Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Twitter Data Sentiment Analysis Model: GuianSpin-Convolutional Network

[1]Supriya Sameer Nalawade* and [2]Dr. Shamala R. Mahadik
[1]Research Scholar, Electronics Department ,Sanjay Ghodawat University Kolhapur
Atigare, Tal- Hatkanagale, Kolhapur, Maharashtra, 416118, India
supriya32119@gmail.com
[2]Professor, Department of Electronics & Communication Engineering, SGU, Kolhapur
Corresponding Author: Supriya Sameer Nalawade
Research Scholar, Electronics Department ,Sanjay Ghodawat University Kolhapur
Atigare, Tal- Hatkanagale, Kolhapur, Maharashtra, 416118, India
supriya32119@gmail.com

ARTICLE INFO

Received: 28 Dec 2024

Revised: 18 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

Millions of people use microblogging to share their thoughts on a variety of subjects and to get feedback on any product, service, or problem. Sentiment analysis is the process of recognizing and categorizing the sentiments conveyed in the source text. Twitter is a well-known microblogging platform where monitoring and summarizing various thinking processes can yield valuable insights into various points of view. Thus, a major difficulty in the current research is to analyses opinions and categories them based on their polarity (positive or negative). In order to improve the effectiveness of sentiment evaluation in addition to other textual analysis, the GuianSpin-Convolutional Network is used in this study to predict the emotions from the tweets. This network combines GuianSpin optimisation with CNN.The TF-IDF technique, which improves sentiment analysis, is used to extract contextual and semantic information from the Twitter data. To get beyond the drawbacks and difficulties associated with feature extraction and data classification when utilising the GuainSpin Convolutional Network, extracted features are chosen and then applied to GuianSpin optimisation, which has the capacity to randomly create the solution and sequentially optimise. By altering the network parameters according to the fitness estimation, the GuianSpin enhances CNN training by making it easier to extract features from text data without changing the structure of the model. It improves the CNN's rate of convergence, optimises sentiment analysis solutions, and adapts dynamically to the context and degree of difficulty.The model achieves high accuracy , sensitivity and specificity of even while considering the TP utilizing the Twitter sentiment database.

Keywords: Sentiment analysis, Twitter reviews, Deep Learning, Evolutionary algorithms, Term-Frequency features.

1. INTRODUCTION

Social networks, microblogs, and blogs have been the primary sources of business development based on analysts and decision-making by business owners during the past several years [1]. Individuals frequently post their thoughts, opinions, and attitudes on social networking sites such as Facebook, Twitter, Instagram, Live Journal, and LinkedIn. Additionally, they select messaging apps such as Whatsapp, Skype, and WeChat [2]. Twitter is one of the most popular websites where users may quickly share information on a specific topic. Along with [3][4] [5] [6], the maximum character limit for a user message on Twitter is 280. Microblogging websites have grown into a vast array of information sources. The nature of microblogs where users post messages in real time about a variety of subjects, current affairs, grievances, and positive sentiments for items they regularly use is the reason behind this. To measure public opinion about their products, companies that produce these items have started to monitor these microblogs [7] [8] [9]. These businesses frequently monitor user feedback and respond to people on microblogs. Sentiment analysis assists in analyzing and separating the subjective information that is offered online. Finding remarks on a product or anything else that indicates neutral, positive, or negative views is the ultimate purpose of sentiment analysis. [10] [11]. Further, the features that are accessible determine how well sentiment analysis works. The sentiment analysis approach initially reads the text present in tweets or any corpus before extracting the necessary content that is included in the initial document[12] [13].

The most challenging problem in sentiment analysis today is to extract specific information because of the prevalence of better accents, spelling, slang terminology, grammatical errors, and other concerns; however, it may be improved by conducting various research[14] [1]. Three categories machine learning, lexicon, and hybrid techniques are used to categorize sentiment analysis [15] [16]. These divisions are predicated on the diverse perspectives that each group contributes to the text view, methodology, and degree of rating. Sentiment detection and classification are accomplished by supervised machine learning algorithms. Accurate feature extraction is accomplished through the use of preprocessing techniques that have been studied by many researchers [17]. Train the classifier using the feature selection approach, which selects the refined features of the subset derived from the original features in the provided dataset [18] [19]. The two main goals of the widely used techniques are to enhance classifier learning, whereby fewer features either result in better prediction or classification performance and to decrease the number of features that are available to decrease the dataset's dimension and reduce the running time for learning process [20] [1]. Text may be regarded of as being similar to the matrix of pixels in an image when it is converted to a matrix. Therefore, we can use the same procedure to the text data to enable the model to be trained in an alternative and efficient manner using the Twitter data. In computer vision applications like image analysis, it is possible to extract a portion of the pixel data information. This is not limited to extracting individual pixels; instead, feature information can be retrieved part by part, with each part including multiple pixels of data information [21][22].

Social media posts uploaded by diverse individuals speaking different languages are dispersed, informal, and lack organization. They also frequently contain nonstandard punctuation, abbreviations, slang, misspelled words, and common words that are unfeasible to describe without the utilization of an intelligent pre-processing structure [23] [2]. RNN is flexible enough to be applied to many text processing applications, but it suffers from exploring gradients and disappearing because of the long-term reliance of Twitter data [24] [25]. Artificial intelligence is applied to machine learning to automatically enhance the system and learn from its experiences, all without the need for programming [2] [16]. Unsupervised machine learning approaches classify the documents according to their commonalities into a predetermined number of clusters [1]. The lexicon-dependent strategies are illustrated using the evaluation of document sentences [27] [28]. Semantic familiarization-based methodology is used to assess the sentiment antithesis [29] [16][30]. When analyzing textual data, the DNN mostly concentrates on learning word embedding or doing tasks like categorizing and clustering the feature vectors it has learned [31][25].

The current work in the enhancement of the GuianSpin convolutional network along with CNN is a great advancement in the field of sentiment analysis and text handling. This Guain Spin convolutional network approach introduces GuianSpin's randomized initialization and iteration refinement way to enhance the training. Therefore, modifying these parameters from the evaluated CNN fitness enables the model to extract the relevant features from the text, which are accompanied by better sentiment classification. In addition to that, the integration accelerates both convergence and optimization of solutions, and it also ensures stability for the results on different datasets and applications.

- ***GuianSpinOptimization:*** GuianSpin achieves higher efficiency due to the iteration process of updating the local and global leaders in response to fitness estimates within a defined search space. By adjusting solution approaches toward the accurate result and employing distilled knowledge from differing solution areas, GuianSpin enhances the speed and quality of convergence. Apart from settling the exploration and exploitation trade-off problem efficiently, this approach also handles the scalability and robustness issues in optimization problems in many fields and thus can be considered a significant improvement in the development of algorithmic optimization strategies.
- ***GuianSpin-Convolutional Network:*** The deep CNN incorporated facilitates effective learning and the potential to extract the crucial features that assist in detecting the contextual information from the text. Further, the weights and biases of the deep CNN classifier are adjusted iteratively utilizing the GuianSpin optimization. In successive generations of the fitness evaluation, parameters are changed to minimize error

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

or loss. This procedure seems to remove much of the noise in the data, preventing overfitting and increasing the accuracy of the classifier and its ability to generalize.

According to its organizational structure, the work is written as follows: The methodologies and difficulties of the current works are described in Section 2. The GuianSpin-Convolutional Network optimization model and the sentiment analysis model's system model are both interpreted in Section 3. In section 4, the mathematical formulas for GuianSpin optimization are given. Results from applying the proposed sentiment analysis model are shown in section 5 along with a conclusion in section 6.

2. MOTIVATION

The goal of sentiment analysis is to determine a writer's viewpoint on a given subject by analysing language reviews. This goal utilised a variety of methods to identify whether an emotion is positive or negative, but they encounter certain challenges when attempting to categorise tweets based on their polarity. As a result, even with the large number of research in the different models, ongoing improvements are needed. In order to gather the necessary data, a number of sentiment analysis research categories are carried out utilising different techniques, as explained in the following section.

2.1 LITERATURE REVIEW

A supervised machine learning technique called Support Vector Machines (SVM) was introduced by Lakshmana Kumar Ramasamy et al. [2] for sentiment analysis. This is closely related to learning algorithms that examine the data within the categorization data set. The SVM model divides and builds a hyperplane. Classification on this hyperplane is faster and requires less time to process the data overall. Unfortunately, this model's performance accuracy was really poor. Analyzing the textual data in the dataset requires a greater focus on sentiment analysis. R. Manjula Devi et al. [11] developed a multi-layer perceptron classifier for the categorization of incoming tweets as either positive or negative, as the tweets contain non-useful characters that complicate sentiment analysis. For a vast amount of data, this approach efficiently increases accuracy. The scalability and high dimensionality of this approach, however, present difficulties. A standard sentiment analysis model was presented by Dr. Shailendra Narayan Singh and Twinkle Sarraf [17]. It consists of three main steps, which are data preparation, review analysis, and sentiment classification using random forest, which enhances sentiment classification performance. On the other hand, the large dimensionality and scalability of this model provide difficulties. To accurately predict the sentiment, Deepali Londhe1 and Aruna Kumari2 [16] presented a hybridized Social Eagle Algorithm-based deep learning model. The process of transliteration identifies the languages and transforms them into a uniform format; features are then retrieved from this standardized data. This model does well in sentiment categorization with high accuracy. However, because of the intricate training procedure, the execution time was lengthy. In [13], Dr. Vedavathi Na and Suhas Bharadwaj presented a methodology based on learner profile creation, clustering, and deep flamingo search-based recommendation systems. This model demonstrated extremely high performance accuracy. Nevertheless, overfitting remains a problem for this particular model. The Attention-based Bidirectional CNN-RNN Model was created by Mohammad Ehsan Basiri et al. [25].In addition, the feature space becomes high dimensional after processing sequences of any length and applying them in the feature extraction layer, which uses pre-trained GloVe word embedding vectors as the starting weights of the embedding layer. Hassonah et al. [1] employed the Support Vector Machines (SVM) classifier to develop a hybrid ML strategy and improve sentiment analysis. This approach yielded the highest accuracy but did not incorporate the filter and wrapper method to determine the appropriate combination for sentiment analysis.

Fanim K.Sufi and Ibrahim Khalil.[32]contributed a sentiment analysis for automated disaster monitoring from the Twitter feeds that utilized Artificial Intelligence(AI) and NLP. The method extracted the location-oriented public sentiments that provided in-depth knowledge about the global disaster. The features extracted from the sentiments were classified utilizing the DL-based CNN for anomaly detection in messages related to disasters. Further, the method analyzed the messages covering a huge set of languages. However, the method was found with a challenge associated with some classifiers was that it has a low capability to extract the local context features. Mohd Usama *et al.*[33]developed an NLP-based sentiment analysis that utilized the attentional model combined with the CNN and RNN that offered the merits of learning the long-term dependencies and learning the high-level features. Further,

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

the attention mechanism is utilized to focus on the features that contribute much to the prediction task and offer effective sentiment prediction.

Hanan T. Halawani et al.[34]created the Harris Hawks Optimization approach, which is used for automated sentiment analysis. The main function was to format the unprocessed social media text into an informative format. Additionally, the word embedding and skip-gram features are used to investigate the reduction of language processing reliance. Finally, the attention model classified the sentiments more effectively. However, the model is limited in identifying and classifying fine-grained emotions including surprise, fear, joy, anger, and so on.

J. Sangeetha and U. Kumaran.[35] developed a trustworthy Harris Hawks Optimization-driven LSTM method for predicting sentiment polarity to enhance marketing plans based on product reviews. The formation of Taylor–HHO, which selects the ideal weights for the hidden layers, helps to enhance the performance of the BiLSTM classifier by integrating Taylor series into HHO. Nevertheless, the loss graph for the validation set is quite large.

2.2 Challenges

The major research challenges are:

i) Conventional word embeddings are unable to manage polysemy since they give all words the same representation regardless of context or meaning. Additionally, sentimental word meanings cannot be captured by traditional word embeddings. The results of sentiment analysis are adversely affected by polysemy, sentiment polarity, and vocabulary gaps [2].

ii) Conventional optimization approaches have a problem of model complexity where overly complex models usually perform well within the training data, but extremely poorly outside the training data and also face challenges related to overfitting [16].

iii) The problem with sentiment analysis is that opinion identification techniques on new resources don't always behave reliably. As a result, sentiment analysis is being used more and more frequently. Any sentiment analysis research should address two challenges. The initial problem is to identify the right subject's positive, negative, or neutral polarity in the documentation, and determining the polarity's strength is the second issue [17].

However, the proposed GuianSpin-Convolutional Network model adopts the Term Frequency-Inverse Document Frequency (TF-IDF) to determine the most relevant terms in the document and finds the words that rank highest to find the contextual information. In addition, a data preprocessing phase involving tokenization, lemmatization, and stop word removal, emoji to text conversion, managing repeated character is carried out to remove structural deflection from the tweets. To reduces the amount of features in the dataset, and reduces its dimensionality by selecting only high score features thereby cutting down on the running time related to the learning process. To overcome the drawbacks and difficulties associated with feature extraction and data classification while utilizing CNN architecture, the proposed work combines GuianSpin with the capacity to randomly generate the solution and eventually optimize the classification results.

3. PROPOSED SENTIMENT ANALYSIS MODEL USING GUIANSPIN-CONVOLUTIONAL NETWORK MODEL:

Figure 1 displays the schematic representation of the proposed GuianSpin-Convolutional Network for sentiment analysis. This research aims to analyze information drawn from Twitter since it is brief as compared to other social sites hence helpful in sentiment analysis precisely. Twitter's dataset consists of millions of tweets that contain text, annotations, images, and various structural patterns. The analytical approach starts with pre-processing where the text is tokenized for splitting into tokens and filtering out unwanted components with stop word removal and then goes through the lemmatization step where words are transformed to their base stem. Words such as 'then,' 'is,' 'was,' and 'that' are excluded while preserving the overall significance of the content. Various methods exist for identifying sarcasm and irony, including the use of cues, contextual information, or external knowledge. Specifically, the proposed model uses TF-IDF method are applied to extract information from the preprocessed Twitter data. Further, the Term Frequency-Inverse Document Frequency (TF-IDF) analyzes the relative importance of a word in a document and determines the top-ranked words in a document to identify the contextual

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

information. The combination of the Guainspin optimization strengthens a proficient deep CNN classifier designed for sentiment analysis. As a result, the developed GuianSpin- Convolutional Network classifier successfully classifies the opinion tweets from Twitter according to their positive or negative or neutral polarities.

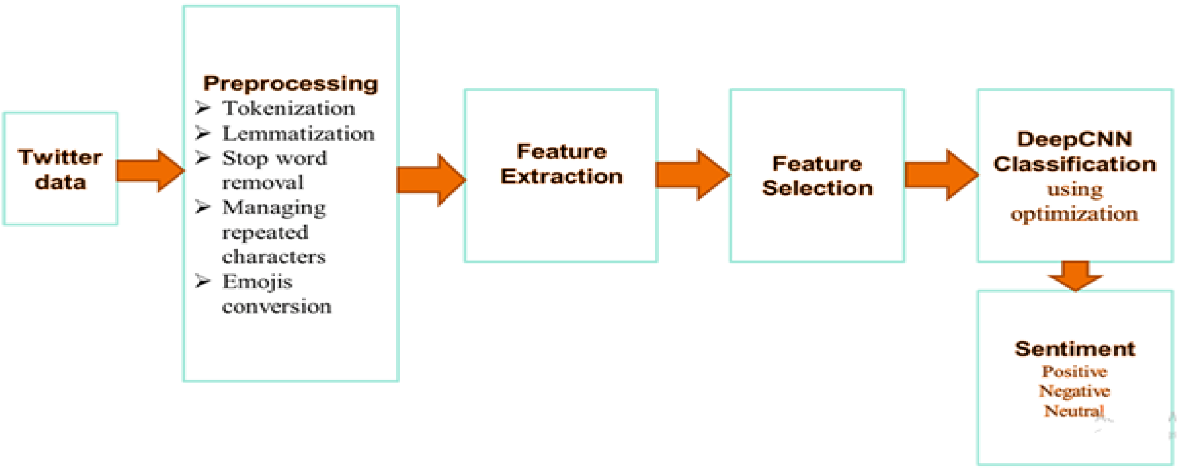

Figure 1. Proposed GuianSpin-Convolutional Network for the sentiment analysis

3.1 Data Processing

The input of sentiment analysis research is sourced from the sentiment 140 the Twitter sentiment dataset which is denoted as D. The sentiment analyzer decides the polarity of the tweets for which the data pre-processing step is the significant step that uncovers the opinion words from the tweets by cleaning the white spaces, punctuation marks, Stop Words, URLs. The pre-processing stages include:

3.1.1 Elimination of Stop word and Irrelevant Elements: All unnecessary content that had no relevance to SA including hashtags (#subject), numerals, special symbols, mentioned usernames (@username), and any URLs that started with "www," "http," or "https." are carefully removed to ensure the precision and reliability of the developed model. Despite being often used in texts, stop words frequently lack strong sentiment polarity and it is preferable to remove the stop words from the datasets rather than storing them there. It's significant to notice that negatives like "not" and "no" should not be eliminated since their removal could alter the meaning of entire phrases.

3.1.2 Lemmatization: Lemmatization is the process of identifying morphemes, or root words, from the resultant phrase by comparing its stemmed words to its dictionary, which identifies the nearest real root word.

3.1.3 Managing Repeated Characters: Some users employ repeating characters to draw attention to strong emotions in their tweets. To standardize such formulations, words that were absent from standard lexicons were converted to their appropriate forms. For instance, "sooooo goooood" became the standard for "so good".

3.1.4 Extending Contractions: It might be challenging to remove punctuation after a contraction, like "isn't" or "don't."To preserve the significance of contractions, they were extended to their complete forms. For example, "isn't" changed to "is not."

3.1.5 Emoji Conversion: Since emojis are frequently used in tweets to convey sentiment and emotion, we decided to convert emojis into their corresponding textual meanings using the 'demojize()' method from Python's emoji module. This improvement was particularly helpful for raising SA's accuracy.

3.1.6 Tokenization: Tokenization is the practice of replacing sensitive data with distinctive, recognizable symbols, without jeopardizing the text's security and returning all the required information. For instance, tokens, like words, phrases, or keywords are separated out of a string sequence during pre-processing and these tokens

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

form the input in various operations. For multi-languages especially, for those that use separation between words, tokenization is typically a simple job. The preprocessed output is denoted as D^*.

3.2 Feature Extraction from the preprocessed data

In tweets with hundreds of sentences, feature selection is the significant step that reduces the feature size through discarding the useless features such that the overfitting issues are discarded. Moreover, the accuracy of the prediction for sentiment analysis can be increased by reducing the number of features. The samples are represented by features, and the algorithm is tuned for a particular feature to correctly classify the polarity. To represent the class attribute in the smaller feature space, the feature selection process chooses the fewest significant features. With the help of feature selection strategies, classification accuracy can be improved dramatically. These techniques also give users a better knowledge of important class features, which helps them better interpret sentiments.

3.2.1 Word Embeddings

In the proposed model, Word embedding is adopted to extract information from the Twitter data. Further, the word embedding strategies employed in the proposed model are outlined in the following subsections.

TF-IDF: In its most basic form, TF-IDF calculates the relative frequency of words in a given document concerning their inverse proportion over the whole document's corpus[48]. This computation, inferred from intuition, establishes the relative importance of a word in a given document. Compared to popular terms like articles and prepositions, words that are frequently used in one or a small number of texts typically have higher TF-IDF scores.

The overall process for implementing TF-IDF is as follows, with a few minor variations across all of its applications. Given an assortment of documents Doc, a word an individual document $d \in Doc$. Further, the TF-IDF features are calculated as

$$w_d = F_{wd} * \log\left(|Doc| \big/ F_{w,Doc} \right) \tag{1}$$

Assume that the size of the corpus is about equal to the frequency of w over Doc, or $|Doc| \sim F_{w,Doc}$. For a very small constant z, if $1 < \log\left(|Doc| \big/ F_{w,Doc} \right) < z$. Even though w_d is less than F_{wd} and d, it is still positive. This suggests that while w is comparatively prevalent across the corpus, it retains some significance throughout Doc. It is said that this word w has a strong discriminating ability. Consequently, returning a document d where w_d is huge will very likely occur when a query contains this w and probably fulfills the user's needs.

The feature vector containing the weights of the individual terms obtained via the TF-IDF method is given by, TI, which forms the input to the suggested model for classification length of the padding is denoted by $\frac{h}{2}$ here h denotes the filter window size, step 1 indicates that from the lookup table $M \in S^v \times |W|$ tokens are mapping to the equivalent word vectors, here v denotes the word vector dimension, W denotes the vocabulary words, a vector that projected each word is $b_j \in S^u$

$$y = \{b_1, b_2, ..b_p\} \tag{2}$$

3.3 Feature Selection:

Each term receives a numerical score when using TF-IDF, which represents: The frequency with which a term occurs in a document. How unique a document is throughout the corpus is known as its inverse document

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

frequency. One way to pick features is to set a minimum TF-IDF threshold. Features (terms) with extremely low TF-IDF scores should be ignored since they can be overly prevalent or meaningless. Choosing the top-N features with the highest TF-IDF average. Sort terms according to their maximum or average TF-IDF score across all documents. Only the features with the top N scores should be kept..

3.4 Sentiment analyzer using GuianSpin-Convolutional Network

GuianSpin-Convolutional Network is adopted to classify the sentiments in the Twitter reviews effectively. Here, the Deep-CNN model's optimal weights are selected by incorporating GuianSpin optimization into the Deep-CNN, creating the GuianSpin-Convolutional Network. Further, the GuianSpin optimization is applied to assist in enhancing the training process, reduce the overfitting problem, and improve the classification accuracy of the model. The deep CNN is adopted in this sentiment classification is due to their ability to discern complex features from the textual data especially useful in determining the sentiment of tweets and other short texts. Firstly, CNN's convolutional layers use filters to slide over the input text and identify local features such as opinion words or expressions. This way the network is capable of learning representations of increasing complexity starting from the individual words used in the input to representations of the meaning and context of the input. Further layers of pooling give a reduction in dimensions and preserve important features that minimize overfitting and improve computational results. Fully connected layers take the abstract features and pass them through a sentiment classification process and with the help of activation functions like softmax, it provides probability distribution. While training the network, the stochastic gradient descent works to improve the parameters of the network which include weights and biases to fit the classification. In general, deep CNNs for sentiment analysis benefit from hierarchical text processing ability and is useful in capturing sentiment from sources such as Twitter.

Architecturally, the convolutional layer receives the feature vector TI as input and uses filters to extract local features from input vectors. The convolutional layer performs the majority of feature computations by utilizing a convolution kernel function. The local sufficient statistics are calculated by the pooling layer, and the pooling layer reduces the feature dimension that makes the deep model to achieve computational time and cost reduction, where the over-fitting issues are avoided. Finally, a probability distribution is employed by the fully connected layer to categorize the tweets under positive and negative polarity.During the convolution operation, different window sizes t with multiple features are applied to the data TI, wherein each available window accompanies the opinion words in the tweets. After applying the filters, the feature map is generated from each layer, and in each filter, a bias term c_d and weight matrix $N_d \in S^{t_v \times iL}$ are learned. Here, t_v denotes the total number of hidden neurons present in the layer. In each word window, local features are extracted by the weight matrix, and the operation of the convolution is mathematically represented as follows.

$$y'_j = i\left(N_d . y_{j\,j+i-1+c_d}\right) \tag{4}$$

here, the hyperbolic tangent function is denoted by $i(.)$, $y_{j:j+t-1}$ denotes the word vector concatenation from position j to $j+i-1$ $y_{j:j+t-1}$. Finally, a fixed size vector for the tweet is obtained and for each feature map, the most essential features are extracted and the hyperparameters are considered as a vector size corresponding to the total hidden units that are determined using the optimization. The output is derived as,

$$y^1 = T\left(N^1 . y' + c^1\right) \tag{5}$$

Softmax is the final layer that provides a probability distribution over the layer, and for training, a stochastic gradient descent algorithm is employed, which shows variations in their directions associated with frequent updates that are computationally expensive due to the processing time. To tackle the challenges of the existing training algorithm, the GuianSpin optimization is proposed and employed that adaptively tunes the hyperparameters, eliminates the overfitting loss, and enhances the training process resulting in high classification accuracy

3.4.1 GuianSpin optimization:

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The GuianSpin optimization method incorporates intelligent guide processes [38] and iterative adjustments [39] to fine-tune the parameters of a GuianSpin-Convolutional Network model. GuianSpin optimization method modifies the neural network's attributes such as weights and learning rates to iteratively carry out every potential solution obtained from the population until it reaches high classification accuracy. Additionally, the GuianSpin optimizer minimizes the overfitting loss and enhances the training process of the GuianSpin-Convolutional Network. It starts with a population distribution function that initially lays solutions equally within the search space, thus guaranteeing the diversification of solutions from the start. Through the guidance and management of the search, GuianSpin constantly enhances decision-making in the context of the optimization framework. At the local and global levels, the leaders make improvements to their solutions based on the evaluations of the fitness function and experiences made in the search space. This dynamic process facilitates both exploration and exploitation, making it possible to quickly converge to a set of optima and at the same time leverage collective intelligence to solve a wide range of problems effectively.

The development of the GuianSpin optimization method is therefore motivated by the need for an optimization method that will solve complex optimization problems effectively. Thus, GuianSpin uses the adaptive foraging behavior of the spider monkeys and the social interactions of the spiders as a source of ideas to develop more efficient optimization algorithms. Spider monkeys do not waste time and energy searching for inadequate food sources, and social spiders are very efficient in making decisions through communication. Combined with all of these characteristics, GuianSpin intends to improve the processes of searching and refining solutions in a search space, and, in turn, achieve the goal of quicker convergence and better solutions. The feature of the method is not only based on the integration of biological concepts with computing methods but also on the solutions to the problem of scalability as well as convergence speed and the quality of the solution in various optimization problems.

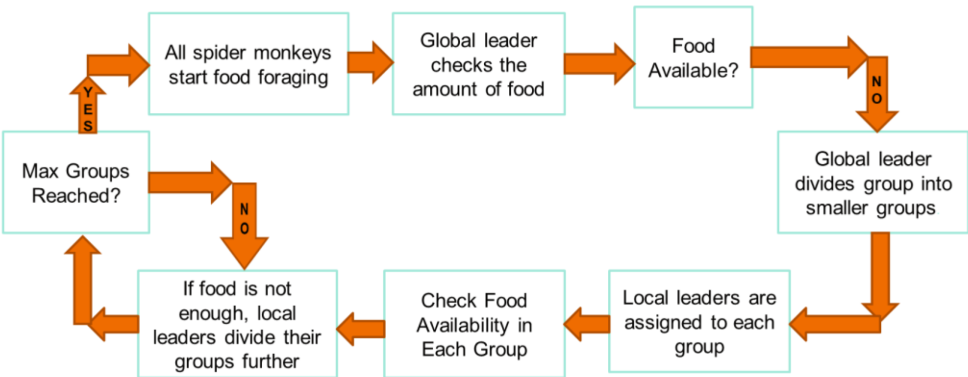

Figure 2 Guina Spin Optimization flow daigram

4. RESULT AND DISCUSSION

The developed sentiment analysis model based on the GuianSpin-convolutional network is implemented and the achievements of the developed model are justified through the result analysis for which the employed tool, metrics, and comparative methods are listed below.

4.1 Experimental setup

The experiment is run on a Windows 10 OS with 8GB of RAM and the Python programming language, and the sentiment 140 and Twitter datasets are used for the study.

4.2 Dataset description

4.2.1 Sentiment 140 dataset [41]: Using the Twitter API, 1,600,000 tweets were extracted for the Sentiment 140 dataset. The tweets have been rated (0 = negatively rated, 4 = positively rated) for deciding the polarities.

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

4.2.2 Twitter sentiment dataset [42]: The dataset contains three classes for sentiment analysis, such as negative (-1), neutral (0), and positive (+1), and possesses two fields, where one field represents the tweet while the other field is the label.

4.3 Performance metrics

These metrics are used to assess the GuianSpin, and these are characterized as follows:

4.3.1 Accuracy: For determining the classification of sentiment using the GuianSpin, accuracy is defined as the percentage of samples that are properly classified by,

$$Accuracy = \frac{E_{TP} + E_{TN}}{E_{TP} + E_{FP} + E_{TN} + E_{FN}} \tag{15}$$

4.3.2 Sensitivity: Sensitivity is the probability that a test result will actually be true positive when the GuianSpin classifies the sentiment, and it is determined by,

$$Sensitivity = \frac{E_{TP}}{E_{TP} + G_{FN}} \tag{16}$$

4.3.3 Specificity: Specificity, which is determined by the GuianSpin is the probability that a test result will actually be negative.

$$Specificity = \frac{E_{TN}}{E_{TN} + E_{FP}} \tag{17}$$

where, E_{TP} and E_{TN} denotes the true positive and true negative respectively, E_{FP} and E_{FN} are the false positive and false negative respectively.

4.4 Experimental results

Figure 3 indicates the experimental results from the sentiment analysis GuianSpin with the sample tweets that were marked positive or negative. The input message is provided and then the polarity of the tweet is marked as either positive or negative.

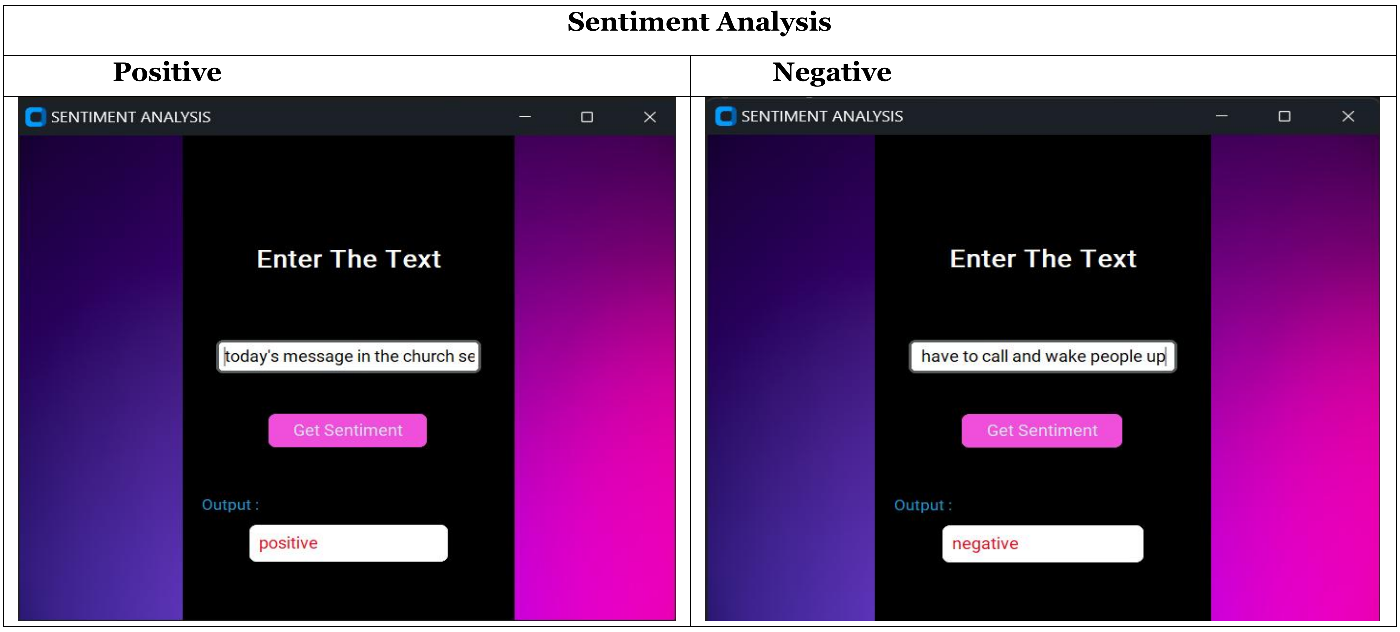

Figure 3: Experimental result obtained using the proposed GuianSpin model

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5. COMPARATIVE DISCUSSION

Comparing the results of the GuianSpin-Convolutional network with other methods including SVM, MLP, RF, deep CNN with Adam optimizer shows that existing models face several challenges. MLP and RF, face challenges of scalability and high dimensionality which reduces the chances of effective feature extraction and classification processes being accomplished. Although Deep CNNs are effective, the model had a slow convergence rate or gets stuck at local optimums, as revealed by Adam optimizer. Further, they may also have problems related to overfitting and generalization on other datasets of the same type. In addition, the SVM method in which processing the texts to extract the crucial information was a complex task that requires further advanced techniques due to the existence of modern accents, slang words, spelling, and other issues that limited the performance[1]. The sentiment analysis with MLP was found with low accuracy and require more feature extraction and selection techniques to enhance the processing time of classification and improve the classification accuracy. The issues are resolved in another GuianSpin-Convolutional network that applies GuianSpin optimization, which includes stochastic initialization to perform a detailed search at the initial stage and iterative improvements at the subsequent stage to escape local optima and increase the convergence rate. It helps in managing the exploration-exploitation trade-off, filters out noise, checks overfitting, and increases the model's capacity to generalize exactly because of the used algorithm, which allows creating of a more accurate and less dataset-specific sentiment classifier.

A comparative analysis is conducted to demonstrate the superiority of the GuianSpin Convolutional Network model over the existing techniques. With the Twitter Sentiment Database taken into account, the model achieves great performance in terms of accuracy of 97.96%, sensitivity of 97.46%, and specificity of 97.61% for k-fold 9. The sentiment 140 dataset and the sentiment dataset from Twitter are used to calculate the following terms; the results are displayed in Tables 1.

Table 1: Comparative discussion table for TP using the sentiment 140 dataset and Twitter sentiment database

Models	TP 90					
	Sentiment 140 dataset			Twitter sentiment database		
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)
SVM	0.687973964	0.6823271	0.733548	0.65742556	0.687837356	0.695849731
MLP	0.72570434	0.760294831	0.765335	0.722270593	0.724979381	0.867130229
RF Classifier	0.784159093	0.771002691	0.852666	0.85036915	0.865636189	0.882193442
DeepCNN With Adam	0.86547481	0.82797117	0.853537	0.870327961	0.879871105	0.896569227
GuianSpin-Convolutional Network	0.950073	0.96754	0.979527	0.979653	0.974619	0.976128

6. CONCLUSION

Finally, the present work on improving sentiment analysis and text processing with the GuianSpin-convolutional network is a promising development toward advancing the field. The proposed method incorporates GuianSpin optimization, which, based on the foraging pattern and guidance, resolves some of the main shortcomings of the existing methods, such as scalability, slow convergence, and overfitting. The random initialization and gradual updates of weights and biases provide extensive coverage of the search space while also improving the feature learning and increasing the model's ability to classify correctly. The proposed GuianSpin-convolutional network shows a good balance between exploration and exploitation, as well as high stability and versatility when tested on different datasets. This makes the presented solution very effective for use in sentiment analysis applications. Several word embedding techniques are applied to extract contextual and semantic information from the Twitter data, such as TF-IDF, which enhances the sentiment analysis. Additionally, decrease the dimensionality of features and minimize the running time associated with the learning process by selecting features which having highest score. Choosing only the top-N features with the highest TF-IDF average and sort terms according to their

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

maximum or average TF-IDF score across all documents. Only the features with the top N scores should be kept. Future work based on this research could lie in refining the GuianSpin-convolutional network utilizing more optimization methods including meta-heuristic hybrid models that would help to further the speed of convergence and the quality of the results acquired.

REFERENCES

[1] Hassonah, M.A., Al-Sayyed, R., Rodan, A., Ala'M, A.Z., Aljarah, I., and Faris, H., "An efficient hybrid filter and evolutionary wrapper approach for sentiment analysis of various topics on Twitter," Knowledge-Based Systems, vol.192, pp.105353, 2020.

[2] Ramasamy, Lakshmana Kumar, Seifedine Kadry, Yunyoung Nam, and Maytham N. Meqdad. "Performance analysis of sentiments in Twitter dataset using SVM models." *International Journal of Electrical and Computer Engineering (IJECE)* 11, no. 3 (2021): 2275-2284.

[3] Phan, H.T., Tran, V.C., Nguyen, N.T., and Hwang, D., "Improving the performance of sentiment analysis of tweets containing fuzzy sentiment using the feature ensemble model," *IEEE Access*, vol.*8*, pp.14630-14641, 2020.

[4] Tang, H., Tan, S., and Cheng, X.,"A survey on sentiment detection of reviews," Expert Systems with Applications, vol.36, no.7, pp.10760–10773, 2009.

[5] Vomfell, L., Härdle, W. K., andLessmann, S. "Improving crime count forecasts using twitter and taxi data," Decision Support Systems, vol.113, pp.73–85, 2018.

[6] Mendon, S., Dutta, P., Behl, A., and Lessmann, S., "A Hybrid approach of machine learning and lexicons to sentiment analysis: enhanced insights from twitter data of natural disasters," Information Systems Frontiers, pp.1-24,2021.

[7] Lin, Chenghua, and Yulan He. "Joint sentiment/topic model for sentiment analysis." In *Proceedings of the 18th ACM conference on Information and Knowledge Management*, pp. 375-384. 2009.

[8] Birjali, Marouane, Mohammed Kasri, and Abderrahim Beni-Hssane. "A comprehensive survey on sentiment analysis: Approaches, challenges, and trends." *Knowledge-Based Systems* 226 (2021): 107134.

[9] Kouloumpis, Efthymios, Theresa Wilson, and Johanna Moore. "Twitter sentiment analysis: The good the bad and the omg!." In *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 5, no. 1, pp. 538-541. 2011.

[10] Whitelaw, Casey, Navendu Garg, and Shlomo Argamon. "Using appraisal groups for sentiment analysis." In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, pp. 625-631. 2005.

[11] Devi, R. Manjula, P. Keerthika, P. Suresh, M. Sangeetha, C. Sagana, S. Savitha, K. Devendran, and B. Nithiesh. "Twitter sentiment analysis using collaborative multi-layer perceptron (MLP) classifier." In *2023 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1-6. IEEE, 2023.

[12] Cambria, Erik. "An introduction to concept-level sentiment analysis." In *Advances in Soft Computing and Its Applications: 12th Mexican International Conference on Artificial Intelligence, MICAI 2013, Mexico City, Mexico, November 24-30, 2013, Proceedings, Part II 12*, pp. 478-483. Springer Berlin Heidelberg, 2013.

[13] Vedavathi, N., and R. Suhas Bharadwaj. "Deep flamingo search and reinforcement learning based recommendation system for e-learning platform using social media." *Procedia Computer Science* 215 (2022): 192-201.

[14] B. G. Patra, D. Das, A. Das and R. Prasath, "Shared task on sentiment analysis in Indian languages (sail) tweets-an overview., in: MIKE," pp. 650–655,2015.

[15] A. Al-Zoubi, J. Alqatawna, H. Faris and M. A. Hassonah, "Spam profiles detection on social networks using computational intelligence methods: The effect of the lingual context," Journal of Information Science, 2019.

[16] Londhe, Deepali, and Aruna Kumari. "Multilingual Sentiment Analysis Using the Social Eagle-Based Bidirectional Long Short-Term Memory." International Journal of Intelligent Engineering & Systems 15, no. 2 (2022).

[17] Singh, Shailendra Narayan, and Twinkle Sarraf. "Sentiment analysis of a product based on user reviews using random forests algorithm." In *2020 10th International conference on cloud computing, data science & engineering (Confluence)*, pp. 112-116. IEEE, 2020.

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[18] Yadav, Ashima, and Dinesh Kumar Vishwakarma. "Sentiment analysis using deep learning architectures: a review." *Artificial Intelligence Review* 53, no. 6 (2020): 4335-4385.

[19] Collomb A, Costea C, Joyeux D, Hasan O, and Brunie L,"A study and comparison of sentiment analysis methods for reputation evaluation," Rapport de recherche RR-LIRIS, 2014.

[20] Kiritchenko, Svetlana, Xiaodan Zhu, and Saif M. Mohammad. "Sentiment analysis of short informal texts." *Journal of Artificial Intelligence Research* 50 (2014): 723-762.

[21] Xu, Guixian, Yueting Meng, Xiaoyu Qiu, Ziheng Yu, and Xu Wu. "Sentiment analysis of comment texts based on BiLSTM." *Ieee Access* 7 (2019): 51522-51532.

[22] Dang, Nhan Cach, María N. Moreno-García, and Fernando De la Prieta. "Sentiment analysis based on deep learning: A comparative study." *Electronics* 9, no. 3 (2020): 483.

[23] Gonçalves, Pollyanna, Matheus Araújo, Fabrício Benevenuto, and Meeyoung Cha. "Comparing and combining sentiment analysis methods." In *Proceedings of the first ACM conference on Online social networks*, pp. 27-38. 2013.

[24] M. S. Akhtar, D. Gupta, A. Ekbal and P. Bhattacharyya, "Feature selection and ensemble construction: A two-step method for aspect-based sentiment analysis," Knowledge-Based Systems, vol.125, pp.116–135, 2017.

[25] Basiri, M.E., Nemati, S., Abdar, M., Cambria, E., and Acharya, U.R., "ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis," Future Generation Computer Systems, vol.115, pp.279-294, 2021.

[26] Tang, Duyu, Furu Wei, Bing Qin, Nan Yang, Ting Liu, and Ming Zhou. "Sentiment embeddings with applications to sentiment analysis." *IEEE Transactions on Knowledge and Data Engineering* 28, no. 2 (2015): 496-509.

[27] M. Song, H. Park, and K.-s. Shin, "Attention-based long short-term memory network using sentiment lexicon embedding for aspect-level sentiment analysis in Korean," Inf. Process. Manage,vol.56,no.3, pp.637–653, 2019.

[28] Saif, Hassan, Yulan He, and Harith Alani. "Semantic sentiment analysis of Twitter." In *The Semantic Web–ISWC 2012: 11th International Semantic Web Conference, Boston, MA, USA, November 11-15, 2012, Proceedings, Part I 11*, pp. 508-524. Springer Berlin Heidelberg, 2012.

[29] Psomakelis E, Tserpes K, Anagnostopoulos D, and Varvarigou T, "Comparing methods for Twitter sentiment analysis, 2015.

[30] Agarwal, Apoorv, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca J. Passonneau. "Sentiment analysis of Twitter data." In *Proceedings of the workshop on language in social media (LSM 2011)*, pp. 30-38. 2011.

[31] F. Xing, E. Cambria, R. Welsch, "Intelligent asset allocation via market sentiment views," IEEE Comput. Intell. Mag., vol.13, no.4, pp.25–34, 2018.

[32] Sufi, F.K. and Khalil, I, Automated disaster monitoring from social media posts using AI-based location intelligence and sentiment analysis, IEEE Transactions on Computational Social Systems. (2022).

[33] Usama, M. Ahmad, B. Song, E. Hossain, M.S. Alrashoud, M. and Muhammad, G, Attention-based sentiment analysis using convolutional and recurrent neural network, *Future Generation Computer Systems*. 113 (2020) 571-578.

[34] Halawani, H.T., Mashraqi, A.M., Badr, S.K. and Alkhalaf, S., 2023. Automated sentiment analysis in social media using Harris Hawks optimization and deep learning techniques. Alexandria Engineering Journal, 80, pp.433-443.

[35] Sangeetha, J. and Kumaran, U., 2023. A hybrid optimization algorithm using BiLSTM structure for sentiment analysis. Measurement: Sensors, 25, p.100619.

[36] Chhabra, A. Hussien, A.G. and Hashim, F.A, Improved bald eagle search algorithm for global optimization and feature selection. *Alexandria Engineering Journal,* 68 (2023)141-180.

[37]Jiang, S. Shang, J. Guo, J. and Zhang, Y, Multi-Strategy Improved Flamingo Search Algorithm for Global Optimization, *Applied Sciences*. 13 (2023) 5612.

[38]James, J.Q. and Li, V.O. A social spider algorithm for global optimization, *Applied soft computing*, *30* (2015) 614-627.

[39] Cuevas, E., Cienfuegos, M. Zaldívar, D. and Pérez-Cisneros, M. A swarm optimization algorithm inspired in the behavior of the social-spider, *Expert Systems with Applications* 40 (2013) 6374-6384.

2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

[40] Sharma, H. Hazrati, G. and Bansal, J.C. Spider monkey optimization algorithm, *Evolutionary and swarm intelligence algorithms* (2019) 43-59.

[41] Sentiment 140 dataset, "https://www.kaggle.com/kazanova/sentiment140", accessed on December 2022.

[42] Twitter sentiment dataset, https://www.kaggle.com/datasets/saurabhshahane/twitter-sentiment-dataset, accessed on December 2022.

[43] Shin, H.-C. Roth, H. R. Gao, M. Lu, L. Xu, Z. Nogues, I. Yao, J. Mollura, D. and Summers, R. M. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, *IEEE transactions on medical imaging* 35 (2016) 1285-1298.

[44]Suthaharan, S. and Suthaharan, S. Support vector machine, *Machine learning models and algorithms for big data classification: thinking with examples for effective learning* (2016), 207-235.

[45] Alboaneen, D.A. Tianfield, H. and Zhang, Y, Sentiment analysis via multi-layer perceptron trained by meta-heuristic optimization, *In 2017 IEEE International Conference on Big Data (Big Data)* (2017) 4630-4635. IEEE.