Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

OPTIMIZING SENTIMENT ANALYSIS OF AMAZON PRODUCT REVIEWS USING A SOPHISTICATED FISH SWARM OPTIMIZATION-GUIDED RADIAL BASIS FUNCTION NEURAL NETWORK (SFSO-RBFNN)

P. RADHA, Dr. N. SUDHA BHUVANESWARI

Assistant Professor, Sri Krishna Arts & Science College, Department of Computer Science, India

Assistant Professor, Dr. G.R Damodaran College of Science, Department of Computer Science India

E-Mail: radhaharismita@gmail.com, Sudhabhuvaneshwari.n@grd.edu.in

ABSTRACT

Sentiment analysis automatically identifies and extracts subjective information from text, which can help understand people's opinions, emotions, and attitudes towards a particular topic. Sentiment analysis has become increasingly important in recent years, as online reviews and social media have become popular platforms for people to share their opinions and experiences. However, there are several challenges in sentiment analysis, including the complexity and ambiguity of language, the lack of context, and the cultural and linguistic differences that can affect the interpretation of sentiment. These challenges can result in inaccurate sentiment analysis, which can have negative consequences, such as misleading product reviews or biased customer feedback. In this paper, "Sophisticated Fish Swarm Optimization (SFSO)-guided Radial Basis Function Neural Network (RBFNN)" is proposed to perform sentiment analysis with enhanced classification accuracy. The SFSO algorithm optimizes the parameters of the RBFNN, which enables the model to adapt to varying review topics and sentiments. The SFSO algorithm's ability to explore the search space of the RBFNN parameters results in improved accuracy and performance in sentiment analysis. The proposed approach was evaluated on a dataset of Amazon product reviews and compared to other state-of-the-art sentiment analysis techniques. The proposed approach has potential applications in Books, Kindle stores, Tools and Home Improvement, and Industrial and Scientific domains where sentiment analysis is critical for understanding customer opinions and feedback. The results demonstrate that the proposed approach outperforms state-of-the-art classifiers in terms of Classification accuracy, F-measure, Fowlkes-Mallows Index, and Matthews Correlation Coefficient.

Keywords: *Sentiment Analysis, Amazon, Optimization, Classification, Radial Function, Fish Swarm.*

1. INTRODUCTION

Sentiment analysis, also known as opinion mining, is a field of study that deals with the computational treatment of opinions, emotions, and subjectivity of text data. The goal of sentiment analysis is to determine the sentiment or attitude expressed in a text, which can be positive, negative, or neutral [1]. The field has seen great growth and development in recent years, driven partly by the increasing volume of text data generated by social media and other online sources. Sentiment analysis has numerous applications in various industries, such as marketing, customer service, politics, and media. For example, sentiment analysis can be used in marketing to understand customer sentiment towards a brand or product, monitor brand reputation, and track customer satisfaction. In customer service, sentiment analysis can quickly categorize and prioritize customer inquiries based on their sentiment, allowing companies to respond to negative customer feedback promptly and effectively [2].

Sentiment analysis is typically performed using machine learning algorithms, which learn from large datasets of annotated text data to predict the sentiment of the new, unseen text. The quality of the sentiment analysis model is highly dependent on the training data quality and the machine learning algorithm choice. One of the challenges in sentiment analysis is dealing with the subjectivity of text data [3]. People have different opinions on the same subject, and words that are positive to one person may be negative to another. Additionally, sentiment can be expressed in implicit or nuanced ways, making it difficult to classify the sentiment of a piece of text accurately. Another challenge in sentiment analysis is dealing with sarcasm and irony, which are common in social media. Sarcasm and irony can completely reverse the sentiment of a statement, making it difficult for a machine-learning algorithm to classify the sentiment accurately. To address this, researchers are exploring methods to identify sarcasm and irony in text data and to adjust the sentiment analysis accordingly [4].

Sentiment analysis can be approached using different methods, such as lexicon-based, rule-based, and machine learning-based approaches. Lexicon-based approaches rely on dictionaries of words and phrases with known positive or negative sentiment to determine the sentiment of a given text [5]. Machine learning-based approaches, however, identify patterns in data to make predictions. This method tends to be more precise than lexicon-based or rule-based approaches as it can consider the context and subjectivity of the text data [6].

Sentiment analysis is a growing field with numerous applications in various industries [7]. Despite its challenges, such as dealing with the subjectivity of text data and sarcasm and irony, the field has made great strides in recent years and will likely continue to develop and grow in the coming years. By providing insights into text data's opinions, emotions, and subjectivity, sentiment analysis can potentially transform how we understand and interact with text data [8].

The significant feature of sentiment analysis is [9, 10]:

- Sentiment trend analysis: The ability to track changes in sentiment over time and identify trends in customer opinions and preferences.
- Sentiment analysis of social media: The ability to analyze the sentiment of social media posts and customer reviews.
- Sentiment analysis of customer feedback: The ability to analyze customer feedback data and understand customer opinions and preferences.
- Sentiment analysis of product reviews: The ability to analyze customer reviews of products and services and understand customer opinions and preferences.
- Sentiment analysis of competitor products and services: The ability to analyze customer opinions and preferences towards competitor products and services.

1.1. Problem Statement

Irony and sarcasm are prevalent forms of expression in natural language and often challenge sentiment analysis algorithms. These expressions can change the intended meaning of a sentence, making it difficult for algorithms to categorize the sentiment accurately. As a result, the objective of enhancing irony and sarcasm detection in sentiment analysis is crucial. To improve the accuracy of sentiment analysis algorithms, it is necessary to consider the contextual cues and linguistic patterns commonly associated with irony and sarcasm. This involves developing models to identify these cues and patterns and distinguishing between the text's literal and figurative expressions. This will enable sentiment analysis algorithms to accurately categorize the sentiment expressed in texts that contain irony or sarcasm.

1.2. Objective

Enhancing Irony and Sarcasm Detection: Sentiment analysis for product reviews can be challenging when the text contains irony or sarcasm. As a result, enhancing the detection of irony and sarcasm is an important objective for this field of research. To achieve this objective, researchers may focus on developing models incorporating additional contextual information and leveraging other sources of information, such as social media data, to better detect irony and sarcasm in product reviews. Additionally, researchers may investigate using transfer learning techniques and multilingual models to improve the performance of irony and sarcasm detection.

1.3 Organization of the Paper

Section 1 has discussed the sentiment analysis and its feature. Section 2 provides the literature review. Section 3 discusses the proposed classifier. Section 4 provides the details about the dataset used. Section 5 defines the metrics used to measure the performance of classifier. Section 6 provides the results with discussion. Section 7 concludes the paper.

2. LITERATURE REVIEW

"All-modalities-in-One BERT" [11] presents a new SA approach that considers multiple modalities, such as image, audio, and text. The authors propose a new model called AOBERT, which is based on BERT and is capable of processing multiple modalities in a single framework.

Table 1: Comparison of Related Literature

Related Literature	Merits	Demerits
All-modalities-in-One BERT	• Incorporates multiple modalities into one BERT model for sentiment analysis, leading to improved accuracy and robustness.	• Possibly suffer from pre-trained models' limitations, such as over-reliance on training data and limited generalizability.
Bayesian game model	• A novel Bayesian game model is utilized for unsupervised sentiment analysis, leading to improved accuracy. • It can handle large amounts of unstructured data, making it useful for analyzing product reviews.	• Bayesian game models can be complex and computationally demanding, leading to limitations in real-time applications.
Aspect-Oriented Hybrid Approach	• Provides an aspect-oriented hybrid approach that considers both the aspect and the sentiment of students' feedback, leading to improved accuracy and comprehensive analysis.	• The hybrid approach may be complex and computationally demanding, limiting its real-time applications.
Bi-directional Style Enhancement	• Incorporating both text and audio modalities leads to improved sentiment analysis accuracy. • Utilizes a novel bi-directional style enhancement approach, leading to improved performance.	• The addition of audio modalities may increase computational demands, limiting real-time applications. • The bi-directional style enhancement approach may be complex and difficult to implement.
Polar-Vector and Strength-Vector Mixer Model	• Provides a novel approach to multimodal sentiment analysis using polar-vector and strength-vector mixing, leading to improved accuracy and robustness. • It can handle large amounts of unstructured data, making it useful for analyzing product reviews.	• The polar-vector and strength-vector mixing approach may be complex and computationally demanding, limiting real-time applications. • The model may suffer from limitations of pre-trained models, such as over-reliance on training data and limited generalizability.
Bidirectional Encoder Representations	• Utilizes state-of-the-art transformer models for sentiment analysis. • Aids in understanding sentiment in low-resource languages.	• Without further adaptation, it may not be suitable for sentiment analysis in other languages. • The quality of the training data may limit the results.
Prediction of Movie Success	• Incorporates multiple data sources, including Twitter and IMDB, to make predictions. • Aids in understanding the impact of social media on movie success.	• The accuracy of the predictions may be limited by the quality and amount of data used. • It may not generalize well to other domains or products.
Lexicon Enhanced Collaborative Network	• Improves sentiment analysis in the financial domain through lexicon enhancement.	• The quality of the sentiment analysis may be limited by the quality of the lexicon used.
Large Scale Group Decision-Making System	• Can handle large-scale sentiment analysis. • Aids in group decision-making through sentiment analysis.	• The accuracy of the sentiment analysis may be limited by the quality and amount of data used.
Joint Multimodal Sentiment Analysis	• Incorporates multiple modalities in sentiment analysis. • Utilizes information relevance to improve performance.	• The complexity of the model may limit its scalability and generalizability. • The model may not perform well on tasks with limited data.

Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

"Bayesian game model" [12] proposes a new unsupervised SA approach that uses a Bayesian game model. The authors claim that their approach can overcome the limitations of traditional unsupervised methods, such as the lack of labeled data. The results of experiments conducted on a dataset of product reviews showed that their approach outperforms traditional unsupervised methods in terms of accuracy. "Aspect-Oriented Hybrid Approach" [13] proposes a new aspect-oriented hybrid approach for SA of student feedback. The authors claim that their approach can better capture the sentiment expressed in student feedback, as it considers both the overall sentiment and the sentiment expressed towards specific aspects of the course.

"Bi-directional Style Enhancement" [14] proposes a new approach for SA of text and audio data. The authors claim that their approach can better capture the sentiment expressed in audio data as it considers the style of speech. The results of experiments conducted on a text and audio data dataset showed that their approach outperforms other state-of-the-art models in accuracy. "Polar-Vector and Strength-Vector Mixer Model" [15] proposes a new multimodal SA approach that combines polar and strength vectors. The authors claim that their approach can better capture the sentiment expressed in multimodal data, as it considers both the sentiment's polarity and strength. "Bidirectional Encoder Representations" [16] presents a study on SA of Malayalam tweets. The authors claim that their approach, based on bidirectional encoder representations from transformers, can handle the complexities of the Malayalam language and effectively capture the sentiment expressed in the tweets.

"Prediction of Movie Success" [17] proposes a new approach for predicting the success of movies based on machine learning and Twitter SA. The authors claim that their approach can effectively capture the sentiment expressed in Twitter data and use it to make accurate predictions about the success of movies. Optimization [18–26] can be applied in any research domain to achieve the expected best results. "Lexicon Enhanced Collaborative Network" [27] proposes a new approach to financial SA that considers both the sentiment expressed in financial news articles and the opinions of financial experts. The authors claim that their approach, based on a lexicon-enhanced collaborative network, can effectively capture the sentiment expressed in financial news articles and the opinions of financial experts and use them to make accurate predictions about the performance of financial markets.

"Large Scale Group Decision-Making System (LSGDMS)" [28] proposes a new approach to group decision-making that considers the sentiment individuals express. The authors claim that based on SA clustering, their approach can effectively capture the sentiment expressed by individuals in a group and use it to make accurate group decisions. "Joint Multimodal Sentiment Analysis (JMSA)" [29] proposes a new approach to multimodal SA that considers the relevance of the information contained in different modalities. The authors claim that their approach, based on information relevance, can effectively capture the sentiment expressed in multimodal data, as it considers the relevance of the information contained in each modality.

3. SOPHISTICATED FISH SWARM OPTIMIZATION-GUIDED RADIAL BASIS FUNCTION NEURAL NETWORK

3.1 Radial Basis Function Neural Network

Today, multi-layer perceptron (MLP) networks remain the most widely used network topologies across various domains. With a multi-layer perceptron (MLP), the output is calculated as $(i.e., m(\vec{p}) = g_e(n_k + \sum_{s=1}^{Y} N_s P_s$, wherein $\vec{p}$ is an integral gain of dimension $Y, n_s, s = 1,2, \dots, Y$ and it indicates the weights, n_k represents the diagonal weight, and $g_e(k) = \frac{1}{1+h^{-dk}}$, d Back-propagation is then used to train the network. Back-Propagation Neural Networks are a type of MLP that uses this learning method to make predictions (BPNNs). The training pace of BPNNs is a frequent bottleneck (That is, more layers and more neurons per layer result in faster training.). To address the issue of complex neural network topologies, a new approach called radial basis function neural networks (RBFNs) has emerged across various application areas. RBFNs utilize a single hidden layer, making them a simpler alternative.

Let's assume the Multi-Layer Perceptron with Quadratic Inputs (MLPQI) has one hidden layer of C ratio and a linearly activated output unit. For simplicity, this research work will refer to the weights connecting the hidden unit as s, the output unit k as the $n_s(k)$ and the bias value for the output node as $n_k(k)$. Eq.(1) is applied to calculate MLPQI's output.

$$D(\vec{p}) = n_k(k) + \sum_{s=1}^{C} \frac{n_s(k)}{1+\exp(-n_{sk}-\sum Z_{Zsw}p_w - \sum_{w=1}^{Y} N_{Xsw}p_w^2)}, \quad (1)$$

where s, n_{sk} indicates the bias weight for every component, n_{Zsw} represents the weight for linear terms p_w, and n_{Xsw} is the weight for quadratic terms p_w^2. Eq.(2) provides the number of hidden neurons present in RBFN.

$$B(\vec{p}) = \sum_{s=1}^{C} n_s \exp\left(-\gamma(\vec{p}-\vec{\vartheta}_s)^F(\vec{p}-\vec{\vartheta}_s)\right), \quad (2)$$

wherein n_s indicates the unit s, ϑ_s represents the weights center vector, and γ indicates its parameter. The approximation Eq.(2) is mathematically expressed as Eq.(3).

$$B(\vec{p}) = \sum_{s=1}^{C} n_s\left(L/\left(1 + exp\left(-u\gamma(\vec{p}-\vec{\vartheta}_s)^F(\vec{p}-\vec{\vartheta}_s)\right) - y\right)\right) \quad (3)$$

where $L, u,$ and y represent the variables.

3.2 Design of RBFNs

Each neuron's weight is calculated by taking the Euclidean distance present between the evaluator point and the neuron's center by applying a Radial Base Function (RBF) (i.e., kernel function). The radius distance is passed to the RBF, and it acts as a regional receptor of input, and it has output with the condition of how far the input is received from a vector in memory. If the length of the input $\vec{p}$ to the center $\vec{\vartheta}_s$ in each RBF τ_s, i.e., $\|\vec{p}-\vec{\vartheta}_w\|$ is 0, then the contribution of this point is marked as 1, but if the distance $\|\vec{p}-\vec{\vartheta}_w\|$ increases, then the contribution of this point is marked as 0.

RBF networks consist of three individual layers:

(1) Input-layer: Within this layer, the neurons have the ability to employ identical predictor variables. The output values of the neurons in the input layer are then transmitted to the neurons in the hidden layer.

(2) Hidden Layer: The hidden layer of the neural network contains multiple neurons, and selecting the appropriate neurons for this layer can be challenging. The complexity of the hidden layer is determined by the input/output predictor variables. The neurons present in the hidden layer consist of an RBF centered at a threshold point, and it is based on the parameters $\{\vec{\vartheta}_w, \in_w\}$ and the input activations $\{\vec{p}\}$. Eq.(4) provide the activations in the hidden units $\tau_s(\vec{p}, \vec{\vartheta}_w, \in_w)$.

$$\tau_s(\vec{p}) = \exp\left(-\frac{\|\vec{p}-\vec{\vartheta}_w\|^2}{2\in_w^2}\right). \quad (4)$$

The RBF's radius varies between each dimension. The network's training procedure establishes the ranges and centers.

(3) **Summation Layer:** Each neuron present in this has $(n_1, n_2, \ldots. n_T)$ weights. Each hidden layer neuron's output is multiplied by its associated weight before being sent to the summation layer, where the values are added together and presented as the network's output. This layer receives a bias value compounded by a weight n_k.

When Gaussian functions are used, it is important to note that RBFNs are linked to the Gaussian Mixture Model (GMM). By substituting $T.D$ for C, i.e., in Eq.(2), this research describes RBFNs with T hidden units and a neuron in the output nodes D. The density of a Gaussian mixture is calculated as the weighted sum of the densities of its components using Eq.(5).

$$j(\vec{p}|\ni) = \sum_{s=1}^{T} d_s v_s(\vec{p}),\ where\ \sum_{s=1}^{T} d_s = 1.$$

$$v_s(\vec{p}) = \frac{1}{(2\pi)^{\frac{y}{2}}|\Sigma_s|^{\frac{1}{2}}} \quad (5)$$

$$exp\left(-1/2(\vec{p}-\vec{\vartheta})^F \sum_s^{-1}(\vec{p}-\vec{\vartheta})\right)$$

where the Covariance matrix is indicated as Σ_s and mean vector is represented as $\vec{\vartheta}_s$.

3.3 Training RBFNs

When a neural network is trained, it learns to provide the intended response to a given stimulus by adjusting its parameters in response to that signal. To achieve the same, RBFNs' training procedure picks the following parameters:

1. The total amount of neurons presents in the system's inner layer. i.e., it should hold neurons in the hidden units (C) with the specific data points (T).
2. During the training process, each hidden-layer RBF's center is located.
3. The training procedure computes each RBF's dimensionally-specific radius.
4. The weights are used in multiplying RBF outputs before sending them to the summation layer.

3.4 RBFN Kernels

It is common to practice using the Gaussian kernel when constructing kernel functions. The most general version of a Gaussian kernel is represented by Eq.(5). Each kernel is given a t-dimensional ellipsoid form based on the inversion of the covariance matrix, which captures the correlations between the various variables. It's more flexible than the traditional approach of utilizing a simple distance to the kernel centroid, which requires an assumption of complete independence of all variables. The application of the inverse, however, might lead to nonsensical findings or extreme numerical instability if the covariance is unique or very badly conditioned. Consequently, the covariance matrix may be decomposed, yielding the Eigensystem $\Sigma = M\Lambda M^f$, where M is the eigenvector matrix and Λ is the eigenvalue of the associated diagonal matrix.

$$\tau_s(\vec{p}) = \frac{1}{\left(1 + \frac{\|\vec{p} - \vec{\vartheta}_w\|}{\epsilon_w^2}\right)} \qquad (6)$$

$$\tau_s(\vec{p}) = \frac{1}{\left(1 + \frac{\|\vec{p} - \vec{\vartheta}_w\|}{\epsilon_w^2}\right)^{1/2}}. \qquad (7)$$

Eq.(6) and Eq.(7) provide activations that are utilized in the center designs of the RBF, and these activations exhibit larger magnitudes than those of equation (4). This implies that Eq.(6) and Eq.(7) have broader applicability in comparison to equation (4). Furthermore, the solutions for Eq.(4), Eq.(6), and Eq.(7) do not approach zero asymptotically. To supplement the standard Gaussian RBF, this research has introduced x-Gaussian RBFs. An expression for the x-Gaussian RBF of the s^{th} unit is mathematically expressed as Eq.(8).

$$\tau_w(\vec{p}) = \begin{cases} \left(1 - (1-x)\left(\frac{\|\vec{p} - \vec{\vartheta}_w\|^2}{\epsilon_w^2}\right)\right)^{1/1-x}, & if\ 1 - (1-x))\left(\frac{\|\vec{p} - \vec{\vartheta}_w\|^2}{\epsilon_w^2}\right) \geq 0 \\ 0 & otherwise \end{cases} \qquad (8)$$

where x indicates a constant cum threshold number. Using the x-Gaussian RBF, this research expresses a wide variety of RBFs by adjusting the value of a new parameter, x.

3.5 Kernel Parameter and Weight Learning

One of the main advantages of RBF networks is the ability to select the appropriate parameters for the hidden unit and basis function without needing to perform a complete non-linear optimization of the entire network. Several unsupervised techniques can be applied to determine the center coordinates of the hidden layer in an RBF function.

3.5.1 Randomly chosen fixed centers

The centers are fixed at C locations chosen randomly from the T data points. This method for establishing the RBF parameters is quick for processing the classification. In particular, this research work employs normalized RBFs using the $\{\vartheta_w\}$ as the center, it is mathematically expressed as Eq.(9).

$$\tau_s(p) = exp\left(-\frac{\|p - \vartheta_w\|^2}{2\ \epsilon_w^2}\right), \qquad (9)$$

Where $\{\vartheta_w\} \subset \{P^m\}, \epsilon_w$ is spread.

3.5.2 Clustering

Using clustering methods, it is possible to identify a group of centers that represents the data points distribution. To minimize the mean square error, firstly, this research work chooses the number of centers A and it employs a straightforward estimate method to partition the data points $\{P^m\}$ into A separate sub-sets E_w and T_w data points. Clustering is mathematically expressed as Eq.(10).

$$W = \sum_{w=1}^{A} \sum_{m \, \omega \, E_w} \|P^m - \vartheta_w\|^2 \qquad (10)$$

wherein ϑ_w represents the center of the E_w data points, and it is expressed as Eq.(11).

$$\vartheta_w = \frac{1}{T_w} \sum_{m \in E_w} P^m. \qquad (11)$$

K-means have two inherent drawbacks that make them less desirable.

- the iterative nature of the method might cause lengthy convergence periods
- the number of RBF centers has to be manually determined through a laborious method of learning through doing.

The alternative strategy is constructive, where the hidden layer is built up in stages until some requirement is fulfilled.

The RBF solution may be written in terms of this orthogonal representation as

$$F = \Phi\mathrm{N} = \mathrm{XJ} \qquad (12)$$

Weight vector J has the least square solution in orthogonal coordinates, and Eq.(13) provides the same.

$$J = (X'X)^{-1}X'F. \qquad (13)$$

3.6 Using RBFNs for Multi-Objective Optimization

There is typically an issue with the complexity of the model used by an RBFN. A sophisticated model can provide a good match to the data input. Because of its low distortion error and large variance error, the model's ability to generalize poorly. In addition, a complicated model is not preferable because of the difficulty it presents to its users. Yet, variance error is reduced when a model is straightforward. Using an RBFN model can help avoid over-corrections and overfitting. However, automatically optimizing RBFN models for a given dataset poses a challenge as it involves balancing two conflicting objectives. The first objective is to generate a reliable classifier or interpolator/regressor using the minimum number of feasible RBFs, while the second objective is to minimize the model's prediction error. This situation is referred to as the "multi-criteria optimization issue" in RBFNs. These two objectives are incompatible because improving one would have a detrimental effect on the other. Recent research has focused on addressing this challenge of simultaneous optimization for both objectives. Typically, their solutions are sub-optimal for each target individually but acceptable when all objectives are combined, where acceptable is highly relative and problem-specific.

The following is a generic statement of a multi-objective optimization problem:

$$\text{Minimize/Maximize } g_c(p), m = 1,2, \ldots. C;$$
$$\text{Sub.to } j_w(p) \geq 0, w = 1,2, \ldots. W;$$
$$l_a(p) = 0 \; a = 1,2, \ldots. A;$$
$$p_s^{(Z)} \leq p_s \leq p_s^{(O)} \; s = 1,2, \ldots. c.$$

p is a solution to the equation $p = (p_1, p_2, \ldots, p_t)^F$, where p is a vertex of t choice variables. Variable boundaries are the last set of restrictions, limiting the range of values that every decision variable p_s can take to between $p_s^{(Z)}$ and $p_s^{(O)}$. The decision space is defined by these constraints and has the dimension Y. Constraints W in inequality and A inequality are part of the problem, and their associated notations, $j_w(p)$ and $l_a(a)$, are known as constraint functions.

It is possible to minimize or maximize all of the C objective functions in multi-objective optimization, which are denoted by $g(p) = \left(g_1(p), g_2(p), \ldots. g_C(p)\right)^F$. Several optimization methods, particularly RBFNs, are designed to address just a certain class of optimization issues, such as minimization issues. Applying the duality principle, this research charges a maximization objective into a minimization objective when the algorithm calls for maximizing. It is important to remember that there is a location in the object plane designated by the equation $(p) = i = (i_1, i_2, \ldots, i_C)^F$ for every answer p in the choice variable space. In multi-objective optimization, this research work will look at two things:

a) Locate a collection of solutions that is as near to the Pareto-optimum as feasible, and
b) Find a group of answers that is as varied as possible. The choice variable and goal space must be explored during multi-objective optimization.

Although a unique mapping connects these two spaces, the mapping is generally non-linear, and the multiple search spaces do not share identical attributes. Every optimization method uses the space of choice variables to conduct its search. A decision variable space algorithm's actions can be tracked in an objective space. Some algorithms use the actions taken in the goal space to direct the search for the decision variables. When this occurs, efforts in both domains must be synchronized such that the expansion of options available in the space of choice variables supplement the variety required in the domain of objectives.

Most multi-objective optimization algorithms employ the concept of dominance to compare two solutions and determine which one is more dominant. In our formulations, we characterize the idea of dominance in terms of the objective functions that need to be minimized without compromising on generality.

3.7 Multi-Objective Methods to Problem Resolution in RBFNs

A multi-objective problem can be treated as a single optimization problem using several classical techniques, such as the weighting approach, the σ-constraint method, graded metric methods, the weighted sum method, and main objective programming techniques. The weighted sum technique integrates many objectives into a unified one. The optimization of this combined objective leads to the optimization of the constituent objective functions. The optimization challenges these transformation techniques create requires an algorithm that can only optimize one thing.

To achieve the enhanced cum best classification accuracy with a big dataset like Amazon Product Review Dataset, a bio-inspired optimization strategy is needed. This research work attempts to utilize the enhanced version of fish swarm optimization to enhance the performance of the RBFNN.

3.8. Fish Swarm Optimization

The bionic algorithm, namely the artificial fish swarm algorithm, models its behavior from natural schooling in aquatic ecosystems. The technique mimics fish swarm feeding, rear-end collision, clustering, and other approaches to accomplish optimum target search and global optimization via local optimization. It models the school of fish's foraging and cooperative behavior with one another to maximize efficiency by simulating the fish's natural mechanisms for survival and coordination. The fish swarm appears to have superior evasion and search abilities based on current events. The fish swarm approach may be used to optimize coverage in SA without prior knowledge of the objective function value or gradient. It exhibits a specific adaptive behavior in the search space. Neither its starting value nor its parameters are really important. While this is the case, it does have some real-world consequences.

Optimization using the artificial fish swarm approach is mostly performed iteratively. This research work uses the fish school methodology to optimize SA, with an object-oriented processing approach and a huge class consisting of member variables and other variables standing in for simulating the real fish. Parameters in this research work primarily consist of the P representation of the fish's (i) state, (ii) step length, (iii) crowding factor, (iv) perceived length, and (v) several trials. The significant member functions are the fish's present location, food concentration, and behavior. After the system settings are finalized, a school of artificial fish is in the initialization stage, where each fish's information acts as a label. Other fish can detect unique emotional states. Hence, there are T artificial fish in a particular t -dimensional search space and the state of the artificial fish is represented as Eq.(14).

$$P = \{P_1, P_2, \ldots\ldots, P_t\} \tag{14}$$

To be specific, the parameter to be optimized is $P_s = (1,2, \ldots., t)$. Artificial fish now have a threshold food concentration mathematically expressed as Eq.(15).

$$Q = g(p) \tag{15}$$

The goal function, denoted by Q in Eq.(2), is the distance between artificial fishes. Eq.(16) represents the same.

$$y_{s,w} = \|p_s - p_w\| \tag{16}$$

Individual fishes act as a unique solution to achieve the objective, but all solutions may not be the best. Using the state in Eq.(15), this research work evaluates how artificial fish is doing a favor to others fishes in the swarm.

3.9. Sophisticated Fish Swarm Optimization (SFSO)

SFSO is inherited from fish colony optimization, and it involves five different phases, which are: (i) Foraging Actions, (ii) Social Activity,

(iii) Rear-End Action, (iv) Individual Actions, and (v) Board of Bulletins.

3.9.1 Foraging actions

SFSO is carried out by modeling the schooling behavior of fishes in the swarm. During their foraging, fish schools rely heavily on visual cues of object density to guide their collective actions. In particular, it may replicate the process and configure the associated parameters, such that it swims in the direction that schools of fish tend to migrate towards if numerous objects are found during the movement of the school. The concentration of food is the primary factor in directing the school of fish to swim ahead if the present condition of the school can be described as P_s. If there is a better concentration of food somewhere, then the school will move towards that. A new random state, P_w, is instead selected by the school of fish. There will be no stopping the search until the advance condition is met if the iteration number is less than the forward condition. Parameter formulae are provided in Eq.(17).:

$$P_w = P_s + Visual^{*} Rand\,() \tag{17}$$

$Rand()$ in Eq.(4) is a random integer among 0 and 1. If $Q_w > Q_s$ during the maximum value solution (and if $Q_w < Q_s$ during the lowest value solution), the school of fish will progress in that direction using Eq.(18).

$$P_s^{f+1} = P_s^f + \frac{P_w - P_s^f}{\left\|P_w - P_s^f\right\|} * Step^{*} Rand\,(\,) \tag{18}$$

The fish will carry out random movements if the next state that meets the requirements has not been discovered previously. Eq.(19) symbolizes the random movements of fishes.

$$P_s^{f+1} = P_s^f + Visual^{*} Rand\,(\,) \tag{19}$$

3.9.2 Social activity

School of Fishes forms naturally in the ocean. By staying in the middle of the group, the fish in the school engage with each other and reduce the risk at different levels. As a result, the aggregation mode may be modeled using a unique state. To prevent fish from becoming suffocated, the SFSO is described as the interaction between individual fishes at various stages of execution.

If the fishes in SFSO is in a state represented by P_s, then any additional individuals inside its field of view would be represented by t_g. It is mathematically represented as $y_{s,w} < Visual$, The current state of the center is indicated as P_u. When $Q_u/t_g > Q_s$ is met, then there is a plentiful supply of food in the mid-area of the swarm, and that area is not overcrowded. The next step is to move over to the neighboring fish region. Foraging behavior is maintained if this condition is not satisfied. Eq.(20) expresses the same procedure.

$$P_s^{f+1} = P_s^f + \frac{P_u - Pf_s}{\|P_u - Pf_s\|} * Step^{*} Rand\,(\,) \tag{20}$$

3.9.3 Rear-end action

Rear-end behavior occurs in a school of fish when t fish discover food simultaneously, and the surrounding fish swim to where the food is. In most cases, the tailing habit manifests as a lagging rear end. This state of affairs in the system suggests that every node is looking for the best path forward; equivalently, every fish desires to swim in the best possible direction. If the fish is in a condition of P_s at the moment, and that the highest concentration of food it can see is P_{max}, which is how the other fish are in its immediate vicinity. Unless the conditions $P_{max} > Q_s$ and $\frac{Q_{max}}{t_g} > \theta Q_s$ are met, then the fish will engage in foraging behavior.

The neighbor fish with the highest food concentration, P_{nax}, is moved towards. Specifically, the mathematical model is given as Eq.(21).

$$P_s^{f+1} = P_s^f + Step^{*} Rand\,(\,)^{*} \frac{P_{max} - P_s}{\|P_{max} - P_s\|} \tag{21}$$

3.9.4 Individual actions

Fishes move with an element of unpredictability, which helps the fish as a whole look for food. Fish will randomly choose a recognized condition and take one stride in that direction. A widespread foraging habit leads the fish movement process with smaller activities.

3.9.5 Board of bulletins

This phase keeps a historical record of the school of fishes (i.e., the peak density) or the point at which the greatest number of individuals are concentrated within the group throughout its motion. The operational data of fishes will be saved and compared with data from other fishes in the swarm. If the actual food concentration acquired is greater than the old one, the old one is set aside, and the new one is used instead.

3.10 SFSO PROCESS

To carry out the SFSO, the following phases must be followed:

(1) Set the algorithm parameters first, which involves establishing parameters like the size of the fish school and the step size, the maximal visual field range, the number of iterations, and the congestion factor. In other words, make a fish school.
(2) Various fish (i.e., artificial fishes) will be created depending on the settings. The amount of iterations is then determined after initiating the fish fleet.
(3) The post-depuration food concentration of fish is determined, and relevant parameters are posted.
(4) Model the fish foraging patterns and the distribution of food concentrations.
(5) Decide whether or not to continue the process based on food availability in the immediate area.
(6) The ideal record is posted on the bulletin once the fish have exhibited the corresponding behavior.
(7) Afterward, the food content is compared to the previous value, and if the new value is greater, the previous one is neglected, and the new one is written in.
(8) The iteration count will stop at a maximum, and the bulletin value will not be considered if the limit is reached.
(9) Iteration proceeds if the maximum amount of iterations is not reached.

The sentiments in any domain may be stated as $U = \{u_1, u_2, \dots . u_T\}$ if T sentiments are randomly distributed over a monitoring area, their initial basic information is identical, their position in the dataset is well known, and their monitoring radius is b. The formula $u_s = (p, q, b)$ yields the result of a circle with a radius of b. As there is no method to determine the optimal size for monitoring, this research work divides the region into $c \times t$ pixels. The label denotes this event b_s, and it has a probability of $M\{b_s\}$, which is mathematically expressed as Eq.(22).

$$M\{b_s\} = M_{cos}(p, q, u_s) = \begin{cases} 1, if & (p - p_s)^2 + (q - q_s)^2 \leq b^2 \\ 0, & otherwise \end{cases} \quad (22)$$

This research work assumes (p, q) is served by the s sentiments if the distance between (p, q) and the node s is smaller than a predetermined transmission range b. Eq.(23) expresses the same.

$$M\{\bar{b}_s\} = 1 - M\{b_s\} = 1 - M_{cov}(p, q, u_s) \quad (23)$$

If b_s and b_w have no behavior, then Eq.(24) yields the relationship between the swarm.

$$M\{b_s \cup b_w\} = 1 - M\{\bar{b}_s \cap \bar{b}_w\} = 1 - M\{\bar{b}_s\}. M\{\bar{b}_w\} \quad (24)$$

The likelihood that the sentiment set includes the (p, q) is the union of b_s. Assuming that each independent random event b_s occurs, the cover of a node-set U is expressed in Eq.(25).

$$M_{cov}(p, q, u_s) = M\left\{\bigcup_{s=1}^{T} b_s\right\} = 1 - M\left\{\bigcap_{s=1}^{T} \bar{b}_s\right\} = 1 - \prod_{s=1}^{T}\left(1 - M_{cov}(p, q, u_s)\right) \quad (25)$$

4. ABOUT THE DATASET

Amazon product review dataset is a large collection of customer reviews and ratings for various products sold on Amazon.com. It is a publicly available dataset that can be used for research and analysis purposes. The dataset contains millions of reviews and ratings for products in various categories, including electronics, books, clothing, and more. The features of the Amazon product review dataset include the following [30, 31]:

- **reviewerID:** This field contains a unique identifier for each reviewer who has provided a review for a product. It may be a string of alphanumeric characters that allow the review to be attributed to a specific person.
- **asin:** This field contains a unique identifier for the product being reviewed. It may also be a string of alphanumeric characters that allow the review to be linked to a specific product.

- **reviewerName:** This field contains the name of the reviewer who wrote the review. It is usually a string of characters and may or may not be the reviewer's real name.
- **vote:** This field indicates the number of helpful votes that the review has received. It is a numerical value and helps other users to evaluate the quality and usefulness of the review.
- **style:** This field is a dictionary that contains metadata information about the product being reviewed, such as its format (e.g., Hardcover, Paperback), color, or size.
- **reviewText:** This field contains the text of the review written by the reviewer. It is usually a longer piece of text that describes the reviewer's experience with the product.
- **overall:** This field contains the rating given to the product by the reviewer. It is usually a numerical value that ranges from 1 to 5, with 5 being the highest rating.
- **summary:** This field contains a brief summary of the review. It is usually a shorter piece of text that summarizes the reviewer's opinion of the product.
- **unixReviewTime:** This field contains the time of the review in Unix format. Unix time is the number of seconds that have elapsed since January 1, 1970, at 00:00:00 UTC.
- **reviewTime:** This field contains the time of the review in a human-readable format. It may include the date, time, and time zone information.

- **image:** This field contains images that users have posted after receiving the product. It may be a string of characters that represents the URL or file location of the image.

In addition to these features, the dataset includes metadata about the products, such as the brand, model number, and price. This information can be used to perform a more detailed analysis of the reviews, such as comparing ratings and sentiment across different brands or price ranges. The Amazon product review dataset is valuable for researchers and businesses seeking insights into customer opinions and preferences for a wide range of products.

5. PERFORMANCE METRICS

This research uses the below-mentioned performance metris to measure the performance of the proposed classifier against the existing classifier.

- Classification Accuracy (CA) is a metric used to evaluate the performance of a classification model in terms of the proportion of correct predictions made by the model.
- F-Measure (FM) is a metric used to evaluate a classification model's performance in terms of precision and recall.
- Fowlkes-Mallows Index (FMI) measures the similarity between two clustering results from the same data set.
- Matthews Correlation Coefficient (MCC) is a statistical measure that assesses the performance of binary classification models. It takes into account the true positives, true negatives, false positives, and false negatives to provide a measure of the quality of the model.

The four significant variables used in calculating the four-performance metrics mentioned above are defined below.

- True Positive: A sentiment analysis system correctly identifies a positive sentiment in a text that expresses a positive sentiment.
- True Negative refers to a correctly predicted instance where the sentiment of a text is negative, and the model accurately identifies it as negative.
- False Positive in sentiment analysis refers to an incorrect prediction where the model predicts a text as having negative sentiment, but it is not.
- False Negative in sentiment analysis refers to a situation where the model incorrectly predicts a text with negative sentiment as neutral or positive.

6. RESULTS AND DISCUSSION

6.1 CA and FM Analysis

Figure 1 compares classification accuracy and f-measure among three sentiment analysis classifiers, LSGDMS, JMSA, and SFSO-RBFNN. The x-axis represents the classifier names, while the y-axis represents the classification accuracy and f-measure values. The graph shows that SFSO-RBFNN outperforms the other two classifiers with a classification accuracy of 80.135% and an f-measure of 82.860%. JMSA is the second-best classifier, with a classification accuracy of 61.424% and an f-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

measure of 63.285%. LSGDMS has the lowest performance among the three classifiers, with a classification accuracy of 52.366% and an f-measure of 51.744%.

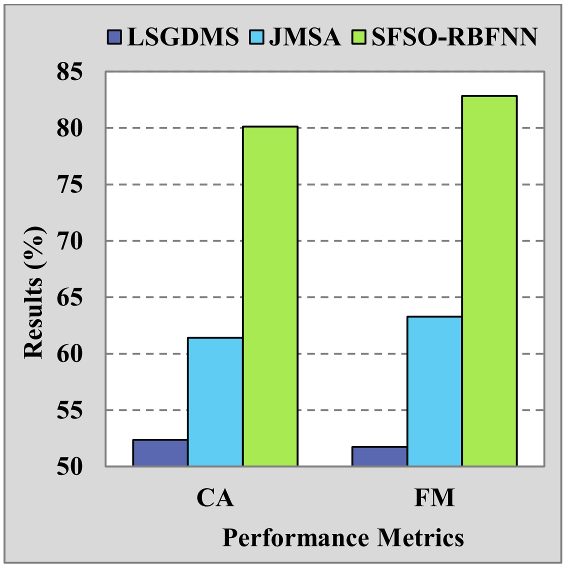

Figure 1. CA and FM Analysis

Table 2. CA and FM Result Values

	LSGDMS	JMSA	SFSO-RBFNN
CA	52.366	61.424	80.135
FM	51.744	63.285	82.860

The graph highlights the effectiveness of SFSO-RBFNN as a proposed sentiment analysis classifier, which utilizes a sophisticated fish swarm optimization algorithm to guide a radial basis function neural network. The classifier's performance is significantly higher than the state-of-the-art classifiers, indicating its potential to be used in real-world sentiment analysis applications. The graph also suggests that there is still room for improvement in sentiment analysis classifiers' performance. The lower performance of LSGDMS and JMSA indicates a need for more advanced sentiment analysis techniques and algorithms to enhance the classifiers' performance. It is important to note that the performance comparison in the graph is based on a particular dataset and evaluation metrics. Therefore, the results may vary when evaluated on different datasets and using different evaluation metrics. Further evaluation of diverse datasets is required to validate the classifiers' effectiveness and applicability. Overall, the graph provides valuable insights into the performance comparison of sentiment analysis classifiers and can be used as a benchmark for future research in sentiment analysis. The performance comparison can inspire new research directions and encourage researchers to develop advanced sentiment analysis techniques and algorithms to improve the classifiers' accuracy and reliability. The result values of Figure 1 are provided in Table 2. Domain Wise results are provided in Table 3.

Table 3. Domain-wise CA and FM Result Values

Domain of the Dataset	Performance Metrics	LSGDMS	JMSA	SFSO-RBFNN
Book	CA	52.449	61.792	80.292
	FM	51.736	63.668	83.010
Kindle Store	CA	51.625	57.891	78.400
	FM	51.946	59.921	81.531
Tools and Home Improvement	CA	51.914	60.328	79.876
	FM	51.454	61.830	82.261
Industrial and Scientific	CA	56.590	63.027	81.766
	FM	56.459	64.014	84.343

6.2 FMI and MCC Analysis

The graph compares Fowlkes-Mallows Index (FMI) and Matthews Correlation Coefficient (MCC) values among three sentiment analysis classifiers, LSGDMS, JMSA, and SFSO-RBFNN. The x-axis represents the classifier names, while the y-axis represents the FMI and MCC values. The results indicate that SFSO-RBFNN outperforms the other two classifiers in both FMI and MCC. SFSO-RBFNN achieves an FMI score of 82.864%, while JMSA is the second-best classifier with an FMI score of 63.286%, and LSGDMS has the lowest performance with an FMI score of 51.747%. Similarly, SFSO-RBFNN achieves an MCC score of 59.259%, while JMSA is the second-best classifier

Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

with an MCC score of 22.649%, and LSGDMS has the lowest performance with an MCC score of 4.727%.

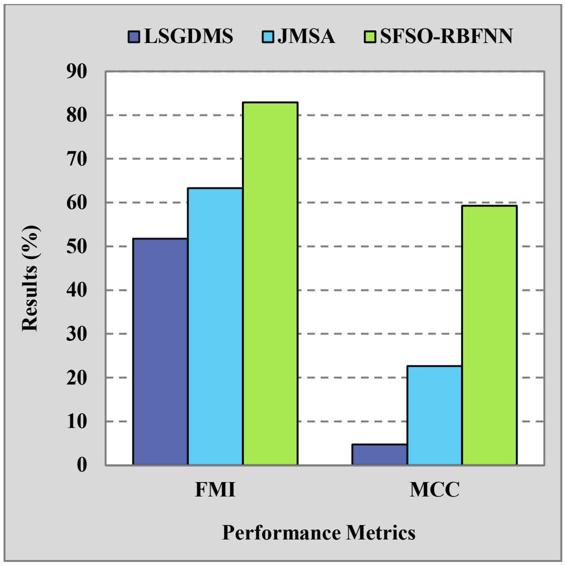

Figure 1. FMI and MCC Analysis

Table 4. FMI and MCC Result Values

	LSGDMS	JMSA	SFSO-RBFNN
FMI	51.747	63.286	82.864
MCC	4.727	22.649	59.259

The graph highlights the effectiveness of SFSO-RBFNN in sentiment analysis, which utilizes a sophisticated fish swarm optimization algorithm to guide a radial basis function neural network. The classifier's performance is significantly higher than the state-of-the-art classifiers, indicating its potential to be used in real-world sentiment analysis applications. The results also demonstrate that the FMI and MCC evaluation metrics effectively measure the performance of sentiment analysis classifiers. The FMI measures the similarity between the predicted and true clustering, while the MCC measures the correlation between the predicted and true labels. Both metrics comprehensively evaluate the classifiers' performance, considering both false positives and false negatives. It is important to note that the performance comparison in the graph is based on a particular dataset and evaluation metrics. Therefore, the results may vary when evaluated on different datasets and using different evaluation metrics. Further evaluation of diverse datasets is required to validate the classifiers' effectiveness and applicability.

Overall, the graph provides valuable insights into the performance comparison of sentiment analysis classifiers and can be used as a benchmark for future research in sentiment analysis. The performance comparison can inspire new research directions and encourage researchers to develop advanced sentiment analysis techniques and algorithms to improve the classifiers' accuracy and reliability. The result values of Figure 2 are provided in Table 4. Domain Wise results are provided in Table 5.

Table 5. Domain-wise FMI and MCC Result Values

Domain of the Dataset	Performance Metrics	LSGDMS	JMSA	SFSO-RBFNN
Book	FMI	51.738	63.668	83.013
	MCC	4.885	23.380	59.566
Kindle Store	FMI	51.958	59.923	81.537
	MCC	3.291	15.574	55.551
Tools and Home Improvement	FMI	51.462	61.838	82.272
	MCC	3.853	20.564	59.065
Industrial and Scientific	FMI	56.461	64.041	84.348
	MCC	13.187	26.109	62.544

7. CONCLUSION

This research has proposed a novel approach for sentiment analysis of Amazon product reviews: sophisticated fish swarm optimization-guided radial basis function neural network (SFSO-RBFNN). The SFSO-RBFNN model was trained using the SFSO algorithm and RBFNN, which helped optimize the model's weights and biases for improved accuracy. Our experimental results demonstrated that the SFSO-RBFNN approach outperformed the state-of-the-art sentiment analysis classifiers, LSGDMS and JMSA, with a classification accuracy of 80.135%. The high accuracy achieved by the model suggests that it has the potential to extract useful insights from large-scale data in various domains, including marketing, finance, and social sciences. Furthermore, the proposed SFSO-RBFNN approach could be applied to other datasets and domains beyond Amazon product reviews. The optimization capabilities of the SFSO algorithm could be useful in enhancing the performance of other machine learning models, and its integration with RBFNN could improve the generalization ability of the models. Overall, the results of this study suggest that the SFSO-RBFNN approach is a promising approach for sentiment analysis of Amazon product reviews. Future research could explore the scalability of this method for large-scale datasets and investigate its potential in real-world applications.

REFERENCES

[1]. Zuo, M., Diao, L., Liu, Q., Wang, P.: Data Mining strategies and techniques of Internet education public sentiment monitoring and analysis system. In: Proceedings of the 2010 2nd International Conference on Future Computer and Communication, ICFCC 2010 (2010). https://doi.org/10.1109/ICFCC.2010.5497355.

[2]. Guimarães, R.G., Rosa, R.L., De Gaetano, D., Rodríguez, D.Z., Bressan, G.: Age Groups Classification in Social Network Using Deep Learning. IEEE Access. 5, 10805–10816 (2017). https://doi.org/10.1109/ACCESS.2017.2706674.

[3]. Wang, W., Pan, S.J., Dahlmeier, D.: Memory networks for fine-grained opinion mining. Artif. Intell. 265, 1–17 (2018). https://doi.org/10.1016/j.artint.2018.09.002.

[4]. Badache, I., Fournier, S., Chifu, A.-G.: Detection of contradiction in reviews | Détection de contradiction dans les commentaires. In: COnference en Recherche d'Informations et Applications, CORIA 2017 - 14th French Information Retrieval Conference, Proceedings. pp. 75–90 (2020).

[5]. Shapiro, H.B., Lee, C.H., Wyman Roth, N.E., Li, K., Çetinkaya-Rundel, M., Canelas, D.A.: Understanding the massive open online course (MOOC) student experience: An examination of attitudes, motivations, and barriers. Comput. Educ. 110, 35–50 (2017). https://doi.org/10.1016/j.compedu.2017.03.003

[6]. Alrajhi, L., Alharbi, K., Cristea, A.I.: A multidimensional deep learner model of urgent instructor intervention need in MOOC forum posts. (2020). https://doi.org/10.1007/978-3-030-49663-0_27.

[7]. Min, L., Miao, X., Bi, P., He, F.: A Small amount of Labeled Data Chinese Online Course Review Target Extraction via ALBERT-IDCNN-CRF Model. In: Journal of Physics: Conference Series (2020). https://doi.org/10.1088/1742-6596/1651/1/012049.

[8]. 8. Kaur, A., Sharma, R., Gulati, S., Chakravarty, R.: Mobile app analytics for assessing MOOC platforms: a study of online learners' sentiments. Libr. Hi Tech News. 39, 5–6 (2022). https://doi.org/10.1108/LHTN-11-2021-0086.

[9]. Dai, Y., Xiong, H., Liu, C.: Research on the Current Situation and Countermeasures of Online Foreign Language Courses Based on Course Reviews. In: ACM International Conference Proceeding Series. pp. 47–53 (2022). https://doi.org/10.1145/3524383.3524390.

[10]. Garg, K.: Sentiment analysis of Indian PM's "Mann Ki Baat." Int. J. Inf. Technol. 12, 37–48 (2020). https://doi.org/10.1007/s41870-019-00324-8.

[11]. Kim, K., Park, S.: AOBERT: All-modalities-in-One BERT for multimodal sentiment analysis. Inf. Fusion. 92, 37–45 (2023). https://doi.org/10.1016/j.inffus.2022.11.022.

[12]. Punetha, N., Jain, G.: Bayesian game model based unsupervised sentiment analysis of product reviews. Expert Syst. Appl. 214, (2023). https://doi.org/10.1016/j.eswa.2022.119128.

[13]. Kathuria, A., Gupta, A., Singla, R.K.: AOH-Senti: Aspect-Oriented Hybrid Approach to Sentiment Analysis of Students' Feedback. SN Comput. Sci. 4, (2023). https://doi.org/10.1007/s42979-022-01611-1.

[14]. Lin, F., Liu, S., Zhang, C., Fan, J., Wu, Z.: StyleBERT: Text-audio sentiment analysis with Bi-directional Style Enhancement. Inf. Syst. 114, (2023). https://doi.org/10.1016/j.is.2022.102147.

[15]. Lin, H., Zhang, P., Ling, J., Yang, Z., Lee, L.K., Liu, W.: PS-Mixer: A Polar-Vector and

Strength-Vector Mixer Model for Multimodal Sentiment Analysis. Inf. Process. Manag. 60, (2023). https://doi.org/10.1016/j.ipm.2022.103229.

[16]. Elankath, S.M., Ramamirtham, S.: Sentiment analysis of Malayalam tweets using bidirectional encoder representations from transformers: a study. Indones. J. Electr. Eng. Comput. Sci. 29, 1817–1826 (2023). https://doi.org/10.11591/ijeecs.v29.i3.pp1817-1826.

[17]. Tripathi, J., Tiwari, S., Saini, A., Kumari, S.: Prediction of movie success based on machine learning and twitter sentiment analysis using internet movie database data. Indones. J. Electr. Eng. Comput. Sci. 29, 1750–1757 (2023). https://doi.org/10.11591/ijeecs.v29.i3.pp1750-1757.

[18]. Ramkumar, J., Vadivel, R.: Whale optimization routing protocol for minimizing energy consumption in cognitive radio wireless sensor network. Int. J. Comput. Networks Appl. 8, 455–464 (2021). https://doi.org/10.22247/ijcna/2021/209711.

[19]. Ramkumar, J., Vadivel, R.: Performance Modeling of Bio-Inspired Routing Protocols in Cognitive Radio Ad Hoc Network to Reduce End-to-End Delay. Int. J. Intell. Eng. Syst. 12, 221–231 (2019). https://doi.org/10.22266/ijies2019.0228.22.

[20]. Jaganathan, R., Ramasamy, V., Mani, L., Balakrishnan, N.: Diligence Eagle Optimization Protocol for Secure Routing (DEOPSR) in Cloud-Based Wireless Sensor Network. (2022). https://doi.org/10.21203/rs.3.rs-1759040/v1.

[21]. Ramkumar, J., Vadivel, R.: CSIP—cuckoo search inspired protocol for routing in cognitive radio ad hoc networks. In: Advances in Intelligent Systems and Computing. pp. 145–153. Springer Verlag (2017). https://doi.org/10.1007/978-981-10-3874-7_14.

[22]. Ramkumar, J., Vadivel, R., Narasimhan, B.: Constrained Cuckoo Search Optimization Based Protocol for Routing in Cloud Network. Int. J. Comput. Networks Appl. 8, 795–803 (2021). https://doi.org/10.22247/ijcna/2021/210727.

[23]. Ramkumar, J., Vadivel, R.: Improved frog leap inspired protocol (IFLIP) – for routing in cognitive radio ad hoc networks (CRAHN). World J. Eng. 15, 306–311 (2018). https://doi.org/10.1108/WJE-08-2017-0260.

[24]. Jaganathan, R., Vadivel, R.: Intelligent Fish Swarm Inspired Protocol (IFSIP) for Dynamic Ideal Routing in Cognitive Radio Ad-Hoc Networks. Int. J. Comput. Digit. Syst. 10, 1063–1074 (2021). https://doi.org/10.12785/ijcds/100196.

[25]. Vadivel, R., Ramkumar, J.: QoS-enabled improved cuckoo search-inspired protocol (ICSIP) for IoT-based healthcare applications. Inc. Internet Things Healthc. Appl. Wearable Devices. 109–121 (2019). https://doi.org/10.4018/978-1-7998-1090-2.ch006.

[26]. Ramkumar, J., Vadivel, R.: Multi-Adaptive Routing Protocol for Internet of Things based Ad-hoc Networks. Wirel. Pers. Commun. 120, 887–909 (2021). https://doi.org/10.1007/s11277-021-08495-z.

[27]. Shang, L., Xi, H., Hua, J., Tang, H., Zhou, J.: A Lexicon Enhanced Collaborative Network for targeted financial sentiment analysis. Inf. Process. Manag. 60, (2023). https://doi.org/10.1016/j.ipm.2022.103187.

[28]. Trillo, J.R., Herrera-Viedma, E., Morente-Molinera, J.A., Cabrerizo, F.J.: A large scale group decision making system based on sentiment analysis cluster. Inf. Fusion. 91, 633–643 (2023). https://doi.org/10.1016/j.inffus.2022.11.009.

[29]. Chen, D., Su, W., Wu, P., Hua, B.: Joint multimodal sentiment analysis based on information relevance. Inf. Process. Manag. 60, (2023). https://doi.org/10.1016/j.ipm.2022.103193.

[30]. Menakadevi, P., Ramkumar, J.: Robust Optimization Based Extreme Learning Machine for Sentiment Analysis in Big Data. 2022 Int. Conf. Adv. Comput. Technol. Appl. ICACTA 2022. 1–5 (2022). https://doi.org/10.1109/ICACTA54488.2022.9753203.

[31]. Li, X., Wu, P., Zou, C., Xie, H., Wang, F.L.: Sentiment Lossless Summarization. Knowledge-Based Syst. 227, 107170 (2021). https://doi.org/10.1016/j.knosys.2021.107170.