Emoji Prediction Using Bi-Directional LSTM

Vinayak Sudhakar Kone[1*], *Atrey Mahadev* Anagal[1], *Swaroop* Anegundi[1], *Priya* Jadekar[1] and *Priyadarshini* Patil[1]

[1]School of Computer Science and Engineering, KLE Technological University, Hubballi (Karnataka), 580031, India

Abstract. Messengers and social media dominate today's internet usage across the globe. For the large population, a typical day starts with messages flooding on mobiles, from simple good morning wishes, business meeting invites, reminders, and schedules for the day and the list is endless. A striking feature of today's digital communication is the variety of emojis used, without which text communication almost look incomplete. Emojis are graphic symbols/logograms used with text communication to enhance the effectiveness of emotions and set an undertone that makes texting a more fun experience for the users. Emojis are the visual language of the new generation. They give consumers a means to communicate their feelings while reducing the quantity of text that needs to be typed by the sender. Every social media and messenger platform like Facebook, Instagram, Twitter, WhatsApp, and many more have its own emoji set. To lure more and more users, many new emojis are added day by day. Predicting and suggesting emojis based on the text, emotion and user patterns to the user is an important feature of today's messengers and social media applications. If you start typing a message, relevant emojis will be displayed from which users can choose an emoji, further enhancing the user texting experience. This process is done using natural language processing and machine learning techniques. In this paper, we study emoji prediction techniques and propose an emoji prediction model using bi-directional LSTMs. We compare emoji prediction NLP techniques, including RNN, LSTM, LSTM networks, and Bi-LSTM. Based on our implementation, we suggest that the bi-directional LSTM model is the most effective technique. Our model outperforms many baseline approaches with an accuracy of 94% when tested on a CodaLab Twitter data set with 60000 rows and two columns. Our study shows the effectiveness and efficiency of bi-directional LSTMs for text-based systems for communication.

1 Introduction

According to the latest survey data from data report [1], more than over 4.76 billion individuals are using social media apps in 2023. In today's digital era, people worldwide are connected over different social networks. For efficient and easy communication, people tend

---

*Corresponding author: vskonenpn@gmail.com

to use emojis that enhance the way to express one's emotions. Some emojis are even added to the Oxford Dictionary, thus proving their importance. Emojis are not only a better way to express feelings but also reduce the number of sentences to be typed by the user. The Emoji Research team discovered that emojis are used online by 92% of users. Emoji usage on Messenger is estimated by Facebook to be five billion per day on average. This shows emojis are an integral part of today's digital communication. Emojis are more than a passing trend, despite research also showing that people and brands use them differently. Emojis are small digital images or icons used to express an emotion or idea in a text-based conversation. They are used in chats, emails, and text messages to convey emotions or add context to a conversation. Emojis are the visual language of the new generation. [2] [3]

The reasons why emojis are becoming so popular are:
• Better Connect: Emojis draw attention, just the way photographs do. Emoji use generates 57% more likes, 33% more comments, and 33% more shares on Facebook. Emojis can show happiness, sadness, anger, surprise, love, and many other emotions. They can also be used to add humour to a conversation or to emphasize a point. Emojis are an incredible way to add personality to a conversation and make it more engaging.
• Give texts a human touch: People like emojis because they give our words emotion. Emojis help to personalise your postings and answers.
• Speak more in less time: A single well-selected emoji can help communicate emotions and conserve space. This is important when telling a narrative on social media, where it is ideal to keep it brief and to the point.

The use of emojis in chats began in 1999 when Japanese mobile carrier NTT DoCoMo released its first set of 176 emojis. The word "emoji" is a translation of the Japanese "e" which means picture and "moji" which means letter. In the following decade, the use of emojis began to grow and spread to other platforms and messaging services, including Apple's iOS in 2011 and WhatsApp in 2013. The increasing availability of emojis to users of different devices has further contributed to the rise of emoji usage in chats. In recent years, the popularity of emojis has exploded, with over 3 billion emojis sent daily on Facebook alone. Emoji support is available on popular social media networks including Facebook – 3300 emojis, Instagram – 3000 emojis, Snapchat – 3000 emojis, Twitter – 3245 emojis, Telegram – 3000 emojis and WhatsApp – 3664 emojis. The most widely used emojis, according to Emojipedia [4], are the following: read heart, heart hands, smiling face with smiling eyes, fire, and sparkles.

User experience with using emojis in chats has been overwhelmingly positive. People of all ages, genders and cultures find emojis a fun and effective way to express emotions and ideas without typing out long messages. Emojis provide a quick and effortless way to express feelings and reactions in text, often humorously or creatively. Additionally, emojis can help break the ice in conversations, making them more enjoyable and engaging for all participants. Emoji prediction using NLP can be applied in day-to-day life in a variety of ways. For example, it can help people to express their emotions better in text messages and emails. It can also be used to help people understand better the feelings of others in conversations. Additionally, it can help people understand the sentiment of online reviews and comments much more efficiently, allowing them to make more informed decisions. Finally, it can help businesses better understand customer sentiment and tailor their products and services accordingly. Hence, the field of automatic emoji prediction and suggestions to enhance user experience has gained attention from researchers.

The main problem with predicting emojis for given statements is that the meaning of a statement can be interpreted differently by different people. For example, one person may interpret a statement as being positive, while another may interpret it as negative. Additionally, the statement's context can also affect how it is interpreted. For example, a statement that is said jokingly may be interpreted differently than if told seriously. Furthermore, the use of slang or colloquialisms can make it difficult to predict the right emoji for a given statement accurately. Today's technology is using machine learning to predict emojis in chats. This technology uses natural language processing and algorithms to process and understand the context of conversations. It then uses this information to recommend relevant emojis based on the exchange. This technology can help people express themselves better and provide a more natural way to communicate. Sample emojis can be seen in Fig.1.

Fig. 1. Emojis

There are different techniques used for prediction. For example, the Recurrent Neural Network [5] commonly called as RNN model is used to process the sequential data format [6] because the output of each layer is fed to the next layer as an input. This technique is also referred to as the feed-forward method. Hence RNNs can recognize sequential characteristics and use these patterns for prediction. An RNN called LSTM (Long short-term memory) [7] is a better approach for predictions because it has a memory that solves the issue of long dependence; i.e., LSTM can store the output of a layer too many farther layers than just an immediate layer, as in RNNs. Another variant of RNN is Bi-Directional LSTM. Adding to memory in LSTM, in a Bi-LSTM, the data can flow in both directions hence accumulating past and future data from nodes for prediction. The paper goes through each of these methodologies in detail in different sections. Initially, emojis were predicted to be used primarily by younger users and to express emotion in digital communication. They were also initially seen to express feelings or emotions without using words. Today, AI-powered technologies are being used to predict emojis by analysing the context and sentiment of a message. For example, natural language processing (NLP) algorithms can find the sentiment of a message, such as whether it is positive or negative, and can suggest an emoji that best reflects the sentiment. AI (Artificial Intelligence) models can also be trained to recognize the context of a message and suggest a suitable emoji. For instance, facial recognition algorithms can find faces in an image and suggest a suitable emoji.

This paper's proposed method is based on Bi-Directional LSTM, using the Twitter data set of 12MB. Dense Hidden layers consisting of neurons with ReLU (Rectified Linear Activation Unit) are appended to Bi-LSTM. The SoftMax layer makes the final prediction score for the output. Depending on the class that gets the highest score is taken as the final predicted output.

The paper's structure begins with a review of existing research on the topic in Section 2. This review helps inform our understanding of the problem, find potential areas for improvement, and highlight gaps in current knowledge. It also highlights the strengths and limitations of different approaches. In Section 3, we outline our method for this study. Our research findings are presented in Section 4, followed by a summary of key takeaways and potential future directions in Section 5.

2 Related Work

The author of this paper [8] explores the use of federated learning to improve emoji prediction in a mobile keyboard. Federated learning is a distributed machine-learning approach which allows multiple data sources to train a machine-learning model collaboratively without sharing the data. The paper specifically focuses on using federated learning to improve the accuracy of emoji prediction. It presents a method which trains a model on data from multiple users in a privacy-preserving way by using federated learning. This method has been tested on a data set having the typing patterns of over a million users and the results prove that federated learning can improve the accuracy of emoji prediction by a significant margin. The paper also analyses the trade-offs between privacy and accuracy when using federated learning. Federated learning is computationally expensive as it requires multiple communication rounds between the clients and the server. The model does not consider the user's earlier usage of emojis, which can help predict the user's future emoji usage. The authors evaluated the model on two publicly available data sets and found that their proposed model was able to predict the output with an accuracy of 86.3% on the Emoji50K data set and an accuracy of 94.2% on the Emoji4K data set.

The author of this paper [9] presents a novel approach for predicting emoticons, or "emojis", using Long Short-Term Memory (LSTM) and Naive Bayes algorithms. The paper first explores the current state of research on emoji prediction and classifies existing techniques into two categories: machine learning-based algorithms and rule-based algorithms. It then proposes a hybrid approach, combining both techniques, that combines the advantages of LSTM and Naive Bayes algorithms to improve emoji prediction. The paper evaluates the proposed approach on a data set of over 10,000 tweets and shows that the hybrid model achieved higher accuracy and F1 scores than the individual models. The paper also reported that the LSTM model performed better than the Naive Bayes model in terms of accuracy, with the LSTM model achieving an accuracy of 0.88, compared to the Naive Bayes model's accuracy of 0.86.

The author of this paper [10] presents a novel approach to the task of detecting emotions in text using deep learning and big data. The paper first outlines the challenges of emotion detection in text and the need for better approaches to the task. It then introduces the proposed method, which is based on deep learning and big data. This approach involves using a convolutional neural network (CNN) to generate contextualized emotion embeddings, which are then used to train a logistic regression classifier. The paper reports experimental results on a range of benchmark data sets, showing that this approach outperforms existing methods in terms of accuracy. The paper also discusses future extensions of the approach, such as incorporating more features and using different models. Finally, the paper concludes by summarizing the advantages of the proposed approach and its potential applications. The model achieved an F1 score of 0.922, a precision of 0.919, and a recall of 0.925 on a test data set consisting of 10,000 text samples.

The author of this paper [11] provides an overview of the advances, challenges, and opportunities in text-based emotion detection. It reviews the current ultramodern, discussing the challenges and opportunities posed by recent advances in natural language processing and machine learning. The authors also discuss the need for more research into the effects of emotion detection on user experience and the ethical implications of using emotion detection in various contexts. Finally, the paper outlines directions for future research into emotion detection in text-based applications. The performance of the model is not explicitly said. However, the paper does supply information on the accuracy of several different models used for emotion detection from text. For example, the paper cites a study which used an SVM (Support Vector Machine) model to achieve an accuracy of 82.3% and another which used a BLSTM-CNN model to achieve an accuracy of 87.8%.

The author of this paper [12] describes a method for predicting emoticons on textual data using a stacked Long Short-Term Memory (LSTM) model. The stacked LSTM model is composed of multiple LSTM layers, which are used to process the text and generate an output corresponding to a particular emoticon. The model is trained using a set of labelled text documents, and the performance of the model is evaluated using a classification accuracy metric. Results show that the stacked LSTM model can accurately predict emoticons on textual data. The performance of the stacked LSTM model was evaluated on the SemEval 2019 Task 3 data set. The model achieved an accuracy of 77.89%, which was the highest accuracy among all the models evaluated on the same data set. In addition, the model achieved an F1-score of 77.80%, which was also the highest among all the models evaluated on the same data set.

2.1 Challenges in the prediction of Emojis-

Prediction of emojis from given sentences in chats accurately is a challenging task due to several reasons. Firstly, the context of the conversation is often very subtle and nuanced, and it can be difficult for a computer to understand the underlying meaning of a sentence. Secondly, the language used in chats is often informal and colloquial; therefore, it can be difficult for a computer to interpret the intended meaning of a sentence. Thirdly, emojis are often used to convey emotions or feelings, and a computer may not be able to interpret these emotions or feelings accurately. Finally, chats often have slang or abbreviations, which can be difficult to interpret by a computer. The primary challenge [13] in emoji prediction is the need for more context-specific user data. Since emojis are often used to express a user's emotional state or to add other information to an existing conversation, it is difficult to accurately predict an emoji without having access to the user's context. Additionally, since the same emoji can have different meanings depending on the user, it is easier to create the right prediction with user-specific data.

3 Proposed Methodology

3.1 Dataset Description-

The Twitter data set from CodaLab having around 60000 tweets has been used for the proposed work.

3.2 Dataset Structure-

The data set has 50000 rows and 2 columns. The first column is the textual data and the second column is the emoji assigned to each tweet. To predict more emojis, 10000 additional tweets were added to the data set, making the total number of emojis 40. It includes a red heart, a smiling face, a disappointed face, thumbs-up, a sleeping face and other emojis. In this data set each sentence is converted into only one emoji. The Total size of the data set is 12MB.

3.3 Proposed Architecture-

The suggested architecture has an embedding layer, a layer of Bi-Directional LSTMs (Long Short-Term Memory), dense hidden layers that pass their activations on to a SoftMax activation function that outputs the final prediction scores, and finally, a layer of dense hidden layers. The sequences produced from the text data are initially sent to the neural network batch-wise. These sequences are sent to an embedding layer, where a constant-length vector is created for each sequence number and updated during training. These embedding vectors are designed to make words connected similarly to one another. In other words, strongly related words will have similar vectors after calculation. The calculated vectors are sent to a layer of Bi-Directional LSTMs, which can keep data in their memory after the embedding layer. These cells store contextual information in their long-term and short-term memories, which are continually refreshed throughout training. The network connections are shown in Fig. 2.
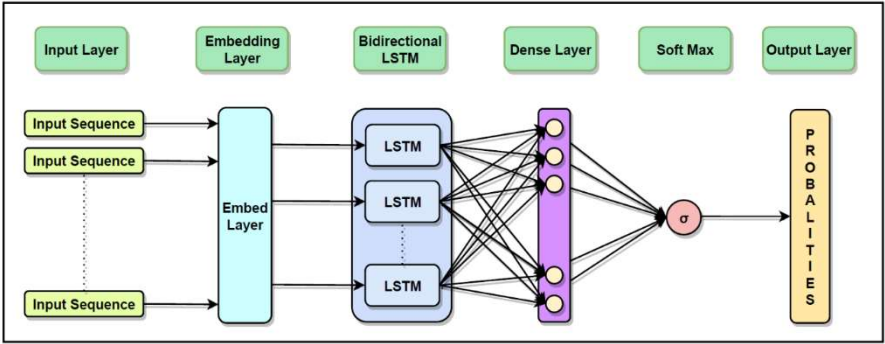
Fig. 2. LSTM Network

The dense layer, which essentially contains neurons with Rectified Linear activation function, receives the output from the LSTMs layer next (ReLU). These cells then send their activation scores to a different layer of neurons. Here, the SoftMax function will be used to activate these neurons. The SoftMax function generates a range of probabilities with one based on the activity of all preceding neurons. The model's final output will be a projected probability for each class. One of these may be picked as the anticipated class, depending on which class has the highest likelihood.

3.4 Techniques

The following Techniques were used to Predict Emojis

3.4.1 RNN –

Recurrent Neural Networks (RNNs) are artificial neural networks used to process sequential data. Unlike traditional neural networks, which are "feed-forward" networks, RNNs are "recurrent" networks, meaning they can remember information from earlier input and use it to influence future outputs. This memory allows them to capture patterns in data that would otherwise be too complex for a single network to detect. In addition, RNNs can process data of varying lengths, making them ideal for text, voice, and other time-series data.

3.4.2 LSTM –

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network useful for predicting sequences. They are commonly used for natural language processing, such as

predicting the next word in a sentence and can also be used for predicting emojis. By training on a data set of text with associated emojis, LSTM networks can learn to recognize text patterns correlated with certain emojis. The network can then use this information to predict the rightest emoji for a given sentence. This type of CNN can be useful for applications such as sentiment analysis, where an exact emoji prediction can help to figure out the sentiment of a given sentence accurately. The diagram of the LSTM Block and LSTM Notations can be seen in Fig.3. and Fig.4. respectively.
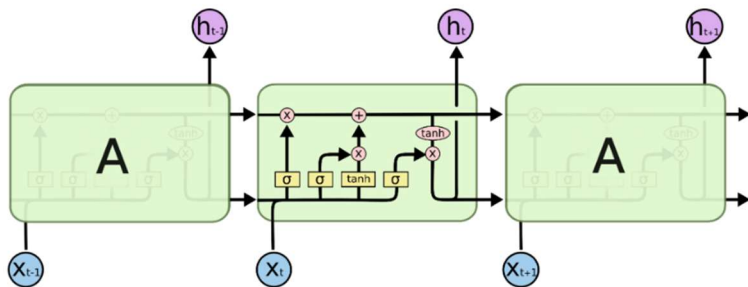
Fig. 3. LSTM Block
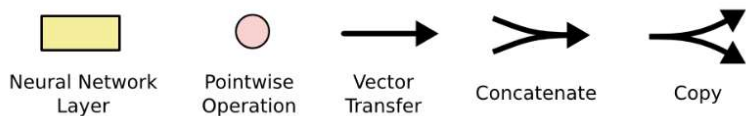
Fig. 4. LSTM Notations

3.4.3 LSTM Networks –

LSTMs are composed of "cells" that hold both an input gate, an output gate, and a forget gate. These gates control the flow of information into and out of the cell and can decide which information to remember and which to discard. These cells allow LSTMs to selectively keep or forget information, which enables them to support a state over time and capture long-term dependencies in data. In addition to the gates, LSTMs also have a cell state, a vector used to store information from the input. The gates change the cell state and can be thought of as a "memory" of the LSTM. The output of the LSTM is also computed using the cell state.

3.4.4 Bi - LSTM –

Bi-Directional Long Short-Term Memory (Bi-LSTM) is a type of recurrent neural network (RNN) with two separate layers of neurons, one running the input sequence forwards and another running it backwards. It allows the network to capture contextual information from both directions in the input sequence, allowing it to understand the data better and make better predictions. The Bi-Directional architecture can also improve the performance of language models, especially for tasks like sentiment analysis. Bi-LSTMs are a powerful tool for natural language processing (NLP) and have been used in many applications such as voice recognition, machine translation, and question answering.

3.5 LSTM Structure

The LSTM holds Memory Cell and three gates: An input Gate, Forget gate and Output Gate.

3.5.1 Memory Cell –

In Long Short-Term Memory (LSTM) networks, the memory cell is a principal part that stores information over time. It is composed of a vector changed by the input, output, and forget gates, which control the flow of information into and out of the cell. The memory cell stores information about the input data that the LSTM processes. This information can be kept for an extended period, allowing the LSTM to capture long-term dependencies in data. The image of a memory cell is shown in Fig.5.
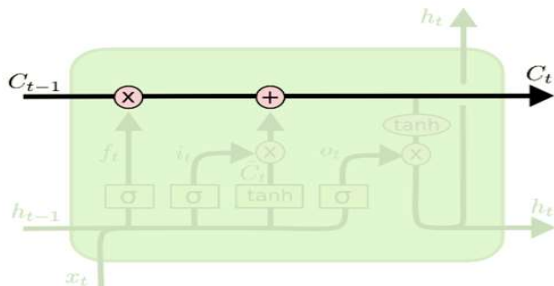
Fig. 5. Memory Cell

The output of the LSTM is also computed using the memory cell. This allows the LSTM to use the information it has stored in the cell to make predictions about the output at each time step.

$$C_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{1}$$

In the context of predicting emojis, the memory cell of an LSTM could be used to store information about the words or phrases that have been processed by the LSTM. This information could then be used to make predictions about the rightest emoji to use in each context.

3.5.2 Input Gate –

In Long Short-Term Memory (LSTM) networks, the input gate is used to control the flow of information from the input data into the memory cell. It does this by deciding which parts of the input data should be added to the cell state and which should be discarded. The image of the Input Gate is shown in Fig.6.
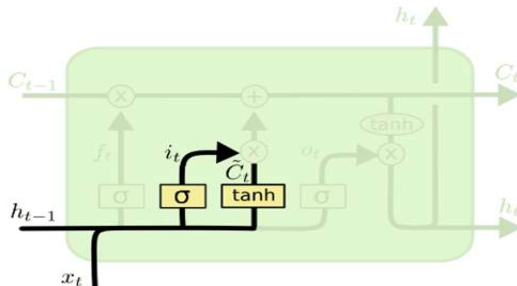

Fig. 6. Input Gate

$$i_t = \sigma(W_i \cdot [\text{h}_{t-1}, x_t] + b_i) \tag{2}$$
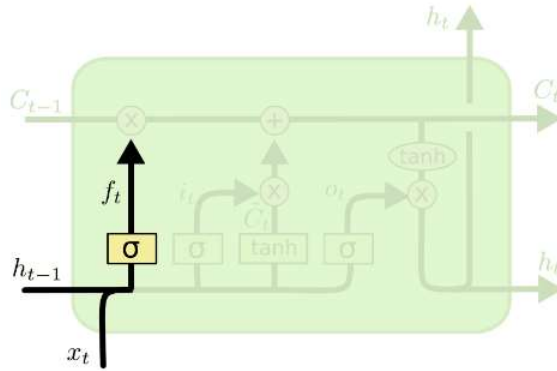
where the sigmoid function is represented as:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

The input gate is implemented as a sigmoid layer, which produces a value between 0 and 1 for each element of the input data. This value represents the "importance" of each element, with a value of 0 showing that the element should be discarded and a value of 1 showing that it should be added to the cell state. The input data is then multiplied element-wise by these important values and added to the cell state. In the context of predicting emojis, the input gate of an LSTM could be used to decide which parts of the input text are most relevant for making a prediction.

3.5.3 Forget Gate –

In Long Short-Term Memory (LSTM) networks, the forget gate is used to control the flow of information from the memory cell to itself. It does this by deciding which parts of the earlier cell state should be kept and which should be discarded. The image of Forget Gate is shown in Fig.7.

$$f_t = \sigma(W_f \cdot [\text{h}_{t-1}, x_t] + b_f) \tag{4}$$

Fig. 7. Forget Gate

The forget gate is implemented as a sigmoid layer, which produces a value between 0 and 1 for each element of the cell state. This value represents the "importance" of each element, with a value of 0 showing that the element should be discarded and a value of 1 showing that it should be kept. The earlier cell state is then multiplied element-wise by these important values and passed on to the current time step. In the context of predicting emojis, the forget gate of an LSTM could be used to decide which parts of the earlier cell state are still relevant for making a prediction.
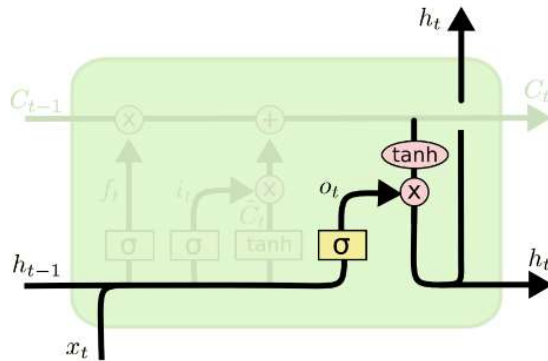
3.5.4 Output Gate –

In Long Short-Term Memory (LSTM) networks, the output gate is used to control the flow of information from the memory cell to the output of the LSTM. It does this by figuring out which parts of the memory cell should be used to compute the output and which should be

discarded. The output gate is implemented as a sigmoid layer, which produces a value between 0 and 1 for each element of the memory cell. This value represents the "importance" of each element, with a value of 0 showing that the element should be discarded and a value of 1 showing that it should be used to compute the output. The memory cell is then multiplied element-wise by these important values and the resulting vector is used to compute the output of the LSTM. In the context of predicting emojis, the output gate of an LSTM could be used to decide which parts of the memory cell are most relevant for making a prediction. The Diagram of the Output Gate is shown in Fig.8.

$$o_t = \sigma\,(W_o \cdot [\mathrm{h}_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = O_t \cdot \tanh C_t \tag{6}$$

Fig. 8. Output Gate

3.6 Understanding Bi-Directional LSTM -

Bi-Directional Long Short-Term Memory (Bi-LSTM) is a type of neural network architecture that is used to process sequences of data. It is a variant of the standard Long Short-Term Memory (LSTM) network, which is a type of recurrent neural network (RNN) that has been developed to process long sequences of data. The main difference between the standard LSTM and the Bi-LSTM is that the latter process the data both forward and backwards. This provides the network with more context and a better understanding of the data. This has made Bi-LSTMs popular in many applications, including natural language processing and time-series forecasting. Recently, Bi-LSTM has been used to predict emojis. Emojis are pictorial emotions used in text messages and other digital communication forms. It is difficult to predict emojis based on the text alone, as the sentiment of the text may not always correlate with the sentiment of the corresponding emoji. Bi-LSTMs, however, can supply a better understanding of the context of the text, allowing for more correct predictions of the right emoji. By processing the text both forward and backwards, the model can gain a better understanding of the sentiment of the text and better predict the right emoji. The image of LSTM Connected Layers can be seen in Fig.9.

Bi-Directional LSTMs (Long Short-Term Memory) are a type of recurrent neural network that is trained to process sequential data. They are particularly useful for natural language processing tasks because they can capture context from both the past and the future of a given sequence. In predicting emojis in chat messages, a Bi-Directional LSTM could be used to analyze the words in the message and the conversation's context to predict which emoji would be most proper.

Here is how a Bi-Directional LSTM might work for this task:

First, the LSTM would process the words in the message, considering the order in which they appear and the relationships between them. This would allow the model to understand the meaning of the message and the sentiment it conveys.
Next, the LSTM would consider the context of the conversation. This might include earlier messages in the conversation, the relationship between the participants, and any other relevant information.
Using this information, the LSTM would then predict which emoji would be most right for the message.

Overall, the role of a Bi-Directional LSTM in predicting emojis in chats is to use its ability to process sequential data and understand the context to make correct predictions about which emojis would be most right for a given message.
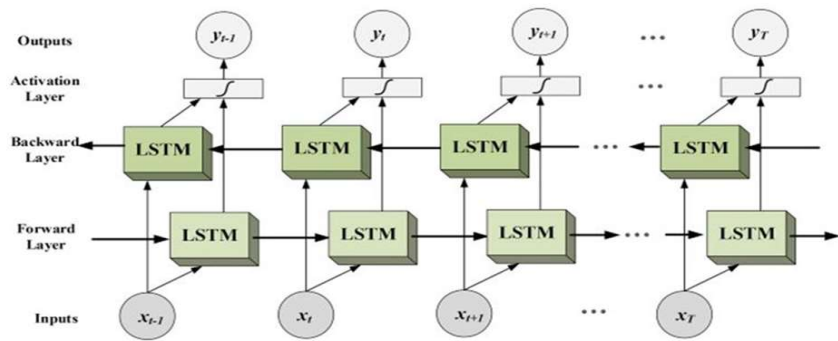
Fig. 9. LSTM Connected Layers

3.7 Implementation Plan-

3.7.1 Pre-processing the dataset -

Pre-processing of a data set is a crucial step in the predictive analysis process. It involves transforming the data into a format suitable for processing by a machine learning algorithm. Pre-processing of a dataset for predicting emoji using an LSTM involves several steps including data cleaning, feature extraction, tokenization, and padding. Data cleaning involves removing any irrelevant or noisy data from the data set, such as duplicated records and incorrect values. Feature extraction involves extracting the most key features from the data set, such as words, phrases, and symbols relevant to the task. Tokenization involves splitting the data into individual tokens, such as words or phrases, for easier processing by the machine learning algorithm. Padding involves adding zeros to the beginning or end of a token sequence to ensure that all sequences are of equal length. These steps are necessary to ensure that the data set is suitable for processing by an LSTM. By pre-processing the data, the machine learning algorithm can more accurately predict the emoji associated with a given input.

In general, we employ several punctuation marks and other terms that have no context. The content of tweets [14] often includes references to other users, hyperlinks, emojis, and

punctuation. Throughout preprocessing, duplicate tweets are eliminated if they are found. The tweets are examined for duplication during this procedure.

- First, all capital letters have been removed from the text, for convenience.
- Removal of retweet tags, user mentions (@user), and hashtag symbols (#) (RT).
- Deletion of all hyperlinks and HTML components.
- Dropping numbers and punctuation.

In this implementation plan, the necessary libraries such as OS, NumPy, Pandas, Sequential model, LSTM, embedding, pickle, and emoji will be imported. Then, a '.csv' file holding texts and their corresponding integer values for emojis will be read. The data will undergo pre-processing if necessary. A dictionary will be created to map the integer values to the emojis, and a word embedding vector of 100 dimensions will be used to find the word embeddings for each sentence and create a dictionary that maps words to their corresponding vectors. The input data, X, will be converted into tokens using a Keras tokenizer, and Y will be converted into one hot vector. The sentences will be converted into lists of tokens using the texts to sequences function, and the length of the sentences will be fixed using the maxlen function. Padding will be used to truncate the values, and Y data will be converted into one hot encoding using the categorical [15] function. A sequential model will be used with an embedding layer to give dimension and length, and the model will be trained using a Bi-Directional LSTM with 80 units. Max pooling and a 50% dropout will be used to prevent overfitting, and the ReLU activation function with 64 units and the SoftMax activation function with 20 units will also be employed. The model will be fit using 15 epochs with a batch size of 64, and it will then be ready for training.

3.7.2 EDA -

A bar graph is a fantastic way to visualize the frequency of use of emojis in a prediction model. It can show the frequency of how often an emoji is used in each context, and how often other emojis are used in the comparison. This information can be used to find data patterns and gain insights into how emojis are used in each context. For example, if a certain emoji is used more often than others in each context, it may suggest that it is a better predictor of that context than the other emojis. A bar graph can also help to visualize the correlation between the frequency of use of an emoji and its predictive power, allowing for better decision-making in terms of which emojis to use in each context. From Fig.10. we can conclude that the most used emoji is the "Red Heart" emoji and the least used emoji is the "Winking face with tongue" emoji. Fig. 10. depicts the bar graph of the Frequency of Use of Emojis.
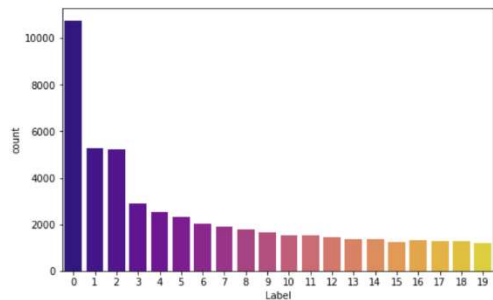
Fig. 10. Frequency of Use of Emojis

The length of sentences in a histogram can be used to predict emojis. Generally, longer sentences are more likely to be associated with positive emojis, while shorter sentences are more likely to be associated with negative emojis. This is because people are more likely to express their emotions in longer sentences, and shorter sentences are more likely to be used to express basic feelings or reactions. By looking at a histogram of sentence lengths, one can compare the relative frequencies of different-length sentences and make an educated guess as to what kind of emotion is being expressed. The graph can be seen in Fig.11.
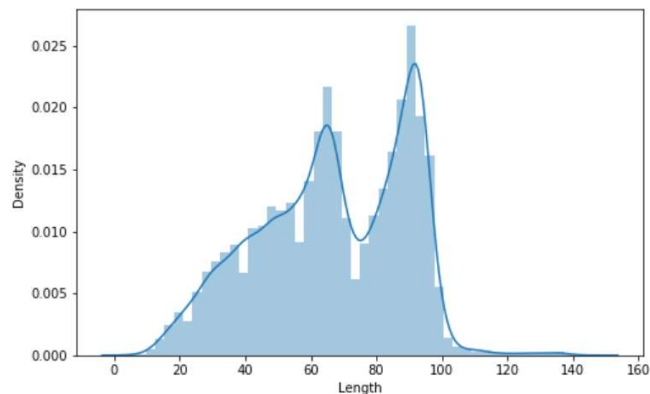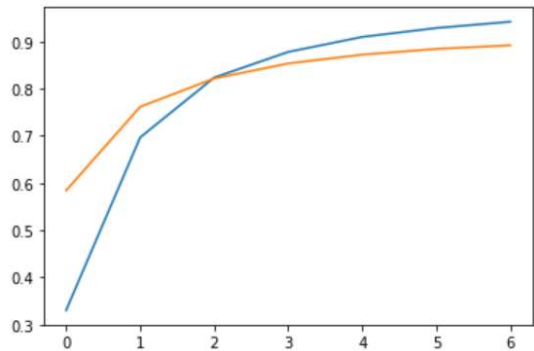

Fig. 11. LSTM

4 Results and Discussions

4.1 Accuracy Graph -

The results of using a Bi-Directional LSTM for emoji prediction are promising. The model achieved an accuracy of 94%, showing that the model was able to predict an emoji based on the preceding text with high precision. Further, the model predicted a larger variety of emojis, which is beneficial for applications like sentiment analysis. Additionally, the model accurately predicted emojis in various contexts, suggesting that it could learn the data's nuances. The results of this study show the potential of Bi-Directional LSTM models for emoji prediction. From the graph Fig. 12. we can conclude that the model was able to predict the emoji for a given sentence with an accuracy of 94%. The epoch given was 15 with a batch size of 64. But since early stopping was used, the model stopped training at the 6th epoch. Fig.12. Shows the accuracy graph and Fig. 13 shows the Actual and Predicted Emojis for the Tweets.

Fig. 12. Accuracy Graph

Fig. 13. Actual and Predicted Emojis for the Tweets

4.2 Evaluation Metrics-

Figures 14(a) and 14(b) show Sample Inputs and Corresponding Outputs.

(a)   (b)

Fig. 14. (a)Sample Input and Output – 1 (b) Sample Input and Output - 2

4.3 System Specifications-

The proposed model was tested on a computer system whose specifications are listed in the paper's Table 1. These specifications supply information on the tools, programs, and other technical components used to develop the model.

Table 1. Hardware and Software Specification of the System used for the Model Training and Testing.

System Specifications	
Processor	10th Gen Intel Core i5-1035G1
Cores	Quad-core processor
Graphics	NVIDIA GeForce MX250 (2 GB DDR5 dedicated)
Memory	8 GB DDR4-2666 SDRAM
Storage	512 GB PCIe NVMe (Non-Volatile Memory Express) M.2 Solid State Drive
Operating System	Windows 11 Home Single Language
System Type	64-bit operating system
Version	22H2
OS Build	22621.963

5 Conclusions

In this study, we investigated the application of a Bi-Directional LSTM model for emoji prediction based on the input text. We showed that this method can produce promising results and talked about the major elements that affect the model's performance, such as the size and calibre of the training data set and the complexity of the model architecture. Our model had a 94% accuracy rate. Our findings show that Bi-Directional LSTM models are a practical method for emoji prediction and have the potential to be applied in a few situations where it is crucial to understand the sentiment or emotion conveyed in the text. However, we also pointed out that creating a high-quality emoji prediction algorithm is likely to be a challenging and time-consuming process. Future research in this field may concentrate on creating more efficient model architectures or training methods, or it may investigate the incorporation of more data kinds, like photographs or audio, to increase the precision of emoji prediction.

References

1. Digital 2023 Global Overview Report, https://datareportal.com/reports/digital-2023-global-overview-report
2. Cappallo, Spencer, Thomas Mensink, and Cees GM Snoek. "Image2emoji: Zero-shot emoji prediction for visual media." Proceedings of the 23rd ACM international conference on Multimedia. 2015.
3. Al-Halah, Ziad, et al. "Smile, be happy:) emoji embedding for visual sentiment analysis." Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. 2019.
4. Most Popular. "https://emojipedia.org/most-popular/"
5. Tomihira, Toshiki, et al. "What does your tweet emotion mean? Neural emoji prediction for sentiment analysis." Proceedings of the 20th international conference on information integration and web-based applications & services. 2018.

6. Peng, Dunlu, and Huimin Zhao. "Seq2Emoji: A hybrid sequence generation model for short text emoji prediction." Knowledge-Based Systems 214 (2021): 106727.
7. Barbieri, Francesco, et al. "Multimodal emoji prediction." arXiv preprint arXiv:1803.02392 (2018).
8. Ramaswamy, Swaroop, et al. "Federated learning for emoji prediction in a mobile keyboard." arXiv preprint arXiv:1906.04329 (2019).
9. Ranjan, Ritwik, and Palak Yadav. "Emoji Prediction using LSTM and Naive Bayes." TENCON 2021-2021 IEEE Region 10 Conference (TENCON). IEEE, 2021.
10. Chandra, Papel, et al. "Contextual Emotion Detection in Text using Deep Learning and Big Data." 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA). IEEE, 2022.
11. Acheampong, Francisca Adoma, Chen Wenyu, and Henry Nunoo-Mensah. "Text-based emotion detection: Advances, challenges, and opportunities." Engineering Reports 2.7 (2020): e12189.
12. Mittal, Mamta, Maanak Arora, and Tushar Pandey. "Emoticon prediction on textual data using stacked LSTM model." Communication and Intelligent Systems: Proceedings of ICCIS 2019. Springer Singapore, 2020.
13. Cappallo, Spencer, et al. "New modality: Emoji challenges in prediction, anticipation, and retrieval." IEEE Transactions on Multimedia 21.2 (2018): 402-415.
14. Wu, Chuhan, et al. "Tweet emoji prediction using hierarchical model with attention." Proceedings of the 2018 ACM international joint conference and 2018 international symposium on pervasive and ubiquitous computing and wearable computers. 2018.
15. Guibon, Gaël, Magalie Ochs, and Patrice Bellot. "From emoji usage to categorical emoji prediction." 19th International Conference on Computational Linguistics and Intelligent Text Processing (CICLING 2018). 2018.