Assignment 2: Sentiment classification

Course Name: Natural Language Processing
Course No: CS60075
Instructor: Prof Sudeshna Sarkar
TA's: Alapan Kuila, Aniruddha Roy, Gautam Sagar, Sai Chaitanya
Assignment Date: 31-08-2022
Due-Date: 14-09-2022 11:59 PM
Platform: Google collab/Kaggle/Machine

Task Definition: Sentiment analysis is the task of classifying the polarity of a given text. For instance, a text-based tweet can be categorized as either "positive," "negative," or "neutral". Given the text and accompanying labels, a model can be trained to predict the correct sentiment. Sentiment analysis techniques can be categorized into machine learning, lexicon-based, and even hybrid methods. Some subcategories of research in sentiment analysis include multimodal sentiment analysis, aspect-based sentiment analysis, fine-grained opinion analysis, and language-specific sentiment analysis.

Sentiment Analysis using LSTM

1. Implement a sentiment analysis using the single layer Bi-LSTM model. You are free to decide the hidden dimension and all other hyper-parameters in the network. Also choose proper error function for this task.
2. You can use FastText/Glove/word2vec and pre-trained embedding to initialize the model.

Evaluation Metrics: F1 score (For evaluating all the models)
Dataset: https://www.kaggle.com/code/lakshmi25npathi/sentiment-analysis-of-imdb-movie-reviews/data
Split the dataset into train(80%), dev(10%) and test(10%) and perform the experiment.
Submission Materials: Python file, Google drive link for the trained model, a doc-file with results, and your observation.