Md Tahmid Yasar, Hawk ID: A20436459
Khalid Saifullah, Hawk ID: A20423546

Project Title: <u>**Sentence Classification using Densely Connected Bidirectional LSTM**</u>

Main Paper:
Densely Connected Bidirectional LSTM with Applications to Sentence Classification [1]
Zixiang Ding, Rui Xia, Jianfei Yu, Xiang Li and Jian Yang
CCF International Conference, 2018

Problem Statement:
Speech and sentences are sequential data. To classify sentences accurately, sequence or order of the word must be taken into account. Recurrent neural network (RNN) is developed to address the problem of classifying sequential data.

Most deep architectures like stacked RNN tend to suffer from the problem of overfitting and vanishing gradients. Because of these, most conventional deep architectures are not easy to train and hence their performance is not so great. As a result, their applications are understudied in natural language processing (NLP) tasks.

Approach:
In this project, we shall build a sentence classifier using densely connected bi directional LSTM which addresses the above-mentioned issues. Our first step is to preprocess the sentence data so that it can be given as input to deep learning architecture. We shall remove the punctuations and convert the words into one-hot encoded vectors using pretainded dictionary. In our model, we will be using the publicly available 300 dimensional Glove vectors [2] which were trained on 42 billion words.

Deep-stacked RNNs [3] are created by stacking multiple RNN hidden layers on top of each other, The output sequence of one layer forming the input sequence for the next. Densely Connected Bidirectional Long Short Term Memory (DC-Bi-LSTM) has been proposed where the input sequence of each layer is the concatenation of the output of the previous layer and inputs of all the previous layers, which alleviates the issues of vanishing-gradient.

We will evaluate the DC-Bi-LSTM on five benchmark datasets mentioned in the paper. Moreover, we will also implement the model on another dataset to verify the performance. We will also evaluate the performance of Bi-LSTM as baseline. We will build the deep-stack Bi-LSTM model (DS-Bi-LSTM) and compare the performance with the proposed model which is not mentioned in the paper.
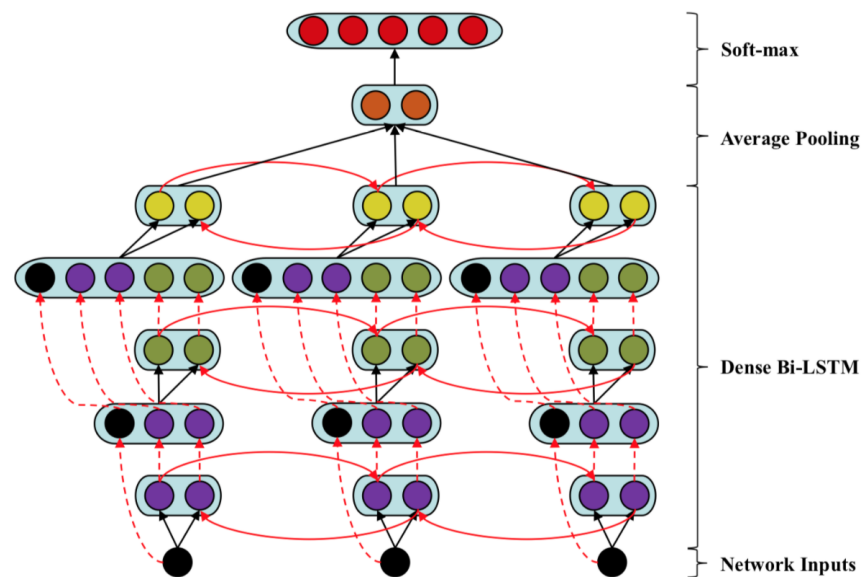
Figure: Architecture of DC-Bi-LSTM

Data:
The model will be evaluated on five benchmark datasets of sentence classification. The last dataset will be an addition to five presented in the paper.
i) MR: Movie Review Data
ii) SST-1: Stanford Sentiment Treebank
iii) SST-2: Same as ii) but used in binary mode without neutral sentences
iv) Subj: Subjectivity Dataset
v) TREC: Question Type Classification Task
vi) CR: Customer reviews of various products

References:
[1] https://arxiv.org/pdf/1802.00889.pdf
[2] Glove Vectors
[3] https://arxiv.org/pdf/1303.5778.pdf

Team Member Responsibilities:
1) Md Tahmid Yasar
   ● Data preparation and preprocessing
   ● Build RNN model (Bi-LSTM, DS-Bi-LSTM, DC-Bi-LSTM)
   ● Report and presentation

2) Khalid Saifullah
   ● Build RNN model (Bi-LSTM, DS-Bi-LSTM, DC-Bi-LSTM)
   ● Performance evaluation and comparison
   ● Report and presentation