Application of Soft Computing

Study on sentiment classification strategies based on the fuzzy logic with crow search algorithm

Mazen Sharaf AL-Deen1 · Lasheng Yu1 · Ali Aldhubri2 · Gamil R. S. Qaid2

Accepted: 16 May 2022 / Published online: 11 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
In recent times, sentiment analysis research has gained wide popularity. That situation causes the importance 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 the 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. This study compares the performance of the proposed model with different machine learning algorithms, such as SVM, maximum entropy, boosting, and SWESA. It tests the model on three famous data sets collected from Amazon, Yelp, and IMDB. Experimental results demonstrate the effectiveness of the proposed model and achieve 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 (Serrano-Guerrero et al. 2015).

As a generic rule, sentiment analysis attempts to determine the author’s behavior on articles, speakers, or other topics through intense emotional or emotional responses to archive, appropriately and communicatively. The study shows the response of people on social media. This response is an evaluation of emotion. In other words, the sentimental state of the creator and speaker is an expectation of enthusiastic responses (that is, the effect is intended by the originator or buyer). Numerous customers surveys or recommendations on all topics are available on the Web. The periods and audits may contain surveys on items such as customers or fault-finding of films. Surveys expand rapidly, given that individuals like to provide their views on the Web. A lot of surveys are accessible for

✉ Lasheng Yu
yulasheng@csu.edu.cn
1 School of Computer Science and Engineering, Central South University, Changsha 410083, China
2 Department of Computer Engineering, Faculty of Computer Science and Engineering, Hodeida University, Hodeidah, Yemen
solitary items, which makes it problematic for customers as they must peruse each to make a choice. Next, mining this information, special customer assessments 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 (Serrano-Guerrero et al. 2015; Al Shboul et al. 2015). This process incorporates various strategies, including computational etymology and information retrieval (IR) (Al Shboul et al. 2015). 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 (Saif et al. 2016; Medhat et al. 2014). Whereas unsupervised learning depends on the glossary of sentiments (Tang et al. 2014; Ravi and Ravi 2015; Poria et al. 2014; Nasim et al. 2021).

A fuzzy rule-based system (FRBS) can extract sentimentally and more accurately through the belonging degree of the fuzzy sets. The FRBS (Jefferson et al. 2017) employs the fuzzy degree to analyze the feasibility of a fuzzy rule-based classifier and to obtain the emotion scores. The output is a binary polarity classification (e.g., positive or negative).

Conversely, an unsupervised sentiment analysis method (Vashishtha and Susan 2019) 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 (López et al. 2015) 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: 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 Vashishtha and Susan (2020a) apply 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 the 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. investigated sentiment analysis techniques and offered machine-based learning approaches, such as supervised and semi-supervised learning. They used more widely techniques in this field, such as lexical and hybrid techniques. They concluded that there were still some issues, such as the current methods, that could not manage complicated phrases. Thus, the research did not include an
investigation of feelings in languages other than English (Madhoushi et al. 2015).

Furthermore, other technologies in this subject were investigated, including computational intelligence which includes the evolutionary algorithm, fuzzy sets, rough sets, neural networks, and deep learning (Katarya and Yadav 2018). Formal concept analysis (FCA) was discussed by Medhat et al., and fuzzy formal concept analysis was confused (FFCA). FCA is a mathematical technique that employs relationships that are slightly arranged. FFCA, on the other hand, works on unverified information (Medhat et al. 2014). To assess and summarize the complaints of customers, Kumar Ravi et al. developed a hybrid model that combines fuzzy formal concept analysis (FFCA) and concept-level sentiment analysis (CLSA). They also compared it to FCA and sentiment-level emotional analysis (Ravi et al. 2017). Using decision-level integration, Zadeh et al. attempted to construct an ambiguous, emotive grammar analysis system for multi-speech text. An examination of multimedia sentiment was conducted on a CMU MOSI dataset (Zadeh et al. 2016). In general, three types of fuzzy reasoning models are available: Sugeno, Mamdani, and Tsukamoto. The rule-based fuzzy Tsukamoto system (Medhat et al. 2014; Ravi et al. 2017) 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 (Jefferson et al. 2017). 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 and is suitable for any glossary and two and three-grade sentiment analysis tasks (Vashishtha and Susan 2020b). 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 Fares et al. (2019) 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 Abboud and Tekli (2020). This model perceives an extensible corpus of six primary human emotions (e.g., fear, anger, love, joy, surprise, and sadness). It is expressed by a MIDI musical file, which creates novel polyphonic, (pseudo) thematic, and diversified musical pieces that express these emotions.

3 Proposed model

This section describes the proposed approach of applying the fuzzy logic system and crow search algorithm to analyze the sentiment. Wherein the first subsection explains the application of the fuzzy logic system (FLS) and the pros and cons of this approach, the second subsection describes 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 and its 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 the form of 1 or 0. This form is very close to natural languages and can be interpreted easily. The task of FLS in this work is to classify the polarity of the sentiment into three classes: positive, negative, and natural. 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, especially 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.

3.1 Fuzzy logic system

The FLSs are the most useful approaches in modeling some complex systems that humans can observe because they use linguistic variables as their antecedents and consequent. These linguistic variables can be represented
naturally 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, Sargent models, and Tsukamoto models (Ross 2010). 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 (Baccianella et al. 2010; Nielsen 2011; Dodds et al. 2015). 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 the fuzzy membership is used and associated with the change of style of the fuzzy membership. \( \mu_S: X \to [0,1] \) 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 Eq. 1 (Vashishtha and Susan 2020a):

\[
\mu_S(x) = \begin{cases} 
0, & x \leq d \\
(x - d)/(e - d), & d < x \leq e \\
(f - x)/(f - e), & e < x \leq f \\
0, & x \geq f 
\end{cases}
\]  

(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.

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
- LowNegative (\( L-N \)) = -2.5
- Natural (\( N-N \)) = 0
- Low Positive (\( L_P \)) = 2.5
- High Positive (\( H_P \)) = 5

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

The dataset LabMT also has five membership functions in range \([0,10]\) (As showing in Fig. 3) and 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
- High Positive (\( H_P \)) = 10

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

According to both (labMT and AFINN) datasets (Figs. 3, 4), each Senti Negative and Senti Positive have an individual three membership functions in range \([0,1]\) as
Low = 0, Medium = 0.5 and High = 1, as shown Figs. 5 and 6.

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, negative, and natural domains are (min, nan, max). The mid-value can be calculated in Eq. (2) as follows Vashishtha and Susan (2020a):

\[
\text{Mid} = \frac{(\text{min} + \text{nan} + \text{max})}{3}
\]

(2)

The required variables for building the triangular fuzzy membership of the fuzzy sets are: Low, Medium, and High. It is shown 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 are graphically presented in Fig. 7.

### 3.1.2 Formulating the rule-base

This step is involved in the sentiment classification shown in Fig. 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 (Serrano-Guerrero et al. 2015; Vashishtha and Susan 2019). 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 \(5 \times 3 \times 3 + 3 \times 3 \times 5 = 90\). Figure 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 dataset. The Mamdani inference system is employed to compute the center of gravity using the following statements:

\[
\text{if \ Positive} \Rightarrow \text{class} = \text{Negative} ;
\]
\[
\text{if \ Negative} \Rightarrow \text{class} = \text{Positive} ;
\]
\[
\text{else \ class \ is \ neutral} ;
\]

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

### 3.2 Crow search algorithm

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 into the concept of the FLS to optimize the output generated by FLS. CSA collects all outcomes of the rules of the fuzzy system and 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 behavior of crows (Askarzadeh 2016). 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 initialing 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 \(d\) are 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 produces 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.

\[
x_{i,\text{iter}+1} = \begin{cases} 
    x_{i,\text{iter}} + r_i \times f_{i,\text{iter}} \times (m_{j,\text{iter}} - x_{i,\text{iter}}), & \text{if } r_j \geq A_{\text{iter}} \\
    \text{arbitrary location, otherwise} & 
\end{cases}
\]

where \(r_i\) is any random number uniformly distributed between 0 and 1. \(m_{j,\text{iter}}\) is a best position of crow \(j\) at iteration \(\text{iter}\). 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 upgradation of memories of crows are computed as follows.

\[
m_{\text{iter}}^{t+1} = \begin{cases} 
    x_{\text{iter}+1}^{t+1} & \text{if } f(x_{\text{iter}+1}^{t+1}) \leq m_{\text{iter}}^{t} \\
    m_{\text{iter}}^{t} & \text{otherwise}
\end{cases} \tag{4}
\]

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 \( M \), \( AP \), \( N \), \( f_1 \), 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, \( f_1 \) 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 \( F_n(y_{j+1}) \)
17: Update the crow\&#39;s memory with memory
18: Set \( t = t + 1 \). {Iteration counter increasing}
19: until \( (t \&lt; t \) 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 dataset, a total of 2049 distinct word tokens are found.

4.2 Data preprocessing

**Data preprocessing:** Data preprocessing is considered as an essential step in machine learning and data mining (Gupta and Malhotra 2015; Verma et al. 2018; Ma et al. 2020; Al-Deen et al. 2021). 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 have 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:

| Table 1 Classification accuracy for AMAZON CELLS dataset |
|----------------------------------------------------------|
| Algorithms | Classification accuracy (%) |
| SVM (Ghosh et al. 2020) | 79.13 |
| Boosting (Kaur 2016) | 90.79 |
| Maximum Entropy (Kabir et al. 2021) | 94.18 |
| SWESA (Sarma and Sethares 2018) | 87.19 |
| Fuzzy Logic and CSA | 98.43 |
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 contain 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.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

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

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

- **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.

\[
\text{Recall} = \frac{TP}{TP + FN}
\]

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

\[
F_1 \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

### 4.4 Experimental settings and baselines

Experiments are performed on a PC with Intel(R) Core (TM) i7-6400 CPU at 2.71 GHZ 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

| Algorithms                          | Classification accuracy (%) |
|-------------------------------------|-----------------------------|
| SVM (Ghosh et al. 2020)             | 76.67                       |
| Boosting (Kaur 2016)                | 88.81                       |
| Maximum Entropy (Kabir et al. 2021) | 97.21                       |
| SWESA (Sarma and Sethares 2018)     | 86.93                       |
| Fuzzy Logic and CSA                 | 97.67                       |

| Algorithms                          | Classification accuracy (%) |
|-------------------------------------|-----------------------------|
| SVM (Ghosh et al. 2020)             | 77.29                       |
| Boosting (Kaur 2016)                | 90.27                       |
| Maximum Entropy (Kabir et al. 2021) | 91.02                       |
| SWESA (Sarma and Sethares 2018)     | 81.04                       |
| Fuzzy Logic and CSA                 | 94.98                       |
applied on each dataset, and the performance is compared based on cross-validation, accuracy, precision, recall, and f-score.

1. Accuracy analysis

The experiment results were obtained using fourfold 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. Table 4 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. Table 1 indicates that the difference between our proposed method and the latest techniques is approximately 4.25% in terms of accuracy for the Amazon dataset. Table 2 shows that the difference between our proposed method and the latest techniques is approximately 0.46% in terms of accuracy for the Yelp dataset. Finally, Table 3 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. 9.

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. Figure 10 shows this superiority.

| Methods/datasets                        | AMAZON CELLS | YELP        | IMDB        |
|-----------------------------------------|--------------|-------------|-------------|
|                                         | Recall | Precision | F-score    | Recall | Precision | F-score    | Recall | Precision | F-score    |
| SVM (Ghosh et al. 2020)                 | 0.850  | 0.855     | 0.850      | 0.705  | 0.700     | 0.700      | 0.760  | 0.755     | 0.755      |
| Boosting (Kaur 2016)                    | 0.840  | 0.840     | 0.840      | 0.680  | 0.690     | 0.670      | 0.690  | 0.730     | 0.690      |
| Maximum Entropy (Kabir et al. 2021)    | 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      |
5 Conclusions and future work

Sentiment analysis is essential for anyone who makes a decision. It is 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 and 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.

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.

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

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.

Declarations

Conflict of interest 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.

References

Abboud R, Tekli J (2020) Integration of nonparametric fuzzy classification with an evolutionary-developmental framework to perform music sentiment-based analysis and composition. Soft Comput 24(13):9875–9925. https://doi.org/10.1007/s00500-019-04503-4

Al-Deen HSS, Zeng Z, Al-Sabri R, Hekmat A (2021) An improved model for analyzing textual sentiment based on a deep neural network using multi-head attention mechanism. Appl Syst Innov. https://doi.org/10.3390/asi4040085

Al Shboul B, Al-Ayyoub M, Jararwehy Y (2015) Multi-way sentiment classification of Arabic reviews. In: 2015 international conference on information and communication systems ICICS
2015, pp 206–211, 2015. https://doi.org/10.1109/IACS.2015.7103228

Askarzadeh A (2016) A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm. Comput Struct 169:1–12. https://doi.org/10.1016/j.compstruc.2016.03.001

Baccianella S, Esuli A, Sebastiani F (2010) “SENTIWORDNET 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: Proceeding of 7th international conference on language resources and evaluation 2010, pp 2200–2204

Dodds PS et al (2015) Human language reveals a universal positivity bias. Proc Natl Acad Sci USA 112(8):2389–2394. https://doi.org/10.1073/pnas.141678112

Fares M, Moufarrej A, Tekli J, Grosky W (2019) Unsupervised word-level affect analysis and propagation in a lexical knowledge graph. Knowl Based Syst 165:432–459. https://doi.org/10.1016/j.knosys.2018.12.017

Ghosh S, Hazra A, Raj A (2020) A comparative study of different classification techniques for sentiment analysis. Int J Synth Emot 11(1):49–57. https://doi.org/10.4018/ijse.202001010a

Gupta G, Malhotra S (2015) Text document tokenization for word frequency count using rapid miner (taking resume as an example). In: International conference on advanced engineering and technology no. ICAET, pp 24–26 2015

Jefferson C, Liu H, Cocea M (2017) Fuzzy approach for sentiment analysis. IEEE Int Conf Fuzzy Syst. https://doi.org/10.1109/FUZZ-IEEE.2017.8015577

Kabir M, Kabir MMJ, Xu S, Badhon B (2021) An empirical research on sentiment analysis using machine learning approaches. Int J Comput Appl 43(10):1011–1019. https://doi.org/10.1080/1206212X.2019.1643584

Katarya R, Yadav A (2018) A comparative study of genetic algorithm in sentiment analysis. In: Proceedings of the 2nd international conference on inventive systems and control, ICISC 2018, no. ICISC, pp 136–141, 2018. https://doi.org/10.1007/978-1-4939-7887-8_9

Kaur P (2016) Design and implementation of boosting classification algorithm for sentiment analysis on newspaper articles. Int J Comput Sci Info Technol 7(4):5–38. https://doi.org/10.1016/j.fss.2014.01.015

Ma T, Al-Sabri R, Zhang L, Marah B, Al-Nabhan N (2020) The impact of weighting schemes and stemming process on topic modeling of arabic long and short texts. ACM Trans Asian Low Resour Lang Inf Process. https://doi.org/10.1145/3405843

Madhoushi Z, Hamdan AR, Zainudin S (2015) Sentiment analysis techniques in recent works. In: Proceedings of 2015 science and information conference SAI 2015, no. March, pp 288–291. https://doi.org/10.1109/SAI.2015.7237157

Medhat W, Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. Ain Shams Eng J 5(4):1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Nasim Z, Rajput Q, Haider S (2017) Sentiment analysis of student feedback using machine learning and lexicon based approaches. In: International conference on research and innovation in information systems, ICRIS, no. March 2021. https://doi.org/10.1109/ICRIS.2017.8002475

Nielsen F (2011) Afinn. Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby. 2011

Porai S, Cambria E, Winterstein G, Bin Huang G (2014) Sentic patterns: dependency-based rules for concept-level sentiment analysis. Knowl Based Syst 69(1):45–63. https://doi.org/10.1016/j.knosys.2014.05.005

Ravi K, Ravi V, Prasad PSR (2017) Fuzzy formal concept analysis based opinion mining for CRM in financial services. Appl Soft Comput J 60:786–807. https://doi.org/10.1016/j.asoc.2017.05.028

Ravi K, Ravi V (2015) A survey on opinion mining and sentiment analysis: tasks, approaches and applications, vol 89, no. November, 2015

Ross TJ (2010) Fuzzy logic with engineering applications, 3rd edn. Wiley, Hoboken. https://doi.org/10.1002/9781119994374

Saif H, He Y, Fernandez M, Alani H (2016) Contextual semantics for sentiment analysis of Twitter. Inf Process Manag 52(1):5–19. https://doi.org/10.1016/j.ipm.2015.01.005

Sarma PK, Sethares WA (2018) Simple algorithms for sentiment analysis on sentiment rich, data poor domains. In: COLING 2018—27th international conference on computational linguistics proceedings, pp 3432–3435, 2018

Serrano-Guerrero J, Olivas JA, Romero FP, Herrera-Viedma E (2015) Sentiment analysis: a review and comparative analysis of web services. Inf Sci (NY) 311:18–38. https://doi.org/10.1016/j.ins.2015.03.040

Tang D, Wei F, Yang N, Zhou M, Liu T, Qin B (2014) Learning sentiment-specific word embedding. In: AcI, pp 1555–1565, 2014

Vashishtha S, Susan S (2019) Fuzzy rule based unsupervised sentiment analysis from social media posts. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2019.112834

Vashishtha S, Susan S (2020a) Inferring sentiments from supervised classification of text and speech cues using fuzzy rules. Procedia Comput Sci 167(2019):1370–1379. https://doi.org/10.1016/j.procs.2020.03.348

Vashishtha S, Susan S (2019) Fuzzy logic based dynamic plotting of mood swings from tweets, vol 939, no. May 2020b. Springer

Verma T, Renu R, Gaur D (2014) Tokenization and filtering process in rapidminer. Int J Appl Inf Syst 7(2):16–18. https://doi.org/10.1109/ICRIIS.2017.8002475

Zadeh A, Zellers R, Pincus E, Morency L-P (2016) MOSI: multimodal corpus of sentiment intensity and subjectivity analysis in online opinion videos, 2016, [Online]. http://arxiv.org/abs/1606.06259

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.