International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I11P132

Volume 72 Issue 11, 333-344, November 2024

Original Article

Twitter Data Sentiment Analysis using a Novel Pairing Scheme of GHO

Neha Sharma[1], Sanjay Tyagi[2]

[1 2]*Department of Computer Science and Applications, Kurukshetra University, Kurukshetra.*

[1]*Corresponding Author: nehadevsharma19@gmail.com*

Received: 03 May 2024 Revised: 01 October 2024 Accepted: 14 October 2024 Published: 29 November 2024

Abstract - *In social media, numerous public networks, namely Twitter, Facebook, YouTube, Instagram, etc., are used to communicate through videotapes, pictures, and posts. Adapting to such information and data mining from such websites will become troublesome in the future. Sentimental analysis, a type of contextual mining, is frequently popular on social media and includes machine learning-based, lexicon-based, and hybrid methods. This paper aims to develop sentimental analysis using a novel optimization technique pairing scheme. The methodology includes the Grasshopper Optimization Algorithm (GOA) devised for sentimental analysis using the Twitter dataset and the Amazon Reviews dataset for sentimental analysis. The proposed method dealt with the pre-processing of the dataset and evaluation of features using different similarity metrics such as Term Frequency (tf), Inverse Document Frequency (idf), Tanimoto Co-efficient, and Cosine similarity. The optimized and extracted features are further classified using the Machine Learning tool for both datasets. The developed technique provides better accuracy, 91.65% for the Twitter dataset and 88.4% for the Amazon Reviews dataset. The results were further compared with the existing techniques for superiority.*

Keywords - *Sentimental analysis, Twitter, Data mining, Grasshopper, Machine learning.*

1. Introduction

By connecting people who have shared their thoughts and experiences with other social media users, SNSs (Social Networking Sites) are quickly expanding the number of industry researchers and educational focus. A group in a social network is made up of nodes or individuals and links between users. Individuals with similar interests come together to form groups on social media. This study uses similarity index techniques called hybrid similarity, cosine similarity, and Jaccard similarity to find similarities among users of social sites [1]. The social network graph, which is made up of nodes and edges, is considered to measure similarity. The set theory that verifies the intersection of test and raw data is the foundation for Jaccard similarity. The likelihood of an intersection is high in this instance, and the likelihood of similarity is high if the intersection value is higher than the similarity value [2]. SNS (Social Networking Sites) are the latest and most trendy, particularly for the young generation. These sites have become trendy due to the availability of the Internet.

1.1. Basics of Social Networking

Researchers with backgrounds in academia and business are drawn to SNS. Social networking was first thought of as a way to communicate globally, but it has since developed into an essential tool for social and professional purposes [3]. It is necessary to discuss and investigate the commercial impact social networking sites will have on consumer purchasing behavior due to the most significant advancements in social networks and their combined use of social and commercial functions [4].

1.2. Definition of Social Network

It is described as a collection of relationships of n individuals and m social media sites represented by an edge-weighted graph G (L, F, X). Here, L is the group of nodes, $|L|$=n, and F is the collection of directed relationships L⊆ $F \times F, |L| = m$, and X is the weight of edges associated with every edge of F". Sentiment analysis has always been an area of interest for the researchers. Due to the increase in users over social media websites and the increasing impact of tweets, the tweets have been analysed for various purposes. Recently, in India, a few years back, during the election campaign of Narender Modi, social media helped to build a good image for the Bhartiya Janta Party (BJP). As a result, Mr. Modi is the current prime minister of India. In a similar fashion, in the recent precedential elections of the USA, Joe-Baiden's campaign gained more attention, and emotions flew into winning directions. Several organizations have also used Twitter to spread hate against specific content or religion; hence, it becomes a crucial issue to justify the emotion from the tweet. This research article focuses on developing a more accurate and precise classification architecture for Twitter sentiment analysis. Specifically, regarding the architecture of the algorithm, the proposed work has been evaluated for multiple datasets and compared with other state of art techniques. Machine Learning (ML) has been used as the primary tool for analysing contextual and analytical data [5]. Though many ML architectures have been proposed earlier and illustrated in the related work section, the contributions to ML architecture and precise classification architecture of sentiment are illustrated as follows.

- Design a novel behavior of the Grasshopper Algorithm for the training and classification data optimisation.
- Validation of the proposed algorithm with advanced ML architecture
- Comparison with other state of art techniques.

2. Related Work

Sentiment analysis in social media is a fundamental issue with extensive, intriguing applications. Most existing social media sentiment classification systems ignore Twitter data on these sites and determine the sentiment polarity primarily based on textual content. Henríquez and Ruz 2018 proposed the Deep Neural Network (DNN) based functional link for sentimental analysis. The authors compare the performance of the Machine Learning classifiers such as Support Vector Machine, Random Forest, and the proposed neural network-based random vector. Further, imbalanced data is handled using the SMOTE, and the sentimental score is computed. The analysis showed that the average precision using the Chilean earthquake dataset is 37.20 while the accuracy is 82.90%. Further, recall and F1 scores for the same dataset are 35.08 and 36.04. The study provides better accuracy but is limited to validating the datasets [4].

Saleena 2018 proposed the ensemble classifier to develop the single classifier by combining the working of basic classifiers. The authors pre-processed the data and represented the features for polarity classification in positive and negative ways. The base classifier such as Naïve Bayes, Random Forest, SVM, and Logistic Regression was considered to develop the ensemble classifier. Accuracy, Recall, Precision, and F1 scores for positive and negative classes are computed using the different classifiers. The analysis results demonstrated that the suggested technique's accuracy is 74.76%, while the average F1 score is 73.33% for Twitter sentimental data. The limitation of the research work is that the authors do not consider neutral tweets and do not present the analysis for social network platforms [5].

Alharbi and Doncker 2019 presented a neural network model that considers user behaviour within a given tweet. Convolutional Neural Network (CNN) was considered the SemEval-2016 Workshop contributed two datasets to evaluate the proposed system. The suggested model surpasses existing baseline methods, such as SVM and Naive Bayes, demonstrating the value of looking beyond a document's content (in this case, a tweet) in sentiment classification because it gives the classifier a thorough understanding of the problem. The outcomes were measured in accuracy, recall, precision, and F1 score. The results showed that Precision using Long Short-Term Memory (LSTM) is 0.87, while recall is 0.88. The study shows that the F1 Score is 0.86, and the study shows exciting results, but the study is limited to exploring the Recurrent Neural Network (RNN) for sentimental analysis [6]. Nagarajan and Gandhi 2019 collected 600 million public tweets for sentimental analysis. The authors pre-processed the tweet data, and ternary classification was done. The hybridization was done using the genetic algorithm and particle swarm optimization and further classified using the decision tree to determine the classification accuracy for sentimental analysis. The outcomes showed that overall accuracy, precision, recall, and F-measure for the Twitter dataset is 90%, 91.5%, 91.7%, and 91.4% respectively. The study is limited to analysing the performance using other classifiers for sentiment classification [7].

Han et al. 2020 this study proposed a Fisher kernel function based on Probabilistic Latent Semantic Analysis for sentiment analysis using the SVM. The problem of having latent semantic features ignored in text sentiment analysis is addressed by this technique, which improves the classification effect for SVM by incorporating probability characteristics and utilizing latent semantic information as the classification feature. The findings demonstrate that, when compared to the comparison approach, the effect of the method suggested in this study is improved. The recall and average precision using the proposed SVM equipped technique are 87.20% and 88.30%. The analysis results are promising but limited in validating the performance by comparing it with existing techniques [8].

Ruz et al. 2020 considered the Bayesian Network classifier for the sentimental analysis of two datasets. The authors computed the Bayes factor by collecting the data and performing the sentimental analysis by pre-processing the data, and then feature representation was carried out in line. Each tweet in the processed information is labelled with positive and negative sentiments. Further, the authors compute the sentimental score and performance metrics in terms of Recall, Precision, Accuracy, and F1 score is computed. The accuracy using the Chilean earthquake dataset with the SVM classifier is 0.812, while recall is 0.936. The overall performance is better using the proposed technique [9]. Bairavel and Krishnamurthy 2020 this research suggested a multimodal sentiment analysis method based on audio, video, and text. The proposed method uses textual audio and video modalities to investigate sentiments extracted from web recordings. The features acquired from various modalities are combined using a feature-level fusion technique. Therefore, the extracted features are optimally selected using a novel Oppositional Grass Bee Optimization

(OGBEE) algorithm to obtain the best optimal feature. The suggested method uses a Multilayer Perceptron-based Neural Network (MLP-NN) to classify sentiment. The experimental research indicates that the proposed method requires less computational time and offers a higher classification accuracy of about 95.2% [10]. Sadiq et al. 2021 deployed the Multilayer Perceptron to analyse the aggressive comments on the Twitter dataset. The authors proposed a joint framework for extracting features and used the dense layer of the Neural Network for behavioural analysis. The authors perform the 10-fold cross-validation process to analyse the performance of the proposed model. The outcomes in accuracy, recall, F1 score and precision are computed. The accuracy of the proposed technique is 92%, while precision, recall, and F1 scores are 90% each. The limitation of the study is the need for detailed analysis for the detection of aggression using the different features [11].

Shekhawat et al. 2021 proposed a mechanism for extracting the feelings from tweets in this research. Positive, neutral, or negative tweets can all be categorised. Due to the arbitrary behaviour of tweets, metaheuristic-based clustering approaches are superior to traditional methodologies. The best cluster heads for the dataset are determined using a hybrid approach called Hybrid Spider Monkey Optimization with K-Means Clustering. The accuracy of the suggested method is assessed using the Twitter dataset. The proposed method is compared to other metaheuristic techniques to determine its veracity. The accuracy of the proposed technique is 99.98%, which is better than that of other optimization techniques. The study is limited to considering the paradox tweets for the Twitter dataset [12].

Naresh and Krishna 2021 this study suggest a machine-learning method based on optimization for Twitter data classification. The proposed study is conducted in three steps, which comprise the procedure. The first stage involves gathering and pre-processing the data, and the second stage involves optimising the data by removing essential characteristics and using different machine learning algorithms. The updated training set is categorized into different groups in the third stage. The results of each algorithm are different. Compared to previous machine learning techniques, the suggested sequential minimal optimization with the decision tree technique offers a good accuracy of 89.47%. The outcomes showed that overall accuracy, precision, recall, and F-measure for the Twitter dataset is 89.47%, 91.6%, 89.5%, and 96.3% respectively. The study is limited to analysing the performance using other optimization techniques, such as Grasshopper, for sentiment classification [13]. Biradar et al. 2022 proposed the Machine Learning based sentimental analysis for the Twitter dataset. The authors consider the customer reviews in which data is pre-processed and clustered, and the term frequency-inverse document frequency is extracted and then classified as the sentiments. The classification model classifies the tweets by considering the positive, negative, and neutral scores. The accuracy of the TP rate and FP rate for the J-48 classifier is better, while precision for OneR is promising. The F-measure for the OneR classifier is 0.97, which is better than Naïve Bayes and other classifiers. But, the limitation of the study is that there is a need to evaluate the performance using the other classifiers for validation [14]. Pandey et al. 2022 proposed the cuckoo search based robust clustering method for sentimental analysis. The authors determine the optimal clusters and then polarity for emotional tweets. The proposed model is tested considering the benchmark functions. A statistical analysis was also carried out to determine the best solution. The proposed work is divided into phases: pre-processing, feature selection, representation, and efficient clustering. The mean precision and recall for the twitter dataset is 85.89% and 85.12%. This is the disadvantage of the current study in that the authors are limited in considering the multi-labelling for feature selection and dynamic clustering technique [15].

M. Hadni and H. Hassane (2023) introduced a new model for selecting important features in Arabic text using a Chaotic Firefly Algorithm (CFA). They changed the firefly algorithm's attraction factor to chaotic results, which helped balance exploration and exploitation better. They tested their model with classifiers like Naive Bayes (NB), Support Vector Machine (SVM), and K-Nearest Neighbours (KNN). The CFA-SVM model had the best accuracy (89%), better than the CFA-KNN (87%). They focused on reducing the number of features, which is especially challenging in Arabic datasets because of the language's complex structure, and there are few methods for feature selection in Arabic [16].

Ismail Shahin et al. (2023) developed the Grey Wolf Optimizer (GWO) combined with KNN to select the best features in emotion classification. They tested this method on three datasets (Emirati-accented speech, RAVDESS, and SAVEE) and found it worked better than other methods like the bat algorithm (BAT) and cuckoo search (CS). The study showed that choosing the right features is crucial for emotion recognition systems, as using irrelevant features can result in poor classification results [17]. Al-shalif et al. (2024) thoroughly studied how to choose important parts of text data for text classification using methods like simulated annealing, genetic algorithms, particle swarm optimization, and ant colony optimization. They found that these methods help improve the classification by picking out the most valuable parts from big data sets. But, they also noticed that these methods do not always work the same way on different data sets or for different text classification jobs [18]. Therefore, by studying the literature and understanding the research gaps, it is clear that tweets are analysed with less accuracy. Moreover, some literature lacks the optimization technique that optimizes the pre-processed data. However, some researchers focussed on analysing the tweets using the decision tree, which is less efficient due to the complex structure of the tree. The present research work is focused on avoiding the limitations of the existing work to provide better simulation results.

3. Proposed Work

The proposed work has two significant sections: optimization in the training architecture and classification architecture. The overall workflow is explicable using the flow diagram given in Figure 1.

The following datasets have been utilized to perform the operations.

- Twitter Sentiment Analysis: The data contains multi-linguistic comments on oversupplied tweets. The data is open-source data and can be downloaded from https://www.kaggle.com/datasets/jp797498e/twitter-entity-sentiment-analysis. The dataset has three emotions and more than 10,000 records to be processed. Its ability to include multiple languages allows researchers to study sentiment analysis across different languages. This involves handling challenges in Natural Language Processing (NLP), such as breaking down text (tokenization) and understanding language-specific grammar rules.
This dataset is perfect for creating and testing different sentiment classification models like Support Vector Machines (SVM), Naive Bayes, or more complex deep learning methods like LSTMs or transformers (like BERT). It can be used in practical tasks like analyzing public opinion, customer feedback, or product reviews.
- Amazon Reviews Sentiment Analysis: The second data is based on the Amazon reviews that are given under the services that are provided by the Amazon company to the users that are associated with the service infrastructure network. The dataset is open source and downloadable from https://www.kaggle.com/datasets/bittlingmayer/amazonreviews. This tool analyses feelings in text, like deciding if a review is positive, negative, or neutral. The data works well for teaching computers to understand these feelings. It can help find out what customers like or do not like, how well different ways of understanding text work, and how to use language to understand what people are saying.

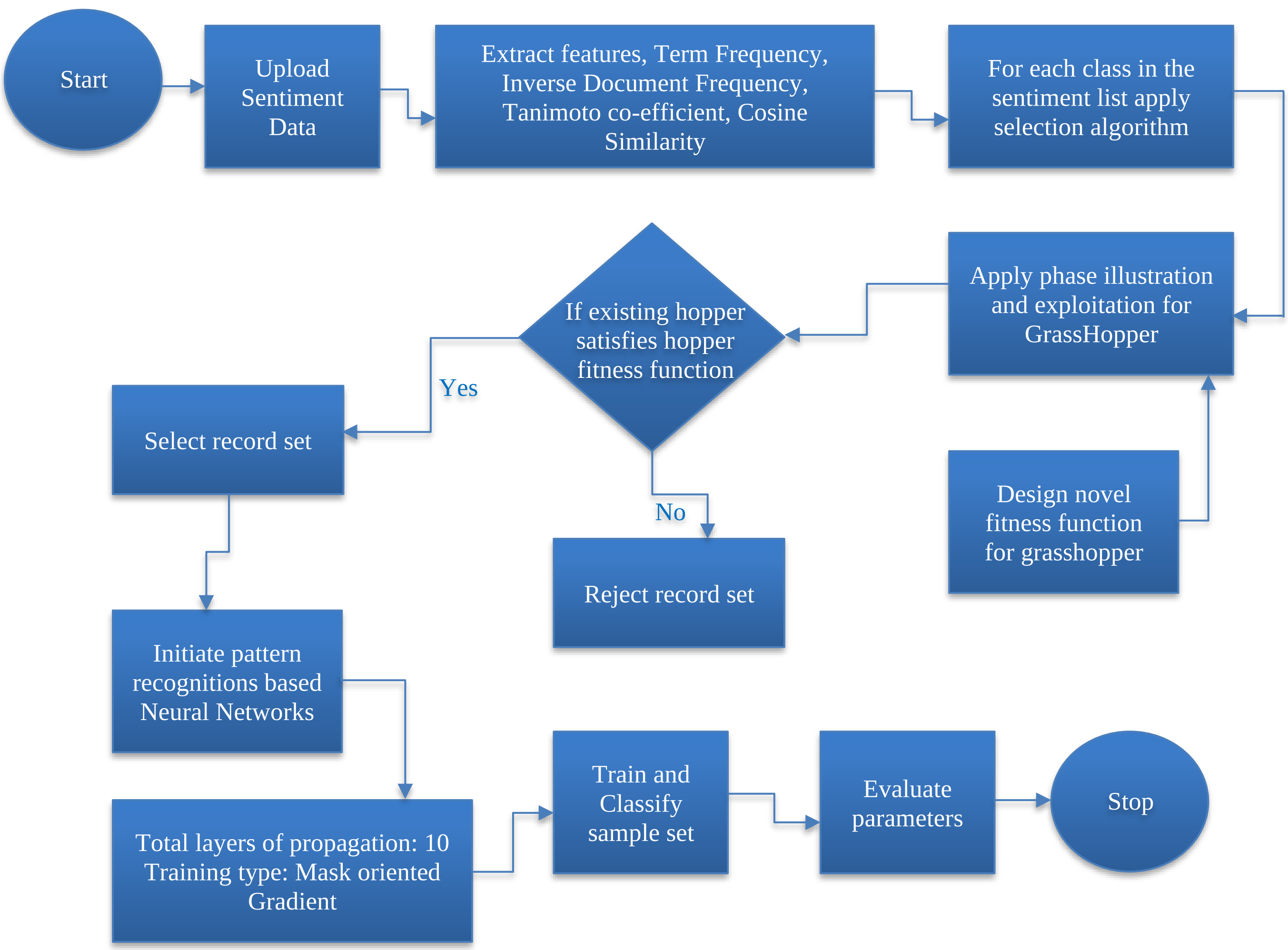

Fig. 1 Proposed work flow

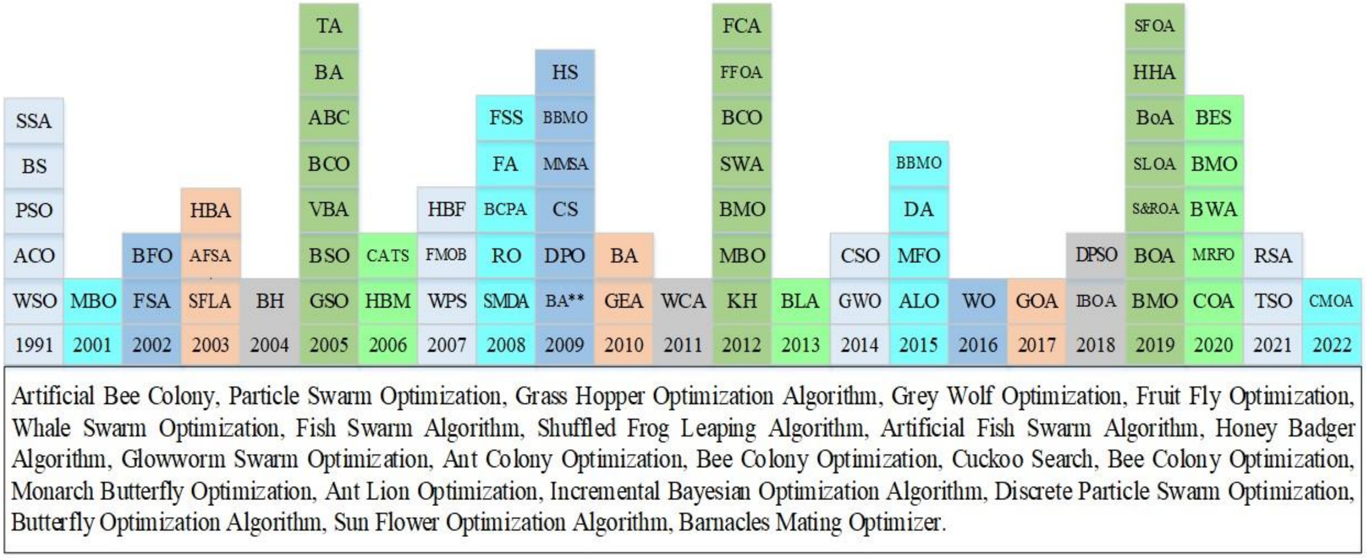

Fig. 2 SI algorithm list

The proposed algorithm architecture uses the following features for the evaluation and processing.

- Term Frequency (TF): The frequency at which a term, viz a keyword, is repeated in the current document.
- Inverse Document Frequency (IDF): The frequency of a keyword repeated in other documents.
- Tanimoto Co-efficient: It is the intersection of two data vector values and represents the angular distance between two vector values.
- Cosine similarity: It is the cosine angle between two vector spaces.

The proposed algorithm architecture illustrates a new behavior of formation and grouping in the grasshopper algorithm architecture. The GHO is a Swarm Algorithm architecture, and the ordinal measures of SI are as follows.

Swarm Intelligence (SI) is a subset of the meta-heuristic algorithm series within the statistical machine learning architecture. The following components are necessary for any SI algorithm: -

- The selectors
- The working identities
- The evaluation method

The identities that contribute to the table are the working identities. SI is essentially based on how various species around the world gather food. Since its initial proposal in 1991, SI has undergone numerous revisions and alterations.

As per the SI Algorithm list (Figure 2), Grasshopper was first proposed in 2017, and by the time the document was written, the algorithmic architecture had advanced to the best of our knowledge. Grass Hopper Algorithm (GHO), also called Grass Hopper Optimization Algorithm (GOA), is what the suggested algorithm seeks to implement in order to be on the updated site [19]. The two Grass Hoppers food collection stages serve as the foundation for a GHO algorithm. Due to the baby hopper's inability to travel, the first stage is known as the exploitation phase.

The second stage is the exploration phase, during which the hopper reaches adulthood and flies to different areas to gather food. The optimized feature set will be validated using ML by applying supervised classification. The post-validated set will be passed for training purposes in order to update or train the knowledge base. A semi-supervised approach will be applied to the classification to evaluate the quantitative parameters. Parameters such as kappa coefficient and accuracy will be computed for evaluation.

The proposed grasshopper is based on the grouping behavior and can be illustrated using the work algorithm. The grouped hopper behavior of the novel-designed GHO algorithm selects the most appropriate set of feature vectors for the training and classification. For this aspect, the proposed algorithm has used ML based learning methods, and ML algorithm architecture can be further segregated as follows.

Algorithm 1 Algorithm Grouped-GHO

Input: Feature Vector as $\mathrm{Grouped_A}$, Emotions of $\mathrm{Grouped_a}$ as $\mathrm{GT_A}$

Output: Selected interval as $\mathrm{Grouped_S}$

$\mathrm{hf = GT_A.lasses;}$ Hopper field is 5 as per dataset

$[\mathrm{h_{id}, h_c}] = \mathrm{k-means}\ (\mathrm{Grouped_A, hf})$ Divide the data into hypothetical fields

where $\mathrm{h_{id}}$ contains the centroid of the hoppers and $\mathrm{h_c}$ is the hypothetical centroid

$\mathrm{gt} = 1$ Starting from firsttground truth class

While $\mathrm{gt \le hf}$ **do**

$\mathrm{h_p = Find(GT_A.Classes == gt)}$Hopper population

$f_s s = \mathrm{Grounded}_A . h_p .$ Find hopper food in $R - R$ intervals
$f_s = 0;$
for i in $f_S s$ **do**
$F\ score = [];$ Initiate Flight Score to Null
$l_f = 5;$ 5 levy flights for exploitation
$p_i = 1;$ Propagated flights
while $p_i \leq l_f$ **do** Exploitation
$n_p = h_p . \mathrm{randomPop}\ ()$ Generate a random Population
$n_p . \mathrm{merge}(i)$Merge the cuGroupedent hopper to the population
$F_S = \mathrm{Grouped}_s . n_p$ Collect the hopper food
$G_f = h_c . n_p .$ Extract the hypothetical centroid value as global food
$G_f = h_c . n_p .$ Extract the hypothetical centroid value as global food
$|f_s| = \mathrm{Hopper} - \mathrm{Fitness} - \mathrm{Exploitation}\ (f_S, G_f);$ Score attained
$p_l + +$
$\mathrm{Fscore}\ [p_l] = f_s;$
f_s is 10 if fitness function is satisfied else 0
$th = 35;$
if $\mathrm{fscore}_i = 50$
$\mathrm{Grouped}_s . \mathrm{Merge}(i);$
End if
end for
Return: $\mathrm{Grouped}_S$

ML is a branch of artificial intelligence and involves two processing types, Statistical Machine Learning (S-ML) and Propagational Machine Learning (P-ML), as shown in Figure 3. S-ML is entirely dependent upon the statistical measures of the data and works specifically on values. For example, k-means is a clustering algorithm termed an S-ML based algorithm as it calculates the Euclidean distance between the data attributes and the formed centroid. Before finally applying the Grouped-GHO algorithm, the feature vectors of Group A and the ground truth classes are analysed, and the statistic description is included in Table 1.

In Table 1, the "Features" column shows the names of the different features in the dataset. Each feature has a "Description" that explains what it measures. The "Mean" is the average value of each feature across all the data points. The "Median" is the middle value when all the feature values are sorted from smallest to largest. The "Standard Deviation" shows how much the feature values differ from the mean, giving an idea of how spread out the data is. The "Minimum" and "Maximum" values tell us the smallest and largest values observed for each feature, which helps us understand the range of data for each feature.

On the other hand, Neural Network (NN) oriented concepts like Feed Forward Neural Networks (FFNN), Conjugate Neural Networks (CBNN) and Deep Neural Networks are some examples of P-ML. Machine Learning (ML) is divided into two segments, namely Statistical Machine Learning (S-ML) and propagation machine learning (P-ML).

3.1. Statistical ML (S-ML)

Statistical ML methods, also known as empirical methods, have been expanded in recent decades and are used to analyse the data of biomedical sensors. ML techniques continue developing, and more data is available to explore the use of these techniques. Statistical ML methods are used for remote sensing, sentiment analysis, etc.

These methods determine the statistical relationship between the sensed data and the variable of interest without knowing the causal relationship. Informed statistical ML methods are simple and similar. However, there is some knowledge about the relationship between the estimated variables. Physics-based ML methods require detailed knowledge about the variables used in architecture modelling. Therefore, selecting an analytical approach entirely depends upon the data available for processing.

For example, a physics-based approach is recommended if there is a better understanding of the physical processes. Otherwise, statistical methods are suitable for processing. The main advantages of S-ML are given as follows: -

- A mathematical rigorous tool to describe the relationship between the sampling and model error.
- Estimate and predict the outcomes of interest and determine the relationships between the variables to quantify uncertainty.
- Test the hypothesis under the specified assumptions.

Table 1. Descriptive statistics

Features	Mean	Median	St. D	Minimum	Maximum
F 1	76.4	75	5.2	60	95
F 2	1.24	1.2	0.3	0.5	2
F 3	0.67	0.7	0.12	0.3	0.9
F 4	3.2	3	0.5	2	5
F 5	150	145	20	100	200
F 6	6.5	7	1.5	2	10
F 7	35	30	10	15	60
F 8	3	2.9	0.4	2.5	4
F 9	4.5	4	1.2	2	8

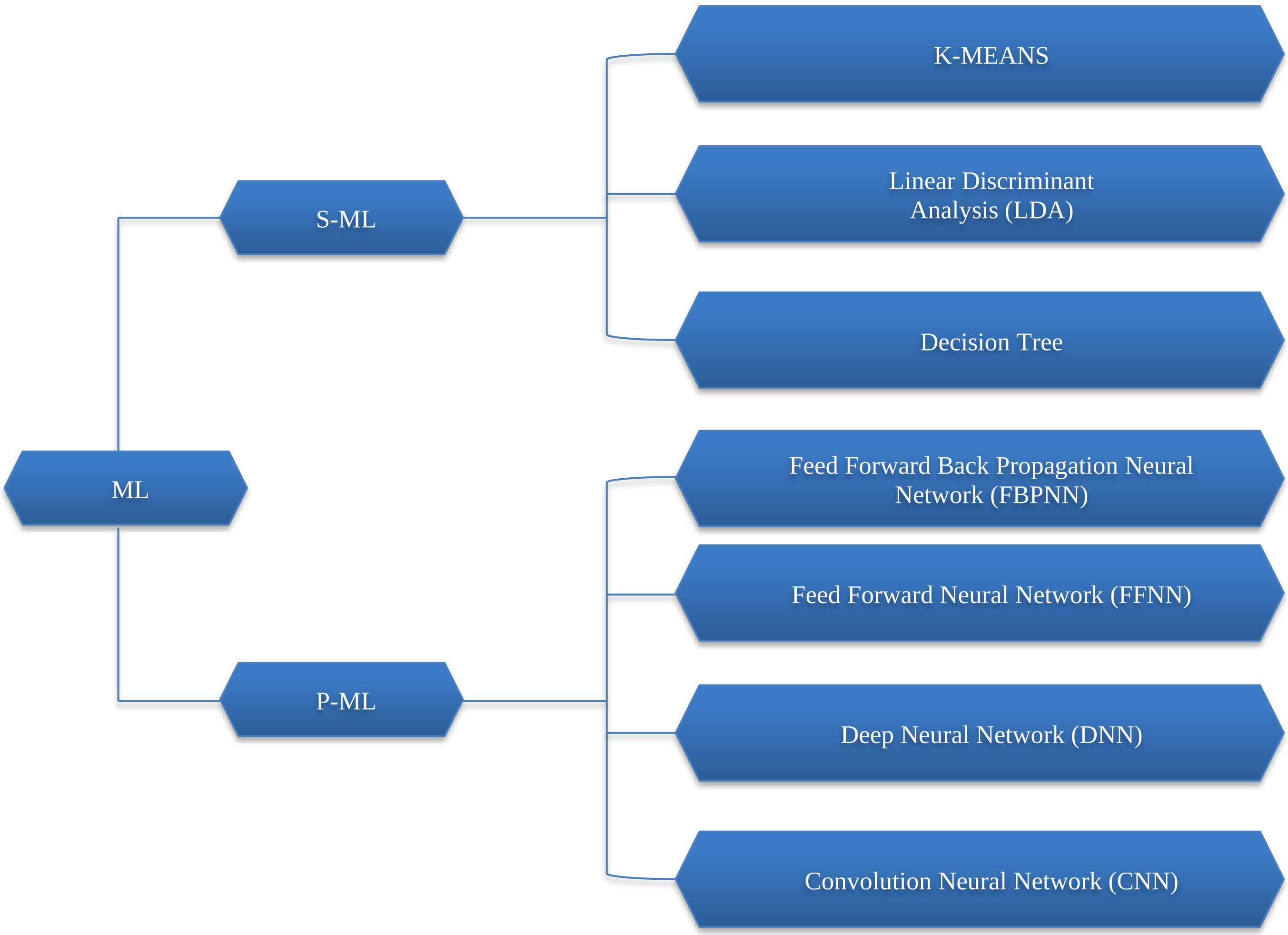

Fig. 3 ML architecture

3.2. Propagation- ML (P-ML)

Propagation-based use of ML techniques is widespread in estimating propagation errors. Various propagation-based ML techniques include Back Propagation Neural Networks (NN), Feed Forward Neural Networks, Deep Neural Networks, and Convolutional Neural Networks. This research used a Conjugate-based neural network to train the ECG samples. However, Back propagation NN is a multilayer Feed forward NN trained as per the error propagated, and it is one of the widely applied NN models. The learning rule of this model is based on using the network's weight value and threshold value for regulation and to compute the Mean Square Error (MSE). The formula of the model is expressed as: -

Output of the hidden layer:

$$y_m = \sum_n x_{mn} w_n - \theta_m \quad (1)$$

Computational Output of the node:

$$z_j = f \sum_m x_{jm} w_m - \theta_j \quad (2)$$

Error of the output node:

$$Error = \frac{1}{2} \sum_j \left(t_j - Z_j\right)^2 \quad (3)$$

Where, w_n is the input node, y_m is the hidden layer, and z_j is the output layer, x_{mn} is the weight of the network between the input and hidden node. x_{jm} is the weight of the network between the hidden and the output node. t_j is the label, and f is the active function.

4. Results and Discussion

The tweets collected using the dataset are pre-processed, optimized, and further classified using the Machine Learning classifier into different classes. The best features from the optimized data are extracted using the GHO algorithm, and then the Machine Learning technique is applied to classify the tweets. The classification was done in three classes: positive, negative, and neutral.

4.1. Performance Metrics

The classification of tweets is shown in the given table, and performance metrics for Twitter sentimental analysis in terms of accuracy, precision, recall, and F1 score are computed for evaluation. The different numbers of samples were considered in line with the different numbers of tweets. The different simulation metrics are explained as follows: -

Accuracy: Accuracy is the ratio of true positive rate, i.e. samples that are classified correctly to the total count of samples.

$$Accuracy = \frac{True\ classified\ samples\ (C_{true})}{Total\ number\ of\ samples\ (T_S)} \quad (4)$$

Precision (P): It is defined as the ratio of True classified samples (C_{true}) to the total number of samples that are classified positively.

$$Precision = \frac{True\ Positive\ (T_P)}{True\ Positive\ (T_P) + False\ Positive\ (F_P)} \quad (5)$$

Recall (R): It is defined as the ratio of true positives to the sum of true positives and false negatives.

$$Recall = \frac{True\ Positive\ (T_P)}{True\ Positive\ (T_P) + False\ Negative\ (F_N)} \quad (6)$$

F-measure: F-measure provides better prediction of results in comparison to Accuracy. It is the weighted average of both precision and recall.

$$F - measure = 2 * \frac{(P*R)}{(P+R)} \quad (7)$$

4.2. Performance Analysis of the Proposed GHO-NN

This section presents the detailed performance analysis of the proposed work against individual scenarios when only GHO or NN are implemented. The proposed technique is developed in this study using two different datasets, Twitter Sentimental Analysis and Amazon Reviews Sentiment Analysis. The performance metrics have been computed considering these datasets.

Table 2 compares different accuracy techniques using the Twitter and Amazon Reviews sentimental analysis dataset. The analysis findings indicated that the average accuracy using the GHO algorithm with the Twitter dataset is 88% and 85.6% using the Amazon Reviews dataset, while accuracy using the NN technique only is 91.65% for the Twitter dataset and 88.4% for the Amazon Reviews dataset. The results using the proposed technique are promising, as 94.7% and 89.98% accuracy are obtained using the Twitter and Amazon Reviews datasets, respectively.

Thus, results using the optimization and Machine Learning techniques are better and improved by 7.6% and 3.4% from GHO only and NN technique only, respectively. This shows that owing to the GHO based optimization, the proposed work presented a better decision-making approach. Therefore, the optimized technique equipped with ML provides a better classification for positive and negative classified tweets.

Table 3 compares different precision techniques using the Twitter sentimental analysis dataset and the Amazon Reviews sentimental analysis dataset. The proposed technique's precision results are 93.64% and 88.98% for Twitter and Amazon Reviews datasets. The analysis revealed that the average precision using the GHO algorithm with the Twitter dataset is 87% and 84.6% for the Amazon Reviews dataset. In comparison, the NN technique is only 89.65% for the Twitter dataset and 86.4% for the Amazon Reviews dataset. Thus, results using the optimization-equipped NN technique are better and improved by 7% and 4% from GOA only and NN technique only, respectively. This clearly reflects the better selection combined with the proposed work's enhanced machine learning based classification architecture. The optimized technique equipped with ML provides a better classification for positive and negative classified tweets. Table 4 compares different recall techniques considering the Twitter sentimental analysis dataset, and Amazon Reviews sentimental analysis dataset. The average recall results using the proposed technique are 92.8% and 89.34% for Twitter and Amazon Reviews datasets. The average recall using the analysis results revealed that the GHO algorithm with the Twitter dataset is 92.8% and 89.34% for the Amazon Reviews dataset, while recall using the NN technique only is 87.65% for the Twitter dataset and 84.32% for the Amazon Reviews dataset. Thus, recall results are improved by 8.5% from GOA and 5.8% from the NN classifier. The GOA optimized technique, which is classified using ML, provides a better classification for positive and negative classified tweets.

Table 2. Comparison of accuracy for different datasets

Number of Samples	Accuracy using the Twitter Dataset			Accuracy using Amazon Reviews Dataset		
	GHO	NN	Proposed (GHO+NN)	GHO	NN	Proposed (GHO+NN)
100	87.7677	91.43	94.5366	85.454	88.2346	89.76
200	87.785267	91.44757	94.554167	85.471567	88.25217	89.777567
300	87.802834	91.46513	94.571734	85.489134	88.26973	89.795134
400	87.890401	91.4827	94.589301	85.506701	88.2873	89.812701
500	87.977968	91.57027	94.676868	85.594268	88.37487	89.900268
600	88.065535	91.65784	94.764435	85.681835	88.46244	89.987835
700	88.153102	91.7454	94.852002	85.769402	88.55	90.075402
800	88.240669	91.83297	94.939569	85.856969	88.63757	90.162969
900	88.328236	91.92054	95.027136	85.944536	88.72514	90.250536
1000	88.415803	92.0081	95.114703	86.032103	88.8127	90.338103

Table 3. Comparison of precision for different datasets

Number of Samples	Precision using the Twitter Dataset			Precision using Amazon Reviews Dataset		
	GHO	NN	Proposed (GHO+NN)	GHO	NN	Proposed (GHO+NN)
100	87.2347	89.43	93.4236	84.454	86.2346	88.76
200	87.303267	89.44757	93.441167	84.471567	86.25217	88.777567
300	87.320834	89.46513	93.458734	84.489134	86.26973	88.795134
400	87.408401	89.4827	93.476301	84.506701	86.2873	88.812701
500	87.495968	89.57027	93.563868	84.594268	86.37487	88.900268
600	87.583535	89.65784	93.651435	84.681835	86.46244	88.987835
700	87.5962917	89.7454	93.739002	84.769402	86.55	89.075402
800	87.6838587	89.83297	93.826569	84.856969	86.63757	89.162969
900	87.7714257	89.92054	93.914136	84.944536	86.72514	89.250536
1000	87.8589927	90.0081	94.001703	85.032103	86.8127	89.338103

Table 4. Comparison of Recall for different datasets

Number of Samples	Recall using the Twitter Dataset			Recall using the Amazon Reviews Dataset		
	GHO	NN	Proposed (GHO+NN)	GHO	NN	Proposed (GHO+NN)
100	85.2347	87.43	92.5808	82.1764	84.1003	89.1236
200	85.303267	87.44757	92.598367	82.193967	84.11787	89.141167
300	85.320834	87.46513	92.615934	82.211534	84.13543	89.158734
400	85.408401	87.4827	92.633501	82.229101	84.153	89.176301
500	85.495968	87.57027	92.721068	82.316668	84.24057	89.263868
600	85.583535	87.65784	92.808635	82.404235	84.32814	89.351435
700	85.5962917	87.7454	92.896202	82.491802	84.4157	89.439002
800	85.6838587	87.83297	92.983769	82.579369	84.50327	89.526569
900	85.7714257	87.92054	93.071336	82.666936	84.59084	89.614136
1000	85.8589927	88.0081	93.158903	82.754503	84.6784	89.701703

Table 5 compares the different F-measure techniques considering the Twitter sentimental analysis dataset, and Amazon Reviews sentimental analysis dataset. The results of the average F-measure with the proposed method are 93.22% for the Twitter dataset and 89.164% for the Amazon Reviews dataset, respectively.

The analysis showed that the average F-measure using the GHO algorithm with the Twitter dataset is 88.64% and 83.52% for the Amazon Reviews dataset. Consequently, the NN technique's F-measure is only 88.65% for the Twitter dataset and 85.38% for the Amazon Reviews dataset. Thus, average F-measure results are better and improved by 7.7% from GOA and 5.1% from the NN classifier.

The GOA optimized technique classified using the NN provides better results and classifies the tweets with better accuracy. Figure 4 compares the performance metrics of Twitter sentimental analysis, and Amazon Reviews sentimental analysis datasets. The analysis showed that performance using the Twitter dataset is better and that promising results have been obtained.

Table 5. Comparison of F-measure for different datasets

Number of Samples	F-measure using Twitter Dataset			F-measure using Amazon Reviews Dataset		
	GHO	NN	Proposed (GHO+NN)	GHO	NN	Proposed (GHO+NN)
100	86.22310374	88.41869	93.0002906	83.29963423	85.15408	88.94142839
200	86.29167995	88.43626	93.01785797	83.31720451	85.17165	88.95899547
300	86.30924931	88.45383	93.03542533	83.33477479	85.18922	88.97656254
400	86.39682805	88.4714	93.05299269	83.35234507	85.20679	88.99412961
500	86.48440677	88.55898	93.14056148	83.43992839	85.29437	89.08169698
600	86.57198546	88.64656	93.22813027	83.52751168	85.38195	89.16926434
700	86.58474386	88.73413	93.31569906	83.61509493	85.46953	89.25683171
800	86.67232253	88.82171	93.40326784	83.70267815	85.55711	89.34439907
900	86.75990117	88.90929	93.49083662	83.79026134	85.64469	89.43196643
1000	86.84747979	88.99687	93.5784054	83.87784449	85.73227	89.51953379

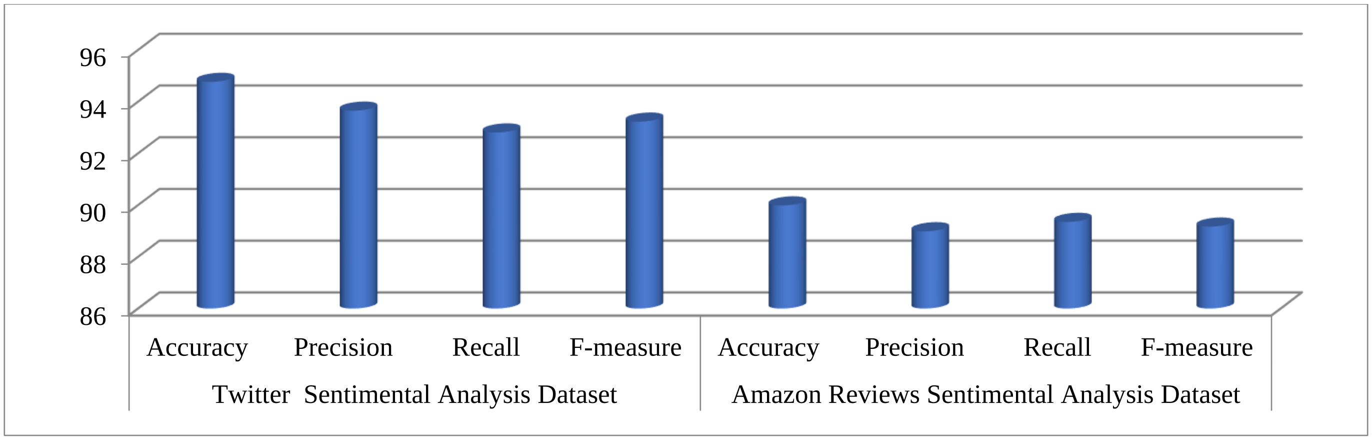

Fig. 4 Comparison of performance metrics for different datasets using the proposed (GHO-NN) technique

4.3. Comparative Analysis

The proposed study for sentimental analysis is compared with the existing techniques for validation. The current work was experimented with using the Twitter and Amazon Reviews datasets for sentimental analysis. The proposed research is compared with the existing work using the combination of Genetic Algorithm, Particle Swarm Optimization, and Decision Tree for Twitter data sentimental analysis presented by Nagarajan and Gandhi 2019 [7]. However, the study conducted by Naresh and Krishna in 2021 used the sequential minimal optimization (SMO) based machine learning technique decision tree (SMODT) for the Twitter dataset [13]. Further, the results are compared with the Multilayer perceptron-based Neural Network technique that Sadiq et al. 2021 proposed for the Cyber troll Twitter dataset [11]. In order to assess the current work, Table 6 presents a comparative analysis of the suggested technique with the existing work. The findings of the analysis demonstrated that the suggested method is enhanced by 5.2% for accuracy, 2.5% for precision, 1.2% for recall, and 2% for F-measure from the hybrid technique proposed using the combination of Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Decision Tree (DT). Furthermore, the study shows that the proposed technique is improved by 5.8% for accuracy, 2.4% for precision, 3.7% for recall, and 2.9% for F-measure from the NN technique. In line with this, the study shows improved results from SMODT and a significant improvement in precision, which is 4.3%. Thus, combining the grasshopper with the NN provides better results in terms of different performance metrics.

Table 6. Comparative analysis of the proposed technique with the existing work

Techniques	Accuracy	Precision	Recall	F-measure
PSO + GA+DT [Nagarajan and Gandhi 2019] [7]	90	91.5	91.7	91.4
SMODT [Naresh and Krishna 2021] [13]	89.47	91.6	89.5	90.3
NN [Sadiq et al. 2021] [11]	92	90	90	90
Proposed Technique	94.732	93.864	92.88	93.22

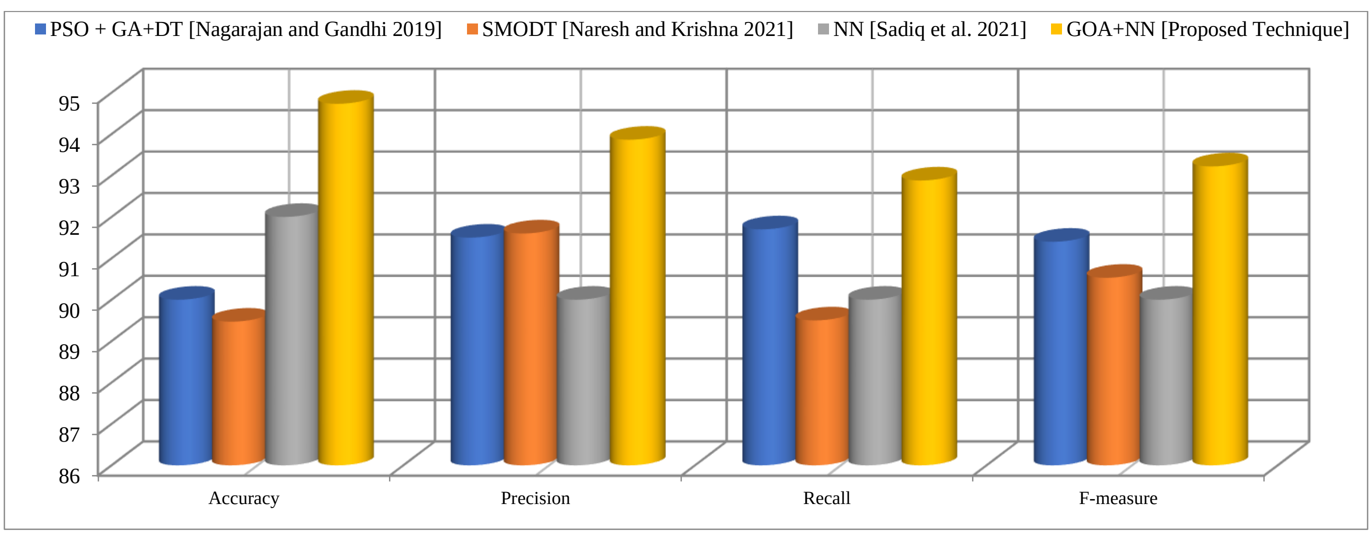

Fig. 5 Comparative analysis of the proposed technique with existing techniques

Figure 5 compares the suggested method with the previous research. The results are improved in terms of different performance metrics. The major improvement is seen in Accuracy from the existing technique due to the better search capability of the proposed grasshopper technique. Moreover, the unique adjustable feature improves the exploration capability for extracting features.The consistent high performance shows that combining the Grasshopper Optimization Algorithm with Neural Networks improves classification accuracy and stability. The PSO + GA + DT method performs less well, especially in recall, which means it struggles to identify important cases. The SMODT technique does well in recall but has lower precision, suggesting it might have problems with false positives. The NN method has average results in all areas but does not reach the same level of overall effectiveness as the proposed method. Overall, these results show that the GOA + NN approach performs better and fixes traditional methods' issues, making it a more dependable choice for classification tasks in complex data sets.

5. Conclusion

The present work introduced the sentimental analysis-based optimization technique for better extraction of features. The proposed method extracts the features and determines the similarity using the different similarity measures for sentimental analysis.

Pre-processing and feature extraction are the key phases in training the suggested model. Machine learning classifiers that offer higher precision, such as F-measure and recall, are used to classify sentiment tweets into negative, positive, and neutral groups.

The optimization was done using the GOA, and when compared to other existing algorithms, the NN, SMODT, and (PSO+GA+DT) combos were demonstrated to perform better. Comparing the proposed work to existing machine learning classifiers and optimization strategies, the overall accuracy achieved is over 95%. Future research could adapt our suggested work to use different classifiers.

References

[1] Auliya Rahman Isnain, Jepi Supriyanto, and Muhammad Pajar Kharisma, "Implementation of K-Nearest Neighbor (K-NN) Algorithm for Public Sentiment Analysis of Online Learning," *Indonesian Journal of Computing and Cybernetics Systems*, vol. 15, no. 2, pp. 121-130, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] N. Vedavathi, and K.M. Anil Kumar, "E-Learning Course Recommendation Based on Sentiment Analysis Using Hybrid Elman Similarity," *Knowledge-Based Systems,* vol. 259, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Shalak Mendon et al., "A Hybrid Approach of Machine Learning and Lexicons to Sentiment Analysis: Enhanced Insights from Twitter Data of Natural Disasters," *Information Systems Frontiers*, vol. 23, pp. 1145-1168, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Pablo A. Henríquez, and Gonzalo A. Ruz, "Twitter Sentiment Classification Based on Deep Random Vector Functional Link," *2018 International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, Brazil, pp. 1-6, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[5] Ankit, and Nabizath Saleena, "An Ensemble Classification System for Twitter Sentiment Analysis," *Procedia Computer Science*, vol. 132, pp. 937-946, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[6] Ahmed Sulaiman M. Alharb, and Elise de Doncker, "Twitter Sentiment Analysis with A Deep Neural Network: An Enhanced Approach Using User Behavioral Information," *Cognitive Systems Research*, vol. 54, pp. 50-61, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[7] Senthil Murugan Nagarajan, and Usha Devi Gandhi, "Classifying Streaming of Twitter Data Based on Sentiment Analysis Using Hybridization," *Neural Computing and Applications*, vol. 31, pp. 1425-1433, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] Kai-Xu Han et al., "Application of Support Vector Machine (SVM) in the Sentiment Analysis of Twitter Dataset," *Applied Sciences*, vol. 10, no. 3, pp. 1-16, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Gonzalo A. Ruz, Pablo A. Henríquez, and Aldo Mascareño, "Sentiment Analysis of Twitter Data During Critical Events Through Bayesian Networks Classifiers," *Future Generation Computer Systems*, vol. 106, pp. 92-104, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] S. Bairavel, and M. Krishnamurthy, "Novel OGBEE-Based Feature Selection and Feature-Level Fusion with MLP Neural Network for Social Media Multimodal Sentiment Analysis," *Soft Computing*, vol. 24, pp. 18431-18445, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Saima Sadiq et al., "Aggression Detection Through Deep Neural Model on Twitter," *Future Generation Computer Systems*, vol. 114, pp. 120-129, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Sayar Singh Shekhawat, Sakshi Shringi, and Harish Sharma, "Twitter Sentiment Analysis Using Hybrid Spider Monkey Optimization Method," *Evolutionary Intelligence*, vol. 14, pp. 1307-1316, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] A. Naresh, and P. Venkata Krishna, "An Efficient Approach for Sentiment Analysis Using Machine Learning Algorithm," *Evolutionary Intelligence*, vol. 14, pp. 725-731, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Shanta H. Biradar, J.V. Gorabal, and Gaurav Gupta, "Machine Learning Tool for Exploring Sentiment Analysis on Twitter Data," *Materials Today: Proceedings*, vol. 56, no. 4, pp. 1927-1934, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Avinash Chandra Pandey, Ankur Kulhari, and Deep Shikha Shukla, "Enhancing Sentiment Analysis Using Roulette Wheel Selection Based Cuckoo Search Clustering Method," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 1-29, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Meryeme Hadni, and Hassane Hjiaj, "New Model of Feature Selection based Chaotic Firefly Algorithm for Arabic Text Categorization," *The International Arab Journal of Information Technology (IAJIT)*, vol. 20, no. 3A, pp. 461-468, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Ismail Shahin et al., "An Efficient Feature Selection Method for Arabic and English Speech Emotion Recognition Using Grey Wolf Optimizer," *Applied Acoustics*, vol. 205, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Sarah Abdulkarem Al-shalif et al., "A Systematic Literature Review on Meta-Heuristic Based Feature Selection Techniques for Text Classification," *PeerJ Computer Science*, vol. 10, pp. 1-45, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Seyedeh Zahra Mirjalili et al., "Grasshopper Optimization Algorithm for Multi-Objective Optimization Problems," *Applied Intelligence*, vol. 48, pp. 805-820, 2018. [CrossRef] [Google Scholar] [Publisher Link]