



# OMLML: a helpful opinion mining method based on lexicon and machine learning in social networks

Mohammadreza Keyvanpour<sup>1</sup> · Zahra Karimi Zandian<sup>2</sup> · Maryam Heidarypanah<sup>3</sup>

Received: 9 April 2019 / Revised: 17 December 2019 / Accepted: 24 December 2019 / Published online: 7 January 2020  
© Springer-Verlag GmbH Austria, part of Springer Nature 2020

## Abstract

Identification of users' polarities and mining their opinions in various areas, especially social networks, has become one of the popular and useful research fields. Although opinion mining and analyzing methods based on machine learning or lexicon have been useful, high training cost based on time or memory used, lack of enriched and complete lexicons, high dimensions of feature space and ambiguity in positive or negative detection of some sentences in these methods are examples of their downsides. To cope with these problems, in this paper a helpful method based on lexicon and machine learning called OMLML is proposed by using social networks. The main superiority of the proposed method compared to other methods is addressing these challenges simultaneously. According to the proposed method, the polarity of the opinions toward a target word is first determined using a method based on lexicon and textual features of words and sentences. Next, having mapped feature space into a 3-D vector, opinions are analyzed and classified based on a new machine learning method. The results of quantitative and qualitative experiments show that mapping data into a new space decreases training cost and that the performance of the proposed method than is acceptable particularly from the perspective of accuracy, *F*-measure and runtime.

**Keywords** Opinion mining · Social network · Lexicon · Machine learning · Classification

## 1 Introduction

With increasing expansion of the Internet, communities, social networks, the rise in their applications and number of users of social networks, the volume of data generated has increased (Chen and Qi 2011; Rahmani et al. 2014). Therefore, it makes relevant information extraction more challenging (Ali et al. 2015). On the other hand, people are more than willing and happy to share their lives, knowledge and experience (Lloret et al. 2012), and the huge amount of information has become an attractive resource for organizations to monitor the opinions of users (Zainuddin et al. 2018), and social networks have been an appropriate framework for expressing users' opinions and ideas in various

applied fields (Lee et al. 2012) and a rich resource for users' opinions mining and sentiment analysis. Hence, mining this kind of data helps extract practical patterns which are useful for business, applications and consumers.

Opinion mining is a research field that deals with information retrieval and knowledge detection from the text (Missen et al. 2013) using data mining and natural language processing methods (Li and Liu 2014; Khan et al. 2009). Data mining is a process that uses data analysis tools to uncover and find patterns and relationships among data that may lead to extraction of new information from a large database (Karimi Zandian and Keyvanpour 2017; Imani et al. 2013; Karimi Zandian and Keyvanpour 2018).

The purpose of opinion mining is research on opinions and thoughts, identification of emerging social polarities based on the views, sentiments, moods, attitudes and expectations of the beneficiary groups or the majority of people (Shandilya and Jain 2009). In general, the objective is to recognize users' attitudes using analysis of their sentences in contents sent to communities. The attitudes are classified according to their polarities, namely positive, neutral and negative. Automatic support from the analysis process is very important, and due to the high volume of information,

✉ Mohammadreza Keyvanpour  
keyvanpour@alzahra.ac.ir

<sup>1</sup> Department of Computer Engineering, Alzahra University, Tehran, Iran

<sup>2</sup> Data Mining Lab, Department of Computer Engineering, Alzahra University, Tehran, Iran

<sup>3</sup> Department of Computer Engineering, Islamic Azad University, Qazvin Branch, Qazvin, Iran

this kind of support is one of the main challenges (Kaiser and Bodendorf 2009). Opinion mining can be considered as an automatic knowledge detection whose goal is to find hidden patterns in many ideas, blogs and tweets.

In recent years, many studies have been performed in different fields of opinion mining in social networks. By investigating the methods proposed in this area is specified that the main challenges are high training cost based on time or memory used, lack of enriched lexicons, high dimensions of features' space and ambiguity in positive or negative detection of some sentences in these methods.

Due to lack of opinion mining methods examining these essential challenges in the same time, in this paper to cope with these challenges a new opinion mining method called OMLML is proposed, which is addressing them simultaneously and is based on lexicon and machine learning.

According to the proposed method, in the first phase, the polarity of the opinions toward a target word is determined using a method based on lexicon and textual features of words and sentences. Next, in the second phase after mapping feature space into a 3-D vector, opinions are analyzed and classified based on a new machine learning method using improved neural-fuzzy network proposed in this paper.

The results of quantitative and qualitative experiments show that mapping data into a new space decreases training cost and that the performance of the proposed method than others is acceptable and remarkable from the perspective of accuracy, *F*-measure and runtime.

The rest of the paper is organized as follows. In Sect. 2, the related works are discussed. In Sect. 3, the proposed method is introduced. Experiments and the evaluation results are presented in Sect. 4, followed by the concluding remarks in Sect. 5.

## 2 Related work

The writing styles used for opinion mining can be divided into formal and informal texts. The formal texts include poems, novels, the scripts, official documents and so on. The latter include chat rooms data, short messages, texts contained in the discussion forums, as well as posts written in social networks such as Facebook and Twitter (Kaur and Saini 2014).

Social networks are useful sources for opinion mining, sentiment analysis and emotion detection. On the other hand, due to length constraints of texts in this area, the classification operation is a challenging task (Kaur and Saini 2014). Therefore, informality and length limitations of the texts are two main challenges of sentiment analysis in social networks. In other words, it is possible that the methods proposed based on formal texts are not suitable to be used for environments containing short or informal texts like social

networks. So far, various methods based on informal texts have been proposed and applied. A look at various methods proposed in opinion mining shows that these methods are based on machine learning or lexicon or combination of them. By investigation of the methods, it is specified that using just lexicon has been rarely proposed for opinion mining.

### 2.1 Machine learning-based methods

Cui et al. (2011) have proposed an opinion mining method to cope with short messages by the analysis of emotion tokens, including emotion symbols, irregular forms of words and combined punctuations. A graph propagation algorithm as a machine learning method has been proposed to label the tokens' polarities, and a multilingual sentiment analysis algorithm is introduced to solve multilingual problem of Twitter. Cho and Kang (2012) have proposed support vector machines (SVM) method to classify tendencies and opinions in texts extracted from Twitter, Facebook and Me2Day. Pang et al. (2002) have used naive Bayes classification, maximum entropy classification and SVM for sentiment classification, and their data set has been obtained from Internet movie database. Akhmedova et al. (2018) have used the fuzzy rule-based classifiers, artificial neural networks (ANN) and SVM for opinion mining. To generate these methods, a modified meta-heuristic method called CORBA has been proposed to solve constrained and unconstrained real or binary parameter optimization problems. In this method, different term weighting schemes have been used as data preprocessing techniques. To evaluate the proposed method, three corpora of The DEFT07 Evaluation Package (Grenoble 2007) have been used: books, video games and debates in parliament. Xia et al. (2011) have proposed a method for sentiment classification that classifies each of the feature sets by three classification algorithms, naive Bayes, maximum entropy and SVM, and then employs three types of ensemble methods, namely the fixed combination, weighted combination and meta-classifier combination for ensemble of the feature sets. They have considered movie review documents introduced in Pang and Lee (2004) and product reviews taken from Amazon.com and reported in Blitzer et al. (2007). Zhang et al. (2011) have used a method that applies standard machine learning techniques naive Bayes and SVM to automatically classify user reviews as positive or negative. They have created a corpus of Cantonese-written reviews by retrieving consumer reviews from a Cantonese site OpenRice to evaluate their method. According to Anjaria and Guddeti (2014), they have studied the sentiment prediction task over Twitter using machine learning techniques,

with the consideration of Twitter-specific social network structure such as retweet. They employed supervised machine learning techniques such as SVM, naive Bayes, maximum entropy and ANN to classify the Twitter data. Se et al. (2016) have proposed a method based on supervised machine learning for classifying the Tamil movie reviews as positive and negative. For analyzing the social media text where the data are increasing exponentially, machine learning algorithms such as SVM, Maxent classifier, decision tree and naive Bayes were used. Poecze et al. (2018) have focused on the content of communications on Facebook to identify significant differences in terms of their user-generated Facebook metrics and commentary sentiments. They have used a grounded theory approach to classify the posts of YouTube. Krishna et al. (2018) proposed a new model for opinion mining and sentiment analysis of the text reviews posted in Twitter. The model proposed in the paper utilizes machine learning techniques and fuzzy approach for opinion mining and classification of sentiments on textual reviews. Kushwaha and Rathod (2016) proposed a novel technique for opinion mining and feature extraction of product reviews. In this method, natural language processing (NLP) technique is used to obtain the polarity of the reviews and AdaBoost classifier is used for review processing from different E-commerce sites. Tan and Na (2017) have proposed a method to mine patterns of semantic labels from domain corpus for sentence-level sentiment analysis of product reviews by integrating PropBank-based semantic parsing and class association rule (CAR) mining. Montejo-Ráez et al. (2012) proposed a novel approach for polarity detection on Twitter posts using extracting a vector of weighted nodes from the graph of WordNet by combining SentiWordNet scores with a random walk analysis of the concepts and a non-supervised solution that was domain independent. Kang et al. (2018) have proposed an opinion mining method based on text-based hidden Markov models for systems such as movie and product reviews. In this method, a sequence of words is used in training texts instead of a predefined sentiment lexicon. To learn text patterns, ensemble text-based hidden Markov models are applied. According to Narayan et al. (2018) for spam review detection, an opinion mining method has been proposed. In this method, different sets of features, LIWC, POS Tags,  $N$ -gram feature and sentiment score, have been used. To classify opinions, six techniques, decision tree, naive Bayes, SVM,  $k$ -nearest neighbors (KNN), random forest and logistic regression, have been applied. In Souza et al. (2018), a novel algorithm for opinion mining and sentiment analysis of the text reviews posted in Twitter has been proposed based on unsupervised clustering. In this method, a hybrid version of particle swarm optimization (PSO) and Cuckoo Search

(CS) has been used. In the preprocessing phase, natural language and  $N$ -gram language models have been applied.

## 2.2 Lexicon-based methods

Ding et al. (2008) have focused on the problem of determining the semantic orientations of opinions expressed on product features in reviews. They have proposed a holistic lexicon-based approach to solve the problem by exploiting external evidence and linguistic conventions of natural language expressions. Palanisamy et al. (2013) proposed a lexicon-based system as sentiment classification for discovering sentiments based on the contextual sentiment orientation of the words in posts of Twitter. Al-Ayyoub et al. (2015) used the lexicon-based approach to determine the polarity of Arabic online reviews in Twitter and built a very large sentiment lexicon and a lexicon-based sentiment analysis tool.

## 2.3 Machine learning- and lexicon-based methods

Akter and Aziz (2016) proposed a method that applies both machine learning approach and lexicon-based dictionary to analyze sentiments of Facebook data. They used naive Bayes as a machine learning method. Mudinas et al. (2012) proposed a concept-level sentiment analysis system that seamlessly integrates into opinion mining lexicon-based and learning-based approaches to mine the opinions in pSenti system including software reviews and movie reviews. The system uses a sentiment lexicon constructed using public resources for initial sentiment detection. The supervised machine learning algorithm used in this system is the linear SVM implementation in LibSVM2 with L2 objective function for optimization and grid-search for parameter tuning. Tan et al. (2008) proposed a novel method based on lexicon and learning. They used a lexicon-based approach to label a portion of informative examples and a learning-based method like centroid classifier to classify sentiments. They used four domain-specific data sets to evaluate their method: movie reviews, computer reviews, education reviews and house reviews. Lima et al. (2015) suggested a polarity analysis framework for Twitter messages, which combines both approaches, lexicon and machine learning based, and an automatic contextual module. Four types of classifiers were considered: naive Bayes, SVM, decision trees, and KNN. According to Dragoni (2018), to propose a new opinion mining method in advertisement industry and based on Twitter posts, a three-phase model has been used. In this model, first aspects discussed by users are generated and the polarity of those opinions is obtained and finally the most interesting aspects based on an advertisement are determined. Najar and Mesfar (2017) have proposed an Arabic opinion mining method from a set of journalistic articles in political field based on a rules-based approach and linguistic approach

using NooJ's linguistic engine to formalize the automatic recognition rules. These rules are used to identify the different political entities and then identify the opinions associated with the extracted named entities. Poria et al. (2014) have proposed a new method for conceptual opinion mining that merges linguistics, commonsense computing and machine learning for improving the accuracy of tasks such as polarity detection. In this work, dependency relation of the input sentence is used to flow from a concept to another. The input sentences have been obtained from two data sets: movie reviews (Pang and Lee 2005) and product reviews (Blitzer et al. 2007). Poria et al. (2016) presented the deep learning approach to aspect extraction in opinion mining on product reviews and used a combination of seven-layer deep convolutional neural network and a developed set of linguistic patterns to tag each word. A central challenge in building sentiment classifiers using machine learning approach is the generation of discriminative features that allow sentiments to be implied. Ortega-Bueno et al. (2018) proposed a new opinion mining method based on lexicon and machine learning. In this paper, effective algorithms have been proposed to build new lexicons of attitude words, especially for Spanish. To classify attitude words, in the first step words are represented based on neural networks and in the second step one classifier for each attitude type and orientation is trained. Inputs of this method are an unlabeled corpus and a lexicon of words annotated with attitude types and orientation. According to Liu et al. (2015a), a fine-grained opinion mining was proposed that involved identifying the opinion holder who expresses the opinion, detecting opinion expressions, measuring their intensity and sentiment and identifying the target or aspect of the opinion. Liu et al. proposed a general class of models based on recurrent neural networks (RNNs) architecture-like Elman-RNN and Jordan-RNN—and word embeddings, which can be successfully applied to fine-grained opinion mining tasks without any task-specific feature engineering effort. To give better initialization to RNNs, they used pre-trained word embeddings from several external sources or lexicons.

### 3 OMLML: the proposed opinion mining method

Given challenges such as high dimensions of features' space, ambiguity involved in recognition of emotion concept of some words and paying attention to only the weight based on word frequency in opinion mining, this paper proposes a method based on neural-fuzzy network. Employing a neural-fuzzy network causes the advantages of neural network and fuzzy logic to be used at the same time. According to the proposed method, the method applied for users' opinions mining combines a machine learning-based method and

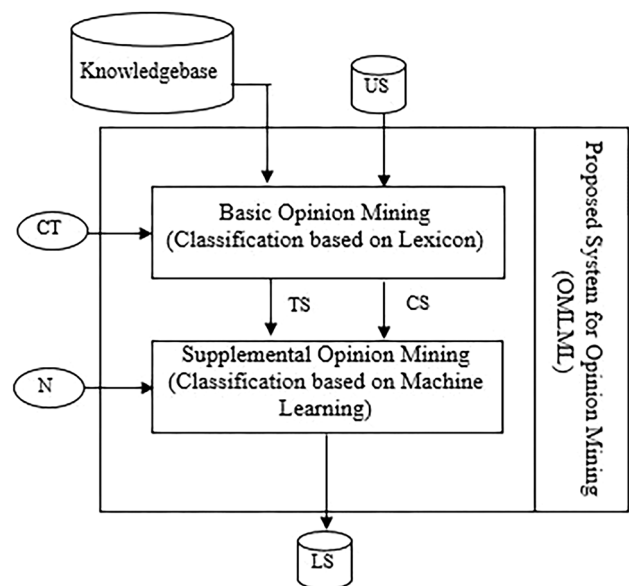


Fig. 1 General structure of the OMLML method

a lexicon-based method in order to classify opinions and sentiments with more accuracy. As shown in Fig. 1, users' opinions in the form of sentences (US), knowledgebase, classification target (CT) and parameter  $N$  used in machine learning phase and determining number of sequential words in a sentence as an expression ( $N$  is used to extract  $N$ -gram) are inputs of the OMLML method, and labeled opinions (LS) constitutes the output.

Accordingly, as specified in Fig. 1, OMLML involves two phases: basic opinion mining and supplemental opinion mining. In the first phase, classification is done based on lexicon and in the second phase, it is done based on machine learning.

#### 3.1 Basic opinion mining

According to Fig. 1, US, knowledgebase and CT are sent as inputs to basic opinion mining phase and its outputs are CS and TS, where CS is a vector of cleaned and refined sentences and TS is a part of the training set.

There are various documents in knowledgebase (Lima et al. 2015) which include

- A list containing words which are frequently repeated and are called stop words such as prepositions and auxiliary verbs.
- A document containing stickers which are frequently used in social networks such as “(:” and “):” and their polarities. In the basic opinion mining phase, if an opinion containing a sticker gets a positive polarity, positive polarity is replaced by the polarity of the sticker in this document. If an opinion containing a sticker gets a nega-

tive polarity, this negative polarity is replaced by the polarity of the sticker in this document.

- A document containing words collected from various lexicons and their polarities.

Lexicon-based classification refers to a classification rule in which documents are assigned labels based on the count of words from lexicons associated with each label (Taboada et al. 2011). For example, suppose that we have opposed labels  $Y \in \{0, 1\}$  and associated lexicons  $W_0$  and  $W_1$ . Then, for a document with a vector of word counts  $x$ , the lexicon-based decision rule is,

$$\sum_{i \in W_0} x_i \geq \sum_{j \in W_1} x_j \tag{1}$$

where the  $\geq$  operator indicates a decision rule. Put simply, the rule is to select the label whose lexicon matches the most word tokens (Eisenstein 2017).

Based on the proposed method, the basic opinion mining is based on lexicon and the opinions are classified using words' features and their location in the sentence. As shown in Fig. 2 and according to the proposed method, the basic opinion mining is divided into two main phases: textual preprocessing phase and opinions classification phase.

### 3.1.1 Textual preprocessing

Available data set based on the users' opinions is in form of textual and unstructured files, which is not stable without initial processing. Therefore, according to Fig. 2, the first step in the basic opinion mining phase is textual preprocessing. US constitutes the input of this step, and KW and CS are the outputs, where KW is an array of all words in all opinions as keywords. In this step, opinions refinement is done, for which tokens are identified, extra characters and symbols like "", '@', '\*', '\$' and '#' are removed, stemming

is done, and stop words like "am", "is" and "can" are deleted (Lima et al. 2015).

### 3.1.2 Opinions classification

In this step, opinions are classified and their labels are predicted. As shown in Fig. 2, Inputs of opinions classification are knowledgebase, CT, KW and CS. According to the proposed method in this paper, to classify opinions and predict their labels, KW is first determined using Knowledgebase. Then, given CT all opinions containing CT are extracted from CS and polarities of their words are determined. In the next step, labels of these opinions are obtained by calculating sum of the distances between the words with positive polarity in the opinion and the CT and sum of the distances between the words with negative polarity in the opinion and the CT. If positive concepts are more frequent in the nearby CT, opinion label is positive. But, if negative concepts are more frequent in the nearby CT, opinion label is negative.

Eventually, these opinions labeled by the proposed method based on the lexicon used in this paper are sent to supplemental opinion mining phase as TS.

### 3.2 Supplemental opinion mining

In the supplemental opinion mining phase, a method based on machine learning is used to classify the opinions. As specified in Fig. 1, TS, CS and  $N$  are inputs of this phase and its output is LS. As shown in Fig. 3, supplemental opinion mining includes two steps: data set repairing for model training and model creation based on machine learning.

#### 3.2.1 Data set repairing for model training

As shown in Fig. 3, data set repairing step receives TS, CS and  $N$  and after providing a suitable data set for model

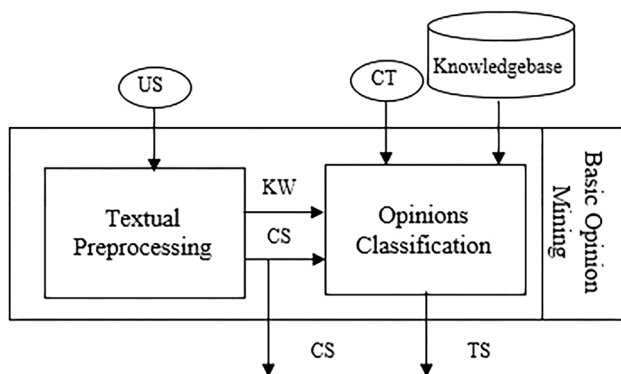


Fig. 2 Block diagram of basic opinion mining

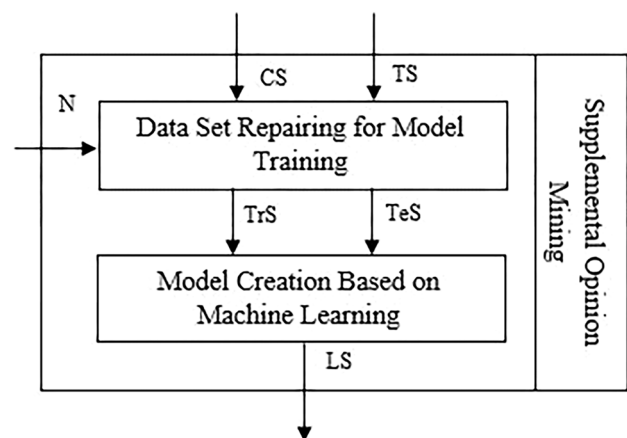


Fig. 3 Block diagram of supplemental opinion mining

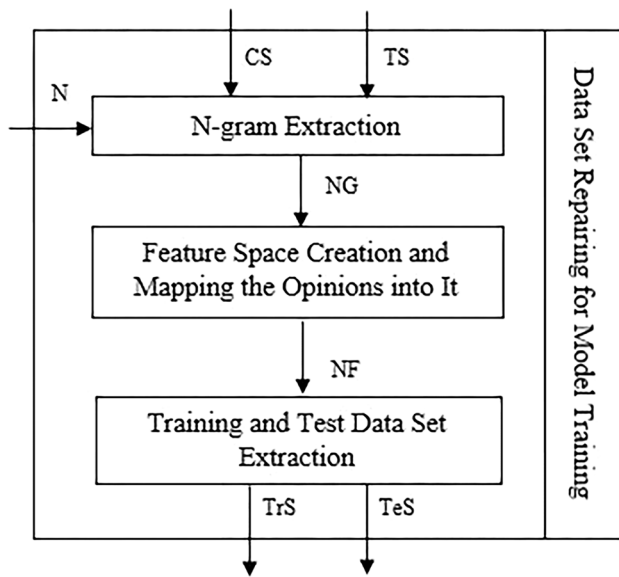


Fig. 4 Steps of data set repairing for model training

training, TeS and TrS are sent to model creation part, where TrS is the training data set and TeS is the test data set. As Fig. 4 depicts, data set repairing for model training are formed from three steps: *N*-grams extraction, feature space creation and mapping the opinions in data set into the created feature space, and training and test data set extraction.

For feature space creation, *N*-grams (NG) that form the words of the opinions are used. Therefore, in the first step, by receiving TS, CS, and N, NG is extracted and as output is sent to the next step. Based on the proposed OMLML method, the feature space contains three features. This feature space maps each opinion into a vector with three components. The features in feature space include:

- Sum of calculated weights of TF.IDF of words in the opinion.
- The number of positive emotions obtained from the opinions that equals the sum of positive emotional weights of all of its words.
- The number of negative emotions obtained from the opinions that equals the sum of negative emotional weights of all of its words.

The procedure to weight the words based on TF.IDF and emotions is described below:

- Weighting the words based on TF.IDF: In this kind of weighting, a popular statistical method called TF.IDF (Hourali and Montazer 2010) is used. Equation 2 shows how to calculate it. In this method, a weight is associated with each word based on its frequency in the opinion

$$TF.IDF_{t,i} = tf_{t,i} \times \log \left( \frac{N}{df_t} \right) \tag{2}$$

where *t* is a word and *i* is an opinion. TF.IDF<sub>*t,i*</sub> is the weight calculated for word *t* in opinion *i*. *tf*<sub>*t,i*</sub> is the frequency of the word *t* in opinion *i* and *df*<sub>*t*</sub> is the number of opinions in which *t* has been shown. *N* is the number of all opinions.

- Weighting the words based on emotion: In the proposed method, for calculation of the number of positive or negative emotions, a statistical method called odds ratio (OR) is used. In this method, the relationship between two features A and B in a population is measured. This relation shows how the existence or absence of feature A influences the existence or absence of feature B (Bland and Altman 2000). In other words, to calculate the relations between two special features A and B, OR is used. To calculate positive and negative weights of words in the opinion, Eqs. 3 and 4 are proposed, respectively

$$POR_i = \log \frac{P(w_i|POS)(1 - P(w_i|NEG))}{(1 - P(w_i|POS))P(w_i|NEG)} \tag{3}$$

$$NOR_i = \log \frac{P(w_i|NEG)(1 - P(w_i|POS))}{(1 - P(w_i|NEG))P(w_i|POS)} \tag{4}$$

where *P*(*w*<sub>*i*</sub>|*POS*) is the probability of the word *w*<sub>*i*</sub> in the positive class. *P*(*w*<sub>*i*</sub>|*NEG*) is the probability of the word *w*<sub>*i*</sub> in the negative class.

Consequently, in the new feature space, each opinion converts to a vector with three dimensions. The first dimension shows the importance of an opinion compared to other opinions. The second dimension is the number of positive emotions, and the third one is the number of negative emotions. The output of the second step in the data set repairing is constituted by new features for each opinion (NF). For training and evaluating learning models, the available data sets are always divided into training data set and test data set. Therefore, in this paper, after mapping the data into the new feature space, in the third step of the data set repairing for model training, 70% of the data set is extracted as TrS and the other 30% is extracted as TeS.

### 3.2.2 Model creation based on machine learning

In this section, in order to improve and increase accuracy and performance of the proposed OMLML method, an improved neural-fuzzy network is proposed. In this network, we use the model proposed in Takagi and Sugeno (1985) as a neural-fuzzy model and Gaussian membership functions in Reddy and Raju (2009) as fuzzifiers.

A regular neural-fuzzy network is a neural network with fuzzy signals and/or fuzzy weights, sigmoidal transfer

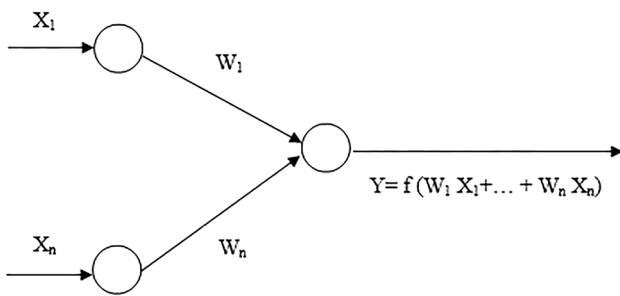


Fig. 5 Regular neural-fuzzy network (Fullér 1995)

function and all the operations are defined by Zadeh’s extension principle (Fullér 1995). Consider a simple regular neural-fuzzy network in Fig. 5.

All signals and weights are fuzzy numbers. The input neurons do not change the input signals, so their output is the same as their input. The signal  $X_i$  interacts with the weight  $W_i$  to produce the product  $P_i = W_i X_i, i = 1, \dots, n$ , where we use the extension principle to compute  $P_i$ . The input information  $P_i$  is aggregated, by standard extended addition, to produce the input

$$\text{net} = P_1 + \dots + P_n = W_1 X_1 + \dots + W_n X_n \tag{5}$$

to the neuron. The neuron uses its transfer function  $f$ , which is a sigmoidal function, to compute the output

$$Y = f(\text{net}) = f(W_1 X_1 + \dots + W_n X_n) \tag{6}$$

where  $f$  is a sigmoidal function and the membership function of the output fuzzy set  $Y$  is computed by the extension principle.

Generally, to use neural-fuzzy network, first, parameters required to create a neural-fuzzy network must be determined using TrS and training models. According to the proposed method in this paper, for training and creating an improved neural-fuzzy network, we use genetic algorithm (GA) and PSO as meta-heuristic algorithms instead of traditional training methods like gradient descent whose problem is convergence to local optimum solutions. Meta-heuristic algorithms are used in order to determine the best and global optimum solutions using global search approaches. Therefore, in the supplemental opinion mining phase and after the data set repairing step, to create improved neural-fuzzy network, optimum values for parameters required are first obtained by a meta-heuristic algorithm. In the next step, the improved neural-fuzzy network based on the model proposed in Takagi and Sugeno (1985) and Gaussian membership functions are modeled. The last step in the model creation phase includes training the model using TrS and labeling and predicting the labels of TeS. As shown in Fig. 3, the output of this phase is LS.

Table 1 Characteristics of the data sets used in the proposed method

Data set	Name	The number of all opinions	The number of opinions with positive polarity	The number of opinions with negative polarity
1	Debate2008	2007	743	1264
2	Sentistrength	4242	3293	949

## 4 Experiments

### 4.1 Data set

The data used in this paper consist of two data sets obtained from Twitter social network that have been collected in 2008–2013 and used in some works (Lima et al. 2015; Yang et al. 2017; Taboada et al. 2011; Eisenstein 2017; Hourali and Montazer 2010; Reddy and Raju 2009; Cambria and Hussain 2012). In these data sets, each user’s opinion has 140 characters at most, and all opinions have label 1 or  $-1$  that are specified by the experts. Label 1 shows that the polarity of the opinion is positive and label  $-1$  shows that the polarity of the opinion is negative. Table 1 shows the characteristics of the data sets used. As shown in Table 1, in debate2008 data set, there are 2007 opinions. The number of opinions with positive polarity is 743, and the number of opinions with negative polarity is 1264. In sentistrength data set, 3293 opinions in 4242 opinions have positive polarity and the rest have negative polarity.

### 4.2 Evaluation criteria

In data mining applications, various criteria are applied for evaluation of the methods proposed and used. In this paper, the following criteria are used: accuracy, precision, recall and  $F$ -measure.

- Accuracy: The most important criterion for evaluation of any classification algorithm is accuracy, which is calculated based on Eq. 7 (Bhattacharyya et al. 2011)

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FN} + \text{TP} + \text{FP}} \tag{7}$$

where TN is the number of the opinions with negative polarity which are labeled negative polarity correctly. TP is the number of the opinions with positive polarity which are labeled positive polarity correctly. FP is the number of the opinions with negative polarity which are labeled positive polarity incorrectly. FN is the number of opinions with positive polarity which are incorrectly labeled negative polarity.

- Precision: As shown in Eq. 8, it is the number of opinions correctly labeled as belonging to the positive class (TP) divided by the total number of opinions labeled as belonging to the positive class (i.e., the sum of true positives and false positives, which are items incorrectly labeled as belonging to the class) (Bhattacharyya et al. 2011)

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (8)$$

- Recall: The number of true positives divided by the total number of opinions that actually belongs to the positive class (i.e., the sum of true positives and false negatives, which are opinions not labeled as belonging to the positive class but should have been) (Eq. 9) (Powers 2011)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (9)$$

- *F*-measure: As stated in Eq. 10, it is the harmonic mean of precision and recall (Powers 2011)

$$F\text{-measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \quad (10)$$

### 4.3 Experiments results

To evaluate the proposed method, three tests have been designed and run. Test 1 investigates the proposed method based on different learning models and criteria mentioned in the previous subsection. Test 2 investigates the runtime of the proposed opinion mining method based on different learning models. Test 3 is the investigation of the effect of different features in feature space on the results of the proposed method based on the criteria mentioned in the previous subsection. In addition to these three tests, to evaluate and compare our method and others comprehensively, the proposed method has been compared with other methods quantitatively and qualitatively.

#### 4.3.1 Test 1: the effect of different learning models in the proposed opinion mining on the evaluation criteria

Test 1 helps to investigate the effect of different learning models in the proposed opinion mining method on the criteria mentioned above. As mentioned in previous sections, in the proposed OMLML method in the supplemental opinion mining phase for training the improved neural-fuzzy network, GA and PSO are used as meta-heuristic algorithms instead of traditional training methods like gradient descent whose problem is convergence to local optimum solutions. Meta-heuristic algorithms are

**Table 2** Values used for GA algorithm parameters

Parameter	Value
The number of population members	150
Crossover rate (%)	40
Mutation rate (%)	9
The number of generations	50

**Table 3** Values used for PSO algorithm parameters

Parameter	Value
The number of population members	150
The number of generations	50
Inertia weight	Adaptive
The particle's experience weight	1
The group's experience weight	3

used in order to determine the best and global optimum solutions using global search approaches. Therefore, this test evaluates the proposed method using neural network, neural-fuzzy network and improved neural-fuzzy network based on either PSO or GA algorithms as a meta-heuristic algorithm.

Before evaluation, it is necessary to determine the values used for GA and PSO algorithms parameters. Tables 2 and 3 show the values used for GA and PSO algorithms, respectively.

To investigate the proposed method, the value of parameter *N* has been initialized with 1–6 as the input of the proposed OMLML method. The results obtained from this test are summarized in Table 4.

As shown in Table 4, neural network has the least performance on average due to lack of fuzzy system in its structure. This result shows that using fuzzy system in the opinion mining method helps to enhance performance. As inferred from Table 4, using combination neural network and fuzzy system usually improves performance more than using just neural network. using PSO and GA as the best solution determiner, that have been proposed in this paper, to train the parameters of a neural-fuzzy model has helped the method to be improved on average. It seems that attention to convergence of neural and neural-fuzzy networks to local optimum solutions and minimum points and applying meta-heuristic algorithms to the improved neural-fuzzy network to solve this challenge causes this method to have better performance and effectively improve the results of this experiment. Comparisons between the results of GA and PSO in Table 4 show that these algorithms have performed similarly with PSO showing slightly better performance. It is necessary to note that the results gained are independent of the data set.



**Table 4** Results of applying different learning models to the proposed opinion mining on the evaluation criteria (Test 1)

N	Supplemental opinion mining method	Debate2008				Sentistrength			
		Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
1	Neural network	63	63	100	77	38	49	41	45
	Neural-fuzzy network	63	63	100	77	52	34	40	36
	Improved neural-fuzzy network using GA	68	67	95	78	71	45	60	51
	Improved neural-fuzzy network using PSO	69	68	95	79	70	45	60	51
2	Neural network	60	62	100	77	57	52	67	58
	Neural-fuzzy network	63	63	100	77	61	69	73	71
	Improved neural-fuzzy network using GA	65	66	94	77	75	50	44	46
	Improved neural-fuzzy network using PSO	66	67	90	77	75	50	43	46
3	Neural network	58	58	100	73	40	21	37	26
	Neural-fuzzy network	62	3	8	76	57	43	41	42
	Improved neural-fuzzy network using GA	62	63	95	76	75	52	63	56
	Improved neural-fuzzy network using PSO	62	63	95	76	76	69	77	73
4	Neural network	64	64	100	78	54	62	59	60
	Neural-fuzzy network	63	63	100	77	52	34	40	36
	Improved neural-fuzzy network using GA	64	64	100	78	71	45	60	51
	Improved neural-fuzzy network using PSO	63	63	100	77	70	45	60	51
5	Neural network	59	59	100	74	37	40	57	47
	Neural-fuzzy network	63	63	100	77	46	51	61	55
	Improved neural-fuzzy network using GA	66	67	90	77	74	55	63	58
	Improved neural-fuzzy network using PSO	67	68	90	78	62	55	63	58
6	Neural network	35	47	34	39	59	51	63	56
	Neural-fuzzy network	63	63	100	77	62	51	57	54
	Improved neural-fuzzy network using GA	66	67	90	77	67	62	66	58
	Improved neural-fuzzy network using PSO	66	67	90	77	64	53	62	50

### 4.3.2 Test 2: the effect of different learning models in the proposed opinion mining on runtime

In Test 2, different learning models were applied to investigate the runtime of the proposed opinion mining method. In this subsection given different values of  $N$ , the runtime was evaluated considering neural network, neural-fuzzy network, improved neural-fuzzy network using GA and improved neural-fuzzy network using PSO. Table 5 indicates the results of this test. As the results obtained are independent of data set, evaluations have been reported based on Debate 2008 data set. Although using the meta-heuristic algorithm in model training increased performance, especially concerning accuracy, it dictates more time overhead on the method. According to Table 5, the improved neural-fuzzy network using GA and PSO needs a duration of approximately 10 and 20 times, respectively, for training and labeling than neural network and neural-fuzzy method. This is because the meta-heuristic algorithm calculates the cost of each member of population repeatedly using a cost function until determining optimum values for the parameters.

As shown in Table 5, runtime in the improved neural-fuzzy network using PSO is two times longer than the runtime in the improved neural-fuzzy network using GA. This is because in PSO algorithm, all members of population are evaluated in each iteration and their costs are calculated, while in GA algorithm the costs of only new members are calculated in each iteration. Therefore, runtime in the proposed method using PSO is longer than when GA is used. It is remarkable that runtime for neural-fuzzy network and neural network is almost equal. In spite of more accuracy for neural-fuzzy network than neural network, it needs equal time to run.

Generally, given improvement in our proposed method based on other criteria, its runtime is not high disregarding others and is acceptable.

**Table 5** Results of applying different learning models to the proposed opinion mining on runtime (Test 2)

$N$	Neural network (s)	Neural-fuzzy network (s)	Improved neural-fuzzy network using GA (s)	Improved neural-fuzzy network using PSO (s)
1	256	265	2380	4606
2	255	267	2452	4707
3	261	254	2452	4445
4	255	259	2235	3846
5	267	291	2301	3800
6	248	227	2340	3860

### 4.3.3 Test 3: the effect of different proposed features in feature space of the proposed method on the evaluation criteria

This test aims to investigate the effect of different features in feature space on the results of the proposed method based on criteria mentioned above.

As stated before, feature space used in training data includes TF.IDF weight, positive emotional weight and negative emotional weight. In this test, the training data set is created based on three different modes: TF.IDF weight and positive emotional weight; TF.IDF weight and negative emotional weight; TF.IDF weight, positive emotional weight and negative emotional weight. Improved neural-fuzzy network was trained by them, separately. Table 6 shows the results obtained from this test.

According to the results reported in Table 6, it is clear that with removing each of the features of positive emotional weight and negative emotional weight, performance criteria introduced in evaluation criteria subsection, including accuracy, were reduced. Since using each of the features causes more complete information to be provided to train data, the performance increases. Therefore, applying the proposed method (using all of the features) helps to improve the results effectively.

### 4.3.4 Comparison between the proposed method and other methods quantitatively

In order to evaluate the proposed method, it is necessary to compare it with other methods. Therefore, in this section the proposed method is compared with methods proposed in Lima et al. (2015). This method is a hybrid approach containing two parts. In the first part, after extracting all keywords,  $n$ -dimensional feature space is created based on TF.IDF weight. In the second part, decision tree, SVM network, KNN model and naive Bayes as learning models are trained and used for opinion classification. Given that accuracy and  $F$ -measure are two main criteria in opinion mining, comparison and evaluation have been done based on them. The results obtained are shown in Tables 7 and 8. In order to evaluate and compare the methods more accurately,

**Table 6** Results of applying different proposed features to feature space of the proposed method on the evaluation criteria (Test 3)

	Improved neural-fuzzy network using GA			Improved neural-fuzzy network using PSO		
	TF.IDF, SW+ (%)	TF.IDF, SW- (%)	TF.IDF, SW+, SW- (%)	TF.IDF, SW+	TF.IDF, SW- (%)	TF.IDF, SW+, SW- (%)
Debate2008						
Accuracy	64	63	68	68	63	69
Precision	65	63	67	67	63	68
Recall	77	100	95	93	100	95
<i>F</i> -measure	73	77	78	78	77	79
Sentistrength						
Accuracy	79	62	81	57	72	80
Precision	52	63	66	40	38	45
Recall	48	57	60	59	56	60
<i>F</i> -measure	49	50	51	50	48	51

each learning model used in Lima et al. (2015) is examined based on accuracy and *F*-measure, separately. Given that the results based on runtime are similar, they are reported as a single method.

*Discussion* As shown in Table 7, the best result of the method proposed in Lima et al. (2015) based on the first data set is related to using naive Bayes with 63% accuracy and 76% *F*-measure. However, in the method proposed in this paper, the best result of our proposed method based on the first data set is related to improved neural-fuzzy network using PSO whose accuracy is 69% and *F*-measure is 79%. These results show that the proposed method in this paper is better than the method proposed in Lima et al. (2015) and improves the performance of opinion mining. As to data set 2, the best result of the proposed method in Lima et al. (2015) is related to naive Bayes model with 78% accuracy and 87% *F*-measure. In contrast, the best result of the proposed method in this paper is related to improved neural-fuzzy network using PSO whose accuracy is 76% and *F*-measure is 73%. The results show that using the proposed method in this paper as an opinion mining method and meta-heuristic algorithms to determine the optimum values of the parameters on data set 1 significantly improved the performance of mining based on accuracy and *F*-measure. However, as to data set 2, the proposed method in Lima et al. (2015) using naive Bayes showed better performance.

On the other hand, as inferred from Table 7, using the naive Bayes method as a learning model in Lima et al. (2015) produced the best result among other learning models proposed in Lima et al. (2015). Also, improved neural-fuzzy network using PSO is the best one among the models considered in this paper. According to Table 8, although the meta-heuristic algorithm which is time-consuming is used in the proposed OMLML method, the runtime of the proposed method used in this paper is lower than that of the methods proposed in Lima et al. (2015). In the OMLML method, first

a new feature space is created with only three dimensions and supplemental opinion mining is done using it. Nonetheless, the new feature space applied in Lima et al. (2015) is created with very high dimensions, which increases the dimensions of feature vectors in all keywords of all opinions. This raises the time of model training. Therefore, using the proposed method boosts the runtime of opinion mining, in addition to other criteria.

As mentioned before, the improved neural-fuzzy network using GA and PSO needs more time than neural network and neural-fuzzy network. This is because the meta-heuristic algorithm calculates the cost of each member of population repeatedly using a cost function until determining optimum values for the parameters. Furthermore, runtime for the improved neural-fuzzy network using PSO is two times longer than the runtime in the improved neural-fuzzy network using GA. Also, in PSO algorithm in each iteration all members of population are evaluated and their costs are calculated, while in GA algorithm in each iteration cost of only new members are calculated. Therefore, runtime of the proposed method using PSO is longer than when GA is used.

Regarding Tables 7 and 8 and considering the mean of the results obtained from the two data sets reported in Table 7, the proposed method has better performance based on accuracy and *F*-measure approximately in all states, while the runtime in our method is less than the methods proposed in Lima et al. (2015). Therefore, this result shows improvement in opinion mining.

### 4.3.5 Comparison between the proposed method and other methods qualitatively

As some methods either have been proposed for opinion mining in a special environment or application or have not been evaluated quantitatively, it is not possible to compare our method with them quantitatively. Therefore, to evaluate

**Table 7** Comparisons between the proposed methods (improved neural-fuzzy network using GA and improved neural-fuzzy network using PSO), basic methods (neural network and neural-fuzzy network) and the methods proposed in Lima et al. (2015) based on accuracy and *F*-measure

N	Data set	Opinion mining method									
		The method proposed in Lima et al. (2015) using Naive Bayes (%)	The method proposed in Lima et al. (2015) using SVM (%)	The method proposed in Lima et al. (2015) using KNN (%)	The method proposed in Lima et al. (2015) using decision tree (%)	Neural network (%)	Neural-fuzzy network (%)	Improved neural-fuzzy network using GA (%)	Improved neural-fuzzy network using PSO (%)		
1	Debate 2008	Accuracy	56	49	46	64	63	63	68	69	
	<i>F</i> -measure	64	53	52	0	77	77	78	79		
	Sentistrength	Accuracy	34	49	54	21	38	52	71	70	
	<i>F</i> -measure	33	59	66	0.9	45	36	51	51		
2	Debate 2008	Accuracy	63	47	46	64	60	63	65	66	
	<i>F</i> -measure	76	52	52	2	77	77	77	77		
	Sentistrength	Accuracy	75	61	55	76	57	61	75	75	
	<i>F</i> -measure	85	72	67	86	58	71	46	46		
3	Debate 2008	Accuracy	63	47	46	51	58	62	62	62	
	<i>F</i> -measure	76	53	52	50	73	76	76	76		
	Sentistrength	Accuracy	77	58	44	21	40	57	75	76	
	<i>F</i> -measure	86	70	52	0.70	26	42	56	73		
4	Debate 2008	Accuracy	42	52	46	62	64	63	66	67	
	<i>F</i> -measure	53	45	52	47	78	77	77	78		
	Sentistrength	Accuracy	77	58	45	21	54	52	71	70	
	<i>F</i> -measure	87	70	52	0.90	60	36	51	51		
5	Debate 2008	Accuracy	40	53	35	62	59	63	66	66	
	<i>F</i> -measure	53	46	52	47	74	77	77	77		
	Sentistrength	Accuracy	78	67	58	21	62	46	74	62	
	<i>F</i> -measure	87	78	70	0.90	47	55	58	58		
6	Debate 2008	Accuracy	40	55	35	48	35	63	66	66	
	<i>F</i> -measure	53	38	52	49	39	77	77	77		
	Sentistrength	Accuracy	70	71	78	21	59	62	67	64	
	<i>F</i> -measure	81	82	87	0.90	56	54	18	2		

**Table 8** Comparisons between the proposed methods (improved neural-fuzzy network using GA and improved neural-fuzzy network using PSO), basic methods (neural network and neural-fuzzy network) and the method proposed in Lima et al. (2015) based on runtime

$N$	Neural network (s)	Neural-fuzzy network (s)	Improved neural-fuzzy network using GA (s)	Improved neural-fuzzy network using PSO (s)	The method proposed in Lima et al. (2015) (s)
1	256	265	2380	4606	4709
2	255	267	2452	4707	4620
3	261	254	2452	4445	4987
4	255	259	2235	3846	4959
5	267	291	2301	3800	5077
6	248	227	2340	3860	4657

and compare our method and others comprehensively in addition to evaluation quantitatively, in this section the proposed method and others are compared qualitatively.

*Comparison between the proposed method and other methods based on existing challenges* Given challenges such as high dimensions of features' space, the ambiguity involved in recognition of emotional concepts of some words, paying attention to only the weight based on word frequency in opinion mining, high training cost based on the time or memory used and the uses of previous methods in particular fields, the superiority of our method compared to other opinion mining algorithms includes using low dimensions of features' space, paying attention to the weights based on other criteria in addition to word frequency in opinion mining and reducing training cost, simultaneously. On the other hand, this method has satisfactory accuracy and is suitable for opinion mining in various fields. Some methods have been proposed to mine opinions in particular applications, environments or fields such as Smart City (Puri et al. 2018; Mishra et al. 2018), Tourism Industry (Bhatnagar et al. 2018), Advertisement (Tudoran 2018; Dragoni 2018), Nutrition Industry (Mostafa 2018), Stock Investment (Jeong et al. 2018), Economy, Commerce and Marketing (Karami et al. 2018; Yun et al. 2018; Rathan et al. 2018; Narayan et al. 2018), Energy (Nuortimo and Härkönen 2018) and Literature Review like Movie Review (Souza et al. 2018). In contrast, this paper proposes a new method called OMLML that is usable in various applications and fields. In comparison with methods using just machine learning or lexicon to mine opinions such as Puri et al. (2018), Bhatnagar et al. (2018), Tudoran (2018), Mostafa (2018), Karami et al. (2018), Yun et al. (2018), Narayan et al. (2018), Nuortimo and Härkönen (2018), Souza et al. (2018), Rozi et al. (2018), Akhmedova et al. (2018), Solanki et al. (2019) and Kang et al. (2018), the method applied for users' opinions mining in this paper combines a machine learning-based method and a lexicon-based method in order to classify opinions and sentiments with more accuracy. In addition, in this paper to create a model of mining based on machine learning and to improve accuracy and performance of the opinion mining method, an improved neural-fuzzy network is proposed.

In this network, we use the model proposed in Takagi and Sugeno (1985) as a neural-fuzzy model and Gaussian membership functions in Reddy and Raju (2009) as fuzzifiers. Employing a neural-fuzzy network brings the advantages of neural network and fuzzy logic at the same time. Some methods proposed recently have paid attention to only the weight based on word frequency in opinion mining and recognition of emotional concepts of some words like (Puri et al. 2018; Mishra et al. 2018; Mostafa 2018; Souza et al. 2018; Akhmedova et al. 2018; Kang et al. 2018) compared to OMLML in which three types of weights are used:

- Sum calculated weights of TF.IDF of words in the opinion.
- The number of positive emotions obtained from the opinions that equals the sum of positive emotional weights of all of its words.
- The number of negative emotions obtained from the opinions that equals the sum of negative emotional weights of all of its words.

In comparison with some methods that have used the polarity of each word in the opinion or many features to analyze and classify the opinion like (Puri et al. 2018; Mishra et al. 2018; Mostafa 2018; Jeong et al. 2018; Yun et al. 2018; Narayan et al. 2018; Souza et al. 2018; Kang et al. 2018), in our method, the polarity of the opinions toward a target word is first determined using a method based on lexicon and textual features of words and sentences. Next, having mapped feature space into a 3-D vector, opinions are analyzed and classified based on a new machine learning method. The decrease in feature space is done to reduce dimensions of the original feature space and training cost. The summary of comparisons between the proposed method and other methods based on existing challenges is presented in Table 9.

*Comparison between the proposed method and other methods based on four criteria, calculation cost, speed, F-measure and dependency on particular field* Before making any comparison, it is necessary to introduce the criteria. Performance evaluation of opinion mining methods is often

**Table 9** Summary of comparisons between the proposed method (OMLML) and other methods based on existing challenges

	Using high dimensions of features' space	Paying attention to only the weight based on word frequency	Training with high cost	Using just lexicon or machine learning	Dependency on particular field
The method proposed in Puri et al. (2018)	✓	✓	✓	✓	✓
The method proposed in Mishra et al. (2018)	✓	✓	✓	×	✓
The method proposed in Bhatnagar et al. (2018)	×	×	×	✓	✓
The method proposed in Tudoran (2018)	×	×	×	✓	✓
The method proposed in Dragoni (2018)	×	×	×	×	✓
The method proposed in Mostafa (2018)	✓	✓	✓	✓	✓
The method proposed in Jeong et al. (2018)	✓	×	✓	×	✓
The method proposed in Karami et al. (2018)	×	×	×	✓	✓
The method proposed in Yun et al. (2018)	✓	×	✓	✓	✓
The method proposed in Rathan et al. (2018)	×	×	×	×	✓
The method proposed in Narayan et al. (2018)	✓	×	✓	✓	✓
The method proposed in Nuortimo and Härkönen (2018)	×	×	×	✓	✓
The method proposed in Souza et al. (2018)	✓	✓	✓	✓	✓
The method proposed in Rozi et al. (2018)	×	×	×	✓	✓
The method proposed in Akhmedova et al. (2018)	×	✓	×	✓	×
The method proposed in Solanki et al. (2019)	×	×	×	✓	×
The method proposed in Kang et al. (2018)	✓	✓	✓	✓	×
OMLML method	×	×	×	×	×

difficult, particularly qualitatively, because different methods with a variety of approaches achieve this goal and applying different criteria to evaluate these methods is not possible. To compare the proposed method and others, the following criteria have been considered and rankings are applied on three different levels: low, medium and high, except the dependency on particular field that is Yes or No.

Given that  $F$ -measure has been defined in the previous section, its definition is ignored here.

- Calculation cost: includes the volume of features' space, the amount of calculations and memory for creating a model, training and achieving the best result.
- Speed: is the amount of run of the opinion mining method in time unit. If the time required to mine opinions

in the document by the opinion mining system is shorter, the speed of the mining process will be higher.

- Dependency on particular field: is the opinion mining approach dependent on a special field and is it useful just for that environment or not?

In this part, our method is compared with those that are comparable based on the proposed criteria.

Dragoni (2018), the researcher, has used two resources called source community and target community, separately for aspect extraction and polarities of users' mining. Consequently, two sources are needed to be prepared as the pre-processing step, in addition to the costs of the main phases of opinion mining. In addition, NLP approach has high calculation cost (Keyvanpour et al. 2018). Therefore, it seems

**Table 10** Summary of comparison between the proposed method (OMLML) and other methods based on the proposed criteria

	Calculation cost	Speed	<i>F</i> -measure	Dependency on particular field
The method proposed in Puri et al. (2018)	Medium	Medium	Medium	Yes
The method proposed in Dragoni (2018)	High	Medium	Low	Yes
The method proposed in Rathan et al. (2018)	Medium	High	High	Yes
The method proposed in Souza et al. (2018)	High	low	Medium	Yes
OMLML method	Medium	Medium	High	No

that the calculation cost of this method is more than that of others. Due to the use of NLP technique to detect the necessary aspects and low speed of running this approach (Keyvanpour et al. 2018) and the necessity of user profile creation in the proposed method, its speed is medium. One of the most common issues in the unsupervised aspect-based approaches is the extraction of false positive aspects (Liu et al. 2015b). Therefore, as this method has used an unsupervised algorithm for aspects extraction and calculating aspects polarities, its *F*-measure is low.

Due to the use of various methods in different phases of the method proposed in Souza et al. (2018), it is necessary to do many calculations to create the opinion mining model. The researchers in Souza et al. (2018) have used the unsupervised-based method in the processing phase of opinion mining, their *F*-measure is not high but regarding the use of different methods based on PSO which is a meta-heuristic algorithm, *F*-measure is improved, and the speed of this method is low.

Puri et al. (2018) proposed an approach using the finite sources and restricted to a limited set of identifying features. Therefore, its runtime decreases. As, in this paper, just pre-defined knowledge has been used for mapping and opinion mining, calculation cost seems to have decreased, but the method's *F*-measure has decreased, as well.

Using different lexicons and supervised learning methods in Rathan et al. (2018) helps to increase *F*-measure. Due to the huge volume of data sets, the researchers in Rathan et al. (2018) have used a feature-level sentiment analysis model. Therefore, it seems that calculation cost to training is less. In this research, a real-time review analysis method is used, which requires a high speed.

In our proposed method, due to a decrease in feature space, dimensions of the original feature space have been reduced. Therefore, training and calculation cost is diminished. Despite dimension reduction in this method, the improved neural-fuzzy network using GA and PSO needs more time. Because the meta-heuristic algorithm calculates the cost of each member of population repeatedly using a cost function until determining optimum values for the parameters, its speed is medium. The method applied for users' opinions mining in this paper combines a machine learning-based method and a lexicon-based

method in order to classify opinions and sentiments with more accuracy and *F*-measure. In addition, in this paper for model creation of mining based on machine learning, in order to improve accuracy, *F*-measure and performance of the opinion mining method, an improved neural-fuzzy network is proposed. According to the obtained results, *F*-measure of OMLML is high.

The summary of comparisons between the proposed method and other methods based on the proposed criteria is presented in Table 10.

## 5 Conclusion

In opinion mining field, although recently some useful methods have been introduced, there are some disadvantages in this area yet. Therefore, in this paper to improve opinion mining methods in social networks, a helpful method based on lexicon and machine learning called OMLML was proposed whose main feature and superiority compared to other methods include addressing challenges like high dimensions of features' space, the ambiguity involved in recognition of emotional concepts of some words, paying attention to only the weight based on word frequency in opinion mining, high training cost based on time or memory used and usability in any particular field simultaneously. In the proposed method, the polarity of the opinions toward a target word was first determined using a method based on lexicon and textual features of words and sentences. Next, having mapped feature space into a 3-D vector, opinions were analyzed and classified based on a new machine learning method. According to the OMLML method, a rich feature space was created focusing on dimension reduction and word weighting. An improved neural-fuzzy network was proposed as well to find optimum solutions instead of using traditional machine learning methods. OMLML was evaluated based on the popular criteria and on two data sets. Given the results, it can be concluded that the proposed OMLML method showed better performance as an opinion mining method in social networks.

## References

- Akhmedova S, Semenkina E, Stanovov V (2018) Co-operation of biology related algorithms for solving opinion mining problems by using different term weighting schemes. In: Madani K, Peaucelle D, Gusikhin O (eds) *Informatics in control, automation and robotics*. Springer, Cham, pp 73–90
- Akter S, Aziz MT (2016) Sentiment analysis on Facebook group using lexicon based approach. In: 3rd international conference on electrical engineering and information communication technology (ICEEICT). IEEE, New York, pp 1–4
- Al-Ayyoub M, Essa SB, Alsmadi I (2015) Lexicon-based sentiment analysis of Arabic tweets. *Int J Soc Netw Min* 2(2):101–114
- Ali F, Kim EK, Kim YG (2015) Type-2 fuzzy ontology-based opinion mining and information extraction: a proposal to automate the hotel reservation system. *Appl Intell* 42(3):481–500
- Anjaria M, Guddeti RMR (2014) A novel sentiment analysis of social networks using supervised learning. *Soc Netw Anal Min* 4(1):181
- Bhatnagar V, Goyal M, Hussain MA (2018) A novel aspect based framework for tourism sector with improvised aspect and opinion mining algorithm. *Int J Rough Sets Data Anal (IJRSDA)* 5(2):119–130
- Bhattacharyya S, Jha S, Tharakunnel K, Westland JC (2011) Data mining for credit card fraud: a comparative study. *Decis Support Syst* 50(3):602–613
- Bland JM, Altman DG (2000) The odds ratio. *BMJ* 320(7247):1468
- Blitzer J, Dredze M, Pereira F (2007) Biographies, bollywood, boomboxes and blenders: domain adaptation for sentiment classification. In: *Proceedings of the 45th annual meeting of the association of computational linguistics*, pp 440–447
- Cambria E, Hussain A (2012) *Sentic computing: techniques, tools, and applications*, vol 2. Springer, Berlin
- Chen L, Qi L (2011) Social opinion mining for supporting buyers' complex decision making: exploratory user study and algorithm comparison. *Soc Netw Anal Min* 1(4):301–320
- Cho SH, Kang HB (2012) Statistical text analysis and sentiment classification in social media. In: *IEEE international conference on systems, man, and cybernetics (SMC)*. IEEE, New York, pp 1112–1117
- Cui A, Zhang M, Liu Y, Ma S (2011) Emotion tokens: bridging the gap among multilingual twitter sentiment analysis. In: *Asia information retrieval symposium*. Springer, Berlin, pp 238–249
- Ding X, Liu B, Yu PS (2008) A holistic lexicon-based approach to opinion mining. In: *Proceedings of the 2008 international conference on web search and data mining*. ACM, New York, pp 231–240
- Dragoni M (2018) Computational advertising in social networks: an opinion mining-based approach. In: *Proceedings of the 33rd annual ACM symposium on applied computing*. ACM, New York, pp 1798–1804
- Eisenstein J (2017) Unsupervised learning for lexicon-based classification. In: *Proceedings of the 31st AAAI conference on artificial intelligence I*, pp 3188–3194
- Fuller R (1995) *Neural fuzzy systems*. CiteSeerX, London
- Grenoble J (2007) *Actes de l'atelier deft'07, plate-forme afia 2007*. <http://delt07.limsi.fr/actes.php>
- Hourali M, Montazer GA (2010) An intelligent approach for constructing domain ontology using Art2 neural network and c-value method. *Comput Eng Syst Appl* 2:79–88
- Imani MB, Keyvanpour MR, Azmi R (2013) A novel embedded feature selection method: a comparative study in the application of text categorization. *Appl Artif Intell* 27(5):408–427
- Jeong Y, Kim S, Yoon B (2018) An algorithm for supporting decision making in stock investment through opinion mining and machine learning. In: *2018 Portland international conference on management of engineering and technology (PICMET)*. IEEE, New York, pp 1–10
- Kaiser C, Bodendorf F (2009) Opinion and relationship mining in online forums. In: *IEEE/WIC/ACM international joint conferences on web intelligence and intelligent agent technologies, WI-IAT'09*, vol 1. IEEE, New York, pp 128–131
- Kang M, Ahn J, Lee K (2018) Opinion mining using ensemble text hidden markov models for text classification. *Expert Syst Appl* 94:218–227
- Karami A, Bennett LS, He X (2018) Mining public opinion about economic issues: Twitter and the us presidential election. *Int J Strateg Decis Sci (IJSDS)* 9(1):18–28
- Karimi Zandian Z, Keyvanpour M (2017) Systematic identification and analysis of different fraud detection approaches based on the strategy ahead. *Int J Knowl Based Intell Eng Syst* 21(2):123–134
- Karimi Zandian Z, Keyvanpour M (2018) MEFUASN: a helpful method to extract features using analyzing social network for fraud detection. *J AI Data Min* 7:213–224
- Kaur J, Saini JR (2014) Emotion detection and sentiment analysis in text corpus: a differential study with informal and formal writing styles. *Int J Comput Appl* 101(9):1–9
- Keyvanpour MR, Karimi Zandian Z, Abdolhosseini Z (2018) A useful framework for identification and analysis of different query expansion approaches based on the candidate expansion terms extraction methods. *Int J Inf Sci Manag (IJISM)* 16(2):19–42
- Khan K, Baharudin BB, Khan A et al (2009) Mining opinion from text documents: a survey. In: *3rd IEEE international conference on digital ecosystems and technologies, DEST'09*. IEEE, New York, pp 217–222
- Krishna BV, Pandey AK, Kumar AS (2018) Feature based opinion mining and sentiment analysis using fuzzy logic. In: Gurumoorthy S, Rao BNK, Gao XZ (eds) *Cognitive science and artificial intelligence*. Springer, Singapore, pp 79–89
- Kushwaha ML, Rathod MSD (2016) Opinion mining of customer reviews based on their score using machine learning techniques. *Analysis* 4:5
- Lee SW, Song YI, Lee JT, Han KS, Rim HC (2012) A new generative opinion retrieval model integrating multiple ranking factors. *J Intell Inf Syst* 38(2):487–505
- Li G, Liu F (2014) Sentiment analysis based on clustering: a framework in improving accuracy and recognizing neutral opinions. *Appl Intell* 40(3):441–452
- Lima ACE, de Castro LN, Corchado JM (2015) A polarity analysis framework for Twitter messages. *Appl Math Comput* 270:756–767
- Liu P, Joty S, Meng H (2015a) Fine-grained opinion mining with recurrent neural networks and word embeddings. In: *Proceedings of the 2015 conference on empirical methods in natural language processing*, pp 1433–1443
- Liu Q, Gao Z, Liu B, Zhang Y (2015b) Automated rule selection for aspect extraction in opinion mining. *IJCAI* 15:1291–1297
- Lloret E, Balahur A, Gómez JM, Montoyo A, Palomar M (2012) Towards a unified framework for opinion retrieval, mining and summarization. *J Intell Inf Syst* 39(3):711–747
- Mishra BK, Sahoo AK, Misra R (2018) Recommendation for selecting smart village in India through opinion mining using big data analytics. In: Saini A, Nayak A, Vyas R (eds) *ICT based innovations*. Springer, Singapore, pp 105–112
- Missen MMS, Boughanem M, Cabanac G (2013) Opinion mining: reviewed from word to document level. *Soc Netw Anal Min* 3(1):107–125
- Montejo-Ráez A, Martínez-Cámara E, Martín-Valdivia MT, Urena-Lopez LA (2012) Random walk weighting over sentiwordnet for sentiment polarity detection on Twitter. In: *Proceedings of the 3rd workshop in computational approaches to subjectivity and sentiment analysis*. Association for Computational Linguistics, New York, pp 3–10



- Mostafa MM (2018) Mining and mapping halal food consumers: a geo-located Twitter opinion polarity analysis. *J Food Prod Mark* 24(7):858–879
- Mudinas A, Zhang D, Levene M (2012) Combining lexicon and learning based approaches for concept-level sentiment analysis. In: *Proceedings of the 1st international workshop on issues of sentiment discovery and opinion mining*. ACM, New York, p 5
- Najar D, Mesfar S (2017) Opinion mining and sentiment analysis for arabic on-line texts: application on the political domain. *Int J Speech Technol* 20(3):575–585
- Narayan R, Rout JK, Jena SK (2018) Review spam detection using opinion mining. In: *Progress in intelligent computing techniques: theory, practice, and applications*. Springer, Berlin, pp 273–279
- Nuortimo K, Härkönen J (2018) Opinion mining approach to study media-image of energy production. Implications to public acceptance and market deployment. *Renew Sustain Energy Rev* 96:210–217
- Ortega-Bueno R, Medina-Pagola JE, Muñiz-Cuza CE, Rosso P (2018) Improving attitude words classification for opinion mining using word embedding. In: *Iberoamerican congress on pattern recognition*. Springer, Berlin, pp 971–982
- Palanisamy P, Yadav V, Elchuri H (2013) Serendio: simple and practical lexicon based approach to sentiment analysis. In: *2nd joint conference on lexical and computational semantics (\*SEM), proceedings of the 7th international workshop on semantic evaluation (SemEval 2013)*, vol 2, pp 543–548
- 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, New York, p 271
- Pang B, Lee L (2005) Exploiting class relationships for sentiment categorization with respect to rating scales. In: *ACL*, pp 115–124
- 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*. Association for Computational Linguistics, New York, pp 79–86
- Poezce F, Ebster C, Strauss C (2018) Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts. *Proc Comput Sci* 130(C):660–666
- Poria S, Cambria E, Winterstein G, Huang GB (2014) Sentic patterns: dependency-based rules for concept-level sentiment analysis. *Knowl Based Syst* 69:45–63
- Poria S, Cambria E, Gelbukh A (2016) Aspect extraction for opinion mining with a deep convolutional neural network. *Knowl Based Syst* 108:42–49
- Powers DM (2011) Evaluation: from precision, recall and *F*-measure to ROC, informedness, markedness and correlation. *J Mach Learn Technol* 2:37–63
- Puri M, Du X, Varde AS, de Melo G (2018) Mapping ordinances and tweets using smart city characteristics to aid opinion mining. In: *2018 companion of the international world wide web conference*. International World Wide Web Conferences Steering Committee, pp 1721–1728
- Rahmani A, Chen A, Sarhan A, Jida J, Rifaie M, Alhadj R (2014) Social media analysis and summarization for opinion mining: a business case study. *Soc Netw Anal Min* 4(1):171
- Rathan M, Hulipalled VR, Venugopal K, Patnaik L (2018) Consumer insight mining: aspect based Twitter opinion mining of mobile phone reviews. *Appl Soft Comput* 68:765–773
- Reddy CS, Raju K (2009) An improved fuzzy approach for COCO-MO's effort estimation using Gaussian membership function. *J Softw* 4(5):452–459
- Rozi M, Mukhlash I, Kimura M et al (2018) Opinion mining on book review using CNN-L2-SVM algorithm. *J Phys Conf Ser* 974:012004
- Se S, Vinayakumar R, Kumar MA, Soman K (2016) Predicting the sentimental reviews in Tamil movie using machine learning algorithms. *Indian J Sci Technol*. <https://doi.org/10.17485/ijst/2016/v9i45/106482>
- Shandilya SK, Jain S (2009) Automatic opinion extraction from web documents. In: *International conference on computer and automation engineering, ICCAE'09*. IEEE, New York, pp 351–355
- Solanki VK, Cuong NHH, Lu ZJ (2019) Opinion mining: using machine learning techniques. In: Agrawal R, Gupta N (eds) *Extracting knowledge from opinion mining*. IGI Global, New York, pp 66–82
- Souza E, Santos D, Oliveira G, Silva A, Oliveira AL (2018) Swarm optimization clustering methods for opinion mining. In: *Natural computing*, pp 1–29 (Online published)
- Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011) Lexicon-based methods for sentiment analysis. *Comput Linguist* 37(2):267–307
- Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans Syst Man Cybern* 15(1):116–132
- Tan SS, Na JC (2017) Mining semantic patterns for sentiment analysis of product reviews. In: *International conference on theory and practice of digital libraries*. Springer, Berlin, pp 382–393
- Tan S, Wang Y, Cheng X (2008) Combining learn-based and lexicon-based techniques for sentiment detection without using labeled examples. In: *Proceedings of the 31st annual international ACM SIGIR conference on research and development in information retrieval*. ACM, New York, pp 743–744
- Tudoran AA (2018) Why do internet consumers block ads? New evidence from consumer opinion mining and sentiment analysis. *Internet Res* 29:144–166
- Xia R, Zong C, Li S (2011) Ensemble of feature sets and classification algorithms for sentiment classification. *Inf Sci* 181(6):1138–1152
- Yang K, Cai Y, Huang D, Li J, Zhou Z, Lei X (2017) An effective hybrid model for opinion mining and sentiment analysis. In: *IEEE international conference on big data and smart computing (BigComp)*. IEEE, New York, pp 465–466
- Yun Y, Hooshyar D, Jo J, Lim H (2018) Developing a hybrid collaborative filtering recommendation system with opinion mining on purchase review. *J Inf Sci* 44(3):331–344
- Zainuddin N, Selamat A, Ibrahim R (2018) Hybrid sentiment classification on twitter aspect-based sentiment analysis. *Appl Intell* 48:1–15
- Zhang Z, Ye Q, Zhang Z, Li Y (2011) Sentiment classification of internet restaurant reviews written in cantonese. *Expert Syst Appl* 38(6):7674–7682

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations'' (in PDF at the end of the article below the references; in XML as a back matter article note).