



# A Proposed Hybrid model for Sentiment Classification using CovNet-Dual LSTM Techniques

Roop Ranjan<sup>a</sup> and A. K. Daniel<sup>a\*</sup>

<sup>a</sup> Madan Mohan Malaviya University, India

[roop.ranjan@gmail.com](mailto:roop.ranjan@gmail.com), [danielak@rediffmail.com](mailto:danielak@rediffmail.com)

\*Correspondence Email: [danielak@rediffmail.com](mailto:danielak@rediffmail.com)

## KEYWORDS

Deep Learning;  
QoS; Word  
Embedding;  
Emotion  
Analysis; Text  
Analysis.

## ABSTRACT

*The fast growth of Internet and social media has resulted in a significant quantity of texts based review that is posted on the platforms like social media. In the age of social media, analyzing the emotional context of comments using machine learning technology helps in understanding of QoS for any product or service. Analysis and classification of user's review helps in improving the QoS (Quality of Services). Machine Learning techniques have evolved as a great tool for performing sentiment analysis of user's. In contrast to traditional classification models. Bidirectional Long Short-Term Memory (BiLSTM) has obtained substantial outcomes and Convolution Neural Network (CNN) has shown promising outcomes in sentiment classification. CNN can successfully retrieve local information by utilizing convolutions and pooling layers. BiLSTM employs dual LSTM orientations for increasing the background knowledge accessible to deep learning based models. The hybrid model proposed here is to utilize the advantages of these two deep learning based models. Tweets for users for reviews of Indian Railway Services have been used as data source for analysis and classification. Keras Embedding technique is used as input source to the proposed hybrid model. The proposed model receives inputs and generates features with lower dimensions which generate a classification result. The overall performance of proposed hybrid model has been compared using Keras and Word2Vec and observed effective improvement in the response of the proposed model with an accuracy of 95.26%.*



# 1. Introduction

The world has evolved rapidly as a result of recent developments. People's ability to use the internet has become more crucial in their life. Social networking services currently provide a convenient form of communication that allows users to connect with one another, discuss daily activities, and express their thoughts on a variety of issues such as companies, commodities, and locations on social forums. Customer evaluations on social media sites may let individuals share questions and thoughts about products before they use them. Data acquisition and analysis are useful to business growth, but only when they are correctly examined and processed. Sentiment analysis is a way of analysing people's feelings and viewpoints from text using text analysis technologies.

The process of text categorization is the foundation of sentiment analysis, and various words contribute differently to classification. In tasks of sentiment classification, it is crucial to learn a non-sparse word vector low-dimensional representation for a word (Zhang et al., 2017). The distributed generation of word vector by Word2vec technique (Zhang et al., 2016a) is an extensively used word representation. The vectors of word have a precise and specific range that contains the semantic information for the given set of word.

Deep Learning (DL) techniques are increasingly being used in NLP applications these days (Lecun et al., 2015). (Bengio et al., 2003) presented a language-building model based on Artificial Neural Networks (ANNs), which conduct word mapping to a lower dimensional space. Word2vec is a Google open source programme that converts words into actual numerical vectors. (Mikolov et al., 2013a, b).

Despite the fact that various research have been conducted in recent decades to suggest effective techniques for the work of sentiment analysis, the development of deep learning as a subset of machine learning has produced a remarkable advance in this domain (Souma et al., 2019). Deep learning algorithms, in fact, may use many processing layers to extract numerous useful characteristics from data without the need for human involvement (Zhang et al., 2018a), (Chowdhury et al., 2019). As a result, these approaches have made significant progress in a range of fields, including recognition of speech (Zhang et al., 2018b), computer vision (Voulodimos et al., 2018) and NLP (Young et al., 2018). The LSTM (Hochreiter and Schmidhuber 1997) is a form of recurrent neural network that has shown a lot of success in tackling a variety of issues. The LSTM network incorporates a self-cycling mechanism, which makes learning long-term dependant information easier than with a basic cycle structure. In a bidirectional based LSTM (BiLSTM) model, the successor and predecessor of the sequence are input into the LSTM network which can efficiently capture the sequence's contextual information (Baziotis et al., 2017).

Deep learning has been widely used in sentiment analysis since it has steadily gained tremendous success in various domains. (Majumder et al., 2017) proposed using CNN to figure out the personality type of author from the given text. CNN is able to retrieve the text's subjective representation; It fails to recognize the characteristics of progression of texts. (Santur, 2019) utilised GRU to analyse sentiment. Despite the fact that GRU may extract context based information, it is a biased model in which words appearing later are valued more than early words. Although (Salur and Aydin, 2020) presented an unique hybrid model on deep learning for categorization of sentiment, complicated models are more prone to overfitting.

The widespread use of deep learning techniques during recent years has aided the advancement of image processing and speech recognition. The deep learning approach has been widely utilized in natural language processing because of its automated learning properties.

Unlike BiLSTM, CNN includes a convolutional layer for extracting vector features and reducing their dimension. To overcome this limitation, the objective of this research work is to develop a innovative deep learning-based text classification model that combines the structure of BiLSTM and CNN (i.e. CovNet and Dual-LSTM). By incorporating a convolution layer in the Convolution Neural Network model, the novel suggested CovNet-Dual LSTM model attempts to overcome the constraints of Dual LSTM. Keras' efficiency and accuracy are used in the suggested approach. The Keras API is built on an architecture that makes it easier to build deep learning models with the Theano or Tensor-Flow APIs (Keras, official website). Keras has a textual data layer of embedding that may be used with recurrent neural networks. It's a versatile model that can be used for a number of tasks, such as learning a word embedding on its own, preserving and restoring it in a subsequent layer, a type of transfer learning. The model is trained using the Roop Ranjan et al. dataset (Ranjan and Daniel, 2020a). The dataset contains 25000 tweets from passengers using Indian Railways services for different days in consideration. According on experimental findings, the suggested CovNet-Dual LSTM model outperformed other models and previous studies.

The key contribution to this work is as below:

Using Keras APIs, an unique hybrid technique for word embedding is suggested in this study..

- To analyse and categorize user reviews, a model based on CNN and DualLSTM is presented..
- Extensive experiments were carried out on the proposed hybrid model utilising Keras and the Word2Vec embedding model, and the model's efficiency was compared to that of current models.
- The suggested model has demonstrated a substantial improvement in accuracy and precision.

Further in Section II literature review of previous works is discussed. Section III presents the working of proposed model for sentiment analysis, while Section IV presents observation of experiments on the dataset and summarizes the findings. Section V and VI contains the conclusions, limitations, and recommendations for future research respectively.

## 2. Related Work

For evaluating the emotional component of text, text-based sentiment categorization and natural language processing techniques are widely used. Deep learning based algorithms have recently attained significant breakthroughs in the area of natural language processing. In a research (Kim et al., 2014) proposed classification of text written in english by taking vectors of word as input provided to CNN to perform sentence level classification. Despite the fact that CNN obtains high text classification results, it concentrates primarily on extracting local characteristics and discards the context of words, that contains a significant influence on the performance of classification results (Liu et al., 2017)(Zhang et al., 2018c). Due to this shortcoming a hybrid model using CovNet-Dual LSTM is proposed.

To improve sentiment categorization performance, (Liu et al., 2016) integrated machine learning and deep learning. The efficacy of the suggested methodology is demonstrated using Turkish and Chinese language datasets in their research. Furthermore, CNN and LSTM based text analysis were used to examine IMDb comments. (Yenter and Verma, 2017). A novel technique for emotion analysis is provided in their suggested method, which uses a high number of layers that includes CNN and LSTM

coupled with kernels. The features of sequential windows feed the LSTM network just after CNN is applied to the texts (Zhou et al., 2015). LSTM learns long term dependencies using higher level features with this technique. As a result, the researcher integrates the advantages of convolutions for obtaining native characteristics with an LSTM layer for determining sentiment long term dependency. A novel deep learning architecture with hybridized CNN and BiLSTM (H2CBI) set of features has been presented in (Wint et al., 2018) which combines the power of CNNs and BiLSTM. The researchers obtained distinct set of vectors of feature supplied as input for LSTM layer using two separate pre-trained vectors of words. A research from (Shi et al., 2017) retrieved the user interest and content based manually obtained characteristics of the given data obtained from social network site Sina Weibo in addition to the word embedding features. The results of the research are sent into an LSTM network for classification. It is clear from their studies that their suggested hybrid approach outperforms basic LSTM or traditional machine learning.

Because of the expeditious expansion in deep learning techniques, deep learning based models for sentiment analysis for textual data have gradually become a pattern these days. (Pang et al., 2002) Using three typical classifiers, experimental research was conducted on a text sentiment analysis task with a good accuracy rate (maximum entropy, Naive Bayes and support vector machine). The classification of text CNN was introduced by (Kim, 2014) and it has since become one of the most significant baselines for sentiment analysis tasks. In a sentiment analysis work, by utilising the bidirectional transmission method, (Brueckner and Schuller, 2014) contributed to the solution of obtaining both previous and future information. For sentiment classification, (Wang et al., 2016) developed a Long Short-Term Memory Network based on attention mechanism. When multiple components of a sentence are accepted as input, the attention mechanism might focus on different sections of the sentence. To address with natural language processing challenges, (Shen et al., 2018) introduced new LSTM based concept called ON-LSTM. The neurons are arranged in this proposed LSTM in a specific way to demonstrate more information. (Du et al., 2019) introduced CRAN, novel network architecture. This model integrates a RNN with a convolutional-based model using attentions and neural model to create a hierarchy based text classification model.

Deep learning is a relatively new branch of machine learning. Its goal is to simulate human brain in terms of learning and thinking. (Feizollah et al., 2019) utilized a layer of deep learning CNN and LSTM algorithms and found that they outperformed other basic approaches. To analyse the character sequence, (Xiao and Cho, 2016) presented an integrated model that comprised of multiple layers of CNN and one RNN layer. CNN performed successfully in character level representations of texts in their research.

(Cui et al., 2016) looked at the issue of n-gram order insensitivity. To increase sentiment analysis quality, their research exploited distributed semantic features of part-of-speech sequences. Convolution Neural Network can learn large-dimensional and complicated classification using the multi-layer perceptron structure in deep learning. Xia et al. (Xia et al., 2019) extracted features from review texts using the conditional random field approach, and a SVM (support vector machine) was utilized for classification of sentiment for the reviews.

To model the syntax and semantic meaning of tweets, as well as the interplay between context and aspect, (Zhang et al., 2016b) used neural network using gated topologies. Interaction based Attention Network) IAN learns the attention relationship among phrases and aspects target by an interactive technique, and it can retrieve features that are in relation with the aspect target using the attention mechanism.

As an important direction of investigation, researchers have integrated LSTM with different network structures. (Wang et al., 2016) proposed the ATAE-LSTM and AE-LSTM neural networks, which are both LSTM networks i.e. attention-based. The LSTM neural network used as Long Shot Term Memory based neural network for modeling the aspect, then paired the hidden state with the context embedding to construct the attention based feature vector, and then used a classifier for determining the sentiment category of the aspect. The ATAE-LSTM network, which is based on AE-LSTM, improved aspect embedding effects even more.

Several researchers have utilized deep learning to improve sentiment classification in recent years, with positive results. Socher et al. (Socher, 2013) introduced the RNTN (Recursive Neural Tensor Network) model, which included a sentiment library of tree that incorporated the meanings on the syntax tree of dual sentiment classification and produced an effective sentiment classification results in a movie review dataset. The CharSCNN (Santos and Gattit, 2014) (Character to Sentence based Convolutional Neural Network) model extracted features of linked words and phrases using two convolutional layers and analyzed semantic knowledge for improving sentiment classification of small texts like tweets. The attention mechanism was incorporated into the LSTM by Baziotis et al. (Baziotis et al., 2017), which performed well in classification of sentiment of Task4 of SemEval-2017 using Twitter. (Amir et al., 2016) created a mode based on deep learning network that prioritizes context based data using lexicon based and grammar based patterns. (Ranjan and Daniel, 2020b) proposed a novel model for detection of emotions of travellers using machine learning and fuzzy inference system.

BiLSTM is one of the RNN techniques used for improving LSTM, that contains text sequence feature limitations. It performs better than LSTM for sequential modeling (Liu et al. 2017), (Niu et al., 2017). Information in LSTM travels from backward to forward, but Bi-LSTM uses two hidden states to flow information in both directions: back to forward and forward to back. (González-Briónes et al., 2018a) proposed an unique multi-agent based strategy for energy consumption minimization induced by inefficient HVAC system utilization on the basis of Artificial Neural Network. A wireless sensor network with a Multiple Agent architecture (MAS) implemented in a Cloud based environment (WSN) achieved 41% of energy savings in an office building. In an interesting research (González-Briónes et al., 2018b) presented a multi-agent based classification of gender and images from the given images. This research also looked at how different filters affected the findings, allowing us to figure out which types of filtering helped the categorization algorithms perform better.

### 3. Proposed Model

The proposed model consists of following sub models for the analysis as shown in Figure 1.

1. Data Collection and Preprocessing
2. Word Vectorization using Keras
3. CovNET-Dual LSTM Model
4. Dense Layer Model
5. ReLU Activation Model

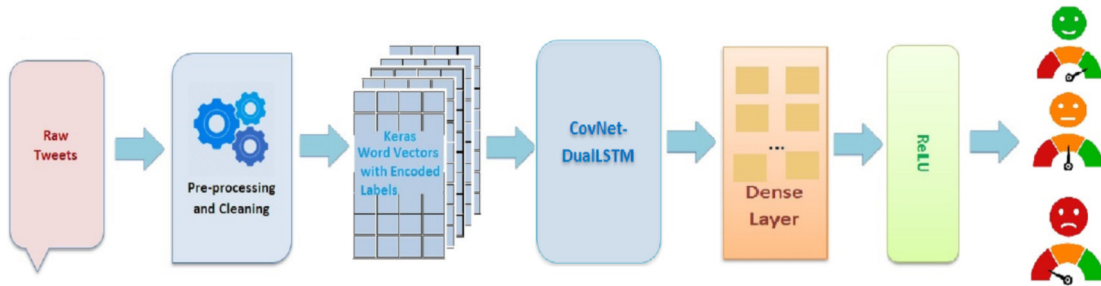


Figure 1: Proposed Architecture

### 3.1. Data Collection and Preprocessing

The reviews were retrieved using Twitter as the data source. The Tweepy API, which is built on Python, is used to gather tweets from a registered developer account on Twitter. Tweets are traveller comments about services provided during travelling through Indian Railways. Because the data collected from tweets contains noise, missing values, and is in an unstructured format, it cannot be directly fed into the model. Therefore before applying the model to the data, it is necessary to pre-process it. Pandas and Numpy, a Python-based toolkit, were used to pre-process the data by identifying and resolving missing values, categorical data encoding, and dataset splitting.

### 3.2. Keras word Vectors with Encoded Labels

Word embedding is a method of expressing written text as a set of numerical vector values. It creates vector representations of the same kind for texts with similar meanings. Each text is converted into a single vector before being fed into a neural network. Keras is the word embedding API utilised here. Keras consists of a Word Embedding module for textual data that is utilized with neural networks. Integer-based encoding of the input data is required. Each word is represented by a different number. Keras comes with a TokenizerAPI is utilised in study. The data in this study is in the form of sentences. When using Keras to perform Word Embedding, the following representation is utilised:

Here every text consists of  $n$  words is which is represented by  $D = \{x_1, x_2, \dots, x_n\}$  and every word in the set is transformed into vector of words of dimension  $p$ . The input text is defined here as follows:

$$D = \{x_1, x_2, \dots, x_n\} \in S^{n \times p} \quad (1)$$

Length of each document is not same always. Since the Keras Embedding layer requires all individual documents to be of same length. Therefore it is require making length of input text of same length (denoted as  $k$ ).Hence inputs are padded for the shorter text with 0. Its length was extended by employing a zero-padding method. The text will be trimmed if it is longer than the predefined length  $l$ . Let us consider a text example of different documents of different length up to maximum length 12.

The documents represented in Table 1 shows that their length is different from each other, therefore padding is required. Table 2 shows the structure of document after padding.



Table 1: Documents with different length

The encoding for document 1 is : [53, 14, 29, 32, 36, 17,32]
The encoding for document 2 is : [17, 12, 32, 33, 34, 32, 42, 23, 13]
The encoding for document 3 is : [25, 15, 18, 22, 33, 34, 12, 9, 3, 4, 23, 38]

Table 2: Documents with same length after padding

The encoding for document 1 is : [53, 14, 29, 32, 36, 17, 0, 0, 0, 0, 0, 0]
The encoding for document 2 is : [17, 12, 32, 33, 34, 32, 42, 23, 130, 0, 0]
The encoding for document 3 is : [25, 15, 18, 22, 33, 34, 12, 9, 3, 4, 23, 38]

The above implementation can be understood by the following Figure 2.

53	14	29		17	12	32		25	15	18
32	36	17		33	34	32		22	33	34
12	0	0		42	23	13		12	9	3
0	0	0		0	0	0		4	23	18
(a)				(b)				(c)		

Figure 2: All documents in 4x3 uniform vector after padding, (a) Document 1 after 5 0's padding (b) Document 2 after 3 0's padding (c) Document 3 , No padding required ( Yellow shaded 0's representing padding)

In the above figure it is observed how all the documents have been uniformed to same size vector after padding process. Therefore dimension of all input sequence will be defined as follows:

$$D = \{x_1, x_2, \dots, x_n\} \in S^{l \times p} \tag{2}$$

### 3.3. CovNet-Dual LSTM Model

This model combines the feature of Covent (CNN) and BiLSTM. The combination of these are applied into the proposed model. The strength of the proposed hybrid model uses CNN convolution layers to extract the maximum amount of information from documents and provides it to the BiLSTM input, allowing the data to be kept in chronological order in forward as well as backward paths. This model can be explained through Covnet Model and BiLSTM Model as follows:

#### 3.3.1 Covnet Model

This layer is used for extracting the most significant words from tweets This model receives an embedded vector matrix

$$D \in S^{l \times p} \tag{3}$$

The set  $D$  is provided as input given by Keras word Embedding Model. The model comprises a list of convolutional kernels with width  $\mu$  and number of filters  $J$  that are convolved with the input text (N-dimensional metrics) to produce an output feature map.

Filter  $J_k$ , where  $1 \leq k \leq K$  produces set of feature maps as shown below:

$$h_i^n = f(x^n \oplus Y_{i:i+x-1} + b^n) \quad (4)$$

For filter  $J_k$  the weight matrix is represented here as

$$a \in S^{e * p} \text{ and}$$

$b^n$  is bias of the given filter  $J_k$ ,

$p$  is representation of dimension of word vector

$\oplus$  represents the execution of convolution.

The term  $Y_{i:i+x-1}$  shows that the feature  $Y_{i:i+x-1}$  from  $Y_i$  is extracted from  $J_k$ .

Nonlinear activation function represented by  $f$

$h_i^n$  is output for the feature map of filter  $J_k$  where each element  $i$  belongs to  $h^n$ .

Here ReLU (Rectified Linear Unit) function is applied to non-linear activation function  $f$ . Following feature maps were obtained for the sentence length  $l$ :

$$h = [h_1, h_2, \dots, \dots, \dots, \dots, \dots, h_n] \quad (5)$$

The sliding process is continued till there are no more values to be covered further.

### 3.3.2 Pooling Model

Pooling Model is used for reducing the dimensions of the feature vectors. The feature vectors which are generated are optimized here using pooling layer. The convolution process output is used as input to the pooling process to execute for extracting the important features. There are various pooling techniques such as Max Pooling, Average Pooling and Global Pooling. The process of Max pooling is considered here for implementation because Max pooling provides the most optimized values for the given region of CovNet convolution.

$$\hat{h} = \max(h) \quad (6)$$

Here  $\hat{h}$  is the set of optimized features extracted from feature vectors  $h = [h_1, h_2, \dots, \dots, \dots, \dots, \dots, h_n]$  provided by Covnet Model.

### 3.3.3 DualLSTM Model

The DualLSTM model deals with optimized feature vectors received from the pooling model. The output of max pooling layer is concatenated to create the input to DualLSTM model. Each layer of DualLSTM contains two hidden states to allow the flow of information dual directions as forward to backward and vice versa. The dual scanning process leads input data from both the previous and the next stages. The forward LSTM state obtains previous information, whereas the reverse LSTM state



obtains subsequent information. The process is iterated from first layer to the nth layer and the output R is received as follows: The process is represented in Figure 3:

$$R = [R_{fwd}, R_{back}] \quad (7)$$

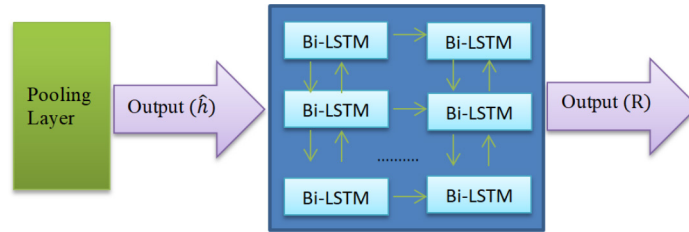


Figure 3: Process of DualLSTM Layer

### 3.4. Dense Model

The dense layer is used to minimize the error of the intermediate layers in backward direction using back propagation technique. A matrix-vector multiplication is performed and the errors in the values are trained and updated using back propagation. Updation at each layer is performed using outputs of all the previous layers.

### 3.5. ReLU Model

There are various activation functions utilized in neural networks like Sigmoid activation, Tanh activation and ReLU activation. An undesirable problem with Sigmoid and Tanh activation is when saturation of neuron's activation occurs at either end of 0 or 1, the values of gradient in these areas is virtually nil which leads to inaccurate calculations of feature vectors. Therefore the proposed model employs ReLU activation function to overcome this problem for generating the optimized feature vector map classification. The result of prediction by ReLU function is shown in Equation (6).

$$f(c) = \max(0, b) = \max \quad (6)$$

This model receives the feature maps and produces optimized classification results.

## 4. Experiment Results

This section discusses about the results achieved from the proposed model and performs a comparative analysis of the performance measurement of the system with other methods implemented in the same field.

Dataset is taken from (Roop Ranjan et al., 2020 a,b). The dataset has been taken and pre-processed for 5000 tweets and categorized in positive, negative and neutral categories. The dataset is divided in three sets namely training dataset, testing dataset and validation dataset in proportion of 60:20:20 percent using Gaussian distribution (Figure 4).

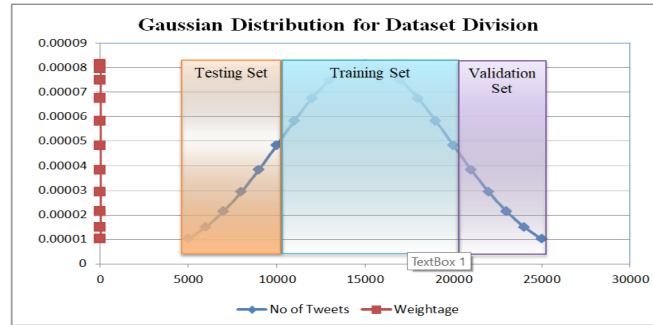


Figure 4: Dataset division using Gaussian Function

The dataset is categorized further in Positive, Negative and Neutral categories as shown in Table 3.

Table 3: Categorized Final Dataset

Date	Positive	Negative	Neutral
08-10-2019	2182	1098	1720
10-10-2019	2154	1074	1772
12-10-2019	2078	1052	1870
14-10-2019	2125	1048	1827
16-10-2019	2156	1080	1764

Table 4 provides the Training set for different days in consideration. Table 5 and Table 6 provides information about Validation set and Testing sets respectively.

Table 4: Training Set

Date	Positive	Negative	Neutral
08-10-2019	1309	659	1032
10-10-2019	1292	644	1063
12-10-2019	1247	631	1122
14-10-2019	1275	629	1096
16-10-2019	1294	648	1058

Table 5: Validation Set

Date	Positive	Negative	Neutral
08-10-2019	436	220	344
10-10-2019	431	215	354
12-10-2019	416	210	374
14-10-2019	425	210	365
16-10-2019	431	216	353

Table 6: Testing Set

Date	Positive	Negative	Neutral
08-10-2019	436	220	344
10-10-2019	431	215	354
12-10-2019	416	210	374
14-10-2019	425	210	365
16-10-2019	431	216	353

#### 4.1. Setting HyperParameter

During the CovNet Model, there was less overfitting and underfitting on output data values. Therefore to avoid the problem of overfitting and underfitting and to obtain high performance of model hyper-parameter settings are used to optimize the performance. To improve accuracy, the randomized search approach was applied. Table 7 provides a brief of the hyperparameters:

Table 7: Setting of HyperParameters

Parameters	Values
Size of Kernel	5
Dimension(Embedding)	Keras(300)
Filter Size	64
DualLSTM Output Size	32
Regularization function	L2
Activation	ReLU
Weight Constraints	Kernel Constraints(max norm is 3)
Batch Size	64
No of Epoch	20
Batch Normalization	Yes
Learning Rate	0.01
Optimizer	Adam

#### 4.2. Performance Measurement

Performance of proposed CovNet-Dual LSTM model with Keras was observed and compared with Word2Vec i.e. CNN, LSTM, BiLSTM, CNN-LSTM Models. The proposed model outperformed the other models with an accuracy of 95.23%. The performance comparison using Word2Vec embedding is shown in Figure 5. The level of accuracy with Word2VEc Embedding achieves 91.8% which is much better than other models.

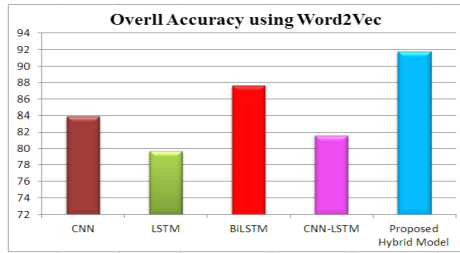


Figure 5: Level of accuracy using Word2Vec Embedding

The above Figure 6 shows that for all the models that have been tested Keras embedding model proves to be a better performer than traditional Word2Vec embedding. It is also observed that when Keras embedding is combined with CovNet-Dual LSTM, the accuracy level increases more and the accuracy level reaches to 95.23%. This clearly shows that the Keras embedding model with CovNet-Dual LSTM outperforms other existing models.

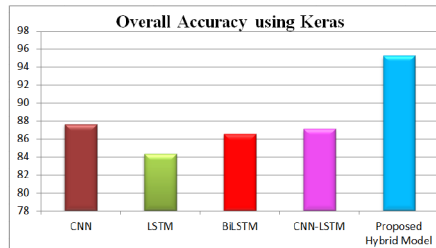


Figure 6: Level of accuracy using Keras Embedding

The performance parameters are validated using the Confusion Matrix with the help of SciKit-Learn API. Figure 7 illustrates parameters of a confusion matrix:

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Figure 7: Parameters of Confusion Matrix

The performance and validation parameters are defined below:

True Positive (TP): *Positive class is accurately predicted by the model* (Tkatek et al., 2020)

True Negative (TN): *Negative class is accurately predicted by the model*

False Positive (FP): *Positive class is inaccurately predicted by the model*

False Negative (FN): *Negative class is inaccurately predicted by the model*

From the given Test Set following outcomes were retrieved as shown in Table 8.

Table 8: Predicted Values of Positive and Negative category

Total Tweets	Total Tweets (Positive+Negative)	Total Positive Predicted	Total Negative Predicted
7500	5894	3209	2685

The proposed model produced the following output using SciKit-Learn Matrix.

Details of parameters of confusion matrix (Figure 8) shows the improved performance of the CovNet-Dual LSTM hybrid model over other models in Figure 7. The proposed hybrid model is tested with the review dataset and on the basis of the study with the optimum hyper-parameter tuning has been compared with other popular deep learning based models. The observation of comparison is presented in Figure 9.

		Actual	
		Positive	Negative
Predicted	Positive	3055 (TP)	183 (FP)
	Negative	104 (FN)	2552 (TN)

Figure 8: Output of Confusion Matrix

Models	Dataset	Keras				Word2Vec			
		Precision	Recall	F1	Accuracy	Precision	Recall	F1	Accuracy
CNN	Tweets Dataset	89.33	87.65	88.48	87.6	84.65	85.78	85.21	84.32
LSTM		90.16	88.35	89.25	84.3	85.14	84.54	84.84	83.52
Bi-LSTM		87.32	91.65	89.43	86.5	81.65	87.63	84.53	82.6
CNN-LSTM		89.12	90.32	89.72	87.1	83.89	86.62	85.23	82.36
Proposed CovNet-DualLSTM		<b>94.34</b>	<b>96.7</b>	<b>95.5</b>	<b>95.23</b>	<b>91.62</b>	<b>93.89</b>	<b>93.36</b>	<b>91.8</b>

Figure 9: Output of Confusion Matrix

Figure 10 shows the performance of the system with Word2Vec Embedding. The above figure clarifies that hybrid model produces much better results as compared with CNN,LST, BiLSTM and Conv-LSTM.

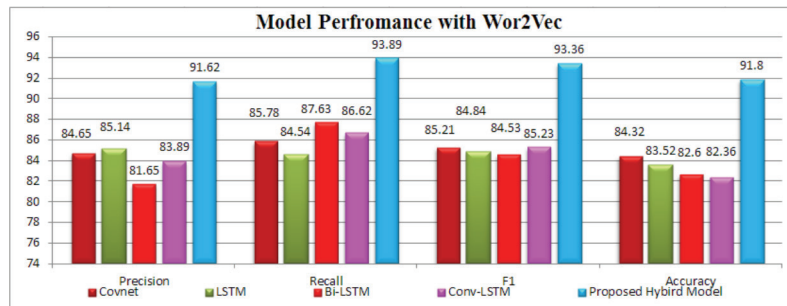


Figure 10: Graphical representation of Model Performance with Word2Vec

The experimental results of performance of the proposed hybrid model with Keras Word Embedding Technique are shown in Figure 11. The hybrid model classification provides a higher level of accuracy and better results than other models. Keras effectively initialises word vectors for datasets, as indicated by the overall correctness of experimental findings. The accuracy level of 95.23% proves that the Keras word vectoring model affected the overall accuracy of proposed model.

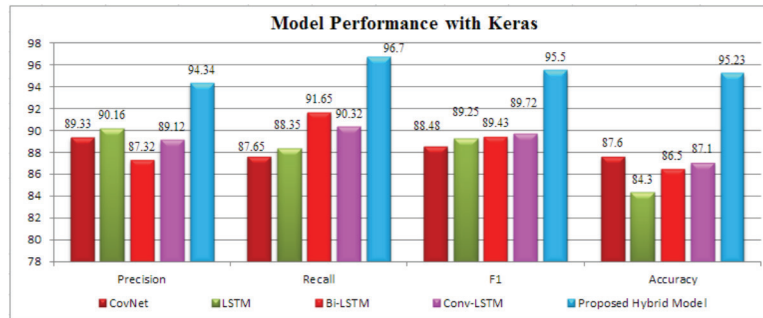


Figure 11: Graphical representation of Model Performance with Keras

Observation in Figure 11 shows that the proposed model used with Keras Embedding has outperformed other models with an improvement in accuracy of 7.63%, 4.18% in precision, 5.05% in recall and 5.05% over all the other models represented in Table 9.

Table 9: Performance Improvement using Keras

Parameter	Performance Improvement
Accuracy	7.63%
Precision	4.18%
Recall	5.05%
F1(F-Measure)	5.78%

Outcomes in Figure 11 show the improvement in performance of the system using Word2Vec embedding as compared to other models. Proposed model with Word2Vec Embedding shows improvement by 1.48% in precision, 1.26% in recall, 2.58% in F1 and 1.65% in accuracy represented in Table 10.

Table 10: Performance Improvement using Word2Vec

Parameter	Performance Improvement
Accuracy	7.48%
Precision	6.48%
Recall	6.26%
F1(F-Measure)	8.13%



It was also observed that is CNN and BiLSTM is used separately on different models, the performance does not provide effective results.

## 5. Conclusion

The training and evaluation of model is performed on the dataset of reviews for Indian Railways. The proposed model performs better utilization and accuracy of the system. Both Keras and Word2Vec embedding were used for word vectorisation separately on proposed hybrid model and the performance of the system has shown a significant improvement in performance. The model was able to effectively classify text sentiment with 7.63% improvement in accuracy using Keras and 7.48% improvement using Word2Vec. The results of the experiment confirmed the model's viability and effectiveness. The proposed model provides a much better performance over other models. The outcome of the study clearly shows that the proposed model works much better than other models using both Word2Vec and Keras embedding.

## 6. Limitation and Future Work

A dataset of 25000 cleaned tweets was used in the study. To test the future performance the proposed model, different datasets can be considered with different sizes. To enhance sentiment classification performance, the suggested Covent-DualLSTM model's structure might be modified in the future. Other pre-processing approaches, such as POS tagging, can be utilised to see if overall accuracy has improved. The studies revealed that word representation has an effect on overall accuracy of the model. As a result, integrating features from various word embedding methodologies may result in increased extraction of feature for the network.

## References

- Amir S., Wallace C. B., Lyu H., Carvalho P., and Silva J. M., «Modelling context with user embeddings for sarcasm detection in social media», in Proc. CONLL 2016.
- Baziotis C., Pelekis N., and Doulkeridis C., «DataStories at SemEval2017 task 4: Deep LSTM with attention for message-level and topicbased sentiment analysis», in Proc. 11th Int. Workshop Semantic Eval. (SemEval), 2017, pp. 747–754
- Bengio Y., Ducharme R., Vincent P., and Janvin C. : «A neural probabilistic language model», J. Mach. Learn. Res., vol. 3, pp. 1137–1155, Feb. 2003.
- Brueckner R. and Schulter B. : «Social signal classification using deep BLSTM recurrent neural networks», in Proc. IEEE Int. Conf. Acoust.,Speech Signal Process. (ICASSP), May 2014, pp. 4823–4827.
- Chowdhury S. M. H., Abujar S., Saifuzzaman M., Ghosh P., and Hossain S. A. : «Sentiment prediction based on lexical analysis using deep learning», in Emerging Technologies in Data Mining and Information Security. Singapore: Springer, 2019, pp. 441–449.
- Cui Z., Shi X., and Chen Y., : Sentiment analysis via integrating distributed representations of variable-length word sequence, Neurocomputing, vol. 187, pp. 126–132, Apr. 2016.

- Du J., Gui L., He Y., Xu R., and Wang X.: «Convolution-based neural attention with applications to sentiment classification», *IEEE Access*, vol. 7, pp. 22983–27992, 2019.
- Feizollah A., Ainin S., Anuar N. B., Abdullah N. A. B., and Hazim M., «Halal products on Twitter: Data extraction and sentiment analysis using stack of deep learning algorithms», *IEEE Access*, vol. 7, pp. 83354–83362, 2019.
- González-Briones A, Prieto J, De La Prieta F, Herrera-Viedma E, Corchado JM. Energy Optimization Using a Case-Based Reasoning Strategy. *Sensors*. 2018a; 18(3):865. <https://doi.org/10.3390/s18030865>
- González-Briones, Alfonso & Villarrubia, G. & De Paz, Juan & Corchado Rodríguez, Juan. (2018b). A multi-agent system for the classification of gender and age from images. *Computer Vision and Image Understanding*. 172. 10.1016/j.cviu.2018.01.012.
- Hochreiter S. and Schmidhuber J. : «Long short-term memory», *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- Kim Y. : «Convolutional neural networks for sentence classification», 2014, arXiv:1408.5882. [Online]. Available: <http://arxiv.org/abs/1408.5882>
- Kim K., Chung B.-S, Choi Y., Lee S., Jung J.-Y, and Park J. : «Language independent semantic kernels for short-text classification», *Expert Syst. Appl.*, vol. 41, no. 2, pp. 735–743, Feb. 2014.
- Lecun Y., Bengio Y., and Hinton G. : «Deep learning», *Nature*, vol. 521, no. 7553, p. 28, May 2015.
- Liu G., Xu X., Deng B., Chen S., and Li L.: A hybrid method for bilingual text sentiment classification based on deep learning in *Proc. 17th IEEE/ACIS Int. Conf. Softw. Eng., Artif. Intell., Netw. Parallel/ Distrib. Comput. (SNPD)*, May 2016, pp. 93–98.
- Liu W., Liu P., Yang Y., Gao Y., and Yi Y. : «An attention-based syntax-tree and tree-LSTM model for sentence summarization», *Int. J. Performability Eng.*, vol. 13, no. 5, pp. 775–782, 2017.
- Majumder N., Poria S., Gelbukh A., and Cambria E. : Deep learning-based document modeling for personality detection from text, *IEEE Intell. Syst.*, vol. 32, no. 2, pp. 74–79, Mar. 2017
- Mikolov T., Chen K., Corrado G., and Jeffrey D. : «Efficient estimation of word representations in vector space», in *Proc. 1st Int. Conf. Learn. Represent*, May 2013a, pp. 1–12
- Mikolov T., Sutskever I., Chen K., Corrado G., and Dean J., «Distributed representations of words and phrases and their compositionality», 2013b, arXiv:1310.4546. [Online]. Available: <https://arxiv.org/abs/1310.4546>
- Niu X., Hou Y., and Wang P. : «Bi-directional LSTM with quantum attention mechanism for sentence modelling», in *Proc. 24th Int. Conf. Neural Inf. Process.* Guangzhou, China: Springer-Verlag, 2017, pp. 178–188.
- Official website of Keras: <https://keras.io/about/>.
- Pang B., Lee L., and Vaithyanathan S., : «Thumbs up?: Sentiment classification using machine learning techniques», in *Proc. ACL Conf. Empirical Methods Natural Lang. (ACL)*, vol. 10, 2002, pp. 79–86.
- Ranjan R., Daniel A.K., Emotion Detection for Travelling Services using Rule-Based Fuzzy Inference System, *Springer International Conference on Machine Intelligence and Smart Systems (MISS2020)*, Rustamji Institute of Technology, Tekanpur, Gwalior, M.P., September 24–25, 2020a
- Ranjan R., Daniel A.K., : Intelligent Sentiments Information Systems Using Fuzzy Logic, *Information and Communication Technology for Intelligent Systems. ICTIS 2020b. Smart Innovation, Systems and Technologies*, vol 195. Springer, Singapore. [https://doi.org/10.1007/978-981-15-7078-0\\_55](https://doi.org/10.1007/978-981-15-7078-0_55)
- Salur M. U. and Aydin I. : «A novel hybrid deep learning model for sentiment classification», *IEEE Access*, vol. 8, pp. 58080–58093, 2020.

- Santos C. D. and Gattit M., : Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts, in Proc. 25th Int. Conf. Comput. Linguistics: Tech. Papers, Dublin, Ireland, 2014, pp. 69–78.
- Santur Y. : «Sentiment analysis based on gated recurrent unit», in Proc. Int. Artif. Intell. Data Process. Symp. (IDAP), Malatya, Turkey, Sep. 2019, pp. 1–5.
- Shen Y., Tan S., Sordoni A., and Courville A., : «Ordered neurons: Integrating tree structures into recurrent neural networks», 2018, arXiv:1810.09536. [Online]. Available: <http://arxiv.org/abs/1810.09536>
- Shi S., Zhao M., Guan J., Li Y., and Huang H. : «A hierarchical lstm model with multiple features for sentiment analysis of sina weibo texts», in Proc. Int. Conf. Asian Lang. Process. (IALP), Dec. 2017, pp. 379–382.
- Socher, R.: Recursive deep models for semantic compositionality over a sentiment treebank, in Proc. Conf. Empirical Methods Natural Lang. Process., Seattle, WA, USA, 2013, pp. 1631–1642.
- Souma W., Vodenska I., and Aoyama H., «Enhanced news sentiment analysis using deep learning methods», J. Comput. Social Sci., vol. 2, no. 1, pp. 33–46, Jan. 2019
- Tkatek S, Belmzoukia A, Nafai S, Abouchabaka J, Ibnou-Ratib Y. Putting the world back to work: An expert system using big data and artificial intelligence in combating the spread of COVID-19 and similar contagious diseases. Work. 2020;67(3):557–572. doi: 10.3233/WOR-203309. PMID: 33164971.
- Voulodimos A., Doulamis N., Doulamis A., and Protopapadakis E., : Deep learning for computer vision: A brief review, Comput. Intell. Neurosci., vol. 2018, pp. 1–13, Feb. 2018.
- Wang Y., Huang M., Zhu X., and Zhao L.,: «Attention-based LSTM for aspect-level sentiment classification», in Proc. Conf. Empirical Methods Natural Lang. Process., 2016, pp. 606–615.
- Wint Z. Z., Manabe Y., and Aritsugi M. : Deep learning based sentiment classification in social network services datasets, in Proc. IEEE Int. Conf. Big Data, Cloud Comput., Data Sci. Eng. (BCD), Jul. 2018, pp. 91–96.
- Xia H., Yang Y., Pan X., Zhang Z., and An W., : Sentiment analysis for online reviews using conditional random \_elds and support vector machines, Electron. Commerce Res., pp. 1–18, May 2019, doi:10.1007/s10660-019-09354-7.
- Xiao Y. and Cho K.: «Efficient character-level document classification by combining convolution and recurrent layers», 2016, arXiv:1602.00367. [Online]. Available: <http://arxiv.org/abs/1602.00367>
- Yenter A. and Verma A. : «Deep CNN-LSTM with combined kernels from multiple branches for IMDB review sentiment analysis», in Proc. IEEE 8th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON), Oct. 2017, pp. 540–546.
- Young T., Hazarika D., Poria S., and Cambria E., «Recent trends in deep learning based natural language processing», IEEE Comput. Intell. Mag., vol. 13, no. 3, pp. 55–75, Aug. 2018
- Zhang Z., Geiger J., Pohjalainen J., Mousa A. E.-D., Jin W., and Schuller B.: «Deep learning for environmentally robust speech recognition: An overview of recent developments», ACM Trans. Intell. Syst. Technol., vol. 9, no. 5, pp. 1–28, Jul. 2018a.
- Zhang D., Hua Xu, Zengcai Su, Yunfeng Xu: Chinese comments sentiment classification based on word2vec and SVMperf., Comput. Sci., vol. 42, no. 4, pp. 1857–1836, Oct. 2016a
- Zhang Q., Yang L. T., Chen Z., and Li P.: A survey on deep learning for big data, Inf. Fusion, vol. 42, pp. 146–157, Jul. 2018b.
- Zhang D. G., Zhou S., and Tang Y.-M.: «A low duty cycle of \_cient MAC protocol based on self-adaption and predictive strategy», Mobile Netw. Appl., vol. 23, no. 4, pp. 828–839, Aug. 2018c.

Zhang Q., Zhang S., and Lei Z.: «Chinese text sentiment classification based on improved convolutional neural networks», *Comput. Eng. Appl.*, vol. 53, no. 22, pp. 111–115, Sep. 2017.

Zhang M., Zhang Y., and Tang D.: Gated neural networks for targeted sentiment analysis, in *Proc. 30th AAAI Conf. Artif. Intell.*, Phoenix, AZ, USA, 2016b, pp. 3087–3093.

Zhou C., Sun C., Liu Z., and Lau F. C. M.: «A C-LSTM neural network for text classification», 2015, arXiv:1511.08630. [Online]. Available: <http://arxiv.org/abs/1511.08630>

## References not cited

Liu S., «Novel unequal clustering routing protocol considering based on network partition & distance for mobile education», *J. Netw. Comput. Appl.*, vol. 88, no. 15, pp. 1–9, 2017.