TWITTER SENTIMENT ANALYSIS USING LONG-SHORT TERM MEMORY (LSTM)

Table of Contents

ABSTRACT:

- The goal of tweet sentiment analysis is to find the positive, negative, or neutral sentiment part in the tweeter data.
- Sentiment analysis can help any organization to find people's opinions of their company and products.
- Twitter is a platform widely used by people to express their opinions and display sentiments on different occasions.
- Sentiment analysis is an approach to analyze data and retrieve sentiment that it embodies.

INTRODUCTION:

- Sentiment analysis gets importance in these days. Every business organization now a days are very interested in knowing their customers emotion or sentiment towards their product or services.
- There are several ways are there to do sentiment analysis.
- Lexicon based and rule-based methods gets outdated because of the arrival of the machine learning methodologies.
- Now machine learning methods in turn gets older because of neural networks.
- We promote a model based on Long-Short Term Memory (LSTM)
- LSTM is better in analyzing long sentence emotions effectively. LSTM is capable of learning long term capabilities.

OBJECTIVE:

The purpose of this study is to first look at sentiment analysis on microblogging to analyse customer feedback on the organization's products. The second is the development of product customer review programs that enable organizations or individuals to parse opinions and mass tweets into useful formats.

LITERATUE SURVEY:

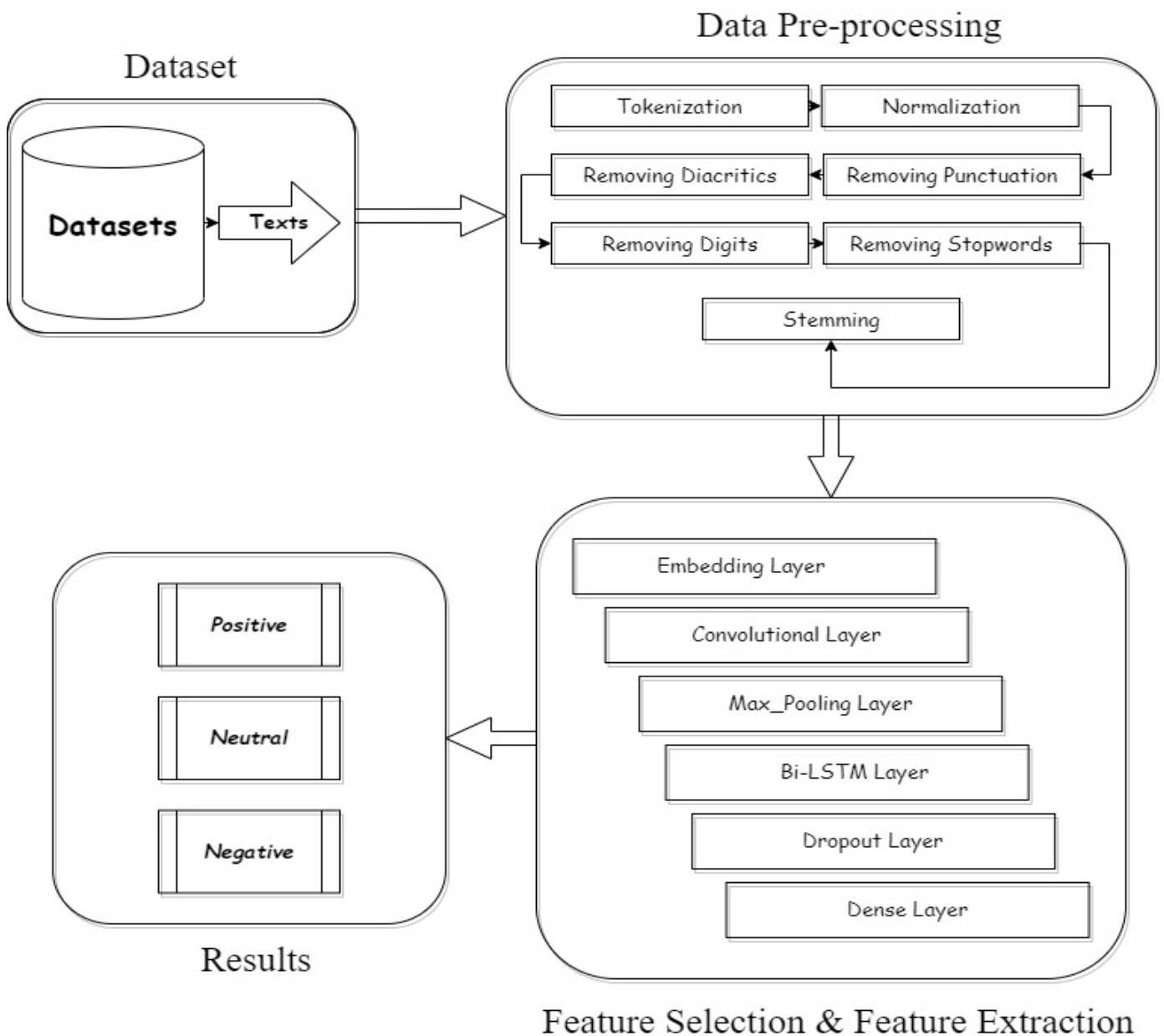| NO	YEAR	TITLE	OBJECTIVE	METHOD USED	ADVANTAGES	DISADVANTAGES
1.	2018	Twitter Sentiment Analysis via Bi-sense Emoji Embedding and Attention-based LSTM	Task of sentiment analysis and achieve the state-of-the-art performance	1.Bi-sense Emoji Embedding 2. Attention-based LSTM	1.Robust 2.Effectively represents Complex semantic, sentiment information	1.Can't tackle tasks involving multi-sense embedding
2.	2018	Advanced Combined LSTM-CNN Model for Twitter Sentiment Analysis	Effectively use lstm-mcnn and achieve the state-of-the-art performance	1. LSTM-MCNN based on LSTM-CNN	1.Makes use of encoder-decoder framework 2.Features learning in CNN are more intrinsic and effective	1.Accuracy is only 76% and need more training
3.	2019	Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data	Evaluate different deep neural network configurations	1.Multiple CNN-LSTM networks	1. Multiple CNN and LSTM networks increases the performance of the system	1. CNN and LSTM networks are combined together they perform better than when used alone
4.	2021	LSTM Based Sentiment Analysis	Model proposes sentiment classifiers that aids in the classification of emotion in text sequences	1. Long Short Term Memory	1.Brands are able to work faster, with more accuracy, toward more useful ends.	1.It gives more accuracy, but features are very few.
5.	2019	US Airlines Sentiment Analysis using LSTM	Sentiment analysis with small amount of tweets of two classes positive and negative	1. RNN language model based on Long Short Term Memory (LSTM)	1.With small amount it is very effective and gives more accuracy	1.It cant handle large data. 2.It gives only two classes
6.	2021	Sentiment Analysis of Nepali COVID19 Tweets Using NB, SVM and LSTM	Detecting sentiments in user opinions centered on a whole phrase and a specific entity	1.NB 2.SVM and 3.LSTM	1. Bernoulli's NB performs best on aspect-based classification 2. LSTM Classifier performs best on entire sentences	1.Model gives less accuracy for small data. 2.Requires more data for higher accuracy
7.	2019	Sentiment Analysis on Product Reviews Using Machine Learning Techniques	Sentimental analysis on product reviews using NB-SVM	1.Dictionary-based approach under lexicon-based 2. NB and SVM	1. NB got98.17%, SVM got 93.54% accuracy for Camera reviews.	1.We cant predict exactly people liked or disliked.
8	2013	Sentiment Analysis Tool using Machine Learning Algorithms	Model proposes sentiment classifiers of emotion in twitter data	1. Naive Bayes, maximum entropy classification	1.High accuracy for classifying sentiment when using this method	1. Twitter messages have unique characteristics, accuracy is less.
9.	2017	Study of Twitter Sentiment Analysis using Machine Learning Algorithms on Python	Analyzing the sentiments of the tweets and feeding the data to a machine learning model	1. Machine learning model-data collection, pre-processing, sentiment detection, classification, training and testing	1.Works efficiently and models reaching the efficiency of almost 85%-90%	1. lacks the dimension of diversity in the data 2. lot of application issues with the slang used and the short forms of words
10.	2014	Twitter Sentiment Analysis	Design of a sentiment analysis, classify customers' perspective via tweets into positive and negative	1.ML model based on web application system.	1.Machine-based learning approach which is more accurate 2.Natural language processing techniques used.	1. Due to limitation of Django which can only work on Linux server or LAMP. Thus, it cannot be realized.

PROBLEM STATEMENT:

Sentiment analysis of a web-based application focusing only on individual tweets. With the rapid growth of the World Wide Web, people are turning to social media such as Twitter, which generate large amounts of opinion text in the form of tweets that can be used for sentiment analysis.

So, we are analysing the needs of companies and needy members, we are giving a solution of providing a model which provides the Sentiment based on the text given to it.
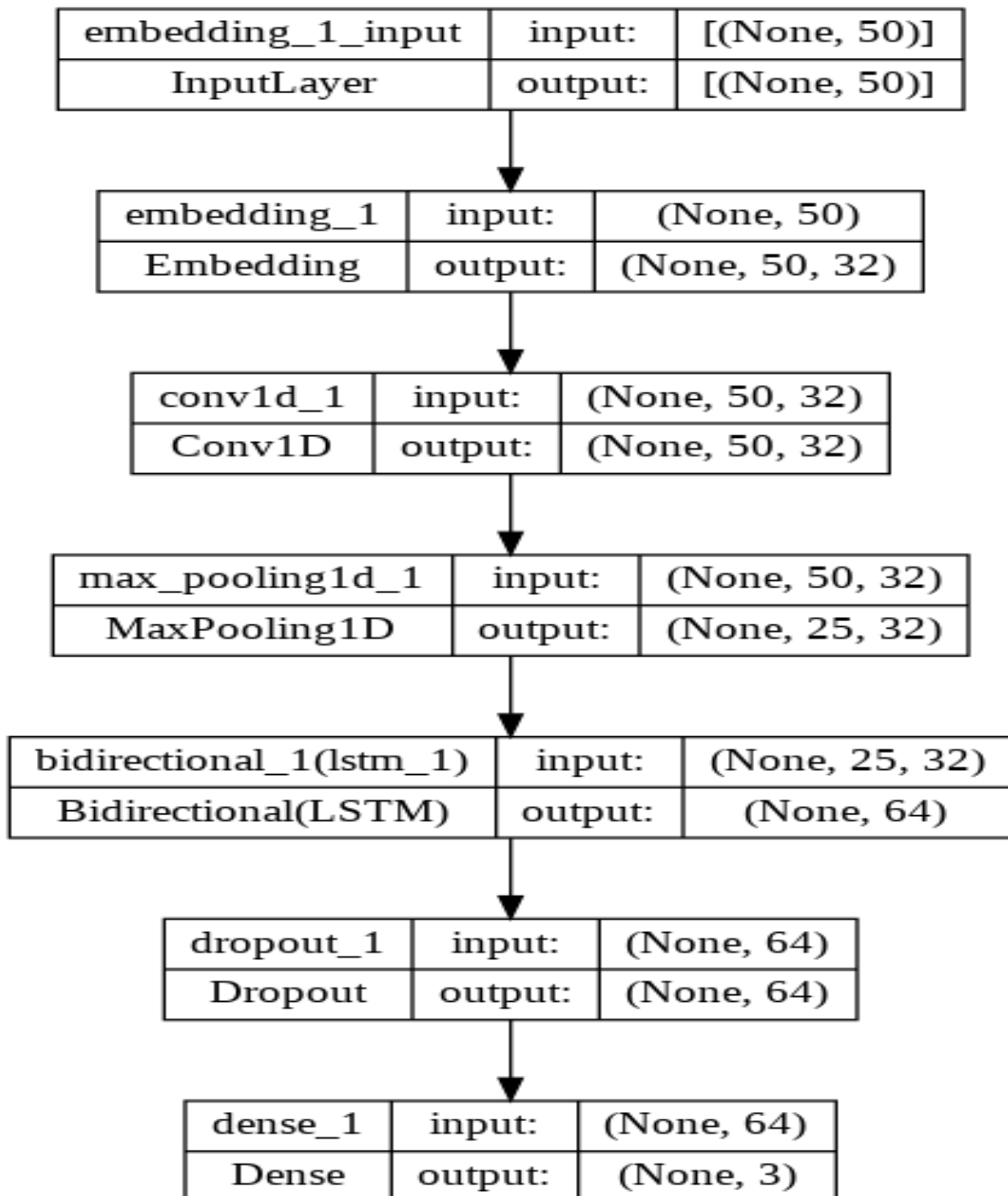
The short form is also widely used in Short Message Service (SMS). Twitter uses abbreviations more commonly to minimize the characters used. This is because Twitter limits the number of characters to 140.

SYSTEM ARCHITECTURE:

Data Pre-processing

Dataset

Datasets → Texts

Tokenization → Normalization

Removing Diacritics ← Removing Punctuation

Removing Digits → Removing Stopwords

Stemming

Embedding Layer

Convolutional Layer

Max_Pooling Layer

Bi-LSTM Layer

Dropout Layer

Dense Layer

Positive

Neutral

Negative

Results

Feature Selection & Feature Extraction

PROPOSED METHODOLOGY:

- Our Methodology is to first pre-process the tweets in the data using the NLP methods. Initially with Tokenization, Normalization and by removing unwanted texts in the tweets.
- Second, we did feature selection and feature extraction with the layering of CNN-LSTM which is Bidirectional LSTM.
- Embedding, Convolutional, Max_Pooling, Bi-LSTM, Dropout and Dense are the layers used.
- Finally, we will predict that whether the text is Positive, Neutral or Negative.

embedding_1_input	input:	[(None, 50)]
InputLayer	output:	[(None, 50)]

embedding_1	input:	(None, 50)
Embedding	output:	(None, 50, 32)

conv1d_1	input:	(None, 50, 32)
Conv1D	output:	(None, 50, 32)

max_pooling1d_1	input:	(None, 50, 32)
MaxPooling1D	output:	(None, 25, 32)

bidirectional_1(lstm_1)	input:	(None, 25, 32)
Bidirectional(LSTM)	output:	(None, 64)

dropout_1	input:	(None, 64)
Dropout	output:	(None, 64)

dense_1	input:	(None, 64)
Dense	output:	(None, 3)

RESULTS:

We have got,
- Accuracy of 91.41%
- Precision of 91.71%
- Recall of 91.05%
- F1 Score of 91.38%
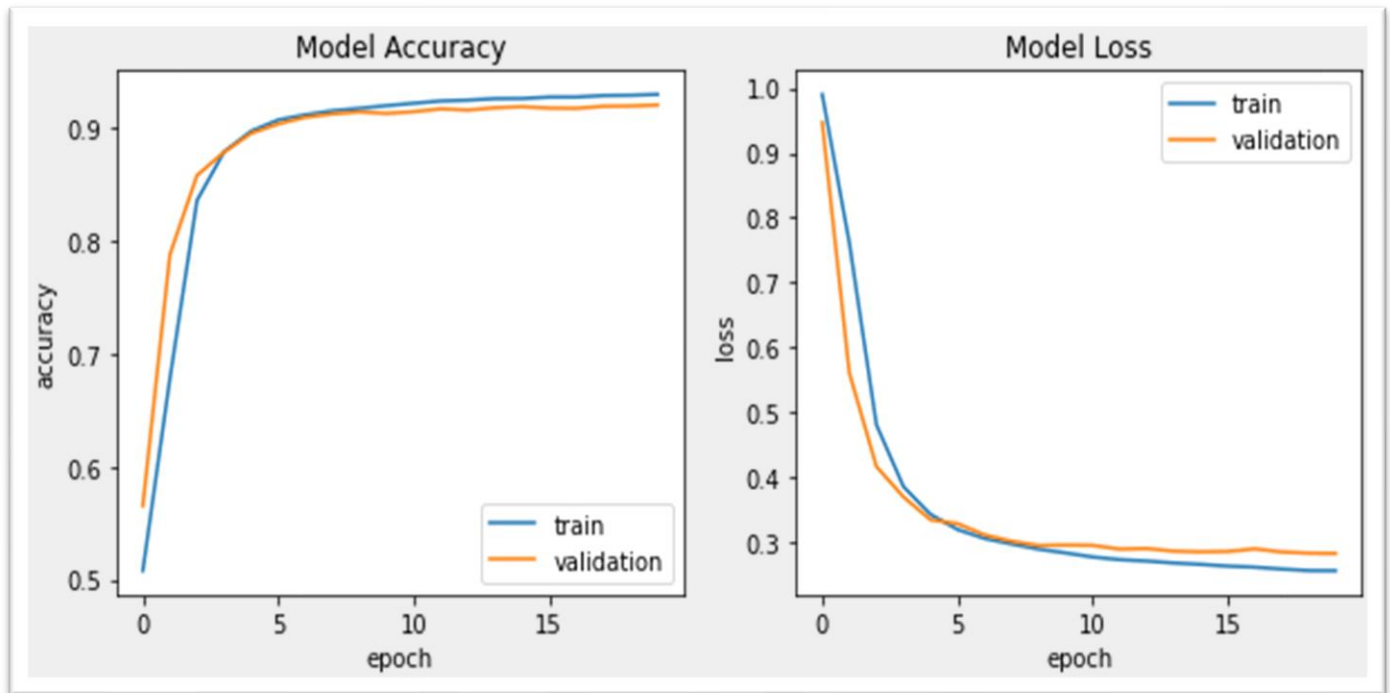
CONCLUSION:

- In Conclusion, I would like to say that our model has performed well in Training and Testing.
- The prediction of the sentence is also done and it gave the accurate results for the given sentence.
- Since our model is performing in a good manner, we can provide our model for the company purposes and it can be used by them for the purpose of sentiment analysis of the customers.

FUTURE SCOPE:

The Future Scope of our project will be, by analysing the previous results of tweets we are initializing the results as positive, neutral or negative. Now, we can scope it for customer reviews and any other reviews that can be scarped or in-built datasets by increasing the accuracy of the Project.