutrosophic cube as shown in Fig. 1, introduced by Dezert (2002), is a useful tool to visualize the concept of Neutrosophic Set. It is drawn in the 3D Cartesian coordinate system where T is the truth axis value ranges from]$^{-}$0 1^{+}[, I is the indeterminacy axis value ranges from]$^{-}$0 1^{+}[and F is the false axis value ranges from]$^{-}$0 1^{+}[.In this figure Neutrosophic cube is taken with ranges [0,1]. The

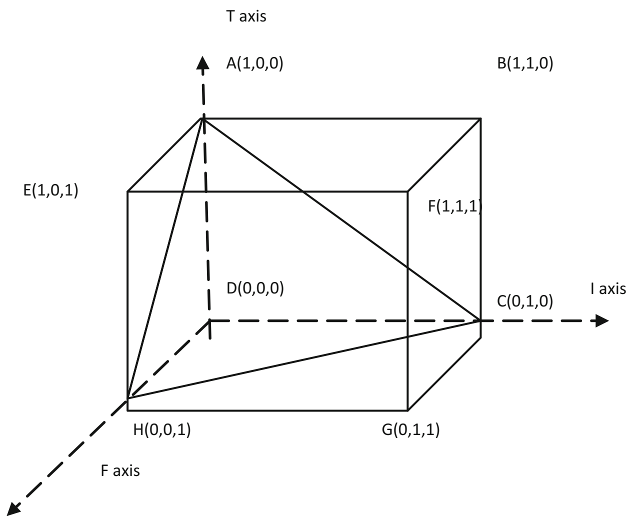

Fig. 1 Neutrosophic cube three-dimensional view

cube can be extended in the more positive and more negative directions to get the range]$^{-}$0 1^{+}[.

For document classification, Neutrosophic classifier is used to classify the document. In the proposed approach, before classification, PSO is used to get the document's optimized Neutrosophic value.

PSO as evolutionary algorithm is chosen as it has two major advantages (as correlated from the literature). First, it is simpler in computation and second, efficient in terms of time and accuracy than other evolutionary group members. In PSO, a swarm of bird decides the path to reach to a specified objective coordinating among them. It is a social meta-heuristic algorithm where each particle moves to its known local best and global best positions towards a solution in the search space. In this iterative method, after every iteration, an updated position leads to a global solution. The stopping criteria in the proposed algorithm follow two basic conditions:

- To reach error value less than a predefined one, or
- To reach for a defined number of iteration.

PSO is very efficient, easy to understand and develop but in certain cases it falls, for instance, in a very large space or on a complex multidimensional dataset, its convergence is slow. The state-of-the-art shows that this problem can be overcome by PSO variants. For multidimensional data, *Cooperative multiple PSO* or for large-scale data *Co-operatively coevolving PSO* are proposed in the literature (Chan et al. 2007; Yang et al. 2008, respectively). *Parallel PSO* (Fan and Chang 2009) is used for parallel processing on large data. To avoid premature convergence, *Attractive Repulsive PSO* (Vesterstroem et al. 2002) uses addition and subtraction operator in attractive and repulsive phase, respectively. Besides these few algorithms, a huge number of variants are available for different needs.

In the sentiment analysis, PSO has been used on selection of optimized features (Ahmad et al. 2015; Liu et al. 2017; Nirmala Devi and Jayanthi 2016). For duel optimization, PSO is used (Basari et al. 2013) for better result of selection of features and SVM is used for classification. In this paper, simple PSO on Neutrosophic value of document texts for polarity detection is used. The concepts and underlying proposed framework is explained in subsequent sections. The proposed algorithm shows a significant improvement after using PSO rather than applying only SVM for positive or negative sentiment selection. Hybrid approaches like PSO with Genetic algorithm for opinion mining also shows increase in accuracy measures over simple PSO application.

5 Proposed framework

In this section, the detailed framework of Senti-NSetPSO is presented. The approach is further explained using a concept diagram and complete flowchart. This section also explains the sample calculation for binary and ternary classifier. Neutrosophic logic gives liberty to express a word beyond duel class, i.e. positive or negative. It enhances the freedom of polarity detection by adding a neutral value to it. In document sentiment classification, fuzzy polarity detection has been used and tested on sentence or document-level. The extension to the fuzzy qualifier is Neutrosophic qualifier.

In the proposed approach, opinion mining algorithm presents a *three-class* and *two-class* problems for the document. In Senti-NSetPSO, purely positive, purely negative and neutral document for ternary classifier and purely positive/negative for binary classifier are considered.

The Neutrosophic values of document are used to create a histogram. The probability score of the histogram is then used for calculations of fitness function as each Neutrosophic value is now considered as *birds in the swarm*. The swarm of points will lead to a best fitted point by using PSO. The value received after optimization is then fitted to the rule-based Neutrosophic classifier to classify the document. SentiWordnet 3.0 is used for extracting the word's neutrosophic value, i.e. positive, negative and neutral. These three-dimensional values are then used for document classification. The positive, negative or neutral values extracted from SentiWordnet 3.0 for each word tokenized from the document has shown that the measure of neutral words are far more than the polarity defined words. So while taking the weight of each component, neutral components are assigned less weight value.

The fitness function for PSO is based on probability distribution taking histogram generated for each component, i.e. *Truth, Indeterminacy and Falsehood.* As for positive and negative document, the neutral words are much more in quantity and then the method suppresses the neutral word taking fraction of indeterminacy value.

PSO is based on the behaviour of particles of swarm and their social interaction. The inspiration for PSO algorithm is based on search of food in a flock of bird. In the beginning, every particle of swarm has its own solution in the 3D space having its coordinate as Truth, Indeterminacy and falsehood. As the time goes, each particle updates its solution while changing its velocity. The velocity is changed towards the resultant of its own experience and the neighbour's experience (factor of velocity). If the i^{th} particle have its position as $X_i\ (t)$ at the time instance t, velocity as $V_i(t)$ and its best known position as $pBest_i(t)$ then the updation of velocity (please refer Fig. 2) at time t + 1 is given from Eqs. 2 and 3:

$$V_i(t+1) = \lambda * (Inertia * C_1 * R_1 * V_i(t) + C_2 * R_2 * (pBest_i(t) - X_i(t)) + C_3 * R_3 * (gBest(t) - X_i(t))) \tag{2}$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \tag{3}$$

where $gBest(t)$ is the best known solution that all the particle experienced up to the time t from the beginning, C_1, C_2, C_3 are the correction factor that controls the contribution of local and global best position vector. C_1 is the inertia and C_2, C_3 are referred as cognitive and social parameters, respectively.

R_1, R_2, R_3 are the random number between [0, 1], and λ is the constriction fraction to control the swarm explosion effect given by Eq. 4.

$$\lambda = 2k/|(2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|\text{where } \ \varphi = c2 + c3 \tag{4}$$

The fitness function includes three components:

- fitness of Truth,
- fitness of Indeterminacy, and
- fitness of Falsehood.

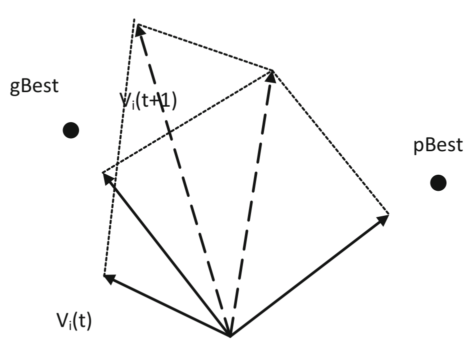

Fig. 2 Particle swarm optimization velocity updation with Local Best(pBest) and Global Best(gBest)

The contribution of fitness of Indeterminacy is suppressed intentionally as for any document whether it is positive or negative document, the neutral words appear hugely. The stop words have been removed from the document, and for classification, Part-Of-Speech as noun, adjective, verb and adverb have been extracted. The score of each derived word is taken from SentiWordNet 3.0 (Baccianella et al. 2010). SentiWordNet 3.0 is used which gives a score of a variable for Negative, Positive and Objective. Each of these three scores lie between [0,1], and sum of these three scores is 1.0. This is in analogous to Neutrosophic logic values. It is a publicly available lexical resource in which each WordNetsynsets is associated with three numerical score Obj(s), Pos(s) and Neg(s). There are two polarity variants which are captured from this resource—*first,* Positive–Negative polarity, *second,* Subjective–Objective polarity for multiclass classification. Figure 3 depicts the overall concept towards the SentiNSetPSO approach to the sentiment analysis.

The algorithm for ternary and binary classification is presented in Fig. 4 and 5. Diving into the details, positive, negative and neutral documents are parsed by Standford parser. The words are then sent to SentiWordNet to get positive and negative sentiment values. Indeterminacy value is calculated as 1-Positive–Negative. The swarms are analysed as a set of Neutrosophic values of words, which is roughly more than 1000 for each document of 100 sentences. Histogram (hist function in MATLAB) is calculated as an interval of *Lmax* specified in the program:

[*n_countT,x_valueT*] = hist(*swarm(:,1,1)*,*Lmax*);
[*n_countI,x_valueI*] = hist(*swarm*(:,1,2),*Lmax*);
[*n_countF,x_valueF*] = hist(*swarm*(:,1,3),*Lmax*);

In this case, obtained histogram intervals are as *x_value* and the count of number of swarms as *n_count* for all the three components Truth, Indeterminacy and False. Probability is generated for *Lmax* number of divisions of Swarms for all the three cases. For each swarm, the velocity "x", "y" and "z" components are taken as Truth, Indeterminacy and Falsehood Neutrosophic values.

Fitness function for each component is calculated as probability distribution of the particular swarm with the Index value in the histogram processing. Fitness function for **Ternary Classifier** and for each swarm of Truth, Indeterminacy and Falsity value is calculated using Eqs. 5 and 6:

$$Fitness = \frac{Probability(1:Index) * x_{value}(1:Index)}{Probability(1:Lmax) * x_{value}(1:Lmax)} \tag{5}$$

$$ResultantFitness = W_1 * Fitness_{Truth} + W_2 * Fitness_{Indeterminacy} + W_3 * Fitness_{False} \tag{6}$$

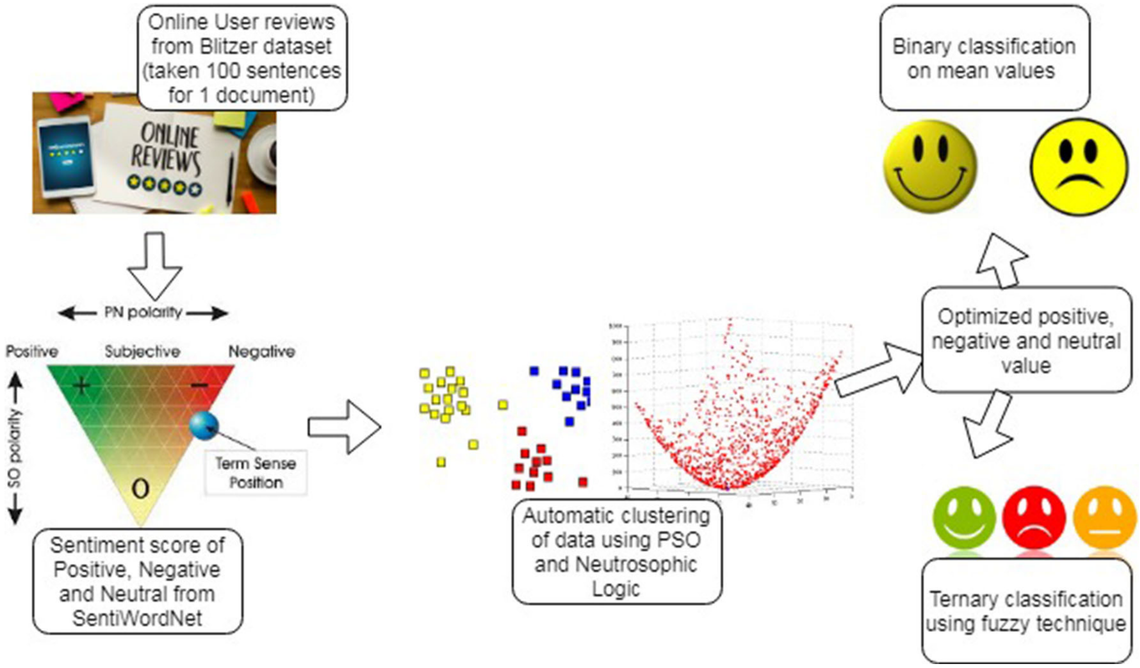

Fig. 3 Concept diagram (binary and ternary classifications) for the proposed Senti-NSetPSO method

Fig. 4 Algorithm for ternary classification using histogram-based PSO and Neutrosophic

1. Input Positive, Negative and Neutral document having more than 100 sentences
2. Parsing and POS tagging
3. Noun, verb,adjective and adverb words are taken
4. Extraction of Neutrosophic Truth, Indeterminancy and Falsehood component as Positive, Neutral and Negative value from SentiWordnet 3.0 of each taken words. SentiWordnet 3.0 gives Positive and Negative value. Neutral value = 1-Positive-Negative.
5. Each Swarm position is now initialized by the Neutrosophic value of the word considered as 3D points.
6. Generating Histogram for each components: Positive, Neutral and Negative
7. Each swarm velocity is initialized by a small number
8. For a predefined number of iteration
    a. Evaluate Fitness for each particle
    b. If fitness is greater than local best
        i. Update Local Best
        ii. If fitness is greater than Global Best
            1. Update Global Best
    c. Update particle velocity
    d. Update particle position
    e. If $\Delta v < \epsilon$, a small value then exit loop
9. Optimized Positive, Neutral and Negative value of positions in 3D obtained
10. Get mean of Positive and Negative swarm values
11. If ((Positive>Negative)||(Mean Positive > Mean Negative)) and |Positive- Negative|>0.15 then the document sentiment is positive
12. Else If((Negative>Positive)||(Mean Negative > Mean Positive)) and |Positive- Negative|>0.15 then the document sentiment is negative
13. Else If |Positive- Negative|<0.15 then the document is neutral

Total Fitness is taken as a percentage combination of Truth fitness, Indeterminacy fitness and False fitness. For each swarm, the fitness is calculated. The maximum fitted swarm for the current iteration is taken as a local best. For all the iteration, the maximum fitted swarm will be the Global best.

For each iteration, the swarm velocity is updated in x, y and z direction as described in Eq. 2 and 3. For the better results, the parameters of Inertia, C_1, C_2 and C_3 are varied. Results of Optimized value of Truth, Falsity determines the positive, negative and neutral inclination of the document.

Fig. 5 Algorithm for binary classification using histogram-based PSO and neutrosophic

1. Input Positive, Negative and Neutral document having more than 100 sentences
2. Parsing and POS tagging
3. Noun, verb,adjective and adverb words are taken
4. Extraction of Neutrosophic Truth, Indeterminancy and Falsehood component as Positive, Neutral and Negative value from SentiWordnet 3.0 of each taken words. SentiWordnet 3.0 gives Positive and Negative value. Neutral value = 1-Positive-Negative.
5. Each Swarm position is now initialized by the Neutrosophic value of the word considered as 3D points.
6. Generating Histogram for each components: Positive, Neutral and Negative
7. Each swarm velocity is initialized by a small number
8. For a predefined number of iteration
   a. Evaluate Fitness for each particle separately for Positive, Neutral and Negative Using Probability Density Function
   b. For Positive, Neutral, Negative do separately
      i. If fitness is greater than local best
      ii. Update Local Best
      iii. If fitness is greater than Global Best
         Update Global Best
      iv. Update particle velocity
      v. Update particle position
9. Points of Positive, Neutral and Negative value of positions in 3D changed according Optimized values
10. Get mean of Positive and Negative swarm values
11. If (Mean Positive> Mean Negative) then the document sentiment is positive
12. If(Mean Negative>Mean Positive) then the document sentiment is negative

Figure 4 describes the proposed algorithm for ternary classification.

The Fitness function for **Binary Classifier** based on Probability Density calculation. For each swarm, fitness is calculated using Eqs. 7, 8 and 9.

$$Fitness_{Truth} = \frac{n_{count} Truth(Index)}{IntegrationundertheHistogramforTruthComponent} \tag{7}$$

$$Fitness_{Indeterminant} = \frac{n_{count} Indeterminant(Index)}{IntegratiomundertheHistogramforIndeterminantComponent} \tag{8}$$

$$Fitness_{False} = \frac{n_{count} False(Index)}{IntegrationundertheHistogramforFalseComponent} \tag{9}$$

Process for optimization of swarms is same as ternary classifier. For predefined iteration, optimization is done as discussed above. Mean of Total Truth components of swarm points and Mean of False components of swarm points are calculated. Results of *Mean_Truth* and *Mean_False* decide the positivity and negativity of the document. Figure 5 shows the algorithm behind the binary sentiment classification.

In this model, stochastic optimization algorithm PSO is used (proposed by Kennedy and Eberhart 1995). The classic population model contains fully connected topology and ring topology. These are also *GBest* and *LBest* model where all particles participate to consider the global best. In the proposed approach, Global topology has been used where all particles are fully connected and every particle is considered to search whole swarm and find the global best fitness value. Inertia is taken as 200.0 and Correction factor 300.0. Constriction factor is calculated from Correction factor which is 0.0033. These parameters kept same for both the ternary and binary classifier.

6 Experimental analysis and results

In this section, the detailed experimental analysis for binary and ternary classifier is presented with the obtained precision and recall values. The proposed framework is analysed based on overall accuracy of their performance.

All the earlier approaches in document classification are feature based. This is the first approach where Neutrosophic values (t, i, f) are used for classification purpose by metaheuristic approach. From earlier approaches, a comparison of machine learning methods like SVM, Naive Bayes, Hybrid SVM and ANN can be seen. Previously, metaheuristic approaches applied on feature selection for categorization. By using PSO, the optimized Truth, Indeterminacy and falsehood value are obtained which are then used on rule-based categorization of positive, negative or neutral document. Algorithm of PSO gives three optimized values of Neutrosophic Truth (NT), Neutrosophic Indeterminacy (NI) and Neutrosophic False (NF). Rules applied on these values give the document sentiment for positive, negative or neutral.

6.1 Settings in training phase

The algorithm is tested on 32-bit operating system with 3 GB RAM. Eclipse: Java programming is used for the POS and SentiWordNet value extraction (Poria et al. 2013). The resultant document is then fed to MATLAB program for classification and optimization.

For effective testing of the proposed methodology, the system uses a multi-domain Blitzer dataset of products of Amazon which is divided in 25 distinct categories. Subjective dataset of movie reviews from Rotten Tomatoes (Ericson and Grodman 2013) is considered as neutral dataset. Blitzer dataset is used in many sentence-level classifiers using machine learning (Liu et al. 2018). Within the scope of the proposed methodology, for designing a document-level classifier, a new dataset is created having 100 sentences from title and review tags of different user review comments. Hundred and thirty such positive documents and 106 such negative documents are created having 100 sentences for each document. Each such file is more than 25 KB size. The dataset has been used mostly for sentence-level classification for Positive–Negative polarity.

The movie snippet dataset consists of plots of movies which are summaries of movies from Internet movie database (Pang and Lee 2004). Plots can be considered as objective dataset or neutral dataset. The objective is to classify neutral classes along with positive–negative dataset. Plot is the dataset considered as neutral sentiment. The rule base applied for ternary and binary classifications is explained in this section.

6.1.1 Rule of thumb for ternary classification

- IF $NT > NForMeanNT > MeanNF$ and absolute difference $(NT - NF) > 0.15$, then the document is classified as positive
- IF $NT\langle NForMeanNF\rangle MeanNT$ and absolute difference $(NT - NF) > 0.15$ then the document is classified as negative
- Else the absolute difference is less than 0.15 $(absolutedifference(NT - NF) < = 0.15)$, the document is classified as neutral.

As a binary classifier to classify, only Positive–Negative polarity gives better result in the generated Blitzer dataset where 100 sentences of title and review comments for each document are taken

6.1.2 Rule of thumb for binary classification

- IF Mean of Truth component of all swarm > Mean of False component of all swarm, then the document is classified as positive
- IFMean of Truth component of all swarm < Mean of False component of all swarm, then the document is classified as negative

To measure the precision, recall and accuracy and F-Measure confusion matrix is created for both the classifier:

1. Positive–Negative classifier or binary classifier
2. Positive–negative–neutral classifier or ternary classifier

Measurement will be based on True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

- True Positive is the positive opinion and classified as positive.
- Conversely, False Negative is positive opinion but classified as negative.
- Similarly, True Negative is negative opinion which is correctly classified as negative.
- False Positive is negative opinion which is categorized wrongly as positive.

Precision, recall, accuracy and F-Measure are the parameters which scale the classifier and their computation is given from Eqs. 10–12.

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

$$F\!-\!Measure = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{13}$$

In Binary classification, out of total 130 positive documents, 130 has come out as positive and no misclassification is observed, whereas 106 negative documents have classified as 95 negative and 11 positive (please refer Table 1). The values obtained for precision, recall and

Table 1 Confusion matrix for positive and negative document in binary classifier

	Positive	Negative	Accuracy (%)
Positive Document = 130	130	0	100
Negative Document = 106	11	95	89.6

F-measure are 0.922, 1 and 0.959, respectively. The overall accuracy has come out to be 95.3%. For multiclass classifier, Eq. 14 and 15 are used to compute the precision and recall.

$$Precision_i = \frac{M_{ii}}{\sum_j M_{ji}} \quad (14)$$

$$Recall_i = \frac{M_{ii}}{\sum_j M_{ij}} \quad (15)$$

For ternary classifier, an overall accuracy of 81.99% is obtained. The same positive and negative documents are used in ternary classifier also. In addition to these, 25 neutral documents of movie plot are used in the system. Out of 130 positive documents, 103 positive documents are classified accurately and 27 documents are wrongly classified. Out of 106 negative documents, 96 classified correctly and ten classified wrongly. For neutral documents, 15 are classified as neutral out of 25 documents.

Table 2 shows precision, recall and F-Measure values for each case and overall accuracy gain is 81.99% for the algorithm used. For ternary classifier, most of the works have been done on Twitter dataset. Twitter comments are very short sentences. But, our work is for large-sized texts which are more than 25 kb size. There is no research work available in the literature for such large-sized text. The plot for positive, negative and neutral accuracy percentage against the iteration of the ternary classifier shows that it is leading to the convergence as the iteration goes up (please refer Fig. 6). It shows that the graphs for positive, negative and neutral are almost parallel to "x" axis and remains almost same.

The proposed model is also used for sentiment analysis on large text (approx 25 KB) collected from aclIMDb and Polarity dataset. Ternary and binary classifier results are shown in Tables 3 and 4. As correlated from previous research works on ternary classifiers which are on short text such as Twitter text or paper review, the comparison of the results is presented in Table 3. Table 4 shows comparison of scientific paper review binary results for machine learning algorithm with binary classifier results for proposed model on large texts.

Ternary or multiclass classification has been a challenging research area in the recent years. A comparative analysis of the results obtained from Senti-NSetPSO with previous ternary classifiers is discussed in Table 3. As

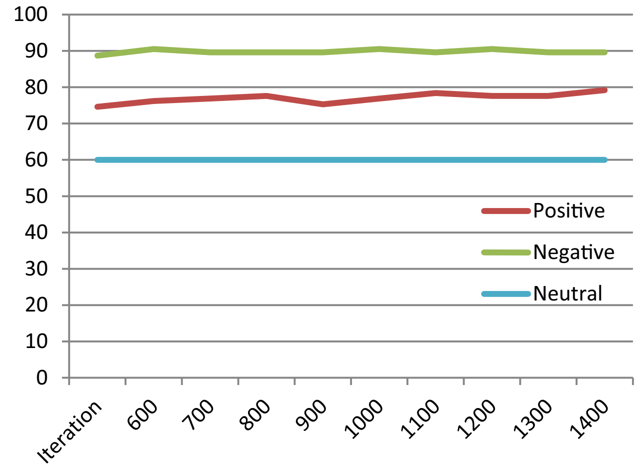

Fig. 6 Efficiency plot of positive, negative and neutral document classification against iteration of the algorithm for ternary classifier

discussed in Sect. 1, earlier methods have been tested on short text or Twitter data. The significance of comparison is to focus on the measuring parameter to give a highlight of current trends of *Ternary classification*. Bouazizi and Ohtsuki (2017) developed a feature-based approach for binary, ternary and multiclass classifier on Twitter dataset. However, Table 3 shows the results of comparison for ternary classification. Bouazizi and Ohtsuki (2017) analysed the neutral class on Twitter data of news or twits which have no emotional polarity. Keith et al. 2017 evaluated scientific reviews for Spanish papers by two algorithms namely, Score algorithm and HS-SVM algorithm. The obtained results were different for binary, ternary and five-point multiclass classification according to the increasing complexity. Keith et al. (2017) have applied their method on review of Spanish papers. As stated earlier, in this paper, the results have been compared with the ternary classification results for both the algorithms.

7 Conclusion and future work

In the present research work, a hybrid composition Senti-NSetPSO based on PSO and Neutrosophic Set is presented. The framework is designed to understand the sentiment analysis of *large-sized text*. The method is novel as it incorporates the advantages of Neutrosophic Set with swarm intelligence. The experimental results (as discussed

Table 2 Confusion matrix for positive, negative and neutral document in ternary classifier

	Positive	Negative	Neutral	Precision	Recall	F-Measure
Positive Document = 130	103	5	22	0.936	0.792	0.858
Negative Document = 106	5	96	5	0.880	0.906	0.893
Neutral Document = 25	2	8	15	0.357	0.6	0.448

Table 3 Comparison of proposed Senti-NSetPSO results versus previous ternary classifiers

Author (s)	Approach and Dataset	Year	Accuracy	Precision	Recall	F-Measure
Bouazizi and Ohtsuki	Pattern-based approach on short Twitter dataset	2017	0.701	0.697	0.701	0.699
Keith et al.	Score Algorithm on Scientific reviews in Spanish Language	2017	0.51	0.58	0.51	0.52
Keith et al.	HS-SVM on Scientific reviews in Spanish	2017	0.56	0.54	0.56	0.54
Keith et al.	NB on Scientific reviews in Spanish	2017	0.46	0.42	0.46	0.41
Keith et al.	SVM on Scientific reviews in Spanish	2017	0.48	0.46	0.48	0.46
Proposed Model	Proposed framework Neutrosophic with PSO	2019	0.819	0.724	0.819	0.733
	Senti-NSetPSO on Large documents (25 KB size) generated from Blitzer dataset					
Proposed Model	Senti-NSetPSO on Large documents (25 KB size) generated from aclIMDdb dataset	2019	0.61	0.73	0.61	0.67
Proposed Model	Senti-NSetPSO on Large documents (25 KB size) generated from Polarity dataset	2019	0.62	0.76	0.62	0.68

Table 4 Comparison of proposed Senti-NSetPSO results for binary classifier vs. previous binary classifiers

Author (s)	Approach and Dataset	Year	Accuracy	Precision	Recall	F-Measure
Keith et al.	Score Algorithm on Scientific reviews in Spanish Language	2017	0.81	0.81	0.81	0.81
Keith et al.	HS-SVM on Scientific reviews in Spanish	2017	0.79	0.8	0.79	0.79
Keith et al.	NB on Scientific reviews in Spanish	2017	0.68	0.67	0.68	0.64
Keith et al.	SVM on Scientific reviews in Spanish	2017	0.7	0.7	0.7	0.69
Proposed Model	Senti-NSetPSO on Large documents (25 KB size) generated from aclIMDb dataset	2019	0.77	0.96	0.56	0.71
Proposed Model	Senti-NSetPSO on Large documents (25 KB size) generated from Polarity dataset	2019	0.76	0.74	0.81	0.77
Proposed Model	Senti-NSetPSO on Large documents (25 KB size) generated from Blitzer dataset	2019	0.953	0.922	1	0.959

in Sect. 5) reveal the consistency of the proposed method. One of the most prominent advantages of the proposed method is that it eliminates the overhead of pre-processing and feature extraction of the text which was a mandatory step in earlier approaches. Further, the para-consistent and incomplete information is efficiently characterized by Neutrosophic Set. This is because there is a notion of absolute truth and relative truth. As there is no restriction on T, I and F (except they lie in the subset of$]^{-}0,1^{+}[$), the non-dualism of input is efficiently handled by Neutrosophic Set.

Moreover, T, I and F are included in the non-standard interval$]^{-}0,1^{+}[$.The powerful construct of Neutrosophic logic gains strength from its characteristics which are discussed in detail in Sect. 2. This also helps in building the Neutrosophic topology which is further computed on unitary non-standard interval. Through Senti-NSetPSO, the effective nature of Neutrosophic Set on text is quite encouraging. The reason is that initially the characterization of Neutrosophic Set was confined to mostly image processing which included image segmentation, de-noising and so on. The flexibility of Neutrosophic Set to form independent sets for truth, falsity and indeterminacy gave researchers a potential motivation to use Neutrosophic Set on text as well.

In this paper, Senti-NSetPSO has been tested on large documents having more than 25 kb of size. The proposed work reached 81.99% for ternary classifier and 95.3% for binary classifier on large text collected from Blitzer dataset. It also gives more than 60% accuracy on large text collected from aclIMDb and Polarity dataset for ternary classifier. Earlier, the techniques are applied on small-sized documents having less than 1 kb size. As for such documents, swarm size is less than 100 which is not leading to good convergence value. For large-sized documents generated from Blitzer review dataset, swarm size is greater than 1000 and has given effective results (please refer Table 3). It is analysed that as the document size increases, the proportion of number of documents decreases. So there is a need for good numbers of large-sized document.

Future work comprises of the implementation of variants of PSO such as self-learning PSO, TRIBES or SPSO 2011 which will overcome the limitation of local optima (Li and Yao 2012; Li et al. 2011; Mekni et al. 2010; Zambrano-Bigiarini et al. 2013). This will work better with small-sized text for ternary classification which has three-dimensional opinion structure. Moreover, Senti-NSetPSO is applicable in cross-domain contents of documents which do not need any feature selection. Senti-NSetPSO can be extended in this direction also. Recent development of deep learning has been used in solving image processing, speech processing and various parts of NLP problem solutions including sentiment analysis. Deep learning or machine learning algorithms have not been tested with Neutrosophic values till now. So this can be taken as an experimental parameter in our future scope.

Further, future research can be applied using various metaheuristic algorithms besides PSO for comparison with the result. Moreover, opinion generation on other review comments such as research paper review can be explored. The working of the proposed framework can be investigated more deeply for classifying documents including review comments on research papers, political analysis on a subject, police verification about any person or happening, news coverage of any event for sentiment polarity finding. Accuracy improvement and more opinion classes will be the other considerations. Further, the Neutrosophic para-consistent logic and its topology can be investigated more deeply on short- and large-sized text.

Compliance with ethical standards

Conflict of interest The authors of the paper declare that we have no conflict of interest.

References

Agarwal B, Poria S, Mittal N, Gelbukh A, Hussain A (2015) Concept-level sentiment analysis with dependency-based semantic parsing: a novel approach. Cogn Comput 7(4):487–499

Ahmad SR, Bakar AA, Yaakub MR (2015) Metaheuristic algorithms for feature selection in sentiment analysis. In: Science and information conference (SAI), 2015. IEEE, pp 222–226

Ali F, Kwak D, Khan P, Islam SR, Kim KH, Kwak KS (2017) Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling. Transpo Res Part C Emerg Technol 77:33–48

Anne C, Mishra A, Hoque MT, Tu S (2017) Multiclass patent document classification. Artif Intell Res 7(1):1

Ansari AQ, Biswas R, Aggarwal S (2013) Neutrosophic classifier: an extension of fuzzy classifer. Appl Soft Comput 13(1):563–573

Anter AM, Hassenian AE (2018) Computational intelligence optimization approach based on particle swarm optimizer and neutrosophic set for abdominal CT liver tumor segmentation. J Comput Sci 25:376–387

Ashbacher C (2002) Introduction to Neutrosophiclogic. Infinite Study

Baccianella S, Esuli A, Sebastiani F (2010) Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: Lrec, vol 10, no. 2010, pp 2200–2204

Bai X, Gao X, Xue B (2018) Particle swarm optimization based two-stage feature selection in text mining. In: IEEE congress on evolutionary computation (CEC). IEEE, pp 1–8

Basari ASH, Hussin B, Ananta IGP, Zeniarja J (2013) Opinion mining of movie review using hybrid method of support vector machine and particle swarm optimization. Procedia Eng 53:453–462

Bing L, Chan KC (2014) A fuzzy logic approach for opinion mining on large scale twitter data. In: Proceedings of the 2014 IEEE/ACM 7th international conference on utility and cloud computing. IEEE Computer Society, pp 652–657

Bouazizi M, Ohtsuki T (2017) A pattern-based approach for multi-class sentiment analysis in twitter. IEEE Access 5:20617–20639

Chan FT, Kumar V, Mishra N (2007) A CMPSO algorithm based approach to solve the multi-plant supply chain problem. In: Swarm intelligence, focus on ant and particle swarm optimization. InTech

Cheung WW, Pitcher TJ, Pauly D (2005) A fuzzy logic expert system to estimate intrinsic extinction vulnerabilities of marine fishes to fishing. Biol Cons 124(1):97–111

Colhon M, Vlăduţescu Ş, Negrea X (2017) How objective a neutral word is? A neutrosophic approach for the objectivity degrees of neutral words. Symmetry 9(11):280

Deli I, Broumi S, Smarandache F (2015) On neutrosophic refined sets and their applications in medical diagnosis. J New Theory 6:88–98

Dezert J (2002) Open questions in neutrosophic inferences. Multiple Valued Log Int J 8(3):439–472

Dragoni M, Tettamanzi AG, da Costa Pereira C (2015) Propagating and aggregating fuzzy polarities for concept-level sentiment analysis. Cogn Comput 7(2):186–197

Ericson J, Grodman J (2013) A predictor for movie success. CS229, Stanford University

Fan SKS, Chang JM (2009) A parallel particle swarm optimization algorithm for multi-objective optimization problems. Eng Optim 41(7):673–697

Gafar MG, Elhoseny M, Gunasekaran M (2018) Modeling neutrosophic variables based on particle swarm optimization and information theory measures for forest fires. J Supercomput 1–18

Hassan A, Mahmood A (2018) Convolutional recurrent deep learning model for sentence classification. IEEE Access 6:13949–13957

Jain A, Nandi BP, Gupta C, Tayal DK (2019) A hybrid framework based on PSO and neutrosophic set for document level sentiment analysis. In: 2nd International conference on information technology and applied mathematics (ICITAM)

Joshi S, Nigam B (2011) Categorizing the document using multi class classification in data mining. In: 2011 International Conference on Computational intelligence and communication networks (CICN). IEEE, pp 251–255

Keith B, Fuentes E, Meneses C (2017) A hybrid approach for sentiment analysis applied to paper. In: Proceedings of ACM SIGKDD conference, Halifax, Nova Scotia, Canada, p 10

Kia PJ, Far AT, Omid M, Alimardani R, Naderloo L (2009) Intelligent control based fuzzy logic for automation of greenhouse irrigation system and evaluation in relation to conventional systems. World Appl Sci J 6(1):16–23

Kumar A, Khorwal R, Chaudhary S (2016) A survey on sentiment analysis using swarm intelligence. Indian J Sci Technol 9(39):1–7

Lee G, Jeong J, Seo S, Kim C, Kang P (2018) Sentiment classification with word localization based on weakly supervised learning with a convolutional neural network. Knowl Based Syst 152:70–82

Li ST, Tsai FC (2013) A fuzzy conceptualization model for text mining with application in opinion polarity classification. Knowl Based Syst 39:23–33

Li X, Yao X (2012) Cooperatively coevolving particle swarms for large scale optimization. IEEE Trans Evol Comput 16(2):210–224

Li C, Yang S, Nguyen TT (2011) A self-learning particle swarm optimizer for global optimization problems. IEEE Trans Syst Man Cybern Part B (Cybernetics) 42(3):627–646

Liu Y, Bi JW, Fan ZP (2017) Multi-class sentiment classification: the experimental comparisons of feature selection and machine learning algorithms. Expert Syst Appl 80:323–339

Liu Q, Zhang Y, Liu J (2018) Learning domain representation for multi-domain sentiment classification. In: Proceedings of the 2018 conference of the North American Chapter of the association for computational linguistics: human language technologies (Long Papers), vol 1, pp 541–550

Maiyar LM, Cho S, Tiwari MK, Thoben KD, Kiritsis D (2018) Optimising online review inspired product attribute classification using the self-learning particle swarm-based Bayesian learning approach. Int J Prod Res 1–22

Mekni S, Châar BF, Ksouri M (2010) TRIBES optimization algorithm applied to the flexible job shop scheduling problem. IFAC Proceedings Volumes 43(4):344–349

Nagarajan SM, Gandhi UD (2019) Classifying streaming of Twitter data based on sentiment analysis using hybridization. Neural Comput Appl 31(5):1425–1433

Nirmala Devi K, Jayanthi P (2016) Sentiment classification using SVM and PSO. Int J Adv Eng Tech 411:413

Pang B, Lee L (2004). A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd annual meeting on Association for Computational Linguistics. Association for Computational Linguistics, p 271

Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? Sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 conference on Empirical methods in natural language processing, vol 10, Association for Computational Linguistics, 2002, pp 79–86

Poria S, Gelbukh A, Hussain A, Howard N, Das D, Bandyopadhyay S (2013) Enhanced SenticNet with affective labels for concept-based opinion mining. IEEE Intell Syst 28(2):31–38

Pu X, Wu G, Yuan C (2019) Exploring overall opinions for document level sentiment classification with structural SVM. Multimed Syst 25(1):21–33

Samui S, Chakrabarti I, Ghosh SK (2017) Improving the performance of deep learning based speech enhancement system using fuzzy restricted Boltzmann machine. In: International conference on pattern recognition and machine intelligence. Springer, Cham, pp 534–542

Sharma R, Nigam S, Jain R (2014) Opinion mining of movie reviews at document level. arXiv preprint arXiv:1408.3829

Smarandache F (2014) Neutrosophic theory and its applications. Collected papers, I. Neutrosophic Theory and Its Applications, 10

Smarandache F (2016) Classical logic and neutrosophic logic. Answers to K. Georgiev. Infinite Study

Smarandache F, Vlădăreanu L (2011) Applications of neutrosophic logic to robotics: an introduction. In: 2011 IEEE international conference on granular computing (GrC). IEEE, pp 607–612

Tripathy A, Anand A, Rath SK (2017) Document-level sentiment classification using hybrid machine learning approach. Knowl Inf Syst 53(3):805–831

Valdez F (2015) Optimization of modular network architectures with a new evolutionary method using a fuzzy combination of particle swarm optimization and genetic algorithms. In: Fuzzy logic augmentation of nature-inspired optimization metaheuristics. Springer, Cham, pp 179–195

Valdez F, Vazquez JC, Melin P, Castillo O (2017) Comparative study of the use of fuzzy logic in improving particle swarm optimization variants for mathematical functions using co-evolution. Appl Soft Comput 52:1070–1083

Vesterstroem J, Riget J, Krink T (2002) Division of labor in particle swarm optimization. In: Proceedings of the IEEE congress on evolutionary computation, pp 1–6

Wang H, Smarandache F, Sunderraman R, Zhang YQ (2005a) Interval neutrosophic sets and logic: theory and applications in computing: theory and applications in computing, vol 5. Infinite Study

Wang H, Smarandache F, Zhang Y, Sunderraman R (2005b) Single valued neutrosophic sets. In: Proceedings of the 10th 476 international conference on fuzzy theory and technology

Yang Z, Tang K, Yao X (2008) Large scale evolutionary optimization using cooperative coevolution. Inf Sci 178(15):2985–2999

Yessenalina A, Yue Y, Cardie C (2010) Multi-level structured models for document-level sentiment classification. In: Proceedings of the 2010 conference on empirical methods in natural language processing. Association for Computational Linguistics, pp 1046–1056

Zadeh LA (1965) Fuzzy sets. Inf Control 8:3

Zambrano-Bigiarini M, Clerc M, Rojas R (2013) Standard particle swarm optimisation 2011 at cec-2013: a baseline for future PSO improvements. In: 2013 IEEE congress on evolutionary computation. IEEE, pp 2337–2344

Zhang M, Zhang L, Cheng HD (2010) A neutrosophic approach to image segmentation based on watershed method. Signal Process. 90(5):1510–1517