International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0184
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1228 - 1234

Fine-Grained Emotion Recognition in Noisy Social Media Text Using Deep Bidirectional Sequence Models

R Kaviyarasi[1], Siddarth R R[2], Puneeth Reddy A S[3], Adwaith Raj V M[4], Muhammed Yaseen[5], V Harathi[6]
[1]Department of Computer Science, Yenepoya University, Mangalore, Karnataka, India.
[2,3,4,5,6]BCA Student, Department of Computer Science and Information Technology, Yenepoya (Deemed to be University), Bengaluru Campus, Karnataka, India.
***Email ID:** arasikavi@gmail.com[1], siddharthrudraswami@gmail.com[2], puneethreddyas@gmail.com[3], adwaith746@gmail.com[4], muhammedyaseen3355@gmail.com[5], harathiraju94@gmail.com[6]*

Abstract
Social media has become an important medium for unbridled expressions of thoughts, opinions, and emotions in short textual formats. The sheer volume of textual content on these platforms makes it imperative to explore automatic emotion detection as an area of research in natural language processing and machine learning. This study aims to develop an efficient model for effective emotion detection based on a deep learning framework for multi-class emotion classification of short-form textual content on social media platforms. This study uses DAIR-AI Emotion Datasets containing tweets with emotion labels to train and test the model. This study uses a Bidirectional Long Short-Term Memory (BiLSTM) model to process contextual dependencies by considering both preceding and following word sequences to improve understanding of emotional contexts. The model aims to classify short-form textual content on social media platforms into six different emotions: sadness, joy, love, anger, fear, and surprise. Preprocessing techniques have been used to process the textual content before training and testing the model. The experimental results show that the proposed model based on the BiLSTM model achieves good classification results and effectively differentiates between emotions with subtle variations. This is evident from the analysis of the confusion matrices used to assess model performance. This study indicates that the proposed model is effective and practical for real-world applications. This study highlights the potential of deep learning models for effective emotion detection in noisy and informal communication on social media platforms.

***Keywords:** Emotion Recognition, BiLSTM, Sentiment Analysis, Deep Learning, Text Classification, Social Media Analysis, NLP*

1. Introduction
In recent years, the usage of social media platforms like Twitter, Instagram, and WhatsApp has increased a lot. Because of this, the way people express their opinions and emotions has also changed. Most users share their thoughts through short text messages, and these messages actually contain a lot of emotional information. So, in this work, automatically identifying emotions from such text can be very useful for applications like customer feedback analysis, public sentiment monitoring, and even mental health related studies. Traditionally, sentiment analysis methods were mainly used to classify text into positive, negative, or neutral categories. This is useful, but it is somewhat limited because it does not fully capture the complexity of human emotions. In real life, emotions are more detailed.[1] So, emotion recognition tries to solve this by identifying more specific emotions like sadness, joy, anger, and fear. However, this task is not that simple, since social media text is usually informal. It contains slang, abbreviations, and sometimes the meaning depends on context, which makes it difficult. Recently, deep learning approaches have shown good performance in emotion detection tasks. Especially, models which are based on recurrent neural networks such as Long Short-Term Memory (LSTM), works well in terms of handling sequential data such as text. Also, Bidirectional LSTM (BiLSTM) improves this further, as it processes the text in both forward and backward directions. So we can see that it helps in understanding the context better. Based on these advantages, in this work we are

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0184
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1228 - 1234

using a stacked BiLSTM architecture which was trained on the DAIR-AI Emotion Dataset to perform emotion classification on social media text.

2. Literature Review

Over the last several years, a considerable amount of research work has been carried out within the realm of sentiment and emotion analysis. This is mainly due to the burgeoning growth of social media usage, whereby individuals express their thoughts and feelings transparently through tweets, posts, and comments.[2][3] Hence, a considerable amount of research work has been carried out to advance the knowledge of this area of research. Traditionally, the focus of sentiment analysis research has been to categorize text as either positive, negative, or neutral. Although this information is useful to a certain extent, it is rather limited. This is because human emotions are not always straightforward and cannot be expressed directly. Recently, researchers have tried to find out the discrete emotions of an individual, such as joy, sadness, anger, and fear. These discrete emotions provide a precise characterization of the sentiment of an individual. It is noteworthy to mention here that emotion and sentiment analysis have become a major area of research within the realm of Natural Language Processing (NLP). This is mainly due to the abundance of user-generated information available on the internet, particularly on Twitter and Instagram. In this work, several research studies have been referred to, which have framed our methodological approach. [1] used emoticons like 😊 and 😡 to automatically label tweets and train models without manual tagging.They trained basic classifiers like Naive Bayes using tweets which had these emojis to detect sentiment. It worked quite well for simple cases but didn't do great when tweets were tricky or sarcastic.[2] made a big list (lexicon) of words and their emotions using crowdsourcing. People gave opinions on what feelings each word brings — like anger, joy, etc. This helped in detecting emotions, but couldn't handle sentence-level context that well. [3] used a deep learning model called RNTN that worked based on the grammar tree of a sentence. They parsed sentences and used a recursive model to understand meaning from bottom-up. They got good results for fine-grained sentiment tasks like "very positive", "somewhat negative" etc. [4] used GRU (which is like RNN) to understand long tweets and full documents.Words are converted into embeddings and passed through GRU to catch long-range meaning. They did better than normal models for long texts and were faster than LSTM also [5] was more like a summary paper (survey) that compared deep learning models like CNN, RNN, and attention. They reviewed many past works and discussed how each model works in sentiment analysis. They found that BiLSTM with pretrained embeddings is doing really well most of the time. [6] created a model called CARER that mixes BiLSTM with emotional knowledge (affect embeddings). They added extra emotional info into the model to understand emotions better.It gave better results than normal BiLSTM on many emotion classification datasets. Similarly, recent research also focuses on the importance of incorporating stacked layers with appropriate preprocessing steps, such as tokenization and padding, in improving the performance of the model. In a similar methodology, we also observed that with a certain level of hyperparameter tuning, the performance of the model was high in terms of accuracy. Therefore, from the synthesis of literature, we can infer that deep learning, specifically the use of BiLSTM architecture, can be used for emotion detection in tweets, which are short texts, with certain improvements through the use of better word embeddings or the use of attention mechanisms.

3. Methodology

The description of this section is a general description of the system's operation from initiation to completion. The methodology adopted is a structured sequential approach. The approach is comparable to a procedure, whereby each step is significant. First, there is selection of the dataset, followed by cleaning, preparation, and, finally, training and evaluation of the model. All procedures are done using Python and relevant deep learning libraries. The discussion of each step is provided in the next sections.

3.1 Dataset Description

For the purpose of this research, the DAIR-AI Emotion Dataset was used. This dataset consists of a set of tweets, with each tweet labeled with a single

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0184
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1228 - 1234

emotion type. The emotions are categorized as: Joy, Sadness, Anger, Love, Fear, and Surprise. Every tweet in the dataset carries a single label for the emotion type. The text in the dataset is in the English language, though not in a formal sense, as the tweets used are in a casual tone, with contractions, emojis, and other informal words used in the tweets. The dataset comprises a large number of tweets, over 20,000, with the emotions not equally represented in the dataset, with some emotions having a greater number of tweets than others. However, the dataset can still be used for training a model that can recognize different emotions expressed in the tweets. Before the tweets are fed into the model, a series of preprocessing steps were taken, which included cleaning the text, where unwanted characters are removed, followed by tokenization, where words are converted into numerical form. This was then followed by padding, which ensures that the length of the sequence is fixed, making the process easier for the model. This dataset is particularly useful for this research because of its simplicity, practicality, and its direct association with Twitter, which gives a glimpse into real-life user-emotion expression. The fact that the dataset is already labeled makes the process easier, as no additional steps are required for labeling the dataset.

3.2 Data Pre-Processing

Before the inputting of the model, preprocessing was done on the tweets to improve the quality of the data. This was necessary due to the casual nature of the information on social media. This included changing the case of the text to lowercase so that words like "Happy" and "happy" would be treated the same. This was a basic normalization and tokenization of the text to account for the casual nature of the information on the social media platforms. This was followed by the tokenization of the tweets into individual words. After this, an encoding procedure was done on the tokens to generate a corresponding numerical form for the model to understand. Padding was also done on the tweets to standardize the size of the input to the model. This is necessary for the model to operate effectively without the possibility of ambiguity.

Figure 1 below shows the size distribution of the tweets from the DAIR-AI Emotion dataset. The figure shows that the majority of the tweets are short, containing less than 20 words.

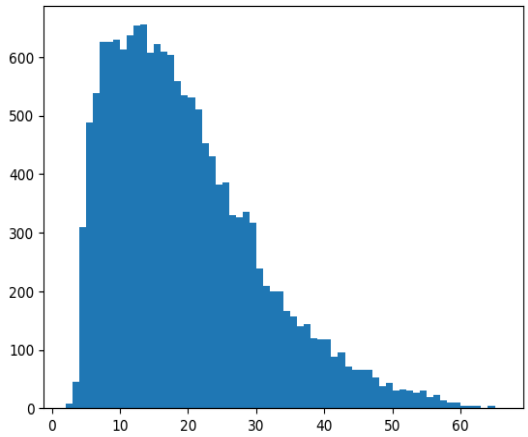

Figure 1 Size distribution of the tweets from the DAIR-AI Emotion dataset

Figure 2 shows the distribution of emotion classes within the dataset. Although certain emotions, for example, joy and sadness have slightly more samples than other emotions, the dataset is balanced to a certain extent, reducing the bias during the training of the model.

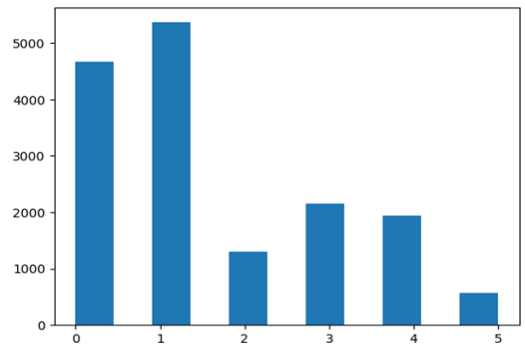

Figure 2 The distribution of emotion classes within the dataset

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0184
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1228 - 1234

3.3 Model Architecture

In this work, we used a deep learning model called Bidirectional LSTM (BiLSTM). Basically, BiLSTM reads the sentence in both directions, that is from left to right and also right to left. Because of this, the model can understand the full context of the sentence better, even if the word order is slightly different. Here, we built a stacked BiLSTM model, which means we used two BiLSTM layers one above the other. This helps the model to learn more deeper features and also capture smaller emotional patterns present in the text. The overall architecture of our model is as follows:

- **Embedding Layer:** First, the input words (which are already converted into numbers) are passed into an embedding layer. This layer converts each word into a dense vector form, which helps in capturing the meaning of the words in a better way.
- **BiLSTM Layers:** After that, we use two BiLSTM layers stacked together. These layers process the text in both forward and backward directions. So, we can see that the model is able to capture context from the entire sentence, which improves understanding of the emotional tone.
- **Dense Layer:** Finally, a dense (fully connected) layer with softmax activation is used. This layer gives the final output by classifying the input text into one of the six emotion categories.

3.4 Model Training and Evaluation

In this work, training the model is somewhat like teaching a small kid to identify emotions from tweets. First, the tweets were converted into numerical form using tokenization. After that, all the sequences were padded to the same length, so that the model does not get confused because of different tweet lengths. The dataset was divided into three sections. Approximately 80% of the tweets were for training, 10% of the tweets were for validation, and the last 10% of the tweets were for testing. The loss function for training was set as categorical cross-entropy. In addition, we set the optimizer as Adam optimizer.[7] Adam optimizer will help in adjusting the learning rate of the model, depending on whether the model is learning fast or slow. We also set early stopping for training. Early stopping will help in stopping the training process once there is no improvement in the validation accuracy.This will help in saving time during training.Accuracy and confusion matrix are the metrics used for testing the performance of the model.Accuracy will help in understanding how correctly the model is performing. Confusion matrix will help in understanding how correctly the model is performing in classifying different emotions.Confusion matrix will also help in understanding how confused the model is, especially when the emotions are similar.

3.5 Baseline Bilstm Model

Firstly, we implemented a baseline model for emotion classification using a Bidirectional LSTM (BiLSTM) architecture. This model is called the baseline model. This model is basically a reference point. We can use this to compare the performance improvement we achieve by using other more complex models. In the baseline model, we use an embedding layer with a fixed vocabulary size. Then we stack two BiLSTM layers on top of this embedding layer with a limited number of hidden units. This is a basic model where the architecture is kept simple and easy to train. This is because the text we are dealing with is short and informal, like tweets.

3.6 Fine-Tuned BiLSTM Model

Apart from the baseline model, we also tried our hands on a fine-tuned BiLSTM model by adjusting some of the hyperparameters. In this, we have tried to adjust some of the hyperparameters, which include the embedding dimension, the number of units in the LSTM, and some of the training hyperparameters, which include the learning rate and the number of epochs. This will help the model learn even better and obtain some fine features from the data. This model has also been trained on the same dataset, with the same pre-processing techniques and evaluation metrics as the baseline model. This is done in order to keep the comparison fair, and we can see how the fine-tuning of the model affects its performance.

4. Results and Discussion

In this work, after training both models on the DAIR-AI Emotion Dataset, their performance was evaluated

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0184
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1228 - 1234

using accuracy and confusion matrix. We will now discuss the results obtained and what they actually represent about how the model is performing.

4.1 Experimental Results

The training process was evaluated by using training accuracy and loss graphs. From these graphs, we can see how the accuracy of our model was improving with time. Also, we can see how the loss of our model was decreasing with time. This actually represents how our model was able to learn from the tweet data. We can also see how our training accuracy was able to reach a very high value after a few epochs. Our validation accuracy was also following a similar trend and was close to our training accuracy. This actually represents how our model was not suffering from overfitting.[8] Looking at our loss graphs, we can see how our training and validation loss was decreasing gradually. This actually represents how our learning process was stable.

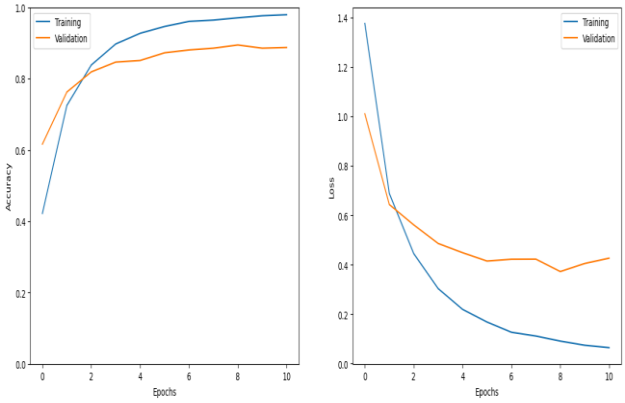

Figure 3 Training and Validation Accuracy and Loss over Epochs

4.2 Baseline vs Fine-Tuned Model Comparison

In this work, we have compared both models based on their accuracy. The accuracy of the baseline model is 92.75%, whereas the accuracy of the fine-tuned model is around 88.85%. From this comparison, we can conclude that the baseline model performed better than the fine-tuned model.[9] This is because, in general, it is expected that a simpler model can perform better for a short and informal text like tweets.

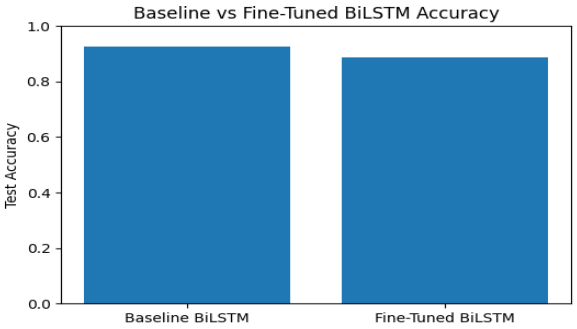

Figure 4 Comparison of Baseline and Fine-Tuned BiLSTM Model Accuracy

4.3 Confusion Matrix Analysis

In order to analyze the performance of the model, we have also used a confusion matrix. The confusion matrix is helpful in understanding how well the model is performing for different emotions. It also helps in understanding how much confusion is there for different emotions. From this confusion matrix, it is clear that emotions like joy and sadness are predicted more accurately. However, there is confusion between emotions like love and joy, and also emotions like fear and surprise. This is because these emotions are described using almost the same words or expressions in tweets. From this confusion matrix, it is clear that the model is performing well in distinguishing different emotions. However, there is some overlap between emotions like love and joy, and also emotions like fear and surprise. This is quite normal.

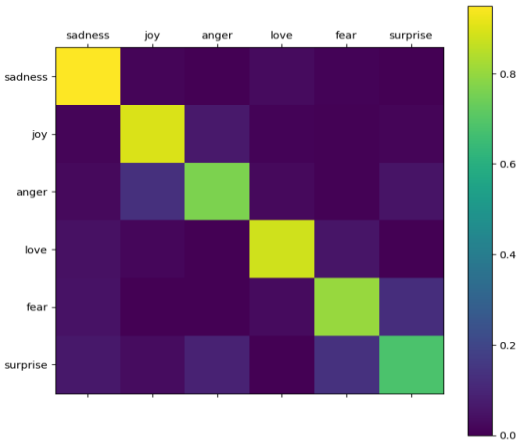

Figure 5 Confusion Matrix for Emotion Classification Model

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0184
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1228 - 1234

4.3.1 Strengths of the Model

In this work, the model is quite simple to design and easy to understand. It does not need very complex architecture, so this model is also suitable for beginners.[10] Training is also very fast, and the model works well even on low-end systems. So, we can say that this model is computationally efficient. This model also works well in detecting common emotions from short text messages such as tweets. Another advantage is that this model does not need heavy preprocessing and tuning techniques. In a way, this model works well even with simple steps.

4.3.2 Limitations and Challenges

However, there are some limitations for this model. This model gets confused sometimes when handling sarcasm, slang, and mixed emotions, which is common in social media platforms. Another limitation is that this model is not very good at handling very complex emotions from text data. This is because the dataset is not very large, which sometimes affects the model's capability to work well with new data.However, this model is good enough for practical use cases.

Conclusion

In this work, we have tried to develop a model using the BiLSTM algorithm to recognize human emotions using tweets with the help of the DAIR-AI Emotion Dataset. Without using complex models or tuning the models, we have tried to achieve good results in classifying emotions such as sadness, joy, anger, love, fear, and surprise.[11][12] While evaluating the model, we have observed that the model has performed well in classifying the emotions. Confusion occurs between similar emotions; otherwise, the model is doing well. Considering that we have used a simple model for emotion detection, the accuracy obtained is quite good. This shows that even a simple deep learning algorithm can perform well in emotion detection. Therefore, in this work, we have tried to show a simple approach for emotion detection using a deep learning algorithm.

Future Scope

There are a number of ways this work can be enhanced in the future. First, advanced models like the ones based on the transformer architecture, such as BERT or RoBERTa, could be used for better understanding and accuracy.[13][14] Second, currently, the model is designed for only six emotions. However, in the future, this range could be increased by adding different types of emotions, such as jealousy, guilt, or pride, through a larger dataset. Third, currently, the model is based only on text data. However, if we include text, emojis, and even speech, we could obtain even better results through multi-modal learning. Lastly, this model could be integrated with social media platforms like Twitter or Instagram.[15] This way, we could analyze the emotions of people in real-time, regarding a product, an event, or a social issue.

References

[1]. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," Stanford University, 2009.
[2]. S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," Computational Intelligence, vol. 29, no. 3, pp. 436–465, 2013.
[3]. R. Socher et al., "Recursive deep models for semantic compositionality over a sentiment treebank," in Proc. Empirical Methods in Natural Language Processing (EMNLP), 2013, pp. 1631–1642.
[4]. B. Felbo et al., "Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm," in Proc. EMNLP, 2017, pp. 1615–1625.
[5]. J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. NAACL-HLT, 2019, pp. 4171–4186.
[6]. M. Abdul-Mageed and L. Ungar, "EmoNet: Fine-grained emotion detection with gated recurrent neural networks," in Proc. ACL (Short Papers), 2017, pp. 718–728.
[7]. F. Calefato, F. Lanubile, and N. Novielli, "EmoTxt: A toolkit for emotion recognition from text," in Proc. Int. Workshop on Emotion Awareness in Software Engineering, 2017, pp. 1–7.
[8]. B. Liu, Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, pp. 1–167, 2012.
[9]. J. Staiano and M. Guerini, "DepecheMood: A lexicon for emotion analysis from crowd-annotated news," in Proc. ACL (Short Papers), 2014, pp. 427–433.
[10]. S. M. Mohammad, S. Kiritchenko, and X. Zhu, "NRC-

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0184
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1228 - 1234

Canada: Building the state-of-the-art in sentiment analysis of tweets," in Proc. SemEval, 2013, pp. 321–327.

[11]. C. N. dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," in Proc. COLING, 2014, pp. 69–78.

[12]. D. Tang, B. Qin, and T. Liu, "Document modeling with gated recurrent neural network for sentiment classification," in Proc. EMNLP, 2015, pp. 1422–1432.

[13]. L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 8, no. 4, p. e1253, 2018.

[14]. Y. Kim, "Convolutional neural networks for sentence classification," in Proc. EMNLP, 2014, pp. 1746–1751.

[15]. J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," Journal of Computational Science, vol. 2, no. 1, pp. 1–8, 2011.