

Study on Sentiment Classification Strategies Based on the Fuzzy Logic with Crow Search Algorithm

Mazen Mohammed

Central South University

Lasheng Yu (✉ yulasheng@csu.edu.cn)

Central South University <https://orcid.org/0000-0001-7078-9068>

Ali Aldhubri

Nanjing University of Science and Technology

Gamil R. S.Qaid

Sana'a University

Research Article

Keywords: Sentiment analysis, Fuzzy rule-based system, Membership function, Crow search algorithm

Posted Date: January 11th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1216679/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Study on Sentiment Classification Strategies Based on the Fuzzy Logic with Crow Search Algorithm

Mazen Sharaf AL-Deen¹, Yu lasheng^{1*}, Gamil R. S. Qaid²

¹School of Computer Science and Engineering, Central South University, Changsha 410083, China.

²Department of computer engineering, Faculty of computer science and engineering, Hodeida University.

Abstract

In recent times, sentiment analysis research has gained wide popularity. That situation is caused by the nature of online applications that allow users to express their opinions on events, services, or products through social media applications such as Twitter, Facebook, and Amazon. This paper proposes a novel sentiment classification method according to the Fuzzy rule-based system (FRBS) with crow search algorithm (CSA). FRBS is used to classify the polarity of sentences or documents, and the CSA is employed to optimize the best output from the fuzzy logic algorithm. The FRBS is applied to extract the sentiment and classify its polarity into negative, neutral, and positive. Sometimes, the outputs of the FRBS must be enhanced, especially since many variables are present and the rules between them overlap. For such cases, the CSA is used to solve this limitation faced by FRBS to optimize the outputs of FRBS and achieve the best result. We compared the performance of our proposed model with different machine learning algorithms, such as SVM, maximum entropy, boosting, and SWESA. We tested our model on three famous data sets collected from Amazon, Yelp, and IMDB. Experimental results demonstrated the effectiveness of the proposed model and achieved competitive performance in terms of accuracy, recall, precision, and the F-score.

Keywords: Sentiment analysis _ Fuzzy rule-based system _ Membership function _ Crow search algorithm

1. Introduction

Given the rapid development of online applications, whether social media or those used in e-commerce (Amazon, Facebook, and Twitter), customer reviews are now readily available. Analyzing these reviews and revealing their opinions are valuable for improving the business process. Reviews can help the customer evaluate a specific service or product quickly. Moreover, reviews offer essential evidence for decision-making regarding service providers. Sentiment analysis is also known as “opinion mining” or “emotion Artificial Intelligence” and alludes to the utilization of natural language processing (NLP), text mining, computational linguistics, and bio measurements to methodically recognize, extricate, evaluate, and examine emotional states and personal information. Moreover, sentiment analysis is generally concerned with the voice in client materials, such as the surveys and reviews on the Web and in Web-based social networks ([1]).

As a generic rule, sentiment analysis attempts to determine the author's behavior of articles, speakers, or other topics through intense emotional or emotional responses to archive, appropriate, or communicative. This action may be a judgment or evaluation full of emotion (in other words, the sentimental state of the creator or speaker) or the expectation of enthusiastic responses (that is, the effect intended by the originator or buyer). Numerous customer surveys or recommendations on all topics are available on the Web. These periods and audits may contain surveys on items such as on customers or fault-finding of films. Surveys are expanding rapidly, given that individuals like to provide their views on the Web. Many surveys are accessible for solitary items, which makes it problematic for customers as they

must peruse each survey to make a choice. Next, mining this information, special customer assessments, and organizing them is a vital pledge. Sentiment mining is a task that takes advantage of NLP and information extraction (IE) approaches to analyze an extensive number of archives for ascertaining the sentiments of comments posed by different authors ([1];[2]). This process incorporates various strategies, including computational etymology and information retrieval (IR) ([2]). Most of the proposed methods, including supervised and unsupervised methods, tackle the task as binary classification. Supervised learning relies on the learning classifier for each label ([3];[4]). Whereas, unsupervised learning depends on the glossary of sentiments ([5];[6];[7];[8]). However, a fuzzy rule-based system (FRBs) can extract sentiments more accurately through the belonging degree of the fuzzy sets. An FRBS ([9]) .employs a fuzzy degree to analyze the feasibility of a fuzzy rule-based classifier to obtain emotion scores. The output is a binary polarity classification (e.g., positive or negative).

Conversely, an unsupervised sentiment analysis method([10].) is in accordance with the fuzzy rules for the Twitter domain. The critical factor involves employing some fuzzy rules to obtain the sentiment of each tweet. That sentiment is determined by the level cross of two input variables (i.e., positive [TweetPos] and negative [TweetNeg] tweets). Unfortunately, this model is unsuitable for a three-class categorization. Meanwhile, a fuzzy logic classifier is used as a sentiment classifier based on emotions ([11]) .and combines deep convolutional neural networks and fuzzy logic.

This paper presents a newly proposed method to analyze the sentiment. The works mentioned above fall under the concept of fuzzy system rules to some extent. That circumstance shows that they help analyze feelings. However, fuzzy logic still has limitations in adapting to exact rules, especially when many variables are connected. Therefore, we added the concept of CSA into FRBS to solve the limitation of this case. Our work consists of two main parts: an application of the FRBS and the crow search algorithm. A novel sentiment classification method that focuses on document-level sentiment analysis tasks is presented in this work. We address these tasks as a three-class classification problem according to fuzzy logic rules. We incorporate the CSA into the concept of a fuzzy logic rules system to optimize the output generated from the fuzzy logic rules. CSA collects all the outcomes of the rules of the fuzzy system and then searches for the best outcomes archived by the system according to specific metrics, namely, precision and recall. Regarding the results, we discuss and verify the accuracy of the performance and outcomes by investigating the validity of the performances of three reference data sentiment analysis data sets, including those from Amazon, Yelp, and IMDB. The experimental results reveal the effectiveness of the proposed model and the possibility of improving performance regarding sentiment analysis in the case of three class predictions. The main contribution of our work can be summarized as follows:

1. The proposed method can combine positive and negative fuzziness and deal with the mystery and blur.
2. This work integrates CSA with a fuzzy logical rules system to optimize the output results of the FRBS.
3. Our proposed model obtains high classification accuracy relative to existing approaches for all cases of experimentation and with efficiency as a critical factor.
4. The crow search algorithm achieves the best result given any ambiguity or overlap between the membership function degrees when applying the fuzzy rules.

2. Related work

This section provides a literary survey of works in this domain using various approaches, such as machine learning techniques and artificially intelligent and soft computing algorithms. One of these techniques is a fuzzy logic system. That system can process doubt or ambiguity in a remarkably productive way because of the presence of interference. The writers of ([12]) .applied the concept of fuzzy rules to reduce their four fuzzy rules to assess the sentiment of reviews. The highlights of their work are as follows. (i) They develop

four fuzzy rules based on a text and speech tick to assess the sentiment of each review. (ii) The suggested decision-level fusion approach performs best, outperforming unimodal and primary feature-level text-speech fusion using a supervised machine learning classifier, namely, support vector machines. (iii) Their work compared the idea of a rule-based system for sentiment analysis with eight modern supervised machine learning techniques. Zohreh Madhoushi et al. explored sentiment analysis techniques and proposed machine-based learning approaches, including supervised and semi-supervised learning, as well as more commonly employed techniques in this field, such as lexical and hybrid techniques. They concluded that open problems such as current techniques could not handle complex sentences, and the research lacked an analysis of feelings in languages other than English ([13]).

In addition, different technologies were surveyed in this field, such as computational intelligence, which involves the genetic algorithm, fuzzy sets, rough sets, neural networks, and deep learning ([14]). Medhat et al. discussed formal concepts analysis (FCA) and confused formal concept analysis (FFCA). FCA is a mathematical approach that uses slightly arranged relationships. Conversely, FFCA deals with unconfirmed data [4]. Kumar Ravi et al. introduced a hybrid model consisting of fuzzy formal concept analysis (FFCA) and concept-level sentiment analysis to analyze and summarize complaints about customer relationship management (CRM). They also compared their approach with FCA and sentiment-level emotional analysis ([15]). Zadeh et al. have tried to build an ambiguous, emotional grammar analysis system for multi-speech text using decision-level integration. Multimedia sentiment analysis was performed on a CMU MOSI dataset([16]). In general, three types of fuzzy reasoning models are available: Sugeno, Mamdani, and Tsukamoto. The rule-based fuzzy Tsukamoto system ([4]).([15]) .was used to analyze feelings. These works employed a trapezoidal blurring organic function to convert numerical values into obscure linguistic terms. The fuzzy logic approach can conduct real-time analysis of tweets to map tweet effects over time([9]). The feelings in these tweets describe the dynamic mood swings of a cricket fan while controlling a cricket match. Mamdani's vague, rule-based system categorizes uncensored feelings for Twitter datasets is suitable for any glossary and two and three-grade emotional analysis tasks ([17]). This work introduces a genius set of mysterious rules for developing a multimedia emotion classification system using the mysterious Sugeno model. A fuzzy logic system (FLS) has been widely used for sentiment polarity identification in the last few years.

A study([18]) .introduced a LISA model, an unsupervised word-level knowledge graph-based LSA framework. The LISA model utilized various variants of shortest-path graph navigation techniques to calculate and propagate effective scores in a lexical-affective graph (LAG), constructed by connecting a typical lexical KB like WordNet, with a reliable effect KB like WordNet-Affect Hierarchy. A new algorithmic framework for autonomous Music Sentiment-based Expression and Composition, called MUSEC, was proposed by Ralph Abboud and Joe Tekli([19]). This model perceives an extensible corpus of six primary human emotions (e.g., fear, anger, love, joy, surprise, and sadness) expressed by a MIDI musical file, and then creates a novel polyphonic, (pseudo) thematic, and diversified musical pieces that express. these emotions.

3. Proposed model

This section describes in detail the proposed approach of applying Fuzzy Logic System and Crow Search Algorithm to analyze the sentiment, wherein the first subsection explains the applying of the FLS, and the pros and cons of this approach, the second subsection describe

es CSA for annealing the gap found in the FLS. This section presented a sentiment analysis framework by discussing the different tasks engaged in sentiment analysis.

The main idea of this proposed method is to apply the FLS and crow search algorithm to process the posts to overcome the above challenges. In this paper, the proposed approach builds on the FLS to process the computational complexity as well as the interpretation. At the same time, the suggested technique can preserve the classification performance in line with the performance of the best algorithms for classifying sentiments. The fuzzy rules can be represented in 1 or 0 forms. This form is very close to natural languages and can be easily interpreted. The task of FLS in this work is to classify the polarity of the sentiment into three classes: positive, natural, and negative. The CSA will solve the overlapping of fuzzy sets, especially when several variables that belong to the same category are involved. Moreover, the CSA will achieve the sentiment polarity to enhance the FLS, particularly when the result produced by FLS is unclear.

The Fuzzy Rules-Based System with Crow Search Algorithm provides an effective method of sentiment analysis based on an FRBS and a CSA to analyze the personality sentiment during the typing of text. The sentiment analysis process is shown in Fig. 1. The text preparation step performs required text pre-processing and Characteristic Selection and Extraction of the dataset, including the removal of stop words. The sentiment identification step determines the sentiment of people expressed in the text and analyzes it based on selecting the features of writing and building the fuzzy rules. The sentiment classification step is conducted to identify the polarity of sentiment to get the results. It is finally implementing the validation and evaluation of produced results to display the effectiveness and performance of the proposed work.

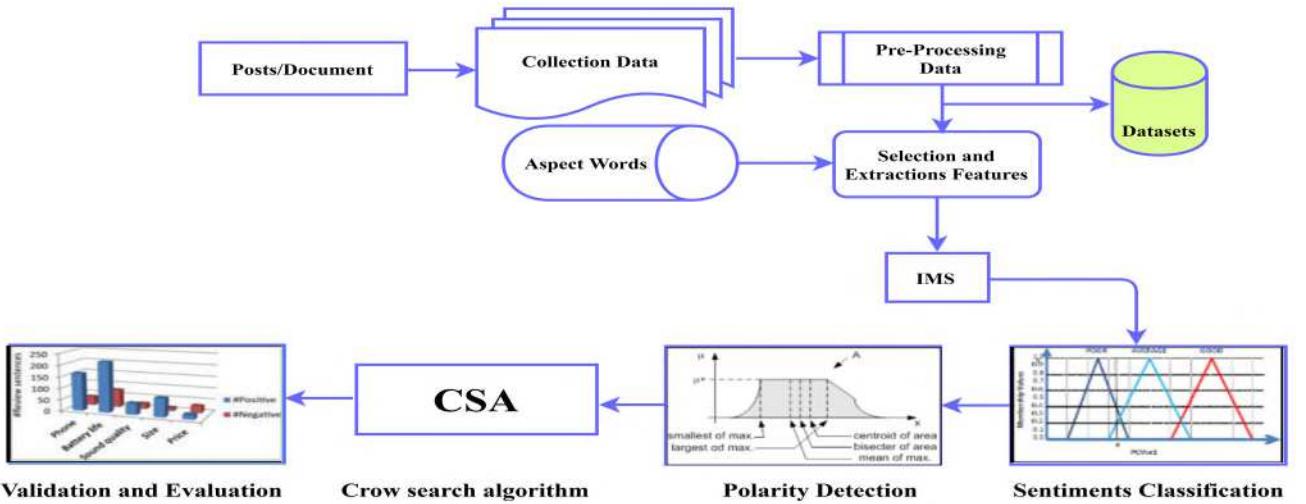


Fig. 1: The structure of the proposed approach.

3.1 Fuzzy Logic System

The FLSs are the most useful approaches in modelling some complex systems that humans can observe because they use linguistic variables as their antecedents and consequent. These linguistic variables can naturally be represented by these sets' fuzzy sets and logical connectives. I employed three standard methods of fuzzy systems based on linguistic rules in practice: the Mamdani systems, Surgeon models, and Tsukamoto models ([20]). In this study, we adopt the Mamdani systems, which entail four steps: fuzzification of input variables, rule evaluation (inference), aggregation of the rule outputs, and

defuzzification. Three different dictionaries have been proposed and investigated for training data: SentiWordNet, AFINN, and LabMT([21] ; [22] ; [23]). SentiWordNet is divided into two parts: Senti Positive and Senti Negative. Thus, four training data sentiments are generated, and we will classify text polarity according to these lexicons. The FLS deals with these lexicons as input data, and each one is classified into the classes listed below.

3.1.1 Fuzzification

This step is included in the sentiment classification shown in Fig. 1. Using the triangular membership function, we fuzzed all data obtained from the second phase's positive, negative, and average scores. All linguistic term T involves three key points, d, e, f, when the triangular fuzzy membership is used and associated with the change of style of the fuzzy membership. $\mu_S:X \rightarrow [0,1]$ defines the membership function (MF) for a fuzzy set S on the universe of discourse X, and each element of X is mapped between the 0 and 1 value. The triangular function equation defined by an upper limit f, an intermediate value e, and a lower limit d, where $d < e < f$, as equation1([12]):

$$\mu_S(x) = \begin{cases} 0, x \leq d \\ (x - d)/(e - d), d < x \leq e \\ (f - x)/(f - e), e < x \leq f \\ 0, x \geq f \end{cases} \quad (1)$$

Fig. 2 illustrates the triangular fuzzy membership, where the fuzzy sets are created using the triangular fuzzy membership for universe variables. This work uses three datasets (AFINN, SentiWordNet, and labMT) as inputs of the fuzzy system. Where SentiWordNet dataset is classified into two classes Senti Positive and Senti Negative. Fig. 2: shows the fuzzy system of these datasets.

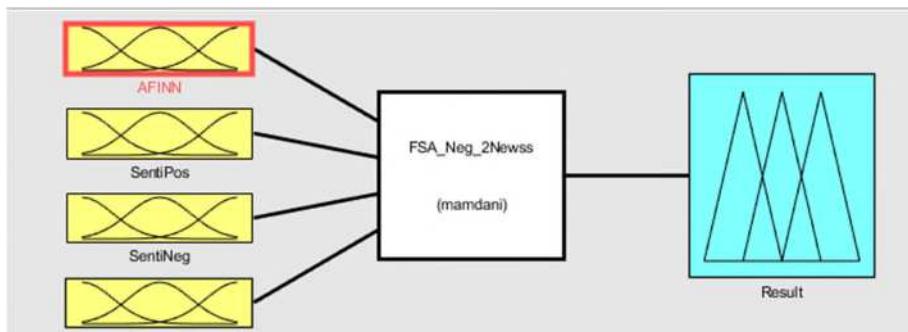


Fig. 2: Fuzzy System of the triangular fuzzy membership

Each one of these datasets has special fuzzy sets; for example, the dataset **AFINN** in this method has the following five fuzzy membership functions in the range[−5, 5]:

$$HighNegative(H - N) = -5, \quad LowNegative(L - N) = -2.5; \\ Natural(N - N) = 0;$$

$$Low Positive (L_P) = 2.5, \text{ and } High Positive (H_P) = 5.$$

The output of this stage is the polarity of sentiment analysis: Positive, Negative or Natural.

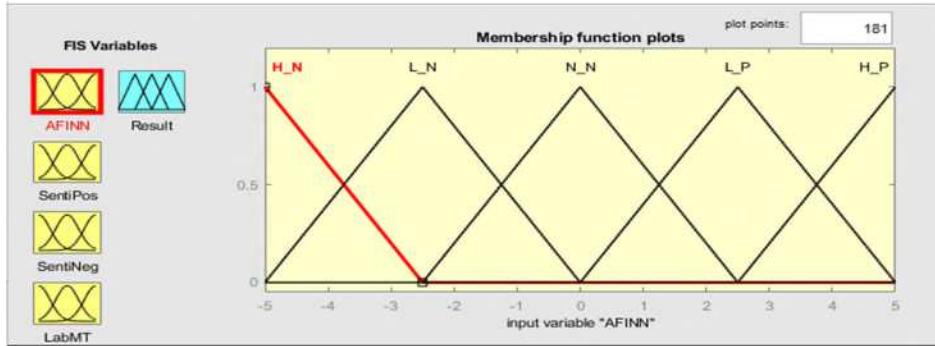


Fig. 3: Membership Function Input of AFINN dataset

The dataset LabMT also has five membership functions in range [0 10] (As showing in Fig. 3) that can be explained as follows: *High Negative (H_N) = 1, Low Negative (L_N) = 3.25;*

$$Natural(N_N) = 5.5;$$

$$Low Positive (L_P) = 7.75, \text{ and High Positive (H_P) } = 10;.$$

The output of this stage is the polarity of sentiment analysis: Positive, Negative or Natural

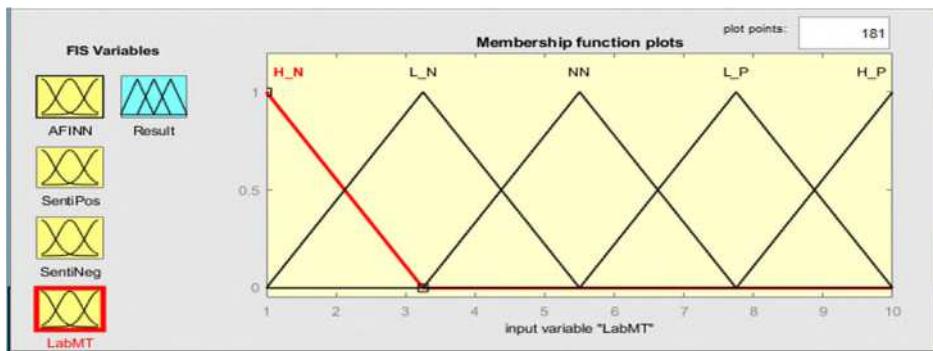


Fig. 4: Membership function of Input of LabMT dataset

According to both (labMT and AFINN) datasets, each Senti Negative and Senti Positive has an individual three membership functions in range [0 1] as *Low = 0, Medium = 0.5 and High = 1*, as shown Fig. 5 and Fig. 6.

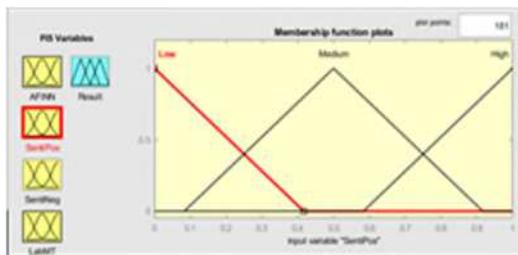


Fig. 5: Membership function of Input of SentiNeg

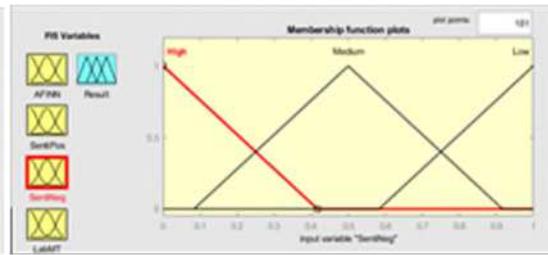


Fig. 6: Membership function of Input of SentiPos

Therefore, the output was classified into three-membership functions, as mentioned in Fig. 7. We calculate the global minimum (min), global medium (nan), global maximum (max) values for

all positive scores, all-natural scores, and all negative scores of all posts in the datasets. The positive, natural, and negative domains are (min, nan, max). Where the mid-value can be calculated in equation (2) as follows ([12]):

$$\text{Mid} = (\text{min} + \text{nan} + \text{max}) / 3 \quad (2)$$

The required variables for building the triangular fuzzy membership of the fuzzy sets are: Low, Medium, and High, as follows.

Low: {min, min, mid}; Medium: {min, mid, max}; High: {mid, max, max}.

Accordingly, the output variable is ranged between [-1, 1], and the parameters for three fuzzy sets (Negative, Neutral, and Positive) that depict the sentiment classes are set as: Negative: {-0.4167 0 0.4167}; Neutral: {0.0833 0.5 0.9167}; Positive: {0.5833 1 1.4167}; The MFs of consequent parts of the proposed rules is graphically presented in Fig. 7.

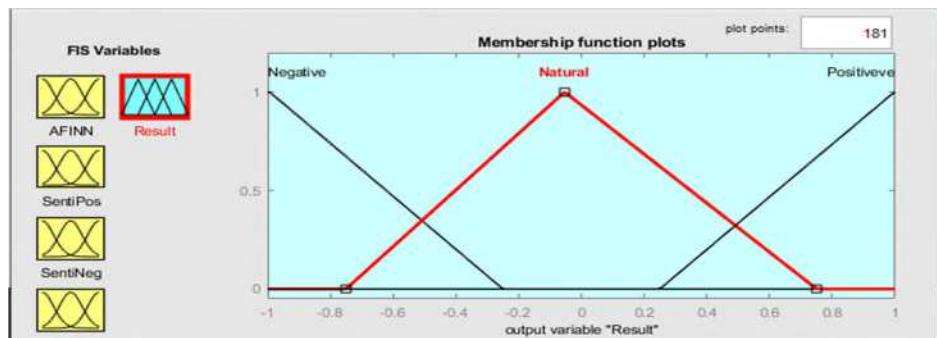


Fig. 7: Membership Function of Output

3.1.2 Formulating the rule-base.

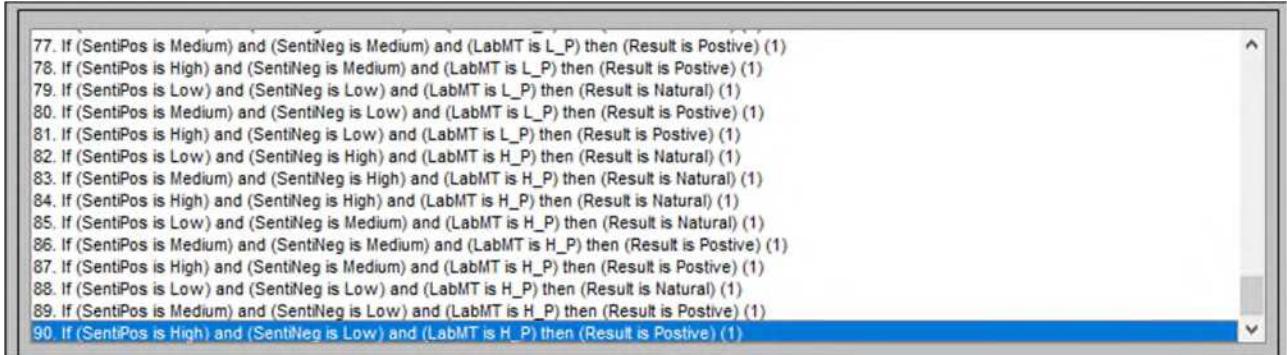


Fig. 8: Part of Proposed Rules

This step is included in the sentiment classification shown in figure 1. We have four inputs as proposed above, and the first input is AFINN lexicons, which have five membership functions in the domain [-5, 5]. The second input is Senti Positive lexicons, which have three membership functions in the domain [1,0]. The third input is Senti Negative lexicons, which also have three membership functions in the domain [1,0]. The fourth input is LabMT Lexicons, which has five membership functions in the range [1,10]. These inputs are implemented through two stages using the logical operation (called min operation and FRBS.

The first stage is implementing the AFINN, Senti Positive, and Senti Negative to fuzzy rules. The second stage implements the Senti Positive, Senti Negative, and LabMT to fuzzy rules. Therefore, the total number of rules becomes ninety rules. These rules are covering all the probabilities and calculated as $5 * 3 * 3 + 3 * 3 * 5 = 90$. Fig. 8 shows the visualization of a part of ninety rules.

3.1.3 Defuzzification

This step is included in the polarity detection shown in Fig.1. Defuzzification is the process of identifying the aggregate polarity of the sentiment of a complete dataset. This process returns the overall sentiment for the selected data set. The Mamdani inference system is employed to compute the center of gravity using the following statements:

$$\begin{aligned} \text{if } (\text{Negative}) > f(\text{Positive}) &\Rightarrow \text{class} = \text{Negative}; \\ \text{if } (\text{Negative}) < f(\text{Positive}) &\Rightarrow \text{class} = \text{Positive}; \text{ else class is neutral.} \end{aligned}$$

This stage entails ascertaining the linguistic term the fuzzy values changed into the so-called crisp values.

3.2 Crow Search Algorithm

As mentioned above, that fall under the concept of fuzzy system rules to some extent, show that they are effective in feelings analysis. Yet, this technique still has limitations in adapting to the exact rules, especially when many variables are connected. Therefore, we focus on document-level sentiment analysis tasks. We addressed this task as a three-class classification problem based on fuzzy logic rules. We integrate CSA to the concept of the FLS to optimize the output generated by FLS. CSA collects all outcomes of the rules of the fuzzy system then looks for the best result archived by the system based on metrics, namely precision and recall.

The CSA is a relatively novel meta-heuristic improvement algorithm based on the intelligent behaviour of crows([24]). CSA is a population-based optimization algorithm with fundamentally adjustable flight length and potential awareness parameters. These properties make the CSA a viable option for complex engineering optimization problems and mathematically complex optimization problems such as solving the feature selection problem, improving text documents classification, estimating the threshold values for magnetic resonance brain images, and solving the DNA fragment assembly problem. The general mechanism of CSA implementation is mentioned as follows.

Firstly, the position of the hiding place of each crow is created randomly, and the memory of each crow is initialized with this position as the best experience. Secondly, the crow evaluates the quality of its position according to the objective function. Finally, the crow randomly selects one of the flock crows and follows it to discover the position of the foods hidden by this crow. If the found position of the food is tasty, the crow updates its position. Otherwise, the crow stays in the current position and does not move to the generated position.

According to the previous hypothesis, in this study, CSA starts by initializing the variables. In this step, the flock population (N), maximum number of iteration (intermix), awareness probability (AP), flight length (fl), and number of decision variable (d) is defined. Second: represents the memory of crows and locations; in this step, the N number of crows are randomly placed in d -dimensional search space. Each position of crows defines a possible solution to the algorithm. Initially, as the crows have no memory, the starting position is the best memory in the problem. That is at the first iteration. The crow has the initial position as a memory location. Third: calculation of objective function. The quality of each crow's position is determined by dropping the variables into the objective function. Fourth: generation of fresh location Crows produce fresh location in the environment as follows: assume crow "I need to produce a fresh location". To do this propose, this crow arbitrarily chooses one of the flock crows (for example, crow j)

and chases it to determine the location of the environment best position by this crow (m_j). The fresh location of crow i found by the cases below.

$$x^{i,iter+1} = \begin{cases} x^{i,iter} + r_i \times f^{i,iter} \times (m^{j,iter} - x^{j,iter}), \\ \text{if } r_j \geq AP^{j,iter} \\ \text{arbitrary location, otherwise} \end{cases} \quad (3)$$

Where, r_i is any random number uniformly distributed between 0 to 1. $m^{j,iter}$ is a best position of crow j at iteration $iter$. Fifth: Calculate the objective function of new position. The objective function value for the new position of each crow is figured out.

Sixth: Update memory up gradation of memories of crows are computed as follows.

$$m^{i,iter} = \begin{cases} x^{i,iter+1} \text{ if } f(x^{i,iter+1}) \leq m^{i,iter} \\ m^{i,iter}, \quad \text{otherwise} \end{cases} \quad (4)$$

Where $f(\cdot)$ depicts the fitness function value, if the fitness function value of memory is poorer to the fitness function value of the new position, the crows update their memory with present position.

In the last: verification of execution of criterion. These steps are reiterated until intermix is achieved. When the implementation task is completed, the best value of the target function is reported as the general solution and the memory information is reported as possible global points of the optimization problem.

Algorithm 1 shows the pseudo-code of the proposed model.

Begin Procedure:

Algorithm 1: Integrating CSA with fuzzy-rules based system

Input: Dataset

Output: Class labels for test dataset load

1: Set the initial values of M, AP, N, f1, and t Max.

Where the M is the Problem dimension (number of decision variables), AP is awareness probability, N is the number of the population size, f1 is the flight length and t max is the maximum number of iterations.

2: Initialize the crow position y randomly.

3: Evaluate the fitness function of each crow F n(y).

4: Initialize the memory of search crow N

5: Set t: = 1. {Counter initialization}.

6: repeat

7: for (j = 1: j < M) do

8: Randomly choose one of crows to follow z

9: if A then

10: = +

11: else

12: = A random position of the search

space

13: end if

14: end for

15: Check the feasibility of

16: Evaluate the new position of crow Fn ($y^{j,t+1}$).

17: Update the crow's memory

18: Set t = t + 1. {Iteration counter increasing}.

19: until (t < t Max). {Termination counter satisfied}.

20: Produce the best solution N. using the fitness function

4. Performance evaluation of the proposed model

4.1 Evaluation datasets.

This subsection describes the data sets and the algorithms used to evaluate the proposed approach. We conduct our experiments on three extensive and well-known datasets. As explained below, these data sets were collected from different fields with varied contexts and vocabularies.

- Amazon: consists of more than 1000 product reviews labeled as 'positive' or 'negative'. This dataset has 1865 distinct word tokens.
- IMDB: IMDB dataset has been used for sentiment classification of film reviews. IMDB includes 50,000 reviews and consists of binary reviews with positive and negative sentiments. It was evenly split into 25,000 reviews for training and 25,000 for testing.
- Yelp: contains more than 1000 restaurant reviews labeled with 'positive' or 'negative'. In this data set, a total of 2049 distinct word tokens are found.

4.2 Data Pre-processing:

Data preprocessing: Data preprocessing is considered as an essential step in machine learning and data mining([25] ; [26] ; [27] ; [28]). The reviews usually contain incomplete sentences, much noise, and weak wording such as words without application with high repetition, imperfect words, and incorrect grammar. Unstructured data also has an impact on sentiment classification results. A series of preprocessing on the

reviews are needed to maintain a stable structure and reduce such problems. Cleaning data with filters, splitting the data into a part for training and testing, and building data sets with favorite words are just a few of the steps employed in our research. Without going into too much depth, we used the following method to prepare the data:

- **Tokenization**

We divided the text into phrases, words, symbols, or other meaningful elements, thus forming a list per comment of individual words. In each comment, we then use each word as a feature for our training classifier.

- **Removing Stop words**

Comment often contains some stop words that have no meaning, such as prepositions, and words that do not add any emotion value like (or, also, able, etc.). The Natural Language Toolkit (NLTK) library provides a stop words dictionary, including a list of words that have meaning neutral and are not suitable for sentiment analysis. To remove the stop words from the comment's text, we check each word in the list against the dictionary and exclude them.

- **Capitalization**

Due to documents and texts containing many sentences and include a diversity of capitalization for a sentence. So, diverse capitalization can be a big problem when classifying extensive documents. The best approach to dealing with inconsistent capitalization is to convert each letter to its lower case. This technique shows all words in the same feature distance to the text and document. Still, it reasons a significant issue in the interpretation of some words (e.g., "US" (United States of America) to "us" (pronoun)).

4.3 Evaluation metrics

The following metrics were used to evaluate the performance of the proposed sentiment analysis framework.

- **Accuracy:** defines the ratio of the true results (both true positives and true negatives) to the total number of documents analyzed by the classifier.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (5)$$

Where TN, FN, FP, and TP represent the true negatives, false negatives, false, and positives true positives, respectively.

- **Precision:** defines the recommended items that are relevant to correct predictions, divided by the total of predicted Recommended items during the testing process

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

- **Recall:** indicates the ratio of a number of documents whose sentiments are correctly classified as positive to the actual number of documents having positive sentiments in the given text corpus.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

- **F1 Score:** This metric represents the harmonic mean of both recall and precision.

$$F_1Score = 2 * \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

4.4 Experimental settings and baselines

Experiments are performed on a PC with Intel(R) Core (TM) i7-6400 CPU at 2.71GHZ with 8 GB of RAM and running on Windows 10. This method was implemented using three data sets collected from Amazon, Yelp, and IMDB and implemented by the applied fuzzy logic R as a classifier for sentiment analysis and with the CSA as optimizer to achieve the best result in finding out by fuzzy rules-based systems. The implementation was an environment of in (MATLAB programming version 2019a).

4.5 Experimental results and discussion

After collecting the datasets, the FLS applying to determine the polarity of the text. After that, the CSA was used to achieve the best polarity. Demonstrating the performing evaluation and experiment results that are based on different machine-learning classifiers. The classifiers were applied on each data-set, and the performance is compared based on cross-validation, accuracy, precision, recall, and f-score.

1. Accuracy analysis

The experiment results were obtained using four-fold cross-validation, computing the accuracy for each fold, and then ascertaining the mean outcome. We compared the performance of our proposed method against four different methods. Tables4 presents the accuracies of various algorithms with the review data sets collected from Amazon, Yelp, and IMDB. The outcomes were evaluated with multiple data sets by comparing the performance of different machine learning models and various methods. Table1 indicates that the difference between our proposed method and the latest techniques is approximately 4.25% in terms of accuracy for the Amazon data set. Table2 shows that the difference between our proposed method and the latest techniques is approximately 0.46% in terms of accuracy for the Yelp data set. Finally, Table3 shows the difference between our proposed method and the latest techniques is approximately 3.96% in terms of accuracy for the IMDB dataset. Therefore, according to the Yelp dataset, no significant difference occurred between our method and the maximum entropy method, as shown in Fig.8.

2. Precision, Recall, and F-score Analysis

The precision, recall, and f -score depicted in Table 4 show the impact of the proposed model compared with baseline models on different review datasets. For each proposed algorithm, the performance has been evaluated in each step Fig. 9 shows this superiority.

Algorithms	Classification Accuracy (%)
SVM ([29]).	79.13
Boosting([30]).	90.79
Maximum Entropy ([31]).	94.18
SWESA ([32]).	87.19
Fuzzy Logic and CSA	98.43

Table 1: Classification Accuracy for AMAZON CELLS dataset

Table 2: Classification Accuracy for YELP dataset

Algorithms	Classification Accuracy (%)
SVM ([29]).	76.67
Boosting ([30]).	88.81
Maximum Entropy ([31]).	97.21
SWESA ([32]).	86.93
Fuzzy Logic and CSA	97.67

Table 3: Classification Accuracy for IMDB dataset

Algorithms	Classification Accuracy (%)
SVM ([29]).	77.29
Boosting([30]).	90.27
Maximum Entropy ([31]).	91.02
SWESA ([32]).	81.04
Fuzzy Logic and CSA	94.98

Table4: Precision, Recall, and F -Score for Different Methods with Review Datasets

Methods/Datasets	AMAZON CELLS			YELP			IMDB		
	Recall	Precision	F-Score	Recall	Precision	F-Score	Recall	Precision	F-Score
SVM ([29]).	0.850	0.855	0.850	0.705	0.700	0.700	0.760	0.755	0.755
Boosting ([30]).	0.840	0.840	0.840	0.680	0.690	0.670	0.690	0.730	0.690
Maximum Entropy([31]).	0.785	0.800	0.785	0.760	0.760	0.755	0.735	0.730	0.730
Fuzzy Logic and CSA	0.902	0.847	0.878	0.901	0.820	0.806	0.904	0.773	0.814

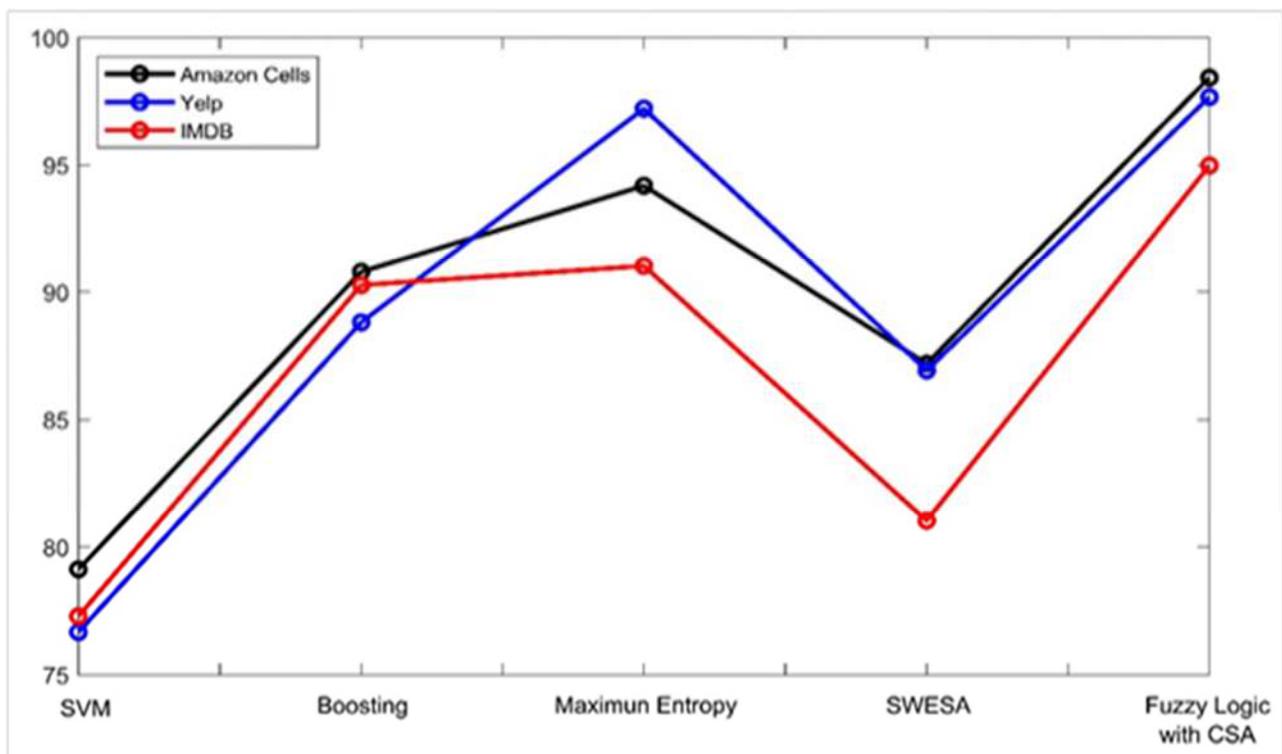


Fig. 8 Accuracy Percentage of Proposed Model and Others

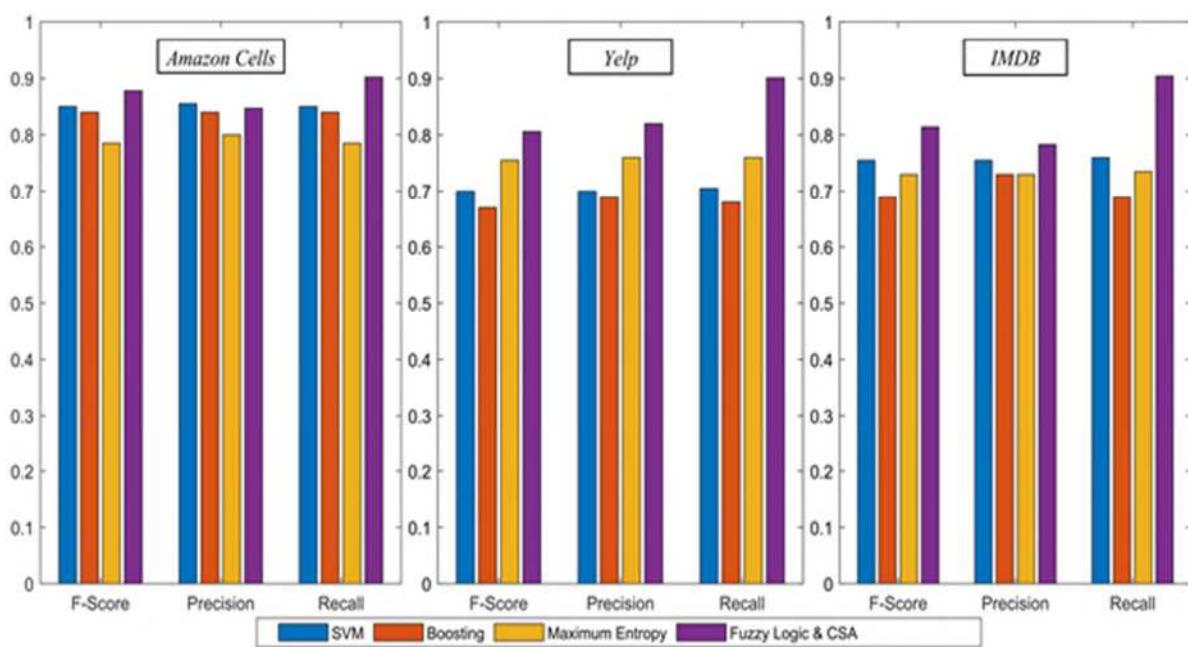


Fig. 9 Accuracy Percentage of Proposed Model and Others

5. Conclusions and future work

Sentiment analysis is essential for anyone who makes a decision. Useful in various areas for measuring feelings, expressing, and identifying emotions. Though the study has achieved interesting results, some changes in our future work will be made to achieve better results and improve the performance. This study provides an empirical analysis of the classification of feelings based on vague rules and crow search algorithms. The results obtained from the classification process showed that our sentiment classification model, which is based on fuzzy vague rules and group search algorithms, achieved great results compared to models based on other various classification techniques such as SVM, Maximum Entropy and Boosting applied to three famous data sets collected from Amazon, Yelp, and IMDB. Consequently, these reviewed data sets are applied to evaluate FLS with CSA models and rank the polarity of sentiments. For each different dataset, the performance, accuracy, and factors (recall, accuracy, and f score) were used to evaluate the results. Finally, the results are analyzed, compared, and presented with a statistical approach.

Statements & Declarations

Funding:

This work was partially supported by the National Natural Science Foundation of China (Z201G10110G20003).

Competing Interests

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property. We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). He is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct.

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by MazenSharaf AL-Deen, Yu lasheng , Gamil R. S. Qaid and Ali Aldhubri. The first draft of the manuscript was written by MazenSharaf AL-Deen and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data Availability

The datasets analyzed during the current study are available in the :

Amazon and Yelp: <https://github.com/ss12345656/FuzzySentiment/tree/master/Data/Test>.

IMDB/ <https://www.kaggle.com/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>

References

- [1] J. Serrano-Guerrero, J. A. Olivas, F. P. Romero, and E. Herrera-Viedma, “Sentiment analysis: A review and comparative analysis of web services,” *Inf. Sci. (Ny.)*, vol. 311, pp. 18–38, 2015, doi: 10.1016/j.ins.2015.03.040.
- [2] B. Al Shboul, M. Al-Ayyoub, and Y. Jararwehy, “Multi-way sentiment classification of Arabic reviews,” *2015 6th Int. Conf. Inf. Commun. Syst. ICICS 2015*, pp. 206–211, 2015, doi: 10.1109/ICACS.2015.7103228.
- [3] H. Saif, Y. He, M. Fernandez, and H. Alani, “Contextual semantics for sentiment analysis of Twitter,” *Inf. Process. Manag.*, vol. 52, no. 1, pp. 5–19, 2016, doi: 10.1016/j.ipm.2015.01.005.
- [4] W. Medhat, A. Hassan, and H. Korashy, “Sentiment analysis algorithms and applications: A survey,” *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, 2014, doi: 10.1016/j.asej.2014.04.011.
- [5] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, and B. Qin, “Learning Sentiment-Specific Word Embedding,” *Acl*, pp. 1555–1565, 2014.
- [6] K. Ravi and V. Ravi, *A survey on opinion mining and sentiment analysis: Tasks, approaches and applications*, vol. 89, no. November. 2015.
- [7] S. Poria, E. Cambria, G. Winterstein, and G. Bin Huang, “Sentic patterns: Dependency-based rules for concept-level sentiment analysis,” *Knowledge-Based Syst.*, vol. 69, no. 1, pp. 45–63, 2014, doi: 10.1016/j.knosys.2014.05.005.
- [8] Z. Nasim, Q. Rajput, and S. Haider, “Sentiment analysis of student feedback using machine learning and lexicon based approaches,” *Int. Conf. Res. Innov. Inf. Syst. ICRIIS*, no. March 2021, 2017, doi: 10.1109/ICRIIS.2017.8002475.
- [9] C. Jefferson, H. Liu, and M. Cocea, “Fuzzy approach for sentiment analysis,” *IEEE Int. Conf. Fuzzy Syst.*, 2017, doi: 10.1109/FUZZ-IEEE.2017.8015577.
- [10] S. Vashishtha and S. Susan, “Fuzzy rule based unsupervised sentiment analysis from social media posts,” *Expert Syst. Appl.*, vol. 138, 2019, doi: 10.1016/j.eswa.2019.112834.
- [11] V. López, S. Del Río, J. M. Benítez, and F. Herrera, “Cost-sensitive linguistic fuzzy rule based classification systems under the MapReduce framework for imbalanced big data,” *Fuzzy Sets Syst.*, vol. 258, pp. 5–38, 2015, doi: 10.1016/j.fss.2014.01.015.
- [12] S. Vashishtha and S. Susan, “Inferring Sentiments from Supervised Classification of Text and Speech cues using Fuzzy Rules,” *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 1370–1379, 2020, doi: 10.1016/j.procs.2020.03.348.
- [13] Z. Madhoushi, A. R. Hamdan, and S. Zainudin, “Sentiment analysis techniques in recent works,” *Proc. 2015 Sci. Inf. Conf. SAI 2015*, no. March, pp. 288–291, 2015, doi: 10.1109/SAI.2015.7237157.
- [14] R. Katarya and A. Yadav, “A comparative study of genetic algorithm in sentiment analysis,” *Proc. 2nd Int. Conf. Inven. Syst. Control. ICISC 2018*, no. Icisc, pp. 136–141, 2018, doi: 10.1109/ICISC.2018.8399051.
- [15] K. Ravi, V. Ravi, and P. S. R. K. Prasad, “Fuzzy formal concept analysis based opinion mining for CRM in financial services,” *Appl. Soft Comput. J.*, vol. 60, pp. 786–807, 2017, doi: 10.1016/j.asoc.2017.05.028.
- [16] A. Zadeh, R. Zellers, E. Pincus, and L.-P. Morency, “MOSI: Multimodal Corpus of Sentiment Intensity and Subjectivity Analysis in Online Opinion Videos,” 2016, [Online]. Available: <http://arxiv.org/abs/1606.06259>.
- [17] S. Vashishtha and S. Susan, *Fuzzy logic based dynamic plotting of mood swings from tweets*, vol. 939, no. May 2020. Springer International Publishing, 2019.
- [18] M. Fares, A. Moufarrej, E. Jreij, J. Tekli, and W. Grosky, “Unsupervised word-level affect analysis and propagation in a lexical knowledge graph,” *Knowledge-Based Syst.*, vol. 165, pp. 432–459, 2019, doi: 10.1016/j.knosys.2018.12.017.
- [19] R. Abboud and J. Tekli, “Integration of nonparametric fuzzy classification with an evolutionary-developmental framework to perform music sentiment-based analysis and composition,” *Soft Comput.*, vol. 24, no. 13, pp. 9875–9925, 2020, doi: 10.1007/s00500-019-04503-4.
- [20] T. J. ROSS, *Fuzzy Logic With Engineering Application*. 2010.
- [21] S. Baccianella, A. Esuli, and F. Sebastiani, “SENTIWORDNET 3.0: An enhanced lexical

- resource for sentiment analysis and opinion mining,” *Proc. 7th Int. Conf. Lang. Resour. Eval. Lr. 2010*, vol. 0, pp. 2200–2204, 2010.
- [22] F. Nielsen, “Afinn. Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby.” 2011.
- [23] P. S. Dodds *et al.*, “Human language reveals a universal positivity bias,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 8, pp. 2389–2394, 2015, doi: 10.1073/pnas.1411678112.
- [24] A. Askarzadeh, “A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm,” *Comput. Struct.*, vol. 169, pp. 1–12, 2016, doi: 10.1016/j.compstruc.2016.03.001.
- [25] G. Gupta and S. Malhotra, “Text Document Tokenization for Word Frequency Count using Rapid Miner (Taking Resume as an Example),” *Int. Conf. Adv. Eng. Technol.*, no. Icaet, pp. 24–26, 2015.
- [26] T. Verma, R. Renu, and D. Gaur, “Tokenization and Filtering Process in RapidMiner,” *Int. J. Appl. Inf. Syst.*, vol. 7, no. 2, pp. 16–18, 2014, doi: 10.5120/ijais14-451139.
- [27] T. Ma, R. Al-Sabri, L. Zhang, B. Marah, and N. Al-Nabhan, “The Impact of Weighting Schemes and Stemming Process on Topic Modeling of Arabic Long and Short Texts,” *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 19, no. 6, 2020, doi: 10.1145/3405843.
- [28] H. S. S. Al-Deen, Z. Zeng, R. Al-Sabri, and A. Hekmat, “An improved model for analyzing textual sentiment based on a deep neural network using multi-head attention mechanism,” *Appl. Syst. Innov.*, vol. 4, no. 4, 2021, doi: 10.3390/asi4040085.
- [29] S. Ghosh, A. Hazra, and A. Raj, “A Comparative Study of Different Classification Techniques for Sentiment Analysis,” *Int. J. Synth. Emot.*, vol. 11, no. 1, pp. 49–57, 2020, doi: 10.4018/ijse.20200101.oa.
- [30] P. Kaur, “Design and Implementation of Boosting Classification Algorithm for Sentiment Analysis on Newspaper Articles,” vol. 7, no. 4, pp. 1767–1770, 2016.
- [31] M. Kabir, M. M. J. Kabir, S. Xu, and B. Badhon, “An empirical research on sentiment analysis using machine learning approaches,” *Int. J. Comput. Appl.*, vol. 43, no. 10, pp. 1011–1019, 2021, doi: 10.1080/1206212X.2019.1643584.
- [32] P. K. Sarma and W. A. Sethares, “Simple algorithms for sentiment analysis on sentiment rich, data poor domains,” *COLING 2018 - 27th Int. Conf. Comput. Linguist. Proc.*, pp. 3422–3435, 2018.