ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2024.0491

AN OPTIMIZED BIDIRECTIONAL CONVOLUTIONAL RECURRENT NEURAL NETWORK ARCHITECTURE WITH GROUP-WISE ENHANCEMENT MECHANISM OF SENTIMENTS FOR THE PERSPECTIVE OF CUSTOMER REVIEW SUMMARIZATION

V. Sriguru and D. Shanmuga Rajathi
Department of Computer Science, RVS College of Arts and Science, India

Abstract

Customer reviews play pivotal roles in consumers' purchase decisions, but the sheer volume of text data can be overwhelming. In existing system, while ensemble methods can enhance performance, the associated computational complexity and resource intensiveness should be carefully considered, and appropriate measures should be taken to address these challenges in the context of Customer Review Summarization. This study introduces a Particle Swarm Optimization (PSO) based optimized architecture for a Bidirectional Convolutional Recurrent Neural Network (BiCRNN) with a group-wise enhancement mechanism tailored for Customer Review Summarization and named as OBiCRNN. This model is designed for the perspective of customer review summarization, aiming to effectively capture sentiments and generate concise summaries. The integration of PSO optimizes the network parameters, enhancing the learning process. Feature extraction is done by Modified Principal Component Analysis (MPCA) which uses correlated feature sets and extracts most informative features for given datasets. BiCRNN utilizes bidirectional LSTM and GRU layers for comprehensive context understanding, while the group-wise enhancement mechanism categorizes sentiment-related features, amplifying essential sentiments and attenuating less relevant ones. With this novel approach, the architecture leverages both PSO and BiCRNN for an advanced framework in customer review summarization where outcomes demonstrate the effectiveness of the deep learning (DL) model in producing coherent and informative summaries, enhancing the accessibility of customer feedback for both consumers and businesses. The study contributes to the field of natural language processing (NLP) and customer sentiment analysis, offering a scalable solution for managing the wealth of information present in online customer reviews.

Keywords:

Customer Review Summarization, Particle Swarm Optimization (PSO), Feature extraction, Bidirectional Convolutional Recurrent Neural Network (BiCRNN)

1. INTRODUCTION

In computational linguistics, sentiment analysis has been a well-researched area of study. Machine learning (ML) communities are now actively researching sentiment analysis of customer evaluations due to the enormous rise of online material. However, traditional approaches don't produce the desired results because of the wide range of items being evaluated online. It is crucial to create a strong sentiment analysis model that can extract product attributes and identify sentiments while adhering to different accuracy metrics as the number of reviews increases. This work develops hierarchical bidirectional recurrent neural network (HBRNN) for characterizing sentiment specific aspects from available DBS Text Mining Challenge's review data. The schema is used for predicting sentiments as vector at review level [1] and optimizes network's parameters. The schema is also compared with Long-Short Term Memory (LSTM) and bidirectional LSTM (BLSTM) for its performce.

Text classification, which aims to assign texts to one or more classes, is one of the most common tasks in the aforementioned topic. This activity can be completed computationally or manually. In this regard, categorizing the feelings of remarks from discussion boards, review websites, and social media has garnered a lot of attention in recent years [2]. Sentiment analyses are computational processes that employ statistics and NLP techniques to identify and classify opinions expressed in text while specifically determining attitudes towards topics or products as positive, negative, or neutral [3]. The work [4], used sentiment and statistical analysis to examine customer feedbacks on women's clothes in e-commerce where dataset's non-text review variables (such as age, the type of dress bought, etc.) were examined in an effort to identify any relationships with product recommendations from customers. Next, a bidirectional Recurrent Neural Network (RNN) using LSTM is used to identify user attitude regarding the product and determine whether a review text promotes the purchased product or not.

CRS has become an essential necessity for entrepreneurs looking to improve their goods and services. The number of people ordering food, electronics, clothing, and other items online has skyrocketed since the Covid-19 scare. As a result, buyers have occasionally left many online reviews. These evaluations include important details about the demands and happiness of the consumer. Most business establishments including hotels and film industry use CRS for better productivity by methodically examining the ratings of individual products. The reviews also help with the thorough examination of rivals. However, the number, diversity, truthfulness, and pace at which reviews are continuously created online make them difficult to mine. The first step in [5] is the aspect-based representation, which is utilized to show prioritized information on aspect opinion that is determined by utilizing opinion strength, polarity, and frequencies. The second stage is reviewing summary creation, which ranks aspects according to their information to automatically generate review summaries. The use of the natural language generation approach makes the resultant summary more logical.

Researchers have spent a lot of time developing modelling frameworks that may be used to mine informational corpora using sentiment analysis, NLP, artificial intelligence (AI), information extraction, and retrieval. Many solutions have been developed during the last decade, but the CRS's enormous volume, accuracy, and scalability present considerable challenges. Sentiment analyses can be executed on words, sentences, phrases and documents [6]. Identifying the sentiment of a person, product or feature, brand, or other entity is the first step in word-level sentiment analysis. For example, the review "I am happy with the iPhone's battery's performance" expresses satisfaction with the iPhone's battery. Sentiment analyses of phrases seek sentiments in phrased words. Sentence level sentiment analysis involves determining its general sentiments. Lastly, document-level analysis uses weighted or averaged techniques on one or more phrases to estimate the overall mood.

CRS employing optimization-based BiCRNN algorithm is the main focus of this research work. The computational complexity of the summarizing process remains the primary issue despite the introduction of several studies and approaches. Inaccurate classification findings are a shortcoming of the current methods. This work's OBiCRNN addresses above detailed issues using Text pre-processes, feature extractions, and classifications, primary contributions, for enhancing overall CRS performances. Using efficient algorithms, the suggested approach yields more precise findings for the provided dataset.

The remainder of this paper is structured as follows: Section 2 provides a concise overview of some of the literature on text pre-processing, feature extraction, and CRS techniques. Section 3 provides specifics on the suggested technique for the OBiCRNN algorithm. Section 4 presents the experimental data and a discussion of the performance analysis. Lastly, Section 5 provides a summary of the findings.

2. RELATED WORK

Nainwal et al. [7] investigated usages of ML, NLP, and DL techniques for text summarizations specifically for product reviews. The methodology used in this study involved applying text summarization techniques to condense product review text into a more manageable and informative format. The major findings of this work suggest text summarizations as useful tools for reducing learning times and increasing the amount of important information from product review data. The implications of this study suggest that text summarization can be a valuable tool for efficiently processing and understanding large amounts of online product review data

Porntrakoon et al. [8] discussed customers' evaluations of products or services that might assist others decide on their purchases. It takes longer to read and comprehend the reviews' perspectives since Thai reviewers usually write lengthy, multifaceted assessments, particularly in the food and beverage industry. Analyses of Thai sentiments can produce review summaries based on cuisine, environments, and service aspects where these summaries can help users comprehend original reviewers' views while saving time. This study presented a technique for analyzing sentiments of Thai foods and producing summaries that include both positive and negative attitudes. The findings demonstrate that our suggested approach may produce 1,876 review summaries, or 46.68%, from 4,000 original reviews. Only 5.13% of the original reviews are contained in the review summaries, which have an accuracy rate of 65.25% for Type 1 and 85.05% for Type 2.

Wu et al. [9] recommended Aspect-based Opinion Summarization (AOS) of product reviews. Aspect extractions and sentiment categorizations are sub-tasks AOS systems need to in practice. Most aspect extraction techniques now in use, such as topic modelling or language analysis, are generic for many items but insufficiently accurate or appropriate for specific ones. Rather, we employ a more specific yet less generic system that explicitly maps every review language into pre-established features. The work suggested the usage of two Convolutional Neural Network (CNN) based techniques namely cascaded CNN and multitask CNN for addressing aspect maps and sentiment classifications where the former used two tiers of convolutional networks in a cascaded CNN which handled sentiment categorizations using single CNN at level 2, while the latter used many CNNs at level 1 for aspect maps. Multiple aspect CNNs and an emotion CNN are also included in multitask CNN; nevertheless, word embeddings are shared across networks. The work found both the CNN techniques with pre-trained word embedding outperformed linear classifiers in experiments with multitask CNNs often outperforming cascaded CNNs.

In [5], the authors suggested a system for Automatic Sentiment Summarization (ASS). There are two stages to this method. The first stage is the aspect-based representation, which uses frequencies, polarity, and opinion strength to express prioritized information on aspect opinion. The second stage is reviewing summary creation, which ranks aspects according to their information to automatically generate review summaries. Using a natural language generation approach makes the resulting summary more logical. Additionally, users may update the resulting summary by adding fresh reviews in the same domain using the suggested ASS system. Benchmarks from the sentiment aspect dataset, including consumer evaluations of Canon, Nikon, and laptop products and services, were employed in the tests. When compared to existing systems that use extractive and abstractive summarizing, the produced summaries from the suggested ASS method perform admirably.

Pecar et al. [11] used neural networks for reviewing sentiments of customers while investigating the impact text pre-processes on sentiment categorizations. They showed that text pre-processing significantly affected sentiment analysis performances and when compared to baseline SVM models, their schema produced extremely encouraging results for model training on a medium-sized dataset.

Mandal et al. [12] cantered on the method for obtaining the information (text) in a condensed or summarized form. The idea of PSO plus sentiment analysis has been applied to accomplish this aim. The results of using PSO in the field of text summarization are astounding. Sentiment analysis (SA), in addition to PSO, has demonstrated its significance in the same field of study.

Bompotas et al. [13] created architecture and put into practice a system that uses certain pre-processing methods at first, as traditional approaches to NLP, such as word embeddings and the TF-IDF bag of words, are used. These methods can also be used as input for LSTM Neural Networks and other classifiers. These classifiers' accuracy is assessed using a dataset that includes several reviews for the widely used assessment metrics of precision, recall, and Fl-measure.

Gautam et al. [14] developed the multi-review summarizing technique to obtain a condensed product review. Extracted summary reviews of e-commerce sites including Amazon and Flipkart were fed to the deep neural network-based model. The product's characteristics were extracted from many reviews using a deep neural network, which then clustered the phrases according to the learnt features. Following clustering, sentences are ranked, and the top n sentences from each cluster are chosen to create an extracted summary.

3. PROPOSED METHODOLOGY

This proposed OBiCRNN enhances performances of CRS using the SemEval-2014 restaurant reviews dataset. Text pre-processing, feature extraction, and classification are this study's primary contributions. The suggested system's overall block diagram is displayed in Fig.1.

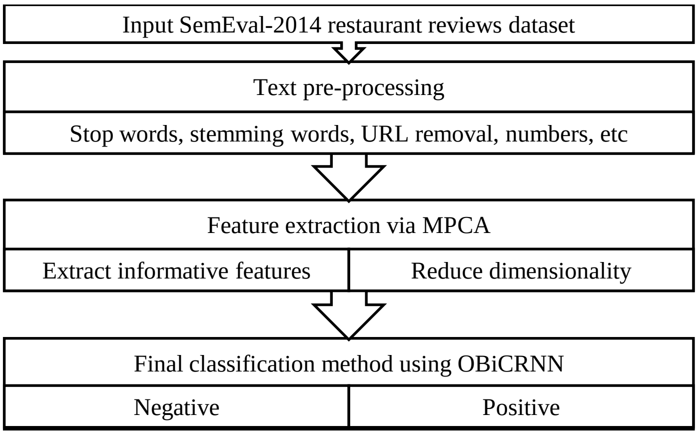

Fig.1. Block diagram of the proposed system

3.1 DATASET COLLECTION

The SemEval-2014 food blogs database is used for evaluating the HAS strategy with 800 exam assessments and 3000 samples for learning. One or more feature keywords have been labelled in each evaluation that is part of this database. The 1.6 million tweets in the Sentiment140 dataset have been sorted by emotional tone, with 0 denoting negativity, 2 neutrality, and 4 positivity. The dataset samples were split into 70% training and the rest for testing for the model.

Initially, datasets from online reviews were obtained and preprocessed with NLP based techniques, including stemming, stop word/ URL eliminations. features were extracted using MPCA. ARF and RRF values were then combined to generate hybrid feature vectors and fed to OBiCRNN for classifying input reviews as either positive or negative. In addition, NLP techniques were used on pre-processed input reviews for semantic analysis of reviews before feature extractions utilizing ARF, RRF and definitional approaches.

3.2 TEXT PRE-PROCESSING

Text messages on Twitter may be rather distracting. Furthermore, Twitter data is unstructured, adding still another degree of difficulty. The major goal of this research is to investigate different pre-processing algorithms for Twitter tweets utilized in the SemEval-2014 restaurant ratings. The primary purpose of this project is to clean up Twitter data in preparation for future investigations. Thus, the purpose of this pre-processing exercise is to simply eliminate all of the useless letters and concentrate on the remaining valuable words.

- **Tokenization:** Tokenization are processes of dividing texts into distinct parts, or tokens. The token might be made up of words, phrases, or anything else that has meaning.
- **Stemming:** This procedure involves removing the term's affixes and suffixes in order to determine the term's foundation or roots.
- Eliminating punctuations, ASCII characters, digits and symbols: Twitter messages frequently include mixes of uppercases, numerals, and symbols. Every single one of them will be eliminated using syntax for matching patterns.
- **Case folding:** Case folds are techniques that transform words into same forms like all lowercase or uppercase characters. At this stage, every word has to be converted to lowercase.
- **Elimination of stop words:** This process entails taking well-known and often used words out of a sentence that don't add anything to its meaning. Using Bahasa Indonesia stop words lists which included words like "dan" (and), "atau" (or), and others, Twitter tweets were cleaned by eliminating these stop words during pre-processing steps.
- **Special characters on Twitter:** Twitter uses special characters in its messages, such the retweet button (RT), the username (@username), and the hashtag (#). These words will be removed while this procedure is being completed. Since words or phrases that best describe conversations are followed by hashtag symbols, this work eliminates hashtag signs leaving only the words or phrases.
- **Managing Emoticons:** Emojis will be able to be translated into the words they stand for. Winks (emot-kedip), grins (emot-senyum), laughs (emot-tawa), love (emot-cinta), sadness (emot-sedih), and sobs (emot-tangis) (emot-ejek) were categorized seven types of emoticons in this study. The development of this aggregate was based on emoticons that were found in the dataset.
- **Eliminating URLs:** Among other things, Twitter tweets frequently contain URLs like http://www.ift.tt/1QBmUt3. Since the focus of this part is on the words that are contained in the tweets, the URLs have been removed.

Algorithm 1: Twitter set Pre-processing

Input R: Twitter reviews

Output PR: pre-processed reviews

```
Obtaining test reviews R
T←getwordtokenization(R)
for (i=1 to size(T))
S(i)←stemming(T(i))
End for
For (i=0 to size(S))
If (S(i)≠symbols)
If (S(i)≠numbers)||length(S(i)>2)
If (S(i)≠ASCII strings)
If (S(i)≠punctuations)
T1←S(i)
End if
End if
```

14: End if
15: End if
16: End for
17: For (i=0 to size(T1))
18: T2←case folding(T1(i)
19: End for
20: For (i=0 to size(T2))
21: if (T2(i)≠stopword)
22: if (T2(i) ≠ specialcharacters ‖ T2(i) ≠ hashtag ‖ T2(i) ≠ username ‖ T2(i) ≠ date ‖ T2(i) ≠ complexcharacter)
23: T3←T2(i)
24: End if
25: End if
26: End for
27: For (i=0 to size(T3))
28: T4←Emoticons Handling(T3(i))
29: End for
30: For (i=0 to size(T4))
31: if (T4(i)≠URL)
32: T5←T4(i)
33: End if
34: End for
35: Return (PR←T5)

3.3 FEATURE EXTRACTION USING MPCA

MPCA technique is presented in this study to extract the more informative features from the provided SemEval-2014 dataset. PCA reduces dimensionalities of data spaces (observed variables) for minimizing feature spaces' dimensionalities (independent variables) resulting in economical definition of data. This is the case when observable variables show a significant relationship with each other. PCA-based feature extraction approach is used in this investigation. A traditional multivariate data analysis technique that works well for linear feature extraction is PCA which efficiently lowers features and presents datasets in a low dimensional subspace by eliminating minor components [15]. These techniques' coefficients serve as feature vectors that effectively depict the SemEval-2014 dataset. Important feature information will be missed by the standard PCA approach when used for feature extractions from tiny datasets and data compressions of PCA cannot be guaranteed. This work handles this issue with a Modified PCA.

MPCA reduces eigenvector influences corresponding to larger eigenvalues by normalizing j^{th} elements y_{ij}, of i^{th} feature vectors y, with respect to their standard deviations, $\sqrt{\lambda_j}$. These resulting new feature vectors y_i' can be rewritten as:

$$y_i' = \left[\frac{y_{i0}}{\lambda_0}, \frac{y_{i1}}{\lambda_1}, \ldots, \frac{y_{i(r-1)}}{\lambda_{r-1}} \right] \tag{1}$$

New feature subspaces are created from Normalized feature vectors and distances between training and testing features are computed by normalizations of these vectors using the square roots of their eigenvalues. Generically, PCA linearly transforms can be expressed as the following equation:

$$Y=TX \tag{2}$$

where T stands for transform matrices, X stands for original vectors and Y implies transformed vectors. Transform matrices T are obtained using:

$$(\lambda I\text{-}S)U=0 \tag{3}$$

where I stands for identity square matrices, S implies original images' covariance matrices U stands for eigenvectors and λ implies eigenvalues. U_j and λ_j (j=1,2,…m) are calculated using Eq.(2) and where $\lambda_1 \geq \lambda_2 \geq \cdots . \geq \lambda_m$ imply ordered eigenvalues. Eigenvectors U can be represented as $U=[U_1,U_2,\ldots,U_m]$.

Training samples from a SemEval-2014 dataset are chosen for the MPCA based on their applicability, and it is from these training examples that the modified matrix T' is derived. The equation that follows is one way to represent it:

$$Y=T'X \tag{4}$$

$$V_N = b_1u_1+b_2u_2+\cdots.+b_Nu_N \tag{5}$$

$$S = \sum_{i=0}^{1} b_1 u_1 ; 1 < N \tag{6}$$

When comparing Eq.(4) and Eq.(6), transform matrices are where the differences are found. Essentially, two distinct samples are used to compute covariance matrices namely training sample matrices, and matrices of the 2014 SemEval dataset.

MPCA primarily minimizes duplicates in information resulting in reducing dimensions without losing required information. Eigen values and Eigen vectors can be comprehended better in PCA using mathematical or statistical tools. MPCA is a mathematical process that employs linear transformations to convert dimensional spaces from high to low. Eigen vectors of covariance matrices can be used to determine low-dimensional spaces and was employed in this study to extract essential, meaningful tweets for data categorizations because to minimized errors and de-correlating capabilities.

Algorithm 2: MPCA for Feature Extraction

1. **Begin**
2. Determine the SemEval-2014 dataset S's mean value, S'.
3. Take S and subtract mean values.
4. Get updated matrices A.
5. Matrices yield covariance $C = AA^T$ where eigen values are derived from covariance matrices V_1 V_2 V_3 V_4…V_N and eigen vectors for C are calculated.
6. Vectors S' can be expressed as linear combinations of Eigen vectors using Eq.(5).
7. Only Eigen values are retained for generating smaller dimensional data spaces.
8. Compare word combinations in provided tweets Eq.(6).
9. Extract more instructive tweets.
10. **End**

3.4 CLASSIFICATION USING OBICRNN ALGORITHM

OBiCRNN method, which can extract both local and global characteristics from text documents, is suggested in this study for customer review categorization and summarization. This work's hybrid model produces better analysis of sentiments for CRS. The CRS-HAR-OBiCRNN model combines OBiCRNN and CRS with hybrid Aspect-Related Features (ARF) and Review-Related Features (RRF).

This study uses the OBiCRNN method for classification. One sort of deep neural network is the RNN where its linked neurons form directed graphs. RNNs are suitable for NLP workloads because they can process input sequences depending on their internal state. To calculate each RNN result, the same code was executed on each instance many times. The computed performances are based on previous computations. Time step durations were determined using input lengths of RNN architectures.

Group-wise enhancements, embeds and layers of bidirectional, convolution, pooling, and fully linked are the six primary modules that make up the suggested scheme. Deep neural language models were initially used the embedding layer extracts dense vector representations from text. The proposed technique extracts high-quality features from text using two bidirectional architectures: bidirectional GRU and bidirectional LSTM, while accounting for both previous and succeeding contexts. Combining forward and hidden layers are efficient methods for sequential modelling. Assignments of higher significance coefficients to GRU and LSTM layers' global features produce group-wise improvements resulting in reduced uninformative features and enhanced significances in features. GRU and LSTM features were split into groups for accentuating important traits in groups and reduce unimportant ones. OBiCRNN extracts high-quality local features using convolution layers and minimizes feature space dimensionalities using pooling layers (maximum-pooling approaches). Finally, a completely connected layer has been utilized to identify the emotion orientation of text using concatenated enriched feature sets. The remainder of this section contains the information for the architecture's modules.

3.4.1 Embedding Layer:

NLP modelling were used to represent tweets and word embeds of tweets were based on pre-training n-gram on embed layers, TF-IDF embedding matrix. In order to produce the word embedding matrix for an n-dimensional text document, it first converts each word into its corresponding V-dimensional word vector.

$$X=[x_1,x_2,..x_n]\in R^{n\times V} \tag{7}$$

3.4.2 Bidirectional layer:

This work uses two synchronous bidirectional designs, bi-LSTM and bi-GRU, in the bidirectional layer module. This module's components are designed to address the issue of disappearing and bursting gradients in traditional RNNs. Furthermore, by processing in both forward and backward directions, bi-LSTM and bi-GRU designs can aid in the acquisition of high-quality context elements. Each unit's output features of a bi-LSTM are made up of context data gathered by forward/backward LSTM [16] where one LSTM processes text documents from front to back while the other from back to front. According to the equations below, contexts therefore encompass both past and future.

$$\vec{c}_t, \vec{h}_{t_{LSTM}} = \vec{LSTM}\left(\vec{c}_t, \vec{h}_{t-1_{LSTM}}, \vec{x}_t\right) \tag{8}$$

$$\overleftarrow{c}_t, \overleftarrow{h}_{t_{LSTM}} = \overleftarrow{LSTM}\left(\overleftarrow{c}_t, \overleftarrow{h}_{t-1_{LSTM}}, \overleftarrow{x}_t\right) \tag{9}$$

$$h_{t_{LSTM}} = \left[\vec{h}_{t_{LSTM}}, \overleftarrow{h}_{t_{LSTM}}\right] \tag{10}$$

where $h_{t_{LSTM}}$ implies hidden states of bi-LSTM vectors while ct stands for their cell states.

Similarly, contexts obtained by forward/ backward GRUs are included in output features of units in bi-GRU [17]. Text documents are processed from front to back by the front one $\left(\vec{GRU}\right)$ and from back to front by the other one $\left(\overleftarrow{GRU}\right)$. According to the following equations, the context therefore encompasses both the past and the future:

$$\vec{c}_t, \vec{h}_{t_{GRU}} = \vec{GRU}\left(\vec{c}_t, \vec{h}_{t-1_{GRU}}, \vec{x}_t\right) \tag{11}$$

$$\overleftarrow{c}_t, \overleftarrow{h}_{t_{GRU}} = \overleftarrow{LSTM}\left(\overleftarrow{c}_t, \overleftarrow{h}_{t-1_{GRU}}, \overleftarrow{x}_t\right) \tag{12}$$

$$h_{t_{GRU}} = \left[\vec{h}_{t_{GRU}}, \overleftarrow{h}_{t_{GRU}}\right] \tag{13}$$

3.4.3 Group-wise Enhancements:

This mechanism which is a hybrid deep neural network module, changes relative relevance of sub features by creating attention factors of spatial positions in semantic categories resulting in individual improvements of learnt expressions in groups while minimizing possible noises. Attention masks used for scaling feature vectors create spatial enhance mechanisms inside feature groups. The goal of this attention mask is to highlight the relevant semantic feature regions while filtering out surrounding noise. Unlike other well-known attention approaches, it builds attention masks based on the similarity of global and local statistical characteristics at each place. To extract informative features, feature weighting algorithms have been used to $h_{t_{LSTM}}$ and $h_{t_{GRU}}$ in the group-wise improvement module. This module is an adaptation of the spatial group-wise enhancement method that was first used to classify images. The mechanism's goal is to make informative features' weight values stronger while weakening those of uninformative characteristics. Obtaining distributed feature sets properly for deep neural architecture-based learning models may be difficult, mainly due to uninformative inputs. Thus group-wise improvements of this work's suggested model focus on semantic knowledge in critical areas to estimate semantic vectors of groups. Group-wise enhancements by hybrid deep neural network module changes relative relevance of sub features by generating attention components for spatial positions in semantic groups. These allow groups to improve their learnt expressions individually while minimizing possible noises. To develop spatial enhancement mechanisms inside feature groups, attention masks are used in scaling positional feature vectors.

Word embedding matrices $X=[x_1,x_2,..x_n]\in R^{n\times V}$ have V-dimensions for n-dimensional text documents. Initially, representation schemes were divided into M groups $M^K = \left[x_1^k, x_2^k, ..., x_n^k\right]$ where $x_i \in R^{n/m}$ and $n = V$ and k denote group numbers. Global statistical features through spatial averages were utilized to obtain semantic vectors of groups based on Eq.(15):

$$g^k = \frac{1}{n}\sum_{i=1}^{n} x_i^k \tag{14}$$

Subsequently, this global feature gets appropriate coefficients of importance for each feature. This is accomplished using straightforward dot products that compute similarities between local features x_i^k and global semantic features g^k using:

$$c_i^k = g^k \cdot x_i^k \tag{15}$$

$$c_i^{*k} = \frac{c_i^k - \mu_c^k}{\sigma_c^k + \varepsilon} \tag{16}$$

$$\mu_c^k = \frac{1}{n}\sum_{j=1}^{n}\left(c_j^k - \mu_j^k\right)^2 \tag{17}$$

$$\left(\sigma_c^k\right)^2 = \frac{1}{n}\sum_{j=1}^{n}\left(c_j^k - \mu_j^k\right)^2 \tag{18}$$

where ε stands for constant parameters (set to $1e^{-5}$) for providing numerical stability.

PSO drawing inspiration from a common wildlife phenomenon like a flock of birds flying out in search of food is used to improve CRS performance. An individual flies under instruction when looking for the location of meals, both from its own experience and from advice from other partners, particularly those who are close to the foods. In PSO, a particle mimics a bird, and the collection of birds is a particle swarm [18]. The way a particle is represented allows it to represent a job scheduler. Finding the optimal scheduler among all particles once they have been examined by a specific time is the primary goal of PSO. Each particle's developing formula is as follows:

$$v_i^{(t+1)} = wv_i^{(t)} + c_1 r\left(pb_i^{(t)} - x_i^{(t)}\right) + c_2 R\left(gb_i^{(t)} - x_i^{(t)}\right) \tag{19}$$

$$x_i^{(t+1)} = v_i^{(t)} + v_i^{(t+1)} \tag{20}$$

$v_i^{(t)}$ and $x_i^{(t)}$ represent speeds and positions of particles i in t^{th} iterations. $pb_i^{(t)}$ and $gb_i^{(t)}$ are global bests of swarms and personal bests of particles i. r and R imply random values in [0,1], and w, c_1 and c_2 stand for weights.

- **Coverage**: It states that practically every important element of the original content should be covered in the summary. Maximum coverage and little information loss are desirable in a summary.
- **Redundancy** is the main emphasis of diversity. The ability to reduce the number of phrases with comparable meanings is a sign of a rich summary. Process variety in summarizing effectively reduces sentence repetition.
- **Balance**: While an unbalanced summary may highlight inaccurate interpretations of the source document's broad concepts, a balanced summary aims to highlight the many significant components of the original work.

3.4.4 Convolution layer:

Filters and convolution operations are first applied to the improved context features that were acquired through the group-wise improvement approach in the convolution module. During this stage, CNNs use convolutions via sliding windows of matrices and slides filters on context features to capture local contexts on the improved features of contexts. Let w be the size of the sliding window and let $F^j \in R^{w\times V}$ be the filter. [19] By concatenating respective vectors' first dimensions, local contextual characteristics for w words were derived using the following equation:

$$m^j = \left[m_1^j, \ldots m_i^j, \ldots, m_{L-w+1}^j\right] \tag{21}$$

$$m_i^j = f\left(x_{i:i+w-1} \otimes F^j + b_0\right) \tag{22}$$

where f corresponds RELU activations and $\otimes$ corresponds to convolution operations. Several identically sized filters were used to extract local CNN architectural characteristics from the text material. The convolution procedure was repeated on increased context to generate the local feature matrix.

3.4.5 Pooling Layer:

In pooling modules, max pooling filters most important features from local feature matrices. The outputs of pooling layers are concatenated and these concatenated feature vectors are fed to fully linked layers.

3.4.6 Fully Connected Layer:

The concatenated feature vectors acquired from previous steps are inputs for the fully connected modules, which are feed forward dense layers. Additionally, this system has three completely linked layers, with last layers using SoftMax activations while priors use RELU activations resulting in identifications of probability distributions of labels for text sentiments.

Algorithm 3: OBiCRNN for CRS classification

Input: SemEval-2014 restaurant reviews dataset

Output: Positive, Negative and Neutral

1. For Twitter dataset, perform text pre-processing
2. Perform feature extraction via MPCA
3. Combine n-gram and TF-IDF to extract more informative features from the tweets
4. Use bi-LSTM and bi-GRU, synchronous bidirectional architectures and obtain contexts of preceding LSTM $\left(\vec{h}_{t_{LSTM}}\right)$ and succeeding LSTM $\left(\overleftarrow{h}_{t_{LSTM}}\right)$ features with contexts of preceding GRU $\left(\vec{h}_{t_{GRU}}, \vec{h}_{t_{LSTM}}\right)$ and succeeding features of GRU $\left(\overleftarrow{h}_{t_{GRU}}, \overleftarrow{h}_{t_{LSTM}}, \overleftarrow{LSTM}\right)$ using Eq.(8)-Eq.(11), respectively
5. Use dual group-wise enhancements on layers to obtain enhanced feature vectors (LSTM and GRU features) using Eq.(14)-(18).
6. Calculate the objective function of PSO
7. Update the gbest and pbest values using Eq.(19) and Eq.(20)
8. Optimize convolutional layer. Apply the convolution layer to the increased context vectors using Eq.(21) and Eq.(22).
9. Use max. pooling to choose the most informative features from the local feature matrix.
10. Concatenate outcomes of pooling layers
11. Feed these concatenated feature vectors to fully connected layers.
12. Use two RELU activation function layers.

13. Apply the complete context strategy to a soft-max classifier to generate sentiment labels.
14. Return Positive, Negative and Neutral results for the given tweets

4. EXPERIMENTAL RESULT

The SemEval-2014 restaurant reviews dataset is taken into consideration in this work in order to evaluate the proposed approach [10]. For the experimental analysis of the proposed model, an I3 CPU and 4 GB of RAM were used together with the MATLAB software running on Windows 10. This study used state-of-the-art methods to analyse the performance of the proposed model on three datasets. Furthermore, the EBRBFSVM makes advantage of state-of-the-art techniques as SAN-ASVM, SVM-PSO, and SVM-RFE. The performances are compared using the following criteria: Accuracy, F1-score, Precision, Recall, and Average Sentiment Analysis Time (ASAT). The F1-score is determined in this manner.:

$$F1-score = 2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \tag{23}$$

$$\text{Precision}=\frac{TP}{FP+TP} \tag{24}$$

$$\text{Recall}=\frac{TP}{FN+TP} \tag{25}$$

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \tag{26}$$

Thus, in emotion classification, FP stands for False Positive, TP for True Positive, and FN for False Negative. The ASAT parameter, which stands for mean processing time, is linked to the computation time, or the average amount of time required to identify attitudes. To estimate the ASAT variable, 100 distinct applications of each classification technique were conducted here.

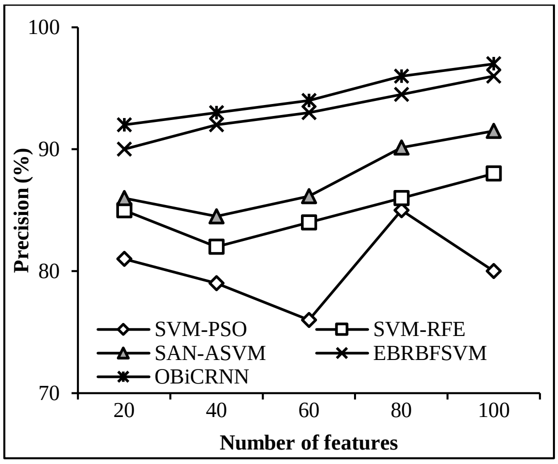

Fig.3 Precision performance comparison

The Fig.3 above depicts comparative precision values of techniques represented on the x-axis and obtained precision values as the y-axis. While the current SVM-PSO, SVM-RFE, SAN-ASVM, and EBRBFSVM algorithms offer lesser precision, the suggested OBiCRNN approach offers superior precision. The goal of the suggested system is to choose the more pertinent data. Accordingly, the outcome shows that the suggested OBiCRNN technique improves the informative characteristics for the CRS procedure with accuracy.

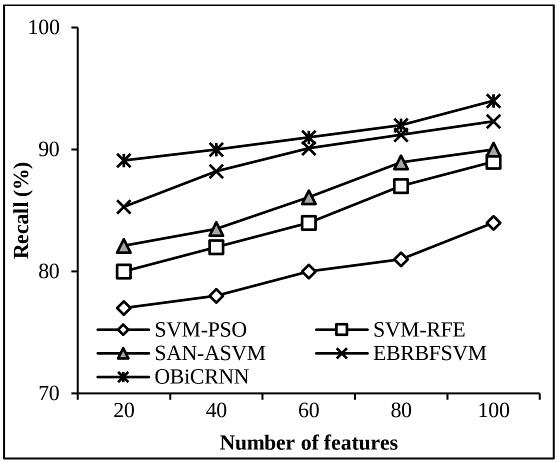

Fig.4 Recall performance comparison

The Fig.4 above depicts comparative recall values of techniques represented on the x-axis and obtained recall values as the y-axis. In contrast to the current SVM-PSO, SVM-RFE, SAN-ASVM, and EBRBFSVM techniques, the suggested OBiCRNN approach offers a greater recall. The goal of the suggested system is to choose the more pertinent data. Accordingly, the outcome shows that the suggested OBiCRNN technique improves the informative characteristics for the CRS procedure with accuracy.

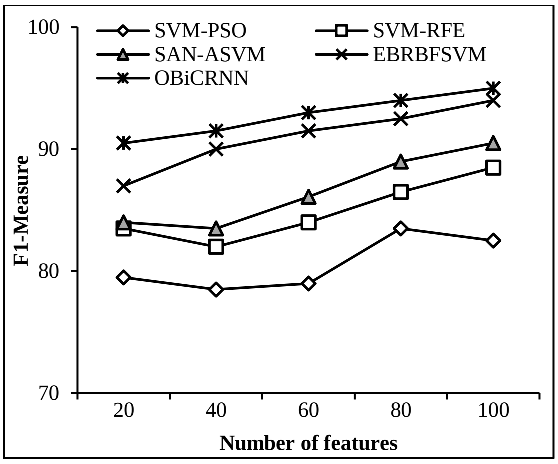

Fig.5. F-measure performance comparison

The Fig.5 above depicts comparative F1-scores of techniques represented on the x-axis and obtained F1-scores values as the y-axis. The suggested OBiCRNN technique yields higher F-measure values by 95% for the 100 features over the specified database, whereas the current SVM-PSO, SVM-RFE, SAN-ASVM, and EBRBFSVM methods provide lower F-measure.

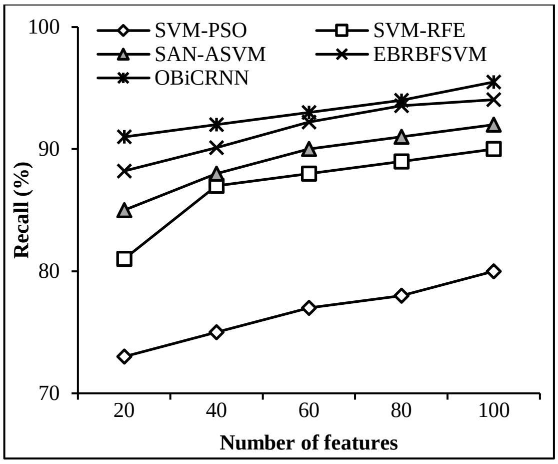

Fig.6 Accuracy performance comparison

The Fig.6 above depicts comparative accuracies of techniques represented on the x-axis and obtained accuracy values as the y-axis. For the provided dataset, the suggested OBiCRNN algorithm offers more accuracy than the current methods, which include the SVM-PSO, SVM-RFE, SAN-ASVM, and EBRBFSVM algorithms. The purpose of text pre-processing is to improve the accuracy of CRS categorization. The CRS performance is improved by the suggested OBiCRNN classification method.

5. CONCLUSION

The OBiCRNN technique is suggested in this study for effective CRS classification using the Twitter dataset. The primary processes in this study are CRS categorization, feature extraction, and text pre-processing. To improve the classifier's performance, text pre-processing is done first. The MPCA algorithm is then utilized to extract the essential keywords using feature extraction. As a result, it chooses well-known CRS keywords from the dataset and uses them to efficiently categorize tweets. OBiCRNN is proposed for obtaining both past and future contexts by combining two hidden layers with opposing orientations to the same context. The features extracted by bidirectional layers were subjected to the group-wise augmentation approach, which divides the features into many groups and strengthens the more significant traits in each group while weakening the less significant ones. Tweets are classed as positive or bad. Consequently, the suggested OBiCRNN algorithm performs better in terms of increased recall, accuracy, precision, and f-measure values. A hybrid optimization approach may be created in the future to enhance the various components of the solution for different Twitter datasets. Future research can also look into ways to modify it in order to identify more precise polarity classifications.

REFERENCES

[1] A. Chaudhuri and Soumya K. Ghosh, "Sentiment Analysis of Customer Reviews using Robust Hierarchical Bidirectional Recurrent Neural Network", *Proceedings of International Conference on Artificial Intelligence Perspectives in Intelligent Systems*, pp. 1-6, 2016

[2] Gagandeep Kaur and Amit Sharma, "Has: Hybrid Analysis of Sentiments for the Perspective of Customer Review Summarization", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 14, No. 9, pp. 11971-11984, 2023.

[3] N. Akhtar, "Aspect Based Sentiment Oriented Summarization of Hotel Reviews", *Procedia Computer Science*, Vol. 115, pp. 563-571, 2017.

[4] A.F. Agarap, "Statistical Analysis on E-Commerce Reviews, with Sentiment Classification using Bidirectional Recurrent Neural Network (RNN)", *Proceedings of International Conference on Machine and Deep Learning*, pp. 1-12, 2018.

[5] Tuan Anh, Jarunee Duangsuwan and Wiphada Wettayaprasit. "A Novel Automatic Sentiment Summarization from Aspect-Based Customer Reviews", *Proceedings of International Conference on Computer Science and Software Engineering*, pp. 1-5, 2018.

[6] Jawad Khan and Byeong Soo Jeong, "Summarizing Customer Review based on Product Feature and Opinion", *Proceedings of International Conference on Machine Learning and Cybernetics*, pp. 89-97, 2016.

[7] H. Nainwal, "Text Summarization of Amazon Customer Reviews using NLP", *Proceedings of International Conference on Computing for Sustainable Global Development*, pp. 45-53, 2023.

[8] Paitoon Porntrakoon, Chayapol Moemeng and Pratit Santiprabhob, "Text Summarization for Thai Food Reviews using Simplified Sentiment Analysis", *Proceedings of International Conference on Computer Science and Software Engineering*, pp. 1-10, 2021.

[9] H. Wu, "Aspect-based Opinion Summarization with Convolutional Neural Networks", *Proceedings of International Conference on Neural Networks*, pp. 1-4, 2016.

[10] Maria Pontiki, Harris Papageorgiou, Dimitrios Galanis, Ion Androutsopoulos, John Pavlopoulos, and Suresh Manandhar, "SemEval-2014 Task 4: Aspect based Sentiment Analysis", *Proceedings of International Conference on Semantic Evaluation*, pp. 112-118, 2014.

[11] Samuel Pecar and Maria Bielikova, "Sentiment Analysis of Customer Reviews: Impact of Text Pre-Processing", *Proceedings of International Conference on Digital Intelligence for Systems and Machines*, pp. 1-8, 2018.

[12] S. Manda and Anita Pal, "PSO-based Text Summarization Approach using Sentiment Analysis", *Proceedings of International Conference on Computing, Communication and Signal Processing*, pp. 241-246, 2019.

[13] A. Bompotas, "A Sentiment-based Hotel Review Summarization using LSTM Neural Networks", *Proceedings of International Conference on Information, Intelligence, Systems and Applications*, pp. 1-12, 2020.

[14] S. Gautam and Gurpreet Singh Josan, "Deep Neural Network Based Multi-Review Summarization System", *International Journal of Next-Generation Computing*, Vol. 12, No. 3, pp. 1-14, 2021.

[15] E.O. Omuya, George Onyango Okeyo and Michael Waema Kimwele, "Feature Selection for Classification using Principal Component Analysis and Information Gain", *Expert Systems with Applications*, Vol. 174, pp. 1-12, 2021.

[16] Shamal Kashid, "Bi-RNN and Bi-LSTM Based Text Classification for Amazon Reviews", *Proceedings of International Conference on Deep Learning, Artificial Intelligence and Robotics*, pp. 1-9, 2022.

[17] L. Zhang, "A Hierarchical Multi-Input and Output Bi-GRU Model for Sentiment Analysis on Customer Reviews", *Proceedings of International Conference on Materials Science and Engineering*, pp. 22-27, 2018.

[18] Nur Hayatin, Gita Indah Marthasari and Lia Nuarini. "Optimization of Sentiment Analysis for Indonesian Presidential Election using Naive Bayes and Particle Swarm Optimization", *Journal Online Informatika*, Vol. 5, No. 1, pp. 1-17, 2020.

[19] Ming Hong and Heyong Wang, "Research on Customer Opinion Summarization using Topic Mining and Deep Neural Network", *Mathematics and Computers in Simulation*, Vol. 185, pp. 88-114, 2021.