HYBRID GAR: A NOVEL APPROACH FOR SENTIMENT ANALYSIS ON TWITTER USING GATED ATTENTION RECURRENT NETWORK

Dr. Manoj Chandak [1], Ravi Thakur[2]

Article History: Received: 05.02.2023 **Revised:** 20.03.2023 **Accepted:** 05.05.2023

Abstract

Sentiment analysis is the most well-known and active field of data mining research. Twitter is one of the most helpful social media tools nowadays for gathering and disseminating people's ideas, emotions, and points of view. thoughts about particular entities. As a result, sentiment analysis in the field of natural language processing became intriguing. Even though a number of approaches for sentiment analysis have been developed, system efficiency and accuracy can always be increased. The suggested architecture is designed to meet it with a deep learning-based sentiment analysis and an efficient and effective feature selection based on optimisation.

The sentiment 140 dataset is used in this study to evaluate the effectiveness of the proposed gated attention recurrent network (GARN) architecture. Pre-processing first purges and filters out the accessible dataset. The Log Term Frequency-based Modified Inverse Class Frequency (LTF-MICF) model is then used to extract the sentiment-based features from the pre-processed data using a term weight-based feature extraction method. In the third phase, a hybrid mutation-based white shark optimizer (HMWSO) is introduced for feature selection. The GARN architecture, which combines recurrent neural networks (RNN) and attention mechanisms, is used to categorise the sentiment classes, such as positive, negative, and neutral, using the selected features.

The performance of the proposed and current classifiers is then compared. The evaluated performance measures are accuracy, precision, recall, and f-measure, and the acquired values for these metrics using the suggested GARN are accuracy, precision, recall, and f-measure, respectively.

Keywords: Twitter sentiment, Deep learning, Gated recurrent attention network , Term weight feature extraction, White shark optimizer ,Twitter sentiment ,Natural language processing, Recurrent neural network

---

[1]Guide (Dean Academics RCOEM), Shri Ramdeobaba College of Engineering and Management (RCOEM), Nagpur,India

[2](MTech CSE RCOEM), Shri Ramdeobaba College of Engineering and Management (RCOEM), Nagpur,India

DOI: 10.31838/ecb/2023.12.s3.495

1. Introduction

Sentiment analysis uses text analysis, natural language processing (NLP), and statistics to evaluate the user's sentiments. SA is sometimes referred to as emotion AI or opinion mining [1]. The word "sentiment" refers to feelings, opinions, or attitudes that have been expressed in relation to a certain individual, situation, or item. To ascertain if the data or information obtained is favourable, neutral, or negative, one NLP approach called SA is applied. It is frequently used by business professionals to monitor attitudes, assess social media data, and gauge brand reputation and customer needs [2] [3].

The volume of material shared or generated online has rapidly increased as a result of the massive growth in Internet users in recent years [4, 5]. Social media platforms like Twitter, Instagram, Facebook, LinkedIn, YouTube, and others have grown in popularity across the globe as a result of technological advancements [6]. People have used social media platforms like Twitter, Instagram, Facebook, LinkedIn, YouTube, and others since the invention of technology [6] to express their ideas or opinions about various products, occasions, or goals.

Twitter [7] is currently the most widely used global microblogging platform allowing users to share their thoughts in the form of brief messages referred to as tweets [7]. Twitter has 152 million daily active users and 330 million monthly active users with 500 million tweets posted everyday [8]. Tweets frequently generate a tonne of sentiment data based on analysis. Twitter is a helpful social media platform for user connection and information sharing. The emotion on Twitter has a significant impact on many aspects of our lives [9]. To extract textual information, sentiment analysis and text categorization further divides polarity into positive (P), negative (N), and neutral (N). Utilising NLP algorithms, information is routinely obtained from text or tweet content. A computer can determine the meaning of each statement that a human has produced using NLP-based sentiment categorization. TSA manual reviewThe (Twitter Sentiment Analysis) process requires a lot of work and more experts. To solve these issues, an automated model is developed. The study of online sentiments has made use of Nave Bayes, Support Vector Machine, Multinomial Nave Bayes, Logistic Regression, and other developments of ML (Machine Learning) techniques [10]. Although these strategies showed good effectiveness, they are very slow and take longer to finish the training process.

The DL model is presented in order to appropriately categorise Twitter sentiments. DL, a subset of ML, uses numerous techniques to solve complicated problems. Due to DL, which uses a succession of progressive events, the system can manage enormous volumes of data with little human input. It can be applied to a wide range of tasks, such as sentiment analysis, product forecasts, product forecasts, and emotion detection [11]. Sentiment analysis with DL yields precise results. As a result of these developments, DL has been used by numerous academics in Twitter sentiment analysis.

Motivation

Sentiment analysis seeks to identify and classify the polarity of the text data. Nowadays, social media is used by many people to convey their ideas and emotions. Due to the massive amount of data that is produced online as a result, it is essential to effectively mine the web data in order to obtain accurate information. Analysis of internet sentiment might result in the creation of a general opinion on specific products. TSA (Twitter Sentiment Analysis) is challenging for a number of other reasons. There is a character limit on short texts, such as tweets, which is a concern. Because misspellings, slang, and emoticons are so common in tweets, there must be an additional pre-processing stage to filter the raw data.

Sentiment categorization would be further impacted by the difficulty of selecting a new feature extraction model. This work seeks to give a new feature extraction and selection process together with a hybrid DL classification model in order to effectively classify twitter sentiment. The current research efforts [12] [13] [14] [15] [16] [17] focus on DL-based TSA, however they haven't yielded significant results because they used fewer datasets and human text tagging took longer. However, incomplete datasets and superfluous data can make the categorization system less successful. A DL approach's performance is further hampered by the dimension occupied by extracted features. In order to address such challenges, the aim of this work is to develop a viable Deep Learning system for doing Twitter Sentiment Analysis.

Pre processing is important in this architecture because it can increase the effectiveness of deep learning by removing unnecessary details from the dataset. The processing time of the feature extraction algorithm is also sped up by this pre-processing. Then, a feature selection process based on optimisation was implemented, which avoids analysing aspects that aren't important in order to save time. In contrast to earlier algorithms, the recommended GARN, however, can investigate the text-based features well. The attention mechanism has been added to the proposed DL algorithm, increasing its overall effectiveness. The attention mechanism has a greater capacity to learn the chosen features as the model's complexity is reduced. Through the integration of the attention mechanism and RNN, this benefit results in efficient performance.

The primary objectives of the proposed study are to:

• Present the Hybrid Mutation-based White Shark Optimizer with a Gated Attention Recurrent Network (HMWSO-GARN), a revolutionary deep model for sentiment analysis on Twitter.
• The feature set is extracted using a novel term weighting-based feature extraction (TW-FE) technique termed Log Term Frequency-based Modified Inverse Class Frequency (LTF-MICF), and it is compared to traditional feature extraction models.
• To identify the polarity of tweets using a bio-inspired feature selection and deep classification model.
• To evaluate the performance using different metrics and compare it to the TSA's standard DL procedures.

Related works

Several studies on DL-based Twitter sentiment analysis include:

Alharbi et al. [12] employed CNN (Convolutional Neural Network), a DNN (deep neural network) based technique, to assess Twitter attitudes. Using a dual method that considered both social interactions and personality variables, tweets were categorised. The sentiment (P, N, or Ne) analysis was shown using the CNN model. Feature lists and pre-trained word embeddings (Word2Vec) are used in the input layer. Processing was done on the dual datasets SemEval-2016_1 and SemEval-2016_2. The current methods performed worse than CNN, which achieved an accuracy of 88.46%. The algorithms with the highest accuracy at the moment are LSTM (86.48%), SVM (86.75%), KNN (k-nearest neighbour) (82.83%), and J48 (85.44%).

Tam et al. [13] developed a Convolutional Bi-LSTM model based on sentiment categorization utilising Twitter data. The CNN-Bi-LSTM integration in this case was defined by the extraction of local high-level features. The input layer receives the text input and tokenizes it. The tokens were changed into numeric values, or NVs. The pre-trained WE (word embedding), including GloVe and W2V (word2vector), was then used to generate the word vector matrix. The essential phrases were extracted using the CNN model, and the feature set was further condensed using the max-pooling layer. Bi-LSTM (backwards, forward) layers were used to grasp the textual context. The dense layer (DeL), which uses weights to link the input data with the output, was added after the Bi-LSTM layer. To test the performance, TLSA datasets were used. The performance was tested using the datasets SST-2 (Stanford Sentiment Treebank) and TLSA (Twitter Label SA). The accuracy for the TLSA dataset was (94.13%), and for the SST-2 dataset it was (91.13%).

Chugh et al. [14] developed an improved DL model for sentiment classification and information retrieval. The hybridised optimisation algorithm SMCA was developed by combining the SMO (Spider Monkey Optimisation) and CSA (Crow Search Algorithm) algorithms. The DeepRNN (presented DRNN) was produced using the SMCA training method. In this case, sentiment categorization was done using DeepRNN-SMCA while information retrieval was handled by FuzzyKNN. Both the telecom tweets dataset and the Amazon mobile reviews dataset were utilised. The first dataset's accuracy for classifying emotions was (0.967), whereas the second dataset's was (0.943). The accuracy of dataset 1's performance utilising information retrieval (IR) increased to (0.831), whereas dataset 2's performance improved to (0.883).Alamoudi et al. [15] carried out successful aspect-based SA and sentiment categorization in relation to WE (word embeddings) and DL. The sentiment categorization algorithm uses both ternary and binary classes.

The YELP review dataset was created and pre-processed before categorization. Glove WE, BoW, and TF-IDF were used to simulate the feature extraction. The first set of features (TF-IDF, BoW features) were originally modelled using the NB and LR, and the glove features for the ternary classification were modelled using a variety of models, including ALBERT, CNN, and BERT. Then, binary SA was conducted using aspects and words. The WE vector for phrase and aspect was produced using the Glove method. The similarity of various features and text vectors was determined using the cosine similarity formula, after which binary aspects were classed. When implemented, the ALBERT model on a YELP 2-class achieved the highest accuracy (98.308%). In contrast, the BERT model improved its accuracy using a YELP 3-class dataset (89.626%).Tan et al. [16] presented the hybrid robustly optimised BERT approach (RoBERTa) with LSTM for the analysis of sentiment data using transformer and RNN. The textual input was analysed using word embedding, and the tokenization of the subword was described using the RoBERTa model. The long-distance Tm (temporal) dependencies were encoded using the LSTM model. The DA (data augmentation) based on pre-trained word embedding was devised to aggregate many lexical samples and depict the minority class-based oversampling. Processing DA corrects an unbalanced data collection with additional lexical training samples. The Adam optimisation technique was used to tune the hyperparameters, and the resulting SA results were better. The implementation datasets were Sentiment140, Twitter US Airline, and IMDb. Accuracy was increased by 89.70%, 91.37%, and 92.96% overall using these datasets.

Hasib et al. [17] recommended a special DL-based sentiment analysis of Twitter data for the US airline business. The Kaggle dataset was used to compile the crowd sourced Twitter US airline sentiment tweet. Convolutional neural network (CNN) and DNN models are both used in the feature extraction process. Before adding the four layers, the tweets are transformed to metadata and tf-idf. The four layers of DNN are input, covering, and output layers. For feature extraction, the stages of data pre-processing, embedded features, CNN, and integration features are utilised. The total precision, recall, and f1-score are, respectively, 85.66%, 87.33%, and 87.66%. Sentiment analysis was used to identify the attitude expressed in text samples. To recognise these feelings, Carvalho and Guedes developed a novel phrase weighting scheme.

Sentiment analysis was used to identify the attitude expressed in text samples. These perspectives were discovered using the unsupervised weighting scheme (UWS), a brand-new phrase weighting approach developed by Carvalho and Guedes in [18]. It has the ability to process data without accounting for the weighting component. Additionally, the SWS (Supervised Weighting Schemes), which use the class information related to the computed term weights, was introduced. It had shown a more promising result when compared to the present weighting methods.

Authors and Years	**Methodology**	**Merits**	**Demerits**
Alharbi et al.2019 [12]	CNN	The behavioural information of the user is added	Difficult to interpret the exact tweet from a group of tweets
Tam et al.2021 [13]	Hybrid CNN-BiLSTM	The performance of the word embedding method is good	Lower retrieval and classification accuracy
Chugh et al.2021 [14]	DRNN	Provides better reviews to take effective decisions	Lower performance accuracy
Alamoudi et al.2021 [15]	CNN, BERT and ALBERT models	Error rate decreased	many mislabelled reviews
Tan et al. 2022 [16]	BERT approach(RoBERTa) with LSTM	Optimization is done using the word embedding method	Lower classification accuracy
Hasib et al.2021 [17]	CNN and DNN	Collected data on the emotions of the airline consumers	Less number of tweets are used

Problem Statement

There are many problems with using DL techniques for sentiment analysis on Twitter.

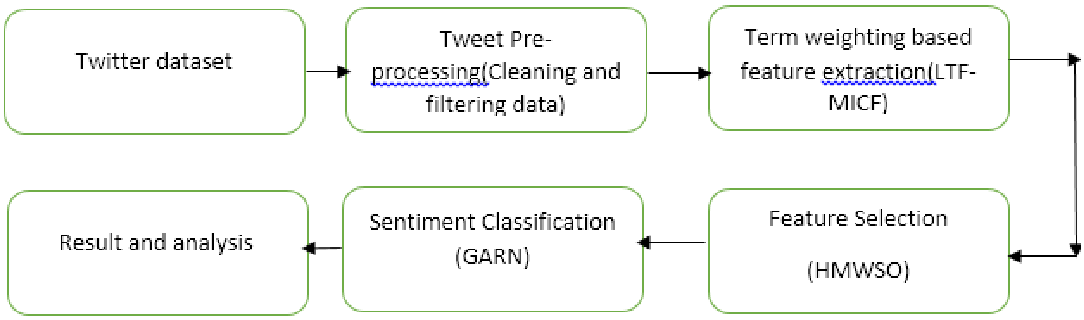
Using the DL model, the author in [16] was able to categorise sentiment from Twitter data. In order to categorise such data, this technique looked at each user's behavioural data. However, this method has had trouble extracting specific tweet words from the vast twitter corpus; as a result, a classification algorithm's performance has been reduced. ConvBiLSTM, a tool for sentiment classification that used word2vec and glove-based features, was introduced in [13]. However, the retrieved features are insufficient to achieve the desired degree of precision. After that, in [14], which uses Deep RNN for sentiment analysis, processing time reduction was given priority. It cannot, however, reduce the size that the extracted features occupy.

It includes two classes as well. Some of the current research struggles to achieve efficient processing speed, complexity, and accuracy due to the availability of large datasets. Furthermore, the extraction of trivial and low-level properties reduces the classifier's efficiency. The utilisation of all extracted attributes also consumes a lot of storage space. These flaws make the earlier algorithms processing-inefficient. Researchers can now concentrate on creating efficient combination algorithms for the processing of Twitter data as a result of these issues. To solve this issue, the proposed solution employs both RNN and an attention technique. The features required for classification are retrieved using LTF-MICF, which provides features for twitter processing. The huge retrieved features' occupied dimension is subsequently minimised using the HMWSO method. This algorithm offers a superior method for choosing the best characteristics and can examine the features more fast. With a smaller misclassification error rate and improved performance across the large dataset, the recommended classifier is now able to accurately classify data.

2. Proposed Methodology

A gated attention recurrent network Deep Learning approach [19]is suggested for classifying the sentiment of Twitter tweets. The public Twitter dataset (Sentiment140 dataset) including sentimental tweets is provided as input. The method then moves on to pre-processing the tweets. Tokenization, number removal, stopword removal, slang and acronym correction, punctuation, stemming and symbol removal, uppercase and lowercase replacement, character &URL, hashtag & user mention removal are all done during the pre-processing stage. The previously processed dataset is now used as input for the next procedure. The Log Term Frequency-based Modified Inverse Class Frequency extraction technique assigns a term weight for each term in the dataset based on term frequency. Next, the best term weight is chosen using the Hybrid Mutation based White Shark Optimizer (HMWSO). The output of the HMWSO is then used to classify sentiment into three categories using the gated attention recurrent network (GARN). A diagrammatic illustration of the suggested methodology is shown in Figure.

Fig. Flow of processing



Tweets Pre-Processing

Pre-processing entails condensing large data to short text in order to carry out subsequent processes like classification, undesired news detection, sentiment analysis, etc. as Twitter users utilise various formatting ways to send their tweets. Some individuals might tweet utilising hashtags, URLs, symbols, abbreviations, and punctuation. Tweets can also contain emoticons, stickers, and emojis to express the user's emotions. Sometimes the tweets will adopt a hybrid format by incorporating URLs, symbols, and abbreviations. Therefore, these types of symbols, abbreviations, and punctuation should be removed from the tweet in order to further categorise the dataset. The elements that need to be removed from the tweet **include** tokenization [28], stop word removal [20], stemming [20], acronym correction [21], number removal, punctuation and symbol removal [21] [22], noise removal, URL[22], hashtags, and replacement [21] [22].

Term weighting‑based feature extraction

The pre-processed data is extracted in text documents based on the term weighting Tw [23] after pre-processing. This study uses a novel term weighting method called Log term frequency-based modified inverse class frequency (LTF-MICF) to extract terms-based features. The method combines the modified inverse class frequency (MICF) and the log term frequency (LTF), two different term weighting systems. Term frequency fT refers to the terms that appear most frequently in the text. However, fT is insufficient on its own because frequently occurring phrases will weigh down the paper. The proposed hybrid feature extraction method can thereby resolve this problem. As a result, fT is combined with MICF, a successful Tw strategy.The ratio of the entire class of terms that occur to the inverse class frequency, ^{f}Ci is the inverse ratio of the total class of terms that occurs on training tweets to the total classes. The algorithm for the TW-FE technique is shown in algorithm [23].

Algorithm: Log term frequency-MICF BASED $\boldsymbol{T_w}$

Input: p-Pre-processed dataset

Output: Weight ranks for each term in the pre-processed dataset

Begin

for each term tp (p=1,2 ……m) and each class, C_r(r=1,2…..m) do

Compute MICF(tp), $w_{sp}.\lfloor C_i(tp)\rfloor$

Compute w_{sp} and $^{f}C_{\mathrm{i}}$ (tp) using equations 3 and 5

End for

for each term, tp (p=1,2 ….m) do

Compute $^{l}T_{\mathrm{f}}$ (tp) using equation 4

Compute ltf-MICF(tp) using equation 2

End for

Return ltf-MICF(tp) (p=1,2….m)

End

ltf lTf must be calculated in two steps. The pre-processed dataset's terms' fTs are calculated in the first phase. The output of the computed fT data is log normalised in the second phase. For each phrase in the manuscript, the MICF, a modified version of (_ f)Ci, is calculated. When MICF is claimed to be executed, each term in the document should have a unique class-specific rank and make a unique contribution to the overall term rank. Different weights must be given to different class-specific ranks. As a result, the overall term rank is determined by adding the weights assigned to each class-specific rank. The following is a representation of the suggested formula for Tw utilising an ltf-based MICF [23];

$$\textbf{LTF MICF(tp)}=\boldsymbol{^{l}T_{\mathrm{f}}} * \sum_{r=1}^{m} \boldsymbol{W_{sp}} - \left[\boldsymbol{^{f}C}\mathbf{i}\ (\mathbf{tp})\ \right]$$

where a specific weighting factor is denoted W_{sp} for each tp for class C_r , which can be clearly represented as;

$$\boldsymbol{w_{sp}}=\log(1+\frac{S_i.\vec{t}}{\max(1,S_i.\overleftarrow{t})}.\frac{S_i.\tilde{t}}{\max(1,S_i\hat{t})})$$

The method used to assign a weight for a given dataset is known as the weighting factor. Where the number of tweets s_i in class C_r which contains pre processed terms tp is denoted as $s_{i\ t}$.

$S_i.\overleftarrow{t}$ means the number of s_i in other classes, which contains tp.

$S_i\hat{t}$ means the number of s_i in-class C_r , which do not possess, tp.

$S_i.\tilde{t}$ means the number of s_i in other classes, which do not possess, tp is denoted as

The constant '1' is used, to eliminate negative weights, In extreme cases, to avoid a zero-denominator issue, the minimal denominator is set to '1' if $S_i.\overleftarrow{t} = 0$ or $S_i\hat{t}$=0. The formula for ${}^{l}T_f$ (tp) and ${}^{f}C$i (tp) can be presented as follows[22].

$$\boldsymbol{{}^{l}T_f}\ \mathbf{(tp)=log(1+{}^{f}T(tp,s_i))}$$

${}^{f}T(tp,s_i)$ means the raw count of tp on s_i or we can say that the total times of tp occurs on s_i.

$$\boldsymbol{{}^{f}C}\mathbf{i\ (tp)\ =log\left(1+\frac{r}{C(tp)}\right)}$$

where r refers to the total number of classes in s_i , and C(tp) is the total number of classes in tp.

Feature Selection

The presence of irrelevant features in the data can lower the classification process' accuracy level and force the model to pick up on such irrelevant features. This problem is known as the optimisation problem. Only by using the best options from the dataset that has been processed can this problem be avoided. In order to accomplish a feature selection procedure, a feature selection method called White shark optimizer with a hybrid mutation strategy is used.

White shark optimizer(WSO)

The WSO is based on the foraging behaviour of the white shark [24]. The great white shark in the sea hunts for meals by manipulating the waves and other oceanic elements to reach prey held in the depths. Since the white shark hunts for prey based on three behaviours, namely: (1) the shark's speed while pursuing the prey, (2) looking for the greatest ideal food source, and (3) other sharks moving towards the shark when it is close to the best food source, the white shark. The first-generation white shark population is shown as;

$$\boldsymbol{o_q^p = lb_q + r \times (up_q - lb_q)}$$

where $\boldsymbol{o_q^p}$ denotes the pth white shark's initial parameters in the q_{th} dimension. The terms up_q and lb_q, respectively, stand for the upper and lower boundaries in the q_{th} dimension. While r stands for a random number between [0, 1].

Gated attention recurrent network (GARN) classifier

A Bi-GRU hybrid network with an attention mechanism is called GARN. Because recurrent neural networks (RNNs) use outdated information rather than current information for classification, many issues arise when they are used. A bidirectional recurrent neural network (BRNN) model that can use both recent and historical data is suggested as a solution to this issue. Therefore, two RNNs are used to carry out the forward and reverse functions. To record the feature sequence, the output will be connected to a comparable output layer. Another bidirectional gated recurrent unit based on the BRNN paradigm.

The (Bi-GRU) model is presented, which swaps out the BRNN's hidden layer for a single GRU memory unit. This hybridization of Bi-GRU and attention is referred to as a generalised attention recurrent network (GARN) [25].

Consider an m-dimensional input data as (y1 , y2 , ..., y m). The hidden layer in the BGRU produces an output P t1 at a time interval t1 is represented as;

$$\vec{P}_{t1}=\sigma\left(W_{e\overrightarrow{yP}}Yt_1+W_{e\overrightarrow{PP}}\vec{P}t_1-1+c_{\vec{P}}\right)$$
$$\overleftarrow{P}_{t1}=\sigma\left(W_{e\overleftarrow{yP}}Yt_1+W_{e\overleftarrow{PP}}\overleftarrow{P}t_1-1+c_{\overleftarrow{P}}\right)$$

$$\mathbf{P}t_1=\vec{P}_t\oplus P_{t1}$$

Where the weight factor for two connecting layers is denoted as w_e , c is the bias vector, σ represents the activation function, positive and negative outputs of GRU is denoted as $\vec{P}_{t1}$ *and* $\overleftarrow{P}_{t1}$ $\oplus$ is a bitwise operator.

Attention mechanism

The attention module in sentiment analysis is crucial for indicating the relationship between the terms in a phrase and the result [26]. In this concept, an attention model called the feed-forward attention model is used for direct simplification. This condensing creates a single vector v from the entire sequence denoted by;

$$E_{t_1}=b(H_{t_1})$$

$$\beta_{t_1}=\frac{\exp(E_{t_1})}{\sum_{s=1}^{R}\exp(E_s)}$$

$$v=\sum_{t1=1}^{R}\beta_{t_1}H_{t_1}$$

where H_{t_1} is used to identify as a learning function(β). By monitoring the average weight of the data sequence H, the attention mechanism determines a fixed length for the embedding layer in a BGRU model for each and every vector v. As a result, the classification's final sub-set is derived from:

$$H^{\#}=\tanh(v)$$

Sentiment classification

Formally, Twitter sentiment analysis is a classification issue. The suggested method divides the sentiment data into three categories: neutral, negative, and positive. The output of the hidden layer H# is classified using the softmax classifier.

3. Results And Discussion

Performance measurements including accuracy, precision, recall, and f-measure are briefly discussed in this section. Pre-processing, feature extraction, feature selection, and classification are all included in the overall analysis of the Twitter sentiment classification, which is also examined and fully presented. Results on comparing term weighting methods with existing and trending classifiers are shown in bar graphs and tables. After briefly discussing the entire methodology, the research was finished by importing the performance metrics that had been assessed. Individuals' expressions of sentiment are dependent on their opinions on many topics. Positive and negative attitudes are primarily identified in tweet-based sentiment analysis. Consequently, it is required to improve the categorization classes so that the datasets include a neutral class.

Dataset

Sentiment 140, a dataset obtained from Kaggle that contains 1,600,000 tweets taken from the Twitter API, is the one used in our suggested research. The score for each tweet is 4, which is the rank value for affirmative tweets. Similar to this, the rank value for negative tweets is 0, and the rank value for neutral tweets is 2. A dataset contains 20832 good tweets, 18318 neutral tweets, 22542 negative tweets, and 12990 irrelevant tweets. A total of 70% of the dataset is utilised for training, 15% is used for testing, and 15% is used for validation.

Performance metrics

In this suggested technique, the performance metrics of precision, f1-score, recall, and accuracy are compared against 4 alternative weight schemes and other existing and new classifiers. True-positive, True-negative, False-positive, and False-Negative are the four notations.

Analysis of Twitter sentiments using GARN

The major objective of the study work is to categorise Twitter feelings into three categories: positive, negative, and neutral. Twitter's api is used to gather the information. Data is gathered and then provided as input for pre-processing. A new pre-processed dataset is produced after the pre-processing technique removes the undesirable symbols. The pre-processed dataset is now provided as input so that the necessary characteristics may be extracted. A unique method known as the log term frequency-based modified inverse class frequency (LTF-MICF) model, which combines the two weight schemes LTF and MICF, is used to extract these characteristics from the pre-processed dataset. The necessary features are extracted in this case, and the extracted features are then used as input to choose the best feature subset.

The hybrid mutation-based white shark optimizer (HMWSO) is used to choose the optimised feature subset. The Cauchy mutation and the Gaussian mutation are two names for the mutation. The feelings are then divided into three classes based on the inputted feature sub-set using a classifier called the gated recurrent attention recurrent network (GARN), a mix of the Bi-GRU and an attention mechanism. It is preferred to categorise the sentiments of Twitter tweets using the evaluated value of the proposed GARN.

The sentiment140 Twitter dataset is used to train the suggested GARN model, which is implemented in the Python environment. The proposed classifier is compared to existing classifiers, including CNN (Convolutional neural network), DBN (Deep Brief Neural Network), RNN (Recurrent Neural Network), and Bi-LSTM (Bi-directional Long Short Term Memory), in order to assess the effectiveness of the classifier. The suggested term weighting scheme (LTF-MICF) is compared to the current term weighting schemes TF (Term Frequency), TF-IDF (Term-frequency-inverse document frequency), TF-DFS (Term-frequency-distinguishing feature selector), and W2V (Word to vector). Both sentiment categorization methods—using an optimizer and not—had their performance assessed. Accuracy, precision, recall, and f1-score are the metrics under consideration.

The currently used, suggested, and implemented (GARU) algorithms are the Bi-GRU, RNN, Bi-LSTM, and CNN. Table 1 discusses the simulation parameters that were utilised to process the suggested and existing approaches. This comparative analysis is done to demonstrate how effective a proposed algorithm is compared to other relevant current algorithms. The accuracy of the GARN and the current classifiers are compared in Table 1. For the LTF-MICF, the accuracy of the existing Bi-GRU, Bi-LSTM, RNN, and CNN is 96.93%, 95.79%, 94.59%, and 91.79%, respectively. In comparison, the proposed GARN classifier is thought to be the most accurate for categorising Twitter sentiments with the LTF-MICF word weight scheme, achieving an accuracy of 97.86%. However, the suggested classifier's accuracy is 97.53%, 97.26%, 96.73%, and 96.12% when compared to various term weighting schemes, such as TF-DFS, TF-IDF, TF, and W2V. The term weight scheme combined with the GARN classifier is thus the ideal approach to solving classification issues.

Method	Layers	Value
GARU (proposed)	Bi-directional GRU	500
	Bi-directional GRU	250
	Attention Layer	Size of Bi_GRU
	Dropout	0.2
	Dense	100
	Dropout	0.2
	Dense	1
	Dense	1
CNN	Input layer	69769*1000
	Covolution layer	Filter=2, Kernel size=2
	Maxpooling layer	Pool size=2
	Flatten	Size of maxpool
	Dense layer	1

Fig: Table1

Table 2: Performance between proposed and existing methods for developed objective

classifier	Accuracy	Precision	Recall	F1-measure
Proposed GARN	97.88	96.65	96.76	95.70
Bi-GRU	95.1951	92.82	93.36	92.7
Bi-LSTM	95.7957	93.6148	94.07	93.613
TCN	95.9302	94.221	94.519	94.275
Transformer	96.007	95.8	95.76	94.89
BERT	95	94.32	94.67	93.62

This is the cumulative measures.

Table 3: Performance comparison between developed existing and proposed methods

Ref no and author name	Technique	Dataset	Performance metric
Proposed	GARN	Sentiment 140 dataset	Accuracy-97.88 %
Alharbi et al. 2019 [12]	CNN	SemEval-2016 1, and SemiEval-2016 2	Accuracy-86.48%,
Tam et al. 2021 [13]	ConvBiLSTM	Retrieved Tweets and SST-2 datasets	Accuracy-91.13%,
Chugh et al. 2021 [14]	DeepRNN-SMCA	Amazon unlocked the mobile reviews dataset, Telecom tweets	Accuracy-97.7%,
Alamoudi et al. 200021 [15]	ALBERT	Yelp Dataset	Accuracy-89.49%,

4. Conclusion

In this study, GARN is used to find the varied viewpoints of Twitter users who use the online platform. The Sentiment 140 dataset was used to carry out the implementation. Four performance metrics—accuracy, precision, f-measure, and recall—along with four-term weighting schemes—LTF-MICF, TF-DFS, TF-IDF, TF, and W2V—are used to compare the performance of the top GARN classifier with that of other DL models, including Bi-GRU, Bi-LSTM, RNN, and CNN. The analysis demonstrates that the top GARN DL approach achieved the desired level for classifying Twitter sentiment. Moreover, by combining the recommended term weighting scheme-based feature extraction technique with the top GARN classifier, an effective result for twitter feature extraction was obtained. The GARN accuracy while using LTF-MICF on the Twitter dataset is 97.88%.

The proposed classifier has the greatest accuracy rating of all the ones already in use. The recommended GARN classifier is also regarded as a useful DL classifier for sentiment analysis applications on Twitter and other platforms. The proposed model has produced acceptable results, but it hasn't reached the necessary level. This is due to the planned architecture's failure to accord the chosen elements equal importance. As a result, a few crucial features are missed, which reduces the effectiveness of the suggested model. Therefore, in the future, we will introduce an efficient DL method with the best feature selection method for identifying visual emotion using all the selected features.

5. References:

S. S. Saberi B, “Sentiment analysis or opinion mining: a review,” vol. 7, no. 5, 2017.

H. A. K. H. Medhat W, “Sentiment analysis algorithms and applications: a survey. Ain Shams Eng J.,” 2014.

K. H. Drus Z, “ Sentiment analysis in social media and its application: systematic literature review.,” *Procedia Comput Sci.,* 2019.

R. J. Zeglen E, “Increasing online information retention: analyzing the effects.,” *J Open Flex Distance Learn. ,* 2018.

D. X. Y. Q. M. B. Y. H. Qian Yu, “On detecting business event from the headlines and leads of massive online news articles.,” *Inf Process Manage,* 2019.

G. Neubaum, “Monitoring and expressing opinions on social networking sites–Empirical investigations based on the spiral of silence theory.,” 2016.

L. M. W. F. D. Y. Karami A, “Twitter and research: a systematic literature review through text mining.,” *IEEE Access. ,* 2020.

F. P. I. S. Antonakaki D, “A survey of Twitter research: data model, graph structure, sentiment analysis and attacks.,” *Expert Syst Appl,* 2021.

K. M. B.-H. A. Birjali M, “A comprehensive survey on sentiment analysis: approaches, challenges and trends. .,” *Knowl-Based Syst,* 2021.

K. O. R. A. G. S. S. A. Yadav N, “Twitter sentiment analysis using supervised machine learning. Intelligent data communication technologies and internet of things.,” *Springer;,* p. p. 631–42., 2021.

K. P. B. G. K. G. Gandhi UD, “Sentiment analysis on Twitter data by using convolutional neural network (CNN) and long short term memory (LSTM).,” *Wirel Pers Commun,* 2021.

d. D. E. Alharbi AS, “Twitter sentiment analysis with a deep neural network: an enhanced approach using user behavioral information,” pp. 50-61, 2019.

S. R. Ö. T. Ö. Tam S, “A ConvBiLSTM deep learning model-based approach for Twitter sentiment classification,” *IEEE*

Access, pp. 83-93, 2021.

S. V. K. S. N. A. Q. B. B. M. J. C. Chugh A, "Spider monkey crow optimization algorithm with deep learning for sentiment classification and information retrieval.," 2021.

A. N. Alamoudi ES, " Sentiment classification and aspect-based sentiment analysis on yelp reviews using deep learning and word embeddings.," pp. 259-81, 2021.

L. C. A. K. L. K. Tan KL, "RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network.," *IEEE Access.*, pp. 17-25, 2022.

K. M. M. A. H. N. A. T. M. I. H. S. Hasib, "A Novel Deep Learning based Sentiment Analysis of Twitter Data for US Airline Service.," in *International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), pp.*, 2021.

G. Carvalho F. Guedes, "TF-IDFC-RF: a novel supervised term weighting scheme.," 2020.

N. C. P. H. B. e. a. Parveen, "Twitter sentiment analysis using hybrid gated attention recurrent network," *J Big Data*, 2023.

C. A. K. S. G. S. A. P. Ahuja R, "The impact of features extraction on the sentiment analysis," *Procedia Comput Sci*, 2019.

N. M. V. K. R. G. B. P. T. B. Gupta B, "Study of Twitter sentiment analysis using machine learning algorithms on Python," *Int J Comput Appl*, vol. 165, no. 9, pp. 29-34, 2017.

K. M. M. G. T. S. A. u. M. L. Ikram A, "12th International Conference on Cloud Computing, Data Science & Engineering(Confluence) pp," *IEEE*, pp. 629-634, 2022.

L. Z. Y. X. Y. Q. Zhao H, "A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. Inf Process Manage," *science direct*, vol. 58, no. 5, 2021.

A. H. J. A. M. A. A.-B. M. A. A. Malik Braik, "White Shark Optimizer: A novel bio-inspired meta-heuristic algorithm for global optimization problems,," *science direct*, vol. 243, 2022.

R. W. S. L. Zeng L, " Attention-based bidirectional gated recurrent unit neural networks for well logs prediction and lithology identification," 2020.

Y. Z. T. W. W. Q. R. M. Niu Z, "Wind power forecasting using attention based gated recurrent unit network," 2020.

Z. D. W. A. Gaye B, "A Tweet sentiment classification approach using a hybrid stacked ensemble technique. Information.," 2021.

O. B, " Information sharing on social media sites," *Comput Hum Behav.*, 2013.

P. AP, "Performance Evaluation and Comparison using Deep Learning Techniques in Sentiment Analysis.," *Journal of Soft Computing Paradigm (JSCP)*, 2021.