

A New Opinion Mining Method based on Fuzzy Classifier and Particle Swarm Optimization (PSO) Algorithm

*Samira Bordbar*¹, *Pirooz Shamsinejad*²

¹*Department of Computer & Electrical Engineering, Shiraz branch, Islamic Azad University, Shiraz, Iran*

²*Department of Computer Engineering & Information Technology, Shiraz University of Technology, Shiraz, Iran*

E-mails: samira.bordbar1987@gmail.com p.shamsinejad@sutech.ac.ir

Abstract: *Opinion Mining or Sentiment Analysis is the task of extracting people final opinion about something through their unstructured sentiments. The Opinion Mining process is as follows: first, product features which are most important to a user are extracted from his/her comments. Then, sentiments will be emotionally classified using their emotional implications. In this paper we propose an opinion classification method based on Fuzzy Logic. Up to now, a few methods have taken advantage of fuzzy logic in opinion classification and all of them have imported fuzzy rules into system as background knowledge. But the main challenge here is finding the fuzzy rules. Our contribution is to automatically extract fuzzy rules and their parameters from training data. Here we have used the Particle Swarm Optimization (PSO) algorithm to extract fuzzy rules from training data. Also, for better results we have devised a mutation-based PSO. All proposed methods have been implemented and tested on relevant data. Results confirm that our method can reach better accuracy than current state of the art methods in this domain.*

Keywords: *Opinion mining, sentiment analysis, Particle Swarm Optimization Algorithm, fuzzy classification algorithm.*

1. Introduction

The importance and severity of the competition between organizations and the speed of changing customer needs has made opinion mining as an interesting subject in Business Intelligence. In recent years with emerge of social networks opinion mining has faced new opportunities because of the huge amount of sentiments in these networks. In the new business world finding new customers is so hard, therefore there is a strong desire to maintain existing customers [1]. Companies are constantly looking for solutions to maintain their communications with their customers. In any organization if stakeholders can reach customer satisfaction and loyalty then they can ensure that their long-term activities will be maintained. To do that, companies should

find out customers opinions about their products directly through Customer Relationship Management or indirectly through opinion mining.

Nowadays, with word wide web extension, humans refer to Internet for buying required items or awareness of various topics. Most internet users, check out opinions of different people when they are buying or searching in the internet [2]. People can share their opinions about different topics on many websites and social networks. As a result, there is a high volume of unstructured and scattered data, which makes extracting proper information very time consuming. Therefore it is so important to provide methods that can make this knowledge available for people in summary form. Opinion mining is introduced as a knowledge mining method from existing comments especially on the web [2].

Opinion mining is a comments, opinions and sentiments analysis process, in which opinions are extracted from reviews and comments on a specific topic [3]. Most people consult with others to make decision, when they want to buy something. Information provided by consuler, can influence the final decision to buy or not to buy the desired product or service [4]. Beside opinion mining audiences such as government and companies, people are very interested to find other people's opinions about different subjects, and some systems were created to predict different events based on opinion mining. People leave comments in different topics like movies, productions and politicians. So we can extract all of these opinions and use them for next goals using opinion mining [5].

Opinion mining approaches are divided into statistical methods and knowledge based methods [2]. In statistical methods, first, a set of different features from text will be extracted, then a subset of computed features will be selected, and finally a classifier will be built based on the training data in new feature space. Knowledge based methods use natural language processing and emotional lexicon networks to derive the most important terms in analysing opinions. Then they use derived features to classify opinions based on emotional load of words in the text.

In the field of designing customer opinion classification systems, few methods have used the fuzzy concepts in their work, these methods incorporate fuzzy rules as background knowledge. But extracting fuzzy rules can be considered as the most important challenge ahead. The aim of this study is providing a fuzzy classification that automatically extracts fuzzy rules and required model parameters from training data.

Fuzzy sets are achieved from generalizing the classical theory of sets. A fuzzy set is a set that its members have degree of membership. Fuzzy set were added to classical theory as adjoint by Lotfi Zadeh in 1965. In classical theory of sets, membership in a set is identified as binary expressions based on binary condition whether a member belongs to a set or not. In contrast, in fuzzy theory relative membership in set is allowed. Basically, fuzzy inference is a formulated process from an input data set to an output data set. Models are interpreted based on fuzzy logic including if-then rules. In fuzzy logic degree of membership reflects membership of an element to a fuzzy set. If degree of membership of an element equals zero, that element is out of set and if degree of membership equals to one, that element is quite on set. If degree of membership of an element is between 0 and 1, this value

represents gradual degree of membership [6]. According to customer satisfaction spectrum, fuzzy logic is an appropriate solution to classify customer's opinions. In fact, the proposed approach in this paper is for opinion analysis and customer satisfaction evaluation based on fuzzy clustering. The most important issue considered in fuzzy classifiers design is identifying fuzzy variables such as membership functions type and place, the antecedent and consequent of rules, and number of optimal rules. In fact mentioned parameters, are structured parameters of fuzzy classifier that try to reach the best performance using finding its optimal values.

This problem is considered as a search problem in high-dimensional space in such a way that each point in solution space, present a rule set with special membership function deployed in special places. According to this description it seems metaheuristic algorithms such as evolutionary and swarm intelligent algorithms, are suitable tool to find best parameters of a fuzzy classifier. The feature that makes these approaches attractive is that they can achieve an optimal estimate for all effective parameters in fuzzy classifier automatically by defining a suitable fitness function. An approach proposed in this paper is for optimal design of a fuzzy classifier by applying Particle Swarm Optimization (PSO). This algorithm can estimate membership function position, required fuzzy rules and required number of rules simultaneously and this should be done without user intervention. We use PSO because of simplicity of this algorithm and its power for fast searching of solution space and its fast convergence to optimal solution.

The rest of this paper is organized as follows, related works are presented in Section 2, proposed model is described in Section 3, Section 4 include the obtained results and conclusion and future works are presented in Section 5.

2. Literature review

Jebaseeli and Kirubakaran [7], presented a statistical approach to classify opinions using a combination of genetic algorithm and artificial neural network. In this paper, they use Genetic Algorithm (GA) to select the best NNs training parameters. Their results show better quality of proposed approach to classify customer's opinions and comments.

Jusoh and Alfawareh [8], proposed an approach for classifying opinions based on fuzzy logic. Customer's comments are different in term of having emotion, and the purpose of classification is to separate sentences. Their method, divide customer's opinions in three categories include positive, negative and neutral. This research shows that fuzzy logic is capable of detection of boundaries between these categories. In addition, degree of positivity or negativity of this opinions also is expressed and in each category difference between different scenarios has been clear.

Bagheri et al. [9], present an unsupervised model to detect different dimensions (aspects) of an opinion. In this model first a method based on multi-word has been proposed to extract different dimensions of an opinion. Afterwards effect of a word to detect different dimensions of an opinion is investigated using a heuristic rule set and then a new metric is introduced based on mutual information and frequency of the specific dimension (aspect). The advantage of this model which is a

statistical is that training data don't need labelling and can apply on different data and languages.

Stylios, Katsis and Christodoulakis [10], proposed a statistical approach based on ant colony to identify user satisfaction. In the approach proposed which is based on ant colony behaviour, each user is assigned to an ant. These ants try to calculate user satisfaction of various products and try to learn this by following user opinion path. Using obtained results of this approach one can predict satisfaction of similar products based on interests of one or more users or similar product clustering. Also how to distribute and update pheromones of a position (product) is determined based on user's satisfaction. The advantage of this approach is to obtain high accuracy in identifying users' satisfaction.

Kalaivani and Shunmuganathan [11], proposed a new method to classify customers' opinions using GA and k-NN. Proposed method uses a statistical approach and probabilistic comparison between GA chromosomes as crossover operation that will speed up convergence to best chromosome. In this paper, k-NN is applied which is known as the best approach to classify English texts. The advantage of this method is its simplicity for implementation. The result show improvements in accuracy.

Dalal and Zaveri [12], proposed a new approach based on knowledge to explore opinions of customers of an electronic store using fuzzy logic. In this approach, they use a mamdani membership function and maps the membership of each opinion to fuzzy sets based on customers' opinions. The advantage of fuzzy logic is ability to classify data in different related categories between both good and bad. Many of recent researches, classify data only in two categories: Good and bad which is suitable for many applications. With this advantage, proposed approach for extracting fuzzy rules are still faced with the problem of generation and extraction of rules which is done manually. Hence, there is a need for methods and mechanisms which first extract fuzzy rules based on existing data, and then fuzzy logic is applied based on these extracted rules.

Sumathi, Karthik and Marikkannan [13], proposed an approach based on swarm intelligence honey bee colony for feature selection. This approach uses a weighting system for feature selection. In this paper, for extracting and selecting features, an operation is conducted in two stages: First, features are weighted using cross information method, then they use honey bee breeding algorithm to select subset of remaining features. They classify documents using k-NN method in which its cluster centres will be found using Honey bee colony algorithm. The advantage of this approach is the high speed of bee colony algorithm. Results show some improvements using this approach.

Rahmath and Ahmad [14], proposed a new knowledge-based method to explore online customers' comments about products based on fuzzy logic and using machine learning techniques. Authors in this paper, conducted a system to extract comments in several stages which consist of pre-processing of a documents data, and present a fuzzy system to extract features based on polarity of them. Proposed approach offers a more accurate system for customers' opinions classification by

utilizing a fuzzy function. Evaluation and experimental results show that proposed system can perform analysing of emotions with an accuracy of 93.85%.

Kalaivani and Shunmuganathan [15], offer a statistical method to classify customers' opinions, emotions and comments. This method uses Genetic Algorithm to reduce feature set. Classification method used in this paper is Bayesian Network. The important advantage of this method is simplicity of Bayesian algorithm that decreases the time complexity of proposed method. In contrast, one disadvantage of this method is that it requires a completely independent features. Thus, if there is correlation between features, performance of this method will be affected.

In 2016 Bilal et al. [16] proposed a new method for classifying sentiment using naïve Bayesian, Decision tree and k-NN. Written reviews in both languages Persian and Urdu were extracted from a weblog. These extracted reviews in textual files have been labelled to provide a training dataset including 150 positive comments and 150 negative comments. Test dataset were presented in three different models and the results were analysed. Presented results by authors show that Naïve Bayesian are superior to decision tree and k-NN in terms of accuracy, precision, recall, and F-measure.

Although, notice that most of works were done with English and Chinese language, and middle eastern languages such as Persian and Arabian attract less attention and there are a few works about them [17].

Bagheri, Saraee and de Jong [17], studied on opinion mining in Persian language. In this paper, they used different feature selection methods such as DF, TFV, MI, and MMI and a new algorithm called MMI is presented to classify emotions. There are three challenges in Persian languages: 1) there are suffix and prefix, 2) there are space and half-space between some words, and 3) there are conversational words. Authors in this paper, these three challenges are investigated by presenting a model to classify emotions at document level. Proposed model is based on rooting and feature selection and it evaluate using Naïve Bayes classifier. The results show that feature selection can increase efficiency in semantic analysis.

3. Proposed method

The method proposed herein is presented in Fig. 1.

According to the flowchart presented for proposed approach, first of all users' opinions set which is the raw dataset is given to system as an input, then the pre-processing operation will be applied to prepare dataset, and important feature of each opinion is extracted and is changed to tabular mode. In next step, feature selection process is performed to achieve optimal features by TF feature selection algorithm. Because of imbalanced class label in dataset, we need to balance our dataset. Our dataset is now ready to enter the proposed algorithm. After separating training and test data, training data is given to PSO algorithm to extract fuzzy rules automatically which will be explained in next sections. Extracted rules from training data are given to fuzzy knowledge base. These rules are applied on test data, and membership amount of each opinion to each category is determined, and after identifying the label of each record of data, accuracy of proposed approach is evaluated. The details of each step will be explained in next sections.

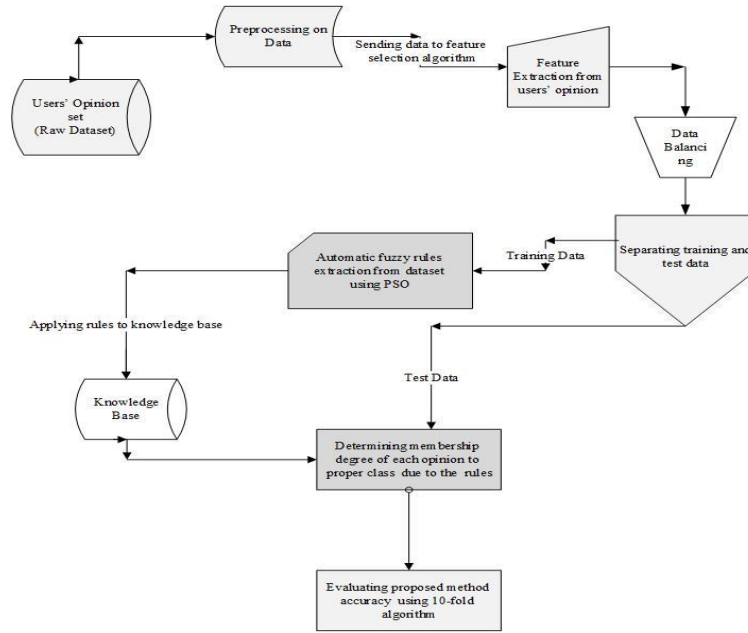


Fig. 1. Proposed method

3.1. Pre-processing

Users' opinions set about electrical product such as mobile phone and camera are given to the system as input.

The pre-processing phase contains extracting features for each opinion and their weights which can be positive or negative. Each opinion belongs to one of the good, very good, bad and very bad classes. Now labelled data are ready for entering to the next phase where feature selection will be performed.

3.2. Feature selection

In this work TF weighting method [3] has been used for feature selection. TF has two types of binary and non-binary. In binary type, weight is between 0 and 1, but in non-binary type, there is no limitation for assigned weights. TF can be computed using Formula (1).

$$(1) \quad TF(t, d) = 0.5 + \frac{0.5 \times F(t, d)}{\text{Max}\{F(w, d): \text{Wed}\}}$$

where $F(t, d)$ is the frequency of term t in document d and $\text{Max}\{F(w, d)\}$ is the number of the most repeated word in document d . It's worth noting that weights will be between 0.5 and 1.

3.3. Balancing data

Due to data imbalance problem, we need to use one of the balancing techniques, because in most of the classification algorithms, it tends for class which has the most number of instances. Hence, it shows less ability to predict minority class data correctly. In such algorithms, instances of minority class are classified incorrectly

compared to instances of majority class. In this paper, we use Synthetic Minority Oversampling TEchnique (SMOTE) [18].

This algorithm creates Synthetic data based on similarity of feature space in instances of minority class. For subset $s_{\min} \in s$, we choose k nearest neighbour samples $x_i \in s_{\min}$ where k is an integer number. The method of choosing k nearest neighbours of samples is based on Euclidian distance in n -dimensional space. To create synthetic sample, we choose one of the k nearest neighbours of x_i sample randomly, then new sample can be obtained by multiplying the difference between these two samples to a random number in $[0, 1]$ and adding the results to x_i sample. Data in dataset, balance in this way and after this step, prepared data for extracting rules go to next step.

3.4. Extracting rules for fuzzy pattern classifying system using PSO

In this section, particle swarm optimization algorithm is presented to extract rule for a fuzzy pattern classifying system. This approach extracts system rules from an uncertain, confused and incomplete training dataset. Proposed approach in this paper is for training uncertain language rules from training data and promote performance of pattern classifying system based on uncertain rule.

In this section, we present PSO to extract and tune rule for classifying pattern based on our proposed uncertain rule based on uncertain dataset.

Step 1. Initialize

In this structure, each rule is treated as a particle and a population consists of a fixed number of classification rules. Each solution is a set of rules which has maximum fitness. Each feature in antecedent section of rule in initial population is selected from dataset number range that is $[-3, 3]$ in fuzzy pattern space randomly. The method of determining of consequent (rules class) is explained completely in following sections. Now compatibility of each processed opinion in dataset with each of rules should be determined. For this purpose, a triangle membership function is determined by three parameters $\{a, b, c\}$. Here lower bound of particle values is shown with a parameter that is in range $[-3, 2]$ and upper bound of each value is shown with c parameter and is between lower bound and 3, these values at the start of work generate randomly by PSO. Required input of triangle function is x which is the same amount of features are included in the dataset.

Hence, the method which is explained in this section is applied in order to create language classifying rules for our uncertain classifying system [19, 20]. Uncertainty in rules is investigated by fuzzy sets in antecedent section of rules. The structure of each rule in this paper for classifying pattern based on uncertain rule for a pattern is as follows.

Here, fuzzy sets, $i = 1, \dots, M$ is number of rules, C_i is cluster label and G_{ci} is certainty degree of rules. This rule has the capacity to consider uncertainties in each of both training dataset and applied mathematical models to measure features in antecedent section of rule. Our proposed method uses training data to extract language rules and manage uncertainty in pattern classifying problems based on rule. Proposed learning rule algorithm is applicable for classifying multidimensional pattern problems with high degree of uncertainty in measurement and language

phrases such as opinion mining program of users about usable productions. The proposed algorithm phases for rule selection are as follows.

Step 2. Determining degree of compatibility of one record with rules

After calculating membership function, compatibility degree of pattern with classifying language rule j , R_j can be calculated with the equation

$$(2) \quad \tilde{A}_i(X_p) = \tilde{A}_{j_1}(X_{p_1}) \cap \tilde{A}_{j_2}(X_{p_2}) \cap \dots \cap \tilde{A}_{j_n}(X_{p_n}),$$

where \cap is one of the fuzzy operations (T -norm) which shows minimum operation. In fact, first degree of compatibility of each feature on a record with each feature of a rule is determined, then compatibility degree of a rule with each record is obtained by calculating the minimum of this compatibility degrees. So we calculate compatibility degree for each record using rules. Total compatibility degree $\tilde{\beta}_{\text{class}h}$ of pattern in each category for classifying rule j , R_j of a fuzzy set is calculated as follows:

$$(3) \quad \tilde{\beta}_{\text{class}h} = \sum_{\forall X_p \in \text{class}h} \tilde{A}_{i1}(X_p).$$

Step 3. Determining of consequent (class) of rules

As mentioned in previous section, antecedent of a rule is generated randomly. Now to obtain consequent (C_j) of rule R_j category with maximum compatibility is selected. It means we use $\tilde{\beta}_{\text{class}h}$ measure that is calculated for one rule with all records of dataset which are member of one class. Class of dataset which maximize $\tilde{\beta}_{\text{class}h}$ is considered as consequent of rule.

If $\tilde{\beta}_{\text{class}h}$ is not exclusive, then two or more category have the same $\tilde{\beta}_{\text{class}h}$ and so consequent of rule set to \emptyset which means empty category. A rule with empty category in consequent section is named dummy rule.

Step 4. determining of certainty degree of rules

Certainty degree G_{c_j} of a dummy rule is 0, and degree of certainty of a non-dummy rule is calculated as follows:

$$(4) \quad G_{c_j} = \frac{\beta_{\text{class}h} - \beta'}{\sum_{k=1}^c \beta_{\text{class}k}},$$

$$(5) \quad \beta' = \frac{\sum_{h \neq k} \beta_{\text{class}k}}{c-1}.$$

At the end of this step, dummy rules is extracted and removed, then remaining rules are optimized with PSO and top rules are extracted.

Step 5. Rule optimization using PSO

Particle swarm optimization algorithm which is used in this paper has lower and upper bound for speed and position because of combining with fuzzy algorithm. The values in global best and personal best in each generation have lower and upper bound as well. Speed and position for particles update are calculated according to equations (6) and (7).

$$(6) \quad v_i^{k+1}.LB = wv_i^k + c_1 \text{rand}_1 \times (pbest_i.LB - x_i^k) + c_2 \text{rand}_2 \times (gbest_i.LB - x_i^k),$$

$$(7) \quad v_i^{k+1}.UB = wv_i^k + c_1 \text{rand}_1 \times (pbest_i.UB - x_i^k) + c_2 \text{rand}_2 \times (gbest_i.UB - x_i^k),$$

$$(8) \quad x_i^{k+1}.UB = x_i^k.UB + v_i^{k+1}.UB,$$

$$(9) \quad x_i^{k+1}.LB = x_i^k.LB + v_i^{k+1}.LB,$$

where: v_i^k .LB, is Lower Bound (LB) of particle in iteration k ; v_i^{k+1} .LB, is lower bound of particle speed in iteration $k+1$; w is inertia weight and c_j is acceleration coefficient that in this paper these two values are set in adaptive way; $j=1, 2, \dots$; rand_i , is a random number between 0 and 1; $i=1, 2, \dots$; x_i^k is current position of particle i in iteration k ; pbest_i .LB, is lower bound of personal best of each particle i ; Gbest .LB, is lower bound of global best particle in population; x_i^{k+1} is position i in iteration $k+1$. Particle swarm optimization algorithm, provides a mapping of all opinion mining features to desired fuzzy logic rules set based on model which is described in next section.

Step 6. Fitness evaluation

A fitness value was assigned to each language classifying rule R_j in current iteration as follows:

$$(10) \quad \text{Fitness}(R_j) = (W_{\text{NCC}} * N_{\text{CC}}(R_j) - W_{\text{NMC}} * N_{\text{MC}} * N_{\text{MC}}(R_j)),$$

where NCC is the number of patterns which were classified by R_j classification rule correctly, NMC is the number of patterns which were classified incorrectly and W_{NCC} , W_{NMC} are weights of classified and unclassified. In fact, this function is PSO goal function that is used for calculating personal best and global best and this function is improved due to algorithm iterations and finally the value of this function is a criterion for selection of a good rule. Function value has to be maximized to find top rules.

Step 7. End

Stop criterion of PSO has reached the desired fitness value. After this step optimal rules are determined. These rules are used for next step and for test data classification.

Step 8. Test data classification using extracted fuzzy rules

In this section, each opinion is classified. To do this, we have to apply extracted rules by PSO on test data. Here we explain this subject:

Step 9. Fuzzy reasoning to classify new pattern

Here fuzzy reasoning is applied to classify new pattern. Suppose, we have a subset S from a generated rule set, a new pattern is classified in rule $R_j \in S$, $j = 1, \dots, M$, which is maximized by class h ,

$$(11) \quad \alpha_j(X_p) = \max\{\tilde{A}_{i1}(X_p) \cdot Gc_j\}.$$

Step 10. Defuzzification

In this step, central defuzzification is used by following defuzzification problem from a fuzzy set to calculate α_j . Accordingly, label of class h with rule $\alpha_j(X_p)$ is defined as follows:

$$(12) \quad \alpha_{\text{class } h} = \max\{(\alpha_j)(X_p) | \text{class } h = C_j \text{ and } R_j \in S\}.$$

This means that each rule which has maximum α is determinant for class of the record in test data. In addition, if two or more category have maximum $\alpha_{\text{class } h}$, set-1 for $\alpha_{\text{class } h}$ which means the pattern X_p can't be classified by rule set S and it is an unclassified pattern. Value of $\alpha_{\text{class } h}$ can be treated as trust criteria for assigning pattern X_p to category.

3.5. Mutated particle swarm optimization

Since PSO may stay in local optimum, we use genetic mutation in this paper to get out from this situation. Three mutation operations in genetic are namely: swap, which exchange position of two random element of candidate solution randomly, reversion, which invert a random selected range of candidate solution, insertion, which inserts a random element after another random element in candidate solution. Any time one of these operation is applied on candidate solution is a mutation.

4. Results and analysis

In this section, all of simulation steps and generated results are reviewed, evaluated and compared. All diagrams plotted using Excel are according to the outputs generated from the simulating done by data mining application such as KMIME and rapid miner and Matlab 2015 and C# programming language.

4.1. Important parameters settings and dataset

The method proposed in this paper is combined from two Fuzzy and PSO algorithms and then we added mutation and we named this method as F-MPSO (Fuzzy-Mutated PSO). There is many parameters in evaluating this method. In this section parameters with more important role are presented and the method is evaluated with different parameters. Parameters of PSO are adapted in the algorithm. First parameter which is important in evaluating of accuracy, is number of particles in PSO which constitute the rules. Second parameter is number of iterations in PSO to reach optimal value of fitness function. Next parameter is threshold value in feature selection and the last parameter is threshold value to detect if data is in a right or wrong class by rules in fitness function. There are weights to calculate fitness function whose optimal values for these functions can be achieved by trial and error approach. In this section, different parameters used in this method are investigated and method accuracy is calculated by 10-Fold validation. This approach works in such a way that in each run data is divided into 10 equal parts, 90% of data for training and 10% of data for test. By repeating this action 10 times and calculating the accuracy each time, finally after averaging all of the calculated accuracies final accuracy can be achieved.

In this work 9 files are used as Datasets (www.cs.uic.edu/~liub/FBS/Reviews-9-products.rar), each of these files consists of information about one of the desired products. Each row in this file expresses one opinion. In each file features of each sentence and the amount of positivity or negativity of each feature has listed before each comment. The positivity and negativity of each feature is between -3 and $+3$ that -3 shows minimum satisfaction and $+3$ is for maximum customer satisfaction.

4.2. Simulation results of proposed method

We use Canon Power Shot SD 500 as dataset and optimal values for parameters were obtained by trial and error and finally this settings are applied to dataset and accuracy is calculated. We compared the proposed method with presented methods in sources [3, 14, 16]. Since we don't have access to dataset used in second and third papers,

presented algorithms are implemented and the results are compared to our proposed algorithm.

As shown in Table 1 the algorithm is evaluated for each different parameter. Best accuracy is obtained about 0.57 for feature selection threshold and 0.4 for correct rule detection in FPSO algorithm. Since the accuracy of 0.4 threshold is higher the F-MPSO is only executed in this threshold and results are shown in Table 1. As shown in Table 1 in the three last rows of results, proposed method has been implemented with different number of particles and different number of iteration. You can see Accuracy comparison in Canon PowerShot SD dataset with F-PSO in Fig. 2.

Table 1. Calculation of accuracy in Canon PowerShot SD dataset

Threshold feature selection	Threshold rule	Number of particle	Iteration	Accuracy F-PSO	Accuracy F-MPSO
0.51	0.1	40	100	90.83%	
	0.3			92.64%	
	0.4			94.05%	94.93%
0.54	0.1	40	100	89.23%	
	0.3			91.08%	
	0.4			92.78%	92.88%
0.55	0.1	40	100	88.97%	
	0.2			91.20%	
	0.4			93.91%	93.97%
0.57	0.1	40	100	91.26%	
	0.3			92.96%	
	0.4			94.91%	95.89%
0.58	0.1	40	100	91.05%	
	0.3			92.93%	
	0.4			93.95%	93.53%
0.57	0.4	30	200	93.23%	93.52%
0.57	0.4	20	40	90.27%	91.16%
0.57	0.4	50	100	94.52%	94.73%

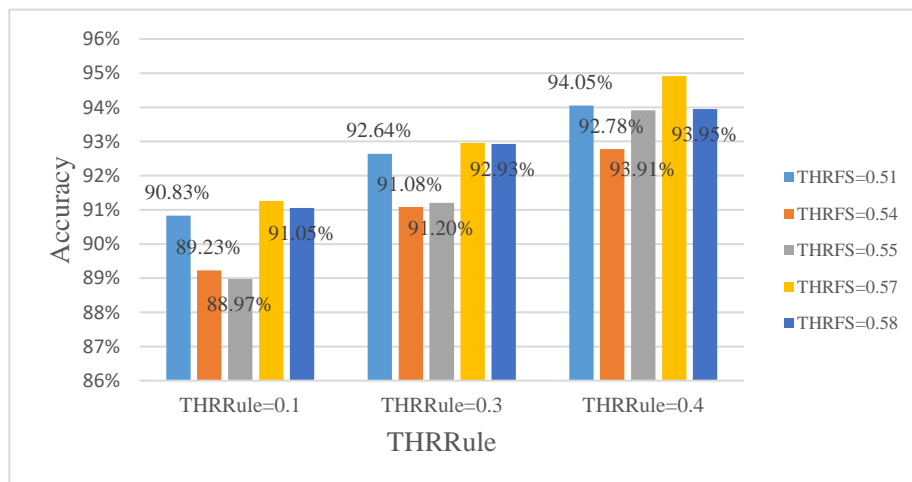


Fig. 2. Accuracy comparison in Canon PowerShot SD dataset with F-PSO

Accuracy values are presented according to the different feature selection threshold and also different threshold for correct identification rules. Maximum accuracy occurred in feature selection threshold 0.57.

In Fig. 3 the advantage of F-MPSO over F-PSO is shown on a sample dataset and compared for different feature selection thresholds.

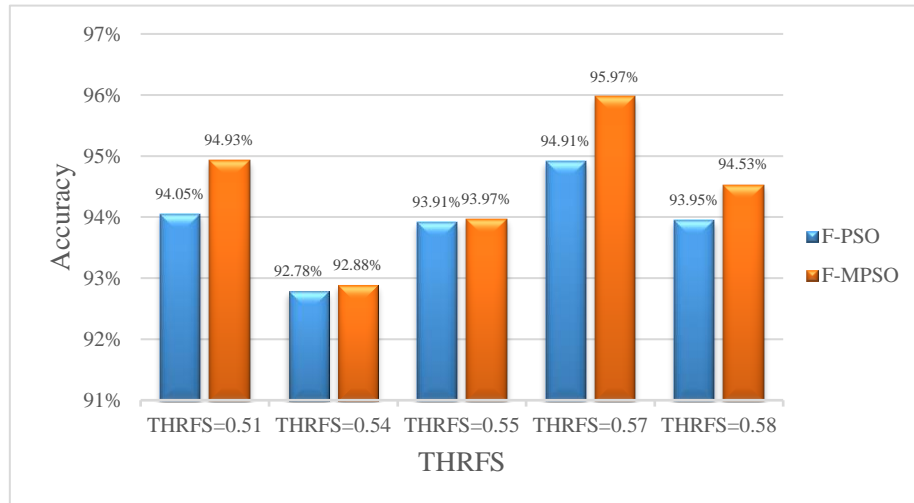


Fig. 3. Comparison of two algorithms F-PSO and F-MPSO on Canon PowerShot SD dataset

Accuracy values are calculated for other datasets with these optimal parameters. These parameters include feature selection threshold 0.57, rules identification threshold 0.4, number of particles 40, and number of iteration in PSO 100. Table 2 indicates the results obtained from implementing proposed algorithm on data after calculating the accuracy. Table 2 shows the results of executing different algorithms compared to the proposed method. Comparison of accuracy of other classifier algorithms with proposed methods are shown in Fig. 4.

Table 2. Comparing the results of implementation of the proposed algorithm with other methods based on accuracy criteria

Dataset	Accuracy k-NN	Accuracy decision tree	Accuracy naïve bayes	Accuracy SVM	Accuracy SVM-PSO	Proposed algorithm F-PSO	Proposed algorithm F-MPSO
Canon powershot SD500	89.05%	91.02%	90.63%	90.68%	91.82%	94.91%	95.89%
Canon S100	85.01%	83.1%	88.6%	82.6%	87.8%	95.83%	97.51%
Diaper champ	87.08%	88.6%	89.2%	86.2%	89.0%	95.78%	96.33%
Hitachi router	80.13%	79.80%	79.81%	78.46%	79.6%	96.42%	97.33%
Ipod	87.89%	85.2%	86.89%	87.93%	88.25%	94.97%	96.64%
Linksys router	82.25%	81.02%	84.56%	83.37%	87.97%	93.86%	93.6%
Micro MP3	65.1%	64.8%	63.8%	67.8%	74.2%	95.72%	95.67%
Nokia 6600	70.02%	75.11%	75.42%	71.25%	72.6%	95.87%	95.80%
Norton	81.90%	80.08%	87.24%	88.14%	89.68%	93.98%	95.89%

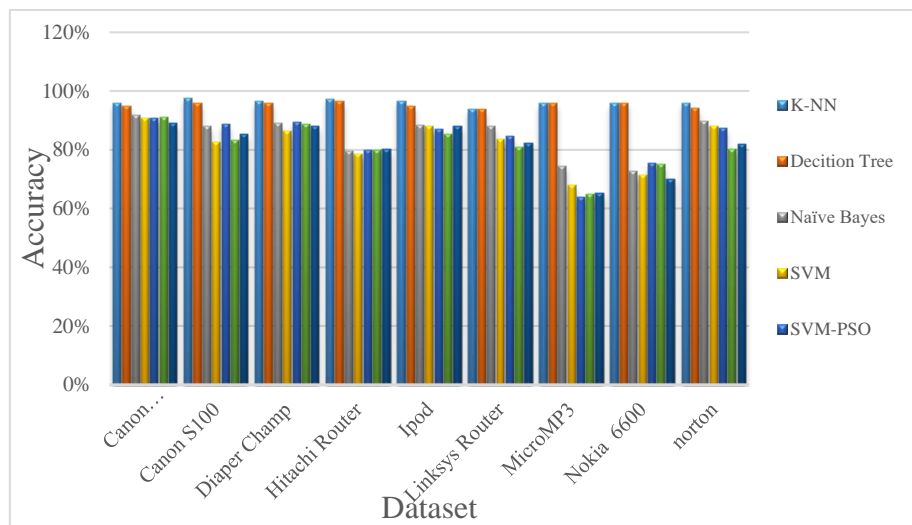


Fig. 4. Comparison of accuracy of other classifier algorithms with proposed methods

The results are a good indication of the advantages of the algorithm being proposed compared to other methods. It is also obvious that maximum accuracy in F-PSO is related to Hitachi Router dataset which is 96.42%, and in F-MPSO is related to Canon 100 dataset which is 97.51%. Generally both algorithm F-PSO and F-MPSO have more satisfying results in comparison with other algorithms.

Fig. 5 illustrates Comparison of average accuracy of proposed method with other methods.

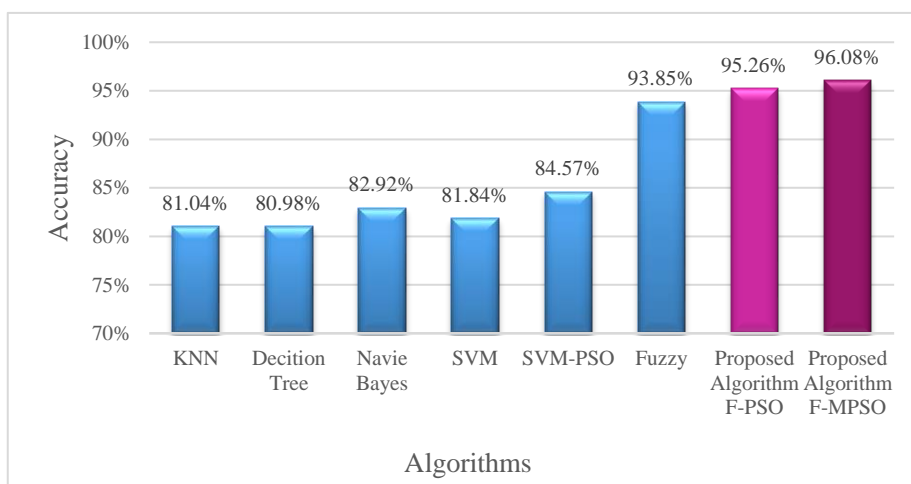


Fig. 5. Comparison accuracy of proposed method with other methods

4.3. Comparison of results of proposed method with other methods

Now we calculate the average accuracy of presented methods to compare the accuracy of proposed method to accuracy of the other methods.

After averaging the accuracy of proposed method and other methods, it can be observed that proposed method (F-PSO) has a greater improvement compared to other presented methods.

Table 3. Comparison of average accuracy of proposed method with other methods

Algorithms	AVG Accuracy
k-NN	81.04%
Decision Tree	80.98%
Naive Bayes	82.92%
SVM	81.84%
SVM-PSO	84.57%
Fuzzy	93.85%
Proposed algorithm F-PSO	95.26%
Proposed algorithm F-MPSO	96.08%

5. Conclusion

The main contribution of this paper is to extract fuzzy rules from opinion mining of data automatically. Previous methods input these rules in the model as background knowledge, while our method being proposed uses a hybrid algorithm from fuzzy and particle swarm optimization algorithms as a new method for classifying opinions data. Proposed method consists of three phases: 1) performing pre-processing on text data, 2) extracting fuzzy parameters and optimal rules, and 3) finally applying extracted rules on test data, data classification and calculating the accuracy of proposed method. In proposed method extracting of fuzzy parameters is performed automatically and without user intervention. Performed simulation indicates acceptable accuracy of proposed method rather than other methods. Some of recommendations which can improve performance of proposed method are as follows:

Presenting a method to estimate features without value in data to improve the performance of this method.

Use of fuzzy logic 2 can also cause improvement in accuracy of proposed method. Combining evolutionary algorithms and covering the issues in one of them by advantages of the other one, can improve performance.

References

1. Ravi, K., V. Ravi. A Survey on Opinion Mining and Sentiment Analysis: Tasks, Approaches and Applications. – Knowledge-Based Systems, Vol. **89**, 2015, pp. 14-46.
2. Balazs, J. A., J. D. Velásquez. Opinion Mining and Information Fusion: A Survey. – Information Fusion, Vol. **27**, 2016, pp. 95-110.
3. Basari, A. S. H., B. Hussin, I. G. P. Ananta, J. Zeniarja. Opinion Mining of Movie Review Using Hybrid Method of Support Vector Machine and Particle Swarm Optimization. – Procedia Engineering, Vol. **53**, 2013, pp. 453-462.
4. Ye, Q., Z. Zhang, R. Law. Sentiment Classification of Online Reviews to Travel Destinations by Supervised Machine Learning Approaches. – Expert Systems with Applications, Vol. **36**, 2009, No 3, pp. 6527-6535.

5. Virmani, D., V. Malhotra, R. Tyagi. Sentiment Analysis Using Collaborated Opinion Mining. – arXiv preprint arXiv:1401.2618., 2014.
6. Zadeh, L. A. Fuzzy Sets. – Information and Control, Vol. **8**, 1965, No 3, pp. 338-353.
7. Jebaseeli, A. N., E. Kirubakaran. Genetic Optimized Neural Network Algorithm to Improve Classification Accuracy for Opinion Mining of m-Learning Reviews. – IJETTCS, Vol. **2**, 2013, No 3, pp. 345-349.
8. Jusoh, S., H. M. Alfawareh. Applying Fuzzy Sets for Opinion Mining. – In: IEEE International Conference on Computer Applications Technology (ICCAT'13), 2013, pp. 1-5.
9. Bagheri, A., M. Saraee, F. de Jong. An Unsupervised Aspect Detection Model for Sentiment Analysis of Reviews. – In: International Conference on Application of Natural Language to Information Systems. Berlin, Heidelberg, Springer, 2013, pp. 140-151.
10. Stylios, G., C. D. Katsis, D. Christodoulakis. Using Bio-Inspired Intelligence for Web Opinion Mining. – International Journal of Computer Applications, Vol. **87**, 2014, No 5.
11. Kalavani, P., K. L. Shunmuganathan. An Improved k-Nearest-Neighbor Algorithm Using Genetic Algorithm for Sentiment Classification. – In: 2014 IEEE International Conference on Circuit, Power and Computing Technologies (ICCPCT'14), March 2014, pp. 1647-1651.
12. Dalal, M. K., M. A. Zaveri. Opinion Mining from Online User Reviews Using Fuzzy Linguistic Hedges. – Applied Computational Intelligence and Soft Computing, 2014, No 2.
13. Sumathi, T., S. Karthik, M. Marikkannan. Artificial Bee Colony Optimization for Feature Selection in Opinion Mining. – Journal of Theoretical & Applied Information Technology, Vol. **66**, 2014, No 1.
14. Rahmath, P., T. Ahmad. Fuzzy Based Sentiment Analysis of Online Product Reviews Using Machine Learning Techniques. – International Journal of Computer Applications, Vol. **99**, 2014, No 17, pp. 9-16.
15. Kalavani, P., K. L. Shunmuganathan. Feature Reduction Based on Genetic Algorithm and Hybrid Model for Opinion Mining. – Scientific Programming, Vol. **12**, 2015.
16. Bilal, M., H. Israr, M. Shahid, A. Khan. Sentiment Classification of Roman-Urdu Opinions Using Naïve Bayesian, Decision Tree and k-NN Classification Techniques. – Journal of King Saud University-Computer and Information Sciences, Vol. **28**, 2016, No 3, pp. 330-344.
17. Bagheri, A., M. Saraee, F. de Jong. ADM-LDA: An Aspect Detection Model Based on Topic Modelling Using the Structure of Review Sentences. – Journal of Information Science, Vol. **40**, 2014, No 5, pp. 621-636.
18. Chawla, N. V., K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer. SMOTE: Synthetic Minority Over-Sampling Technique. – Journal of Artificial Intelligence Research, Vol. **16**, 2002, pp. 321-357.
19. Mendel, J. M. Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions. Upper Saddle River, Prentice Hall PTR, 2001, pp. 131-184.
20. Ishibuchi, H., Y. Nojima. Pattern Classification with Linguistic Rules. – In: Fuzzy Sets and Their Extensions: Representation, Aggregation and Models, 2008, pp. 377-395.

Received 25.11.2017; Second Version 26.01.2018; Accepted 15.02.2018