An Optimization of Fuzzy Rough Set Nearest Neighbor Classification Model Using Krill Herd Algorithm for Sentiment Text Analytics

Sujatha Krishanmoorthy, Yunsheng Chai, Yewei Du, Xiangnan Zhou, Huimin Zhang, Hongyu Liu, and Changjiang Zhang

Abstract In recent years, social networking sites like Twitter and Facebook are more popular through which emotions and thoughts of its users are shared. Detection and classification of emotions, expressed in social media content, is beneficial for several applications in the domains like e-commerce, politics, social welfare and so on. Several works have been conducted earlier focusing sentiment and emotion analysis. These works primarily concentrated on single-label classification, leaving beside the correlation among different emotions expressed by an individual. In this view, the current research article presents a new sentiment and emotion classification model using Fuzzy Rough Set Nearest Neighbor (FRSNN) with Krill Herd (KH) algorithm i.e., FRSNN-KH. The proposed FRSNN-KH algorithm involves preprocessing, feature extraction and classification. Initially, preprocessing is executed to remove the unwanted words present in the tweet. Next, the features are extracted from the pre-processed tweet by following Bag of Words (BoW) model. Afterwards, FRSNN-based classification process is carried out to segregate the instances under different class labels. Finally, soft computing-based KH algorithm is applied to optimize the

S. Krishanmoorthy (✉) · Y. Chai · Y. Du · X. Zhou · H. Zhang · H. Liu
Department of Computer Science, Wenzhou Kean University, Zhejiang Province, Wenzhou, China
e-mail: sujatha@wku.edu.cn

Y. Chai
e-mail: Chaiyu@kean.edu

Y. Du
e-mail: Duy@kean.edu

X. Zhou
e-mail: zhouxian@kean.edu

H. Zhang
e-mail: zhanghui@kean.edu

H. Liu
e-mail: liuhon@kean.edu

C. Zhang
Division of Science and Technology, Hongkong Baptist University, UIC, Zhuhai, China
e-mail: changjiangzhang@uic.edu.hk

D. Gupta et al. (eds.), *Soft Computing for Data Analytics, Classification Model, and Control*, Studies in Fuzziness and Soft Computing 413,
https://doi.org/10.1007/978-3-030-92026-5_1

rule generation sets by FRSNN model. The presented FRSNN-KH algorithm was experimentally validated using benchmark dataset. The simulation outcome inferred the goodness of the proposed FRSNN-KH algorithm over compared methods under several dimensions.

Keywords Classification · Emotions · Fuzzy rough set theory · Sentiment analysis · Soft computing

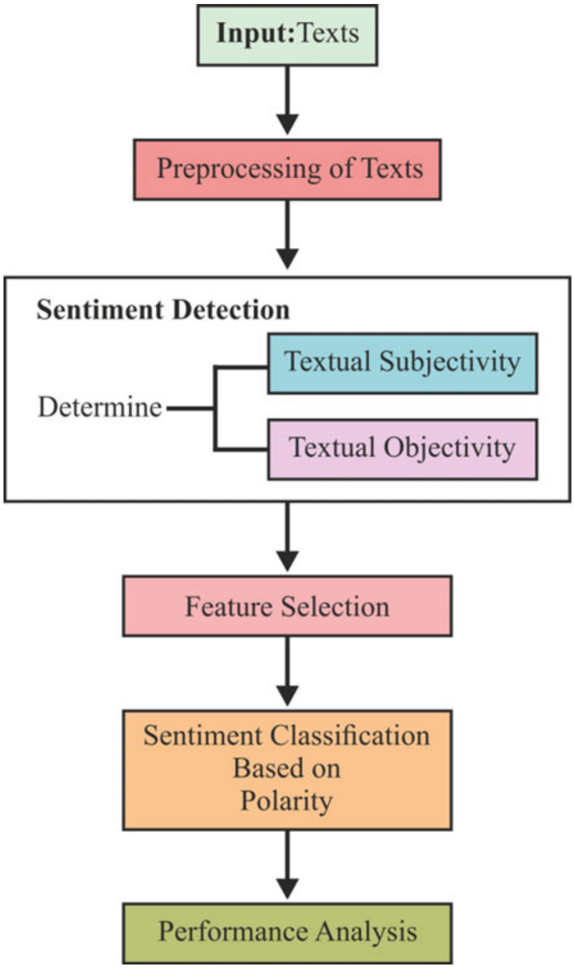
1 Introduction

Sentiment Analysis (SA) study established that there is a frequent development of web by means of volume, velocity and variety of opinion-rich data available in Internet. SA is a tendency for real-time functions and serves as decision support system. It further submits the required data to analysts. Specifically, big data analytics is an effective technique which leverages highly unstructured data and makes a significant data analysis with the help of contemporary tools. In addition, text-mining methods describe the transformation and replacement of unstructured data into structured one for the purpose of Knowledge Discovery (KD). Extensive studies have been carried out to find the best text mining process. Here, SA is referred to as a general text classification operation [1]. It studies the people's suggestions, approach, and sentiments towards an object or any kind of actions. The people's responses might be views and suggestion polarities such as positive, negative or neutral for textual context, accessible by social medium.

In typical SA process, few steps are followed such as data collection, Feature Selection (FS) [2], sentiment classification and sentiment polarity detection. Figure 1 shows the general structure of SA. An effective FS process is said to be a complicated one. It plays a vital role in computing accurate sentiment classification. Additionally, enhanced dimensionality, difficulty, and fuzziness in user-based Twitter data promote the requirement for best sentiment classification models. There are few prominent studies conducted earlier in search of novel concepts that can handle uncertainty, inaccuracy, estimation, incomplete truth and fuzziness. These concepts are demanded for options to enable the replication of human intelligence for personalized as well as common results. Soft Computing (SC) is another application which tends to integrate fresh processing approaches that reflect the consciousness and cognition in diverse dimensions, a knowledge acquired through experience; it can be universalized into applications that lack direct knowledge; and by parallel computer structures which accelerates the biological task. Soft computing is capable of performing rapid mapping of inputs to outputs when compared with inherent serial analytical presentations.

Followed by, a programmatic study on "Sentiment Analysis of Twitter Using SC Methods" has been presented in the literature. The main aim of this model is to collect empirical witness, examine the results from previous studies to provide better outcomes than previous studies, find the gaps in recent search and offer future

Fig. 1 General structure of SA



prospects with respect to the methods developed so far. SC models are assumed to be optimizing approaches which tend to deploy tedious real-time issues that can accomplish strong and minimum cost schemes. Such models are classified into five classes such as Machine Learning (ML), Neural Networks (NN), Evolutionary Computation, Fuzzy Logic (FL) and Probabilistic Reasoning (PR): Bayesian Networks (Bayesian probability).

Sentiments or emotions are mainly used to convey the feelings and opinions among human beings. Online social media, like Twitter and Facebook (FB), are offered in different languages depending on the people. Recently, an update has been given in which the users can make short interactions regarding facts, opinions, emotions, and its intensities on diverse kinds of themes. Authors are fascinated towards the determination of emotions in social media content using natural language processing.

It has been extensively deployed in commerce, public health, social welfare, and so on. For example, it has been applied in public health, public opinion prediction regarding political movements, brand consultation, and stock market observation. SA is defined as the estimation of attitude for the topic under study. The attitude may be either positive or negative or sometimes it may be even an emotional state like happy, sad and anger [3]. Currently, multi-label classification issues have achieved applicable results from a vast range of applications, such as text, scene and video classification, as well as bioinformatics [4]. On the contrary, in classical single-label classification problem, the sample is related with a single label from certain labels. However, in case of multi-label classification problem, a sample is combined with a subset of labels.

Several ML models have been presented for conventional emotion classification and multi-label emotion classification so far [5–8]. Al-Smadi et al. [9] presented two ML techniques like Deep Recurrent Neural Network (DRNN) and SVM. The researchers focused on the assessment of three different operations like aspect class examination, aspect Opinion Target Expression (OTE), and aspect sentiment polarity identification. When compared, the attained results implied that DRNN is an effective performer compared to SVM. Thus, DRNN is said to be robust and rapid in terms of implementation during training and testing. Furthermore, Al-Smadi et al. [10] employed two deep Learning Long Short-Term Memory (LSTM) NN for aspect-based SA. Initially, a character-level bidirectional LSTM was applied in aspect OTE extraction with a definite classification. Alternatively, an aspect-based LTSM was presented in this study for aspect sentiment orientation classification. The final outcome shows that the aspect-OTE extraction as well as orientation classifier function were better when related to baseline study.

In line with this, Al-Ayyoub et al. [11] employed a supervised ML model to extract different aspects and categorize the sentiments for hotels' Arabic reviews. The model has three major operations such as finding the aspect classes, obtaining the opinion targets, and exploring the sentiment polarity. The simulation outcome pointed that the newly-developed method performs well using a similar dataset, compared to other applications. Besides, García-Pablos et al. [12] provided W2VLDA, an aspect sentiment classifier model which needs minimum examination and no necessity of specific languages. It is capable of differentiating opinion words from aspect terms in an unsupervised manner. The model's efficiency was estimated by applying SemEval 2016 dataset. The analysis implied competing outcome for diverse languages. Likewise, Dragoni et al. [13] deployed an aspect-based unsupervised model for opinion tracking which maintains data visualization.

Rathan et al. [14] implied an ontology-based SA method for 'Smart phones' in case of tweets and messages. Here, the model was designed with diverse features like spelling correction, emoji and sentiment detection. Lexicon-driven technology was utilized for automatic training of data labels. A 'Smartphone'-based application, loaded with SVM, was used to enhance the classification accuracy. Cambria et al. [15] presented an ensemble model of symbolic and sub-symbolic AI. In this model, LSTM network was used to find the verb-noun primitives through lexical replacement and connect to common-sense paradigm using 3-level knowledge representation for

SA named 'SenticNet'. Alternatively, Xianghua et al. [16] recommended an unsupervised method to estimate the aspects and emotions in Chinese social reviews. The developers applied Latent Dirichlet Allotiment (LDA) in social reviews to explore the multi-aspect global sections.

Chen et al. [17] boosted LDA to arrive at Automatic Knowledge LDA with complete automation instead of applying the previous domains to find novel aspects. This model was capable of producing aspects and resolving the issues relevant to inferior knowledge through application and improvement of Gibbs sampler approach. Poria et al. [18] arranged an unsupervised rule centric technique to get explicit aspects and Implicit Aspects Clues (IAC) from goods and restaurant reviews. In this model, the researchers applied unsupervised rule centric technique to explore IACs in review. Then, it is divided into aspects with common-sense knowledge and dependency structure of a sentence, under the application of WordNet and SenticNet.

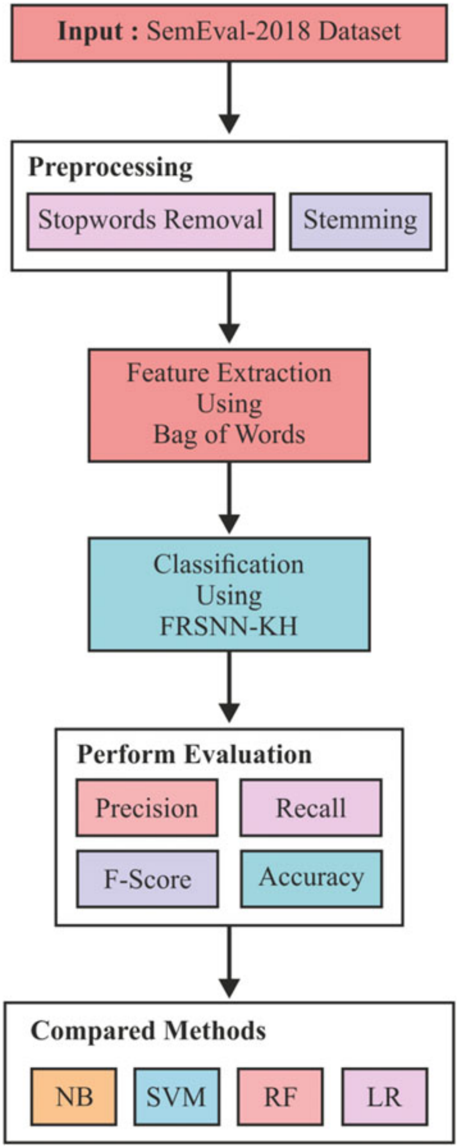
Pham and Le [19] projected a multi-layer structure to point the hotels reviews of users. Learning methods like word embeddings and compositional vectors were employed in the study to collect words, sentences, aspect-based graphs, etc. These elements were used to arrive at productive knowledge. Such representations were integrated with NN to develop a method that can predict the entire hotel reviews at the time of extracting aspects' ratings and weights. Liu et al. [20] concatenated supervised and unsupervised models in which automated rule-based approaches were used to enhance double propagation. Hence, double propagation considers those opinion words with a target. There is a syntactic relationship between the opinion word and a target with similar sentence. Ma et al. [21] integrated LSTM network with hierarchical process. The simulation outcome shows that the integration of planned attention design and Sentic LSTM performs quite well than aspect-based SA process. If the application is composed of autonomous parameters to define the link between aspects and corresponding opinion representations, it is capable of capturing aspect-driven emotions with few data requirement.

The authors, in current study, presented an efficient sentiment and emotion classification model using FRSNN and KH Algorithm, named as FRSNN-KH. The presented model involves preprocessing, BoW-based feature extraction, FRSNN-based classification and KH-based rule set optimization. The application of soft computing-based KH algorithm helps in the optimization of number of rules generated by FRSNN model. Further, it assists in improving the classification performance. The performance of the presented FRSNN-KH algorithm was validated using benchmark dataset and the results were assessed under several aspects.

2 The Proposed FRSNN-KH Model

Figure 2 shows the overall working principle of FRSNN-KH approach. The figure reveals that the input data undergoes preprocessing to remove the unwanted words present in the tweet. Followed by, BoW method is applied as a feature extractor, to

Fig. 2 Block diagram of FRSNN-KH algorithm



acquire the useful features from pre-processed tweet. Then, FRSNN-based classification is carried out to identify and segregate the instances under different class labels. Finally, soft computing-based KH algorithm is applied to optimize the rule generation sets by FRSNN model.

2.1 Fuzzy-Rough Set Theory

A study based on hybridization of fuzzy sets and rough sets was presented with maximum efficiency. It mainly concentrated on fuzzifying the equations for lower and upper approximations [22]. The study followed two basic principles as provided below:

- Set A is generalized to fuzzy set in X that enables the objects which come under the class of applied degrees.
- Instead of using objects with indiscernibility, it is better to estimate the appropriate quality. Finally, the objects are classified into classes or granules along with 'soft' boundaries according to similarity measure. In line with this, sudden transactions occur among these classes by moderate ones with the activation of a component that possess massive number of classes.

In particular, approximate equality is assessed among the objects with frequent attribute values by means of fuzzy relation R in X. This has been allocated for every object based on its similarity value. Basically, R is a minimum fuzzy tolerance formula in which the fuzzy set A of X is provided with lower and upper approximations of A alike how R is manufactured in various points. A common description as given herewith

$$(R \downarrow A)(p) = \underset{q \in X}{inf \ell} (R(p, q), A(q)) \tag{1}$$

$$(R \uparrow A)(p) = \underset{q \in X}{\sup \mathcal{T}} (R(p, q), A(q)) \tag{2}$$

where, ℓ implies an implicator and $\mathcal{T}$, a t-norm. Here, A defines a set and R denotes an equivalent relation in X. Fortunately, it is an effective advantage for specific complexities. For this purpose, the method of Vaguely Quantified Rough Sets (VQRS) is deployed in this study. It applies the linguistic quantifiers like 'most' and 'some' which are against traditional crisp quantifiers such as 'all' and 'at least one' since the latter selects the class of objects to both lower and upper approximations. A couple of fuzzy quantifiers $\left(Q_u, Q_j\right)$ is applied which makes the 'most' and 'some', with lower and upper approximations of A by R as described herewith.

$$\left(R \downarrow^{Q_\mu} A\right)(q) = Q_u\left(\frac{|Rq \cap A|}{|Rq|}\right) = Q_u\left(\frac{\sum_{p \in X} \min\,(R(p, q), A(p))}{\sum_{p \in X} R(p, q)}\right) \tag{3}$$

$$\left(R \uparrow^{Q} A\right)(q) = Q_j\left(\frac{|Rq \cap A|}{|Rq|}\right) = Q_j\left(\frac{\sum_{p \in X} \min\,(R(p, q), A(p))}{\sum_{p \in X} R(p, q)}\right) \tag{4}$$

The intersection of a fuzzy set is represented by min t-norm. The instances of fuzzy quantifiers can be generated with respect to given parameterized equation, for $0 \leq \alpha < \beta \leq$ and p in [0, 1],

$$Q_{!\alpha,\beta)}(p) = \begin{cases} 0, & p \leq \alpha \\ \dfrac{2(p-\alpha)^2}{\beta-\alpha} & \alpha \leq p \leq \dfrac{\alpha+\beta}{2} \\ 1 - \dfrac{2(p-\alpha)^2}{(\beta-\alpha)^2} & \dfrac{\alpha+\beta}{2} \leq p \leq \beta \\ 1, & \beta \leq p \end{cases} \tag{5}$$

Here, $Q_{(0.1,0.6)}$ and $Q_{(0.2,1)}$ are applied to show vague quantifiers (VQ) '*some*' and '*most*' simultaneously. The vital difference between (8) and (9) is that the VQRS approximations do not expand the conventional FR approximations. Here, A and R are crisp and are represented herewith.

$$Q_{>p_1}(p) = \begin{cases} 0, \; p \leq p_1 \\ 1, \; p > p_1 \end{cases} Q_{\geq p_u}(p) = \begin{cases} 0, \; p < p_u \\ 1, \; p \geq p_u \end{cases} \tag{6}$$

With $0 \leq p_1 < p_u \leq 1$ being applied as quantifiers, Ziaroko's variable precision rough set model is recovered and it applies the following equation.

$$Q_{\exists}(p) = \begin{cases} 0, \; p = 0 \\ 1, \; p > 0 \end{cases} Q_{\forall}(p) = \begin{cases} 0, \; p < 1 \\ 1, \; p = 1 \end{cases} \tag{7}$$

Pawlak's standard rough set method is applied in case of VQRS model in which R is considered as a crisp equivalence formula. Then, VQRS approach resolves the noisy data using a novel perception: it acquires the elasticity of VPRS to solve the classifier errors, releases the membership constraints for lower approximation, fix the upper approximation and fuzzy sets to express incomplete constraint convenience. This method was applied to perform the FS process.

2.2 *Fuzzy-Rough Set Nearest Neighbours (FRSNN)-Based Classification*

In this work, FRSNN method was presented in which NN was used to develop fuzzy lower and upper approximations of decision categories. Further, the test samples were divided according to the membership of such approximations. This technology depends upon fuzzy tolerance relation, R. Here, R is developed using the group of conditional attributes A, R as shown by,

$$R(p,q) = \min_{a \in \mathbb{A}} R_0(p,q) \tag{8}$$

where $R_a(p,q)$ implies the degree of objects, p and q similar to attribute a. There are various feasible options and it is selected as follows [22].

$$R_a(p,q) = 1 - \frac{|a(p) - a(q)|}{|a_{\mathrm{mx}} - a_{\mathrm{mn}}|} \tag{9}$$

where σ_a^2 denotes the variance of attribute a, whereas a_{mx} and a_{mn} denote the maximum and minimum attribute values. The main aim of this approach is to quantify the lower and upper approximations of a decision class with NN of test object q, and offer the best clues for membership prediction of specific class. Specifically, when $(R \downarrow C)(q)$ is maximum, it shows every q's neighbours come under class C. In case of higher value for $(R \uparrow C)(q)$, then it implies that at least one neighbour should be in the class. The classification task is often computed for y as the initialization of τ to 0.

In order to execute a better classification task, the models result in decision class with optimal fuzzy lower and upper approximation memberships. The complexity of this model is $(|\mathcal{C}| \cdot (2|X|))$. It is implied as a zero order Takagi–Sugeno manager which is used as a rule, and the membership of the maximum test object. R_d shows fuzzy tolerance correlation for decision feature, d. Here, similar relationship is applied for conditional features. It is evident that with $\ell = \ell_M$ and $\mathcal{T} = \mathcal{T}_M$, condition $\tau_2 = 0$ is satisfied while $R(q,z) = 1$ for each neighbour z in N (total affinity of test object as well as NN). However, $R_d(z_1, z_2) = 0$ for each z_1, z_2 in N(total dissimilarity between the decision values of two neighbours).

As per relevant approximations of remarkable FR set theory, the methods are provided to resolve the presence of noise. This is because Sup and Inf are exploited and generalized with existential as well as universal quantifiers, correspondingly. On the other hand, it was not influenced by the choice of I, instead it was validated against crisp decisions with individual NN applied during classification task. It is advantageous in parameter election process which defines the classification decisions based on single object and it tends to produce maximum noisy data.

For this purpose, VQNN (Vaguely Quantified Nearest Neighbours) model, a FRSNN approach, is presented in which $R \downarrow C$ and $R \uparrow C$ were substituted by $R \downarrow^{Q_u} C$ and $R \uparrow^{Q_1} C$, correspondingly. Analogously, VQNN2 is an variant of FRSNN2, where $R \downarrow R_d z$ and $R \uparrow R_d z$ were exchanged by $R \downarrow^{Q_u} R_d z$ and $R \uparrow^{Q_1} R_d z$, respectively. For FRSNN, the application of K is not significant. In FRSNN, the effect has become minimal since $R(p,q)$ acquires smaller and p results in minimum influence on $(R \downarrow C)(q)$ and $(R \uparrow C)(q)$.

2.3 Optimal Rule Set Generation Using KH Algorithm

KH algorithm is applied to optimize the rule set generated by FRSNN model. KH is defined as a new meta-heuristic Swarm Intelligence (SI) optimization model. This model is used to resolve the optimization issues on the basis of herding nature outcomes attained from krill members [23]. Time-based position of a krill in 2D plane can be computed as three major events:

- Movement made by alternate krill members;
- Foraging behavior;
- External diffusion.

KH method has applied Lagrangian algorithm which is as follows.

$$\frac{dX_i}{dt} = N_i + F_i + D_i, \tag{10}$$

where N_i denotes the motion persuaded by alternate krill individuals; F_i represents the foraging action; and D_i shows external diffusion of ith krill individuals. The process involved in KH algorithm is demonstrated in Fig. 3.

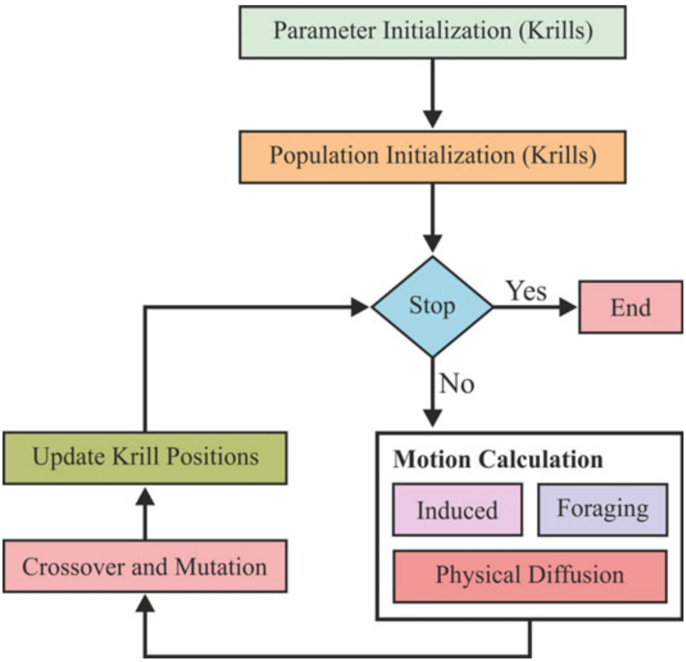


Fig. 3 Flowchart of KH algorithm

2.3.1 Motion Induced by Other Krill Individuals

In case of krill individual, the motion is encouraged by additional krill members which is implied as follows

$$\begin{aligned} N_i^{new} &= N^{\max}\alpha_i + \omega_n N_i^{old}, \\ \alpha_i &= \alpha_i^{1oca1} + \alpha_i^{target}, \end{aligned} \tag{11}$$

where $N^{\max}$ denotes high induced speed, ω_n shows inertia weight of motion persuaded in [0, 1], N_i^{old} signifies final motion provided, α_i^{1oca1} defines the local effect offered by neighbors, and α_i^{target} refers to the target dimension induced by optimal krill members. Consequently, the impact of neighbors could be estimated by,

$$\begin{aligned} \alpha_i^{1oca1} &= \sum_{j=1}^{NN} \widehat{K}_{ij}\widehat{X}_{ij}, \\ \hat{X}_{ij} &= \frac{X_j - X_i}{\|X_j - X_i\| + \varepsilon}, \\ \hat{K}_{ij} &= \frac{K_j - K_i}{K^{worst} - K^{best}}, \end{aligned} \tag{12}$$

where K_i implies the fitness measure of ith krill individual.K^{best} and K^{worst} are superior and inferior fitness rates of krill individuals. K_j defines the fitness of jth$(j = 1, 2, \ldots, NN)$ neighbor. NN mimics the count of neighbours. X indicates the relevant position. The sensing distance of every krill individual is computed as given below:

$$d_{s,i} = \frac{1}{5N}\sum_{j=1}^{N}\|X_i - X_j\|, \tag{13}$$

where N corresponds to the value of krill members. The consequence of KH with optimal fitness on ith individual krill is assumed as given in this equation.

$$\partial_i^{target} = C^{best}\widehat{K}_{i,best}X_{i,best}, \tag{14}$$

where the measure of C^{best} is determined as:

$$C^{best} = 2\left(rand + \frac{I}{I_{\max}}\right) \tag{15}$$

2.3.2 Foraging Motion

This metric is computed with respect to two major parameters. Food location followed by the experience regarding first one are the two parameters involved in this metric. It can represent ith krill individual as given below.

$$\begin{aligned} F_i &= V_f \beta_i + \omega_f F_i^{old}, \\ \beta_i &= \beta_i^{food} + \beta_i^{best}, \end{aligned} \tag{16}$$

where y_f implies foraging speed, ω_f denotes the inertia weight of foraging action from 0 and 1, and F_i^{old} refers to consequent foraging action. β_i^{food} depicts the attraction of food and β_i^{best} shows the impact of better fitness of ith krill. Here, $y_f = 0.02$.

The virtual center of food concentration can be determined on the basis of fitness distribution of KH which is persuaded from the *center of mass* and is represented below.

$$X^{food} = \frac{\sum_{i=1}^{N} (1/K_i) X_i}{\sum_{i=1}^{N} (1/K_i)}. \tag{17}$$

Hence, the food concentration for ith KH is evaluated as given below.

$$\beta_i^{food} = C^{food} \widehat{K}_{i,food} \widehat{X}_{i,food}, \tag{18}$$

where C^{food} denotes the food coefficient using the given function.

$$C^{food} = 2\left(1 - \frac{I}{I_{\max}}\right). \tag{19}$$

The impact of optimal fitness of ith krill member is managed under the application of given formula:

$$\beta_i^{best} = \widehat{K}_{i,ibest} \widehat{X}_{i,ibest}, \tag{20}$$

where $\widehat{K}_{i,ibest}$ implies the previously-visited position of ith krill individual.

2.3.3 Physical Diffusion

The arbitrary diffusion of KH is assumed to be an additional random task. It is defined with respect to high diffusion speed and random directional vector as given herewith.

$$D_i = D^{\max}\delta, \tag{21}$$

where $D^{\max}$ denotes heavy diffusion speed, and δ defines random directional vector, and corresponding arrays are arbitrary measures in $[-1, 1]$. When the position is the best, the random motion becomes minimum. The impacts of motion encouraged by other krill individuals and foraging motion are reduced with improving iterations. Hence, an alternate Eq. (21) is substituted instead of Eq. (22). It gets linearly decreased with random speed and time and it functions according to geometrical annealing schedule.

$$D_i = D^{\max}\left(1 - \frac{I}{\mathrm{I}_{\max}}\right)\delta \tag{22}$$

2.3.4 Crossover

It is managed using crossover probability Cr, while mth component of $X_i\left(x_{i,m}\right)$ is described herewith.

$$x_{i,m} = \begin{cases} x_{r,m} & rand_{i,m} < Cr \\ x_{i,m} & else, \end{cases}$$
$$Cr = 0.2\hat{K}_{i,best} \tag{23}$$

where $r \in \{1, 2, \ldots, i-1,\ i+1, \ldots, N\}$.

2.3.5 Mutation

It is handled by a mutation probability, Mu and an adaptive mutation approach is evaluated as given herewith.

$$x_{i,m} = \begin{cases} x_{gbest,m} + \mu\left(x_{p,m} - x_{q,m}\right) & rand_{i,m} < Mu \\ x_{i,m} & else, \end{cases}$$
$$Mu = 0.05\widehat{K}_{i,bets}, \tag{24}$$

where $p, q \in \{1, 2, \ldots,\ i-1,\ i+1, \ldots, K\},\ \mu \in [0, 1]$.

2.3.6 Processes Involved in KH Algorithm

In general, the predetermined actions are primarily modified with position of krill members to attain optimal fitness. Foraging action, encouraged by alternate krill individuals, possesses two global and two local principles. Simultaneously, KH becomes an effective model. Under the application of diverse parameters, the motion at time and the position vector of a krill individual at interval t to $t + \Delta t$ is provided herewith.

$$X_i(t + \Delta t) = X_i(t) + \Delta t \frac{dX_i}{dt}. \tag{25}$$

It is evident that Δt is a vital constant and should be fixed appropriately on the basis of optimization problem. Δt can be attained easily with the following function.

$$\Delta t = C_t \sum_{j=1}^{NV} \left(UB_j - LB_j\right), \tag{26}$$

where NV denotes the variables and LB_j, UB_j denote the lower and upper bounds of jth variable $(j = 1, 2, \ldots, NV)$ correspondingly. Here, C_t is a constant value in [0, 2].

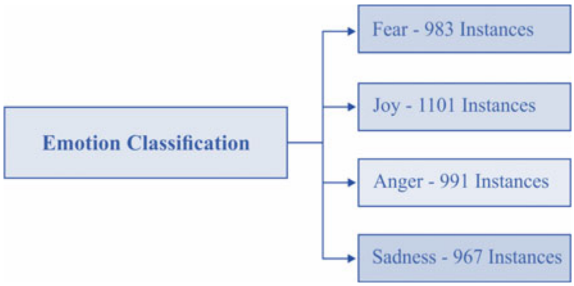
3 Performance Validation

To confirm the efficiency of the newly presented model, a sequence of processes was carried out using Python Programming language. The applied dataset and accomplished outcomes were examined under various scenarios.

3.1 Dataset

In order to verify the simulation outcome of the presented LSTM-EISC approach, a benchmark dataset i.e., SEMEVAL2018 Task-1Emotion Intensity Ordinal Classification dataset was employed in this study [24]. It is comprised of 4042 tweets with four instances like joy, fear, anger, and sadness. Among the tweets considered, 1074 tweets are categorized under Joy, 650 tweets belong to the class of fear, 991 tweets are grouped under anger and finally 555 tweets denote sadness. The dataset was classified into training and testing data in the proportion of 7.5:2.5. Figure 4 shows the data relevant to a dataset.

Fig. 4 Class distribution of the Sem-Eval 2018 dataset



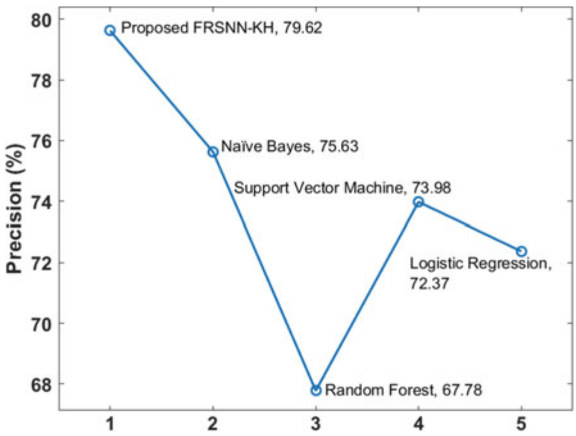
3.2 Results

Table 1 shows the results of the analysis of existing methods against the proposed FRSNN-KH model in terms of Precision, Recall, F-Measure and Accuracy. Figure 5 shows the performance results of FRSNN-KH algorithm in terms of precision. The figure states that the RF model produced an insignificant classifier outcome with a precision of 67.78%. The LR model attained 72.37% precision, somewhat better outcome to a certain extent. Besides, the SVM model achieved a

Table 1 Comparison of results attained by existing models and the proposed FRSNN-KH model under several measures

Methods	Precision	Recall	F-measure	Accuracy
Proposed FRSNN-KH	79.62	85.40	82.84	83.59
NB	75.63	74.47	75.05	75.19
RF	67.78	83.18	74.70	71.76
SVM	73.98	79.18	76.49	75.61
LR	72.37	78.30	75.21	74.15

Fig. 5 Precision analysis



moderate performance with a precision of 73.98%. Moreover, the NB model exhibited better results with 75.63% precision. Furthermore, the FRSNN-KH algorithm accomplished supreme performance and offered the highest precision of 79.62%.

Figure 6 reveals the results attained in the functioning of FRSNN-KH model with respect to recall. The figure reveals that the NB model is an ineffective classifier with a recall value outcome of 74.47%. The LR model exhibited moderate result to definite extent and reached 78.30% recall. On the other end, the SVM method produced better performance with a recall of 79.18%. Additionally, the RF technique implied gradual results with a recall of 83.18%. However, the presented FRSNN-KH approach accomplished the best performance with an optimal recall of 85.40%.

Figure 7 refers to the performance results of the presented FRSNN-KH model by means of F-measure. From the figure, it is clear that the RF model produced inferior classifier outcome i.e., 74.70% F-measure. Followed by, the NB model produced considerable results and achieved F-measure of 75.05%. On the other hand, the LR

Fig. 6 Recall analysis

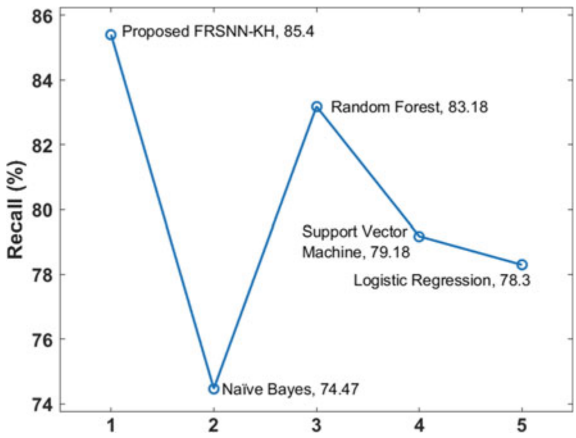
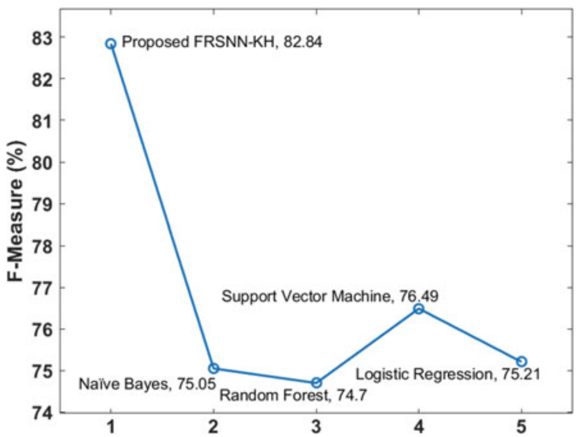


Fig. 7 F-measure analysis



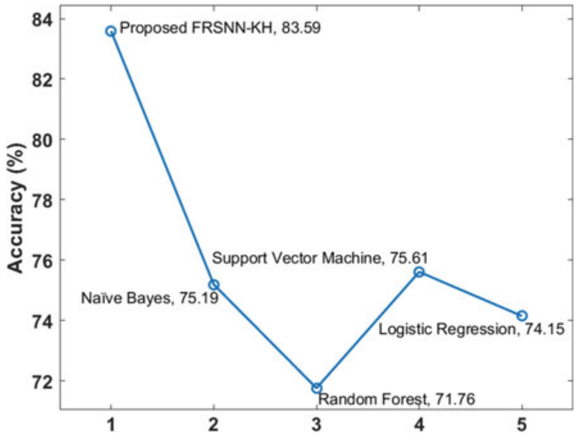


Fig. 8 Accuracy analysis

model accomplished better performance with an F-measure of 75.21%. Additionally, the SVM model implied moderate results with 76.49% F-measure. However, the presented FRSNN-KH technique achieved an optimized performance and provided the best F-measure of 82.84%.

Figure 8 depicts the performance results of the presented FRSNN-KH method in terms of accuracy. The figure refers that the RF model is an ineffective classifier as it achieved 71.76% accuracy. Additionally, the LR model depicted a manageable outcome and achieved an accuracy of 74.15%. Followed by, the NB model implied a gradual performance increment with 75.19% accuracy. The SVM model produced an accuracy of 75.61%, appreciable than previous models. However, the FRSNN-KH model accomplished superior performance and high accuracy of 83.59%.

From the above-mentioned tables and figures, it is inferred that the proposed FRSNN-KH algorithm is an excellent tool for classification of sentiments since it outperformed other models in terms of precision (79.62%), recall (85.40%), F-measure (82.84%) and accuracy (83.59%).

4 Conclusion

The current research work developed an efficient sentiment and emotion classification model named FRSNN-KH algorithm. The proposed FRSNN-KH algorithm has three steps namely preprocessing, feature extraction and classification. The input data was first pre-processed and then BoW method was applied as a feature extractor to extract the useful features from pre-processed tweet. Then, FRSNN-KH-based classification was carried out to segregate the instances under different class labels. Soft computing-based KH algorithm was applied in the optimization of number of rules generated by FRSNN model. This further helped in improving classification performance. The performance of the FRSNN-KH algorithm was validated using the benchmark dataset and the results were assessed under several aspects. The results inferred that the proposed algorithm is an excellent tool for classification of sentiments. In future we will try to involve recurrent model driven emotion and sentiment classification model with gated inputs and outputs.

References

1. Kumar, A., Sebastian, T.M.: Sentiment analysis on twitter. Int J Comput Sci Issues. **9**(4), 372–378 (2012)
2. Dave, K., Lawrence. S., Pennock, D.M.: Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In: Proceedings of the 12th international conference on World Wide Web. Budapest, Hungary (2003)
3. Reilly, T.: Web 2.0 Compact Definition: Trying Again. Sebastopol, CA: O'Reilly Media. http://radar.oreilly.com/2006/12/web-20-compactdefinition-tryi.html (2017). Accessed 24 April 2017
4. Alfouzan, H.I.: Introduction to SMAC-social mobile analytics and cloud. Int J Sci Eng Res. **6**, 128–130 (2015)
5. Shankar, K., WahabSait, A.R., Gupta, D., Lakshmanaprabu, S.K., Khanna, A., Pandey, H.M.: Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. Pattern Recogn. Lett. **133**, 210–216 (2020)
6. Pustokhina, I.V., Pustokhin, D.A., Gupta, D., Khanna, A., Shankar, K., Nguyen, G.N.: An effective training scheme for deep neural network in edge computing enabled Internet of Medical Things (IoMT) systems. IEEE Access **8**(1), 107112–107123 (2020)
7. Raj, J.S., Jeya Shobana, S., Pustokhina, I.V., Pustokhin, D.A., Gupta, D., Shankar, K.: Optimal feature selection based medical image classification using deep learning model in internet of medical things. IEEE Access **8**(1), 58006–58017 (2020)
8. Shankar, K., Lakshmanaprabu, S.K., Khanna, A., Tanwar, S., Rodrigues, Joel J.P.C., Roy, N.R.: Alzheimer detection using group grey wolf optimization based features with convolutional classifier. Comput. Electr. Eng. **77**, 230–243 (2019)
9. Al-smadi, M., Qawasmeh, O., Al-ayyoub, M., Jararweh, Y., Gupta, B.: Deep recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews. J. Comput. Sci., 386–393 (2017). https://doi.org/10.1016/j.jocs.2017.11.006
10. Al-Smadi, M., Talafha, B., Al-Ayyoub, M., Jararweh, Y.: Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. Int. J. Mach. Learn. Cybern., 1–13 (2018). https://doi.org/10.1007/s13042-018-0799-4
11. Al-ayyoub, M., Al-smadi, M., Al-ayyoub, M., Jararweh, Y., Qawasmeh, O.: Enhancing aspect-based sentiment analysis of Arabic hotel' reviews using morphological, syntactic and semantic features. Inf. Process. Manage. 308–319 (2018). https://doi.org/10.1016/j.ipm.2018.01.006
12. García-Pablos, A., Cuadros, M., Rigau, G.: W2VLDA: almost unsupervised system for aspect based sentiment analysis. Expert Syst. Appl. **91**, 127–137 (2018)
13. Dragoni, M., Federici, M., Rexha, A.: An unsupervised aspect extraction strategy for monitoring real-time reviews stream. Inf. Process. Manage. 1103–1118 (2018)
14. Rathan, M., Hulipalled, V.R., Venugopal, K.R., Patnaik, L.M.: Consumer insight mining: aspect based twitter opinion mining of mobile phone reviews. Appl. Soft Comput. J. **68**, 765–773 (2018). https://doi.org/10.1016/j.asoc.2017.07.056
15. Cambria, E., Poria, S., Hazarika, D., Kwok, K.: SenticNet 5: discovering conceptual primitives for sentiment analysis by means of context embeddings. In: Thirty Second AAAI Conference on Artificial Intelligence, pp. 1795–1802 (2018). https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16839. Accessed 26 Oct 2019

16. Xianghua, F., Guo, L., Yanyan, G., Zhiqiang, W.: Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon. Knowl. Based Syst., **37**, 186–195 (2013). https://doi.org/10.1016/j.knosys.2012.08.003
17. Chen, Z., Mukherjee, A., Liu, B.: Aspect extraction with automated prior knowledge learning. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 347–358 (2014)
18. Poria, S., Cambria, E., Ku, L.-W., Gui, C., Gelbukh, A.: A rule-based approach to aspect extraction from product reviews. In: Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP), pp. 28–37 (2014)
19. Pham, D., Le, A.: Data & knowledge engineering learning multiple layers of knowledge representation for aspect based sentiment analysis. Data Knowl. Eng. **114**, 26–39 (2018). https://doi.org/10.1016/j.datak.2017.06.001
20. Liu, Q., Gao, Z., Liu, B., Zhang, Y.: Automated rule selection for aspect extraction in opinion mining. In: IJCAI, pp. 1291–1297 (2015)
21. Ma, Y., Peng, H., Cambria, E.: Targeted aspect-based sentiment analysis via embedding common sense knowledge into an attentive LSTM, In: Thirty Second AAAI Conference on Artificial Intelligence, pp. 5876–5883 (2018). https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16541. Accessed 14 Sept 2019
22. Jensen, R., Cornelis, C.: Fuzzy-rough nearest neighbour classification and prediction. Theoret. Comput. Sci. **412**(42), 5871–5884 (2011)
23. Gandomi, A.H., Alavi, A.H.: Krill herd: a new bio-inspired optimization algorithm. Commun. Nonlinear Sci. Numer. Simul. **17**(12), 4831–4845 (2012)
24. Source: https://competitions.codalab.org/competitions/17751