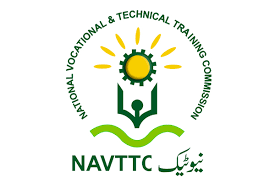
**Twitter Sentiments Analysis by using Deep Neural Network and LSTM.**



**Proposal for Project**

**Al-Khawarizmi Institute of Computer Science**

**University of Engineering & Technology Lahore**

**Muhammad Umar Tariq**

**E-mail:Umartariq7487@gmail.com**

Contents

[1. Introduction 3](#_Toc64461855)

[2. Literature Review 3](#_Toc64461856)

[5. Aim and Objectives 3](#_Toc64461857)

[7. References 3](#_Toc64461858)

# Introduction

Twitter sentiments analysis is a textual mining techniques of twitter that are used for extracting the important information that are help for understanding the social sentiments about his specific brand and also helpful for product quality control or business improvements. It help for taking the opinions person to person about his specific product. Every brand monitoring and product reputation are also depend on the customer feedback. It also helpful for getting the customer positive and negative response and natural emotions about his specific subjects.

Basically, twitter is an American social media network in which people put his post, retweet and also reads the different posts. Twitter with his 319 million monthly active user are now most valuable for individual and organization also make the strong relationship between politician and increase the social and economic interest than enhance and maintain it. Twitter is a most popular platform that are express our comment opinion and continuous issues that are accord in all resources.

Now a days, people use the twitter for extraction huge information about social media, blogs and networking. In twitter peoples are express his opinion and thought about the various events and issues. In a business the hidden information are used for taking opinion and decision.

In twitter the people are generates the short tweet because the tweet size is reduced to 140 characters. The techniques of opinion and sentiments mining are also helpful for making user opinion and sentiments analysis about the specific subjects.

The sentiments analysis provide the service of organization to real time surveying the different social media sides. Text sentiments analyses is an automatic techniques that are used for determining the text segment contain opinionated and subjective content and also determine the sentiments polarity.The main goal of the sentiments classification is to automatically determine the polarity of sentiments is positive or negative.

Most of the peoples in old days for twitter sentiments classification are used the methods that are proposed by pang et al. and also apply the algorithms of machine learning for making classifier of tweets that are manually annotated the label of polarity sentiments. In a recently few years, the interest of using deep learning techniques is increase for enhancing the twitter sentiments classification accuracy increase.

In this articles, our main objective is to classifying the twitter sentiment in different categories. For sentimental classification we use the recurrent neural network (RNN) is a type of neural network of deep learning.

The Recurrent neural network (RNN) is theoretically maintain the long term dependences but practically fall it addresses by Bengio et al. information is needed at long period that are make calculations of distance between nodes. Multiple products of Jacobian matrix the distance calculation is done. This is also prompt the normal vanishing gradient and less continuous exploding gradient.

To overcome this specific issue the Hochreiter et a1. Are present the Long Short Term Memory Network that are also called the LSTMs. It provide the unbelievable performance at NLP and image processing applications.

LSTM is a type of RNN that are equipped for learning long term conditions. Basically, RNN is a combinations of repeating neural network modules. LSTM is a specialized RNN when it contain four network layers that are uniquely interacted. The equation1 is show the outcome of sigmoid function is:

𝑡 = 𝜎 (𝑊𝑓. [ℎ𝑡−1, 𝑥𝑡] + 𝑏𝑓) (1)

In equation2 using the tanh functions that are represent the outcomes of cell state is:

𝑡 = 𝑡𝑎𝑛ℎ (𝑊𝐶. [ℎ𝑡−1, 𝑥𝑡] + 𝑏𝐶) (2)

In each training and testing steps these parameters are updated. The make Schuster et al. are portrayed the bidirectional RNN that are also utilizes the distinctive hidden layers, feed forward to find a similar output layer. The LSTMs model has improve the performance of speech recognition, entity and synthesis recognition. For sentiments analysis deep learning base techniques are used.

In this research work we are used the LSTMs model and base approaches of neural network for tweeter sentiments analysis. For experimental analysis we are used the twitter dataset.

# Literature Review

* Related work

This section deals with the related methods used for sentiment analysis. We have mainly concentrated on deep learning-based approach. Patel et. al. [4], have described a model using Recursive RNN to classify the movie reviews. The proposed system does sentiment analysis based on the review text.

Jan Deriu et. al. [6] have developed a Twitter sentiment classification using deep learning. In the proposed model they have used a Two-layer convolutional neural network. The main task is divided into 3 subtasks.

Anand et.al. [5] have posed a system to do the same task with CNN. In this approach, sentiment analysis is done using Natural Language Processing (NLP) as most text present on Twitter is in the form of emotions or opinions. The proposed model gives a comprehensive view of the emotions in terms of positive, negative, or neutral. Fenna Miedema et al. [7] have proposed a system that used LSTM networks to develop a model that has to turn out to be most efficient among other approaches. Zainuddinet al. [8] have posed a sentiment classification approach using Support Vector Machines. This model lack accuracy in the case of long sentences obtained from a variety of resources. Mathieu Cliché [9] has implemented a CNN and LSTM based approach that is trained on SemEval2017 twitter dataset. This approach uses a large set of unlabelled dataset and pre-trained word embedding. This hybrid approach has shown better improvement in the accuracy level of the classification. The proposed approach included 5 steps – Reading CSV file (Twitter data), Preprocessing, Feature extraction, and classification. In the paper, the tasks are accomplished in two ways in the first approach. ML- based approach is applied to the dataset whereas in the second approach deep neural network-based methods are used. The mode which is incorporating the LSTM approach shows improved performance as compared to other approaches. In the previous few years, the deep learning approaches have demonstrated critical improvement in the performance of NLP mechanisms. [10, 11-14].

1. **Research gap**

**Short coming in the literature work**

In the past the convolution neural network models were facing gradient disappearance problems which were causing no convergence. This gap was covered in recent researches when residual networks were proposed.

**Thing that will you improving**

1. **Problem statement**

Imagine yourself as a young graduate student in Stanford’s Artificial Intelligence Lab, building a system to sentiment analysis. we have used LSTM based approach to analyze the Twitter post. Our deep learning model is built using Keras Python libraries.

# Aim and Objectives

Our main objective is to classify sentiments into different categories. Recurrent Neural Network (RNN) theoretically claimed to maintain the long-term dependencies but practically fall as addressed by Bengio et al. Information correlation over a long period needed distance calculation between nodes. The distance calculation is done based on multiple products of Jacobian Matrix. This prompts all the more normally vanishing gradient and less continuous exploding gradients. To overcome this issue Hochreiter et al. presented Long Short-Term Memory networks typically called LSTMs. It has shown an unbelievable performance in NLP and Image Processing applications.

1. **Methodolog**y

Twitter sentiment dataset is collected from a Kaggle repository and analyzed for negative and positive sentiments. Figure 3 shows the negative sentiment in the red color bar whereas positive sentiments using a green color bar. Dataset is read using the panda library. The data preprocessing includes a set of operations to be performed on the dataset to extract sentiments. It cleans the tweets by removing stop words, spacing, or any special symbols. The step also lemmatizes and stems words. The processed data is further used for the model training and testing. The read dataset after cleaning is classified into training (80%) and test dataset (20%). The Train\_and\_Test() function is used to split the dataset. The trained model is tested further to check model efficiency

**Work flow of system**

First we can use randome forest to predict the results of our text data and second we use ocr for pitcher data and also predict in live automation.

# References

[1] Hochreiter, Sepp & Schmidhuber and Jürgen Long Short-term Memory 1997 Neural Computation 9 1735-80

[2] Bengio Y, Simard P and Frasconi P 1994 Learning long-term dependencies with gradient descent is difficult IEEE Transactions on Neural Networks 5 2 157- 66

[3] Mike Schuster and Kuldip K 1997 Bidirectional recurrent neural networks IEEE Transactions on Signal Processing 45 11 2673–2681

[4] Patel P, Patel D and Naik C 2021 Sentiment Analysis on Movie Review Using Deep Learning RNN Method

[5] Satapathy S, Zhang Y D, Bhateja V, Majhi R 2021 Intelligent Data Engineering and Analytics J Advances in Intelligent Systems and Computing 2 1177

[6] Anand M, Sachin S 2017 Proc. 5 th Int. Conf. on Mining Intelligence and Knowledge Exploration: MIKE 2017 (Hyderabad, India) Sentiment Analysis of Tweets in Malayalam using Long Short-Term Memory Units and Convolutional Neural Nets

[7] Deriu, Jan & Cieliebak, Mark 2016 Sentiment Analysis using Convolutional Neural Networks with Multi-Task Training and Distant Supervision on Italian Tweets

[8] Fenna Miedema and Sandjai Bhulai 2018 Sentiment Analysis with Long Short-Term Memory network Vrije Universiteit Amsterdam

[9] Zainuddin, Nurulhuda & Selamat, Ali 2014 Proc. Of 1st Int. Conf. on Computer, Communications, and Control Technology Sentiment analysis using Support Vector Machine pp 333-37