**UNIVERSITY OF MISSOURI, KANSAS CITY**

**PYTHON & DEEP LEARNING – CS 5590**

**PART – 2 - DEEP LEARNING**

**LAB – 3 REPORT**

**BHAVYA TEJA GURIJALA**

**STUDENT ID: 16220446**

**CLASS ID: 14**

**TASK:**

Implement the Text Classification with three different models CNN, RNN and LSTM models with new dataset using in python. And also, to compare all the three models and should figure out which is the best model among the three.

1. **INTRODUCTION:**

The dataset chosen for performing the three different models is the IMDB movie review dataset where we would be performing the classification of the text based on the sentiment of the review whether it belongs to positive or negative. We will be applying three different models CNN, RNN and LSTM on this dataset and we calculate the test accuracy of the model. The model which gets the most is said to be the best model of all.

RNNs are specifically unfolded to a finite sequence of hidden units, and are dependent on both hidden unit and the input. LSTMs specialize the RNNs by augmenting the units with the LSTM. LSTMs are generally used for the time series which defines the past and the present data. Unlike RNNs and LSTMs, CNNs excel in the situations where the data is sampled based on the dimensions. The output parameters are confined to convolutions instead of per unit basis.

1. **OBJECTIVE:**

The main objective of the task is to perform the text classification on the same dataset by three different models and should decide which is the best model of the three. We divide the dataset into test and train samples and perform the modeling on them. After training the model, we test the accuracy of the model. And based on the test accuracies we can decide that which model is the best fitted model for the text classification.

1. **APPROACH/METHOD:**

The approach for deciding which performs better for the text classification is to apply the model to the same IMDB review dataset. If we use the same dataset we can easily identify the model which is the best among the three. Mapping of every movie review into a real vector domain which is called as the word embedding. This approach is followed here while building the model. These real vectors are encoded in higher dimensions and based on the similarity of the words they are arranged closer to each other.

1. **WORKFLOW:**

The workflow of the task is as follows.