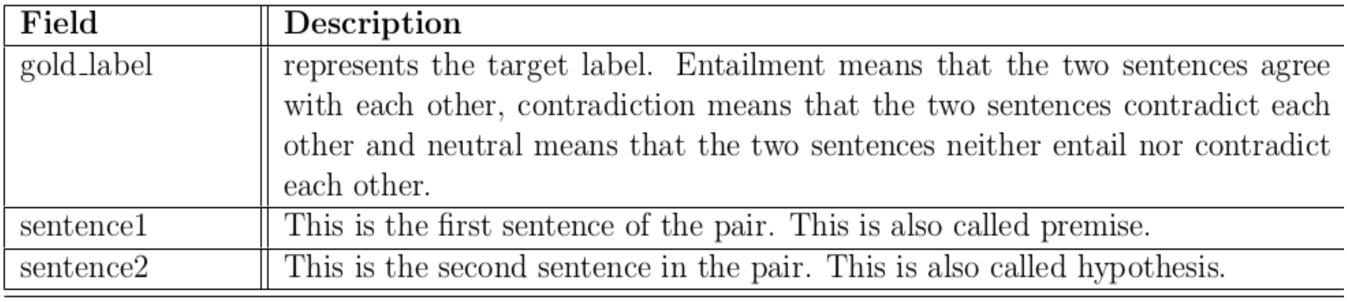
Deep Learning Assignment-3
Shubham Sharma (16013)

Task: In this task we worked on the task of natural language inference in which we would be given two sentences called premise and hypothesis and we have to determine whether hypothesis id true(entailment), false(contradiction) or undetermined(neutral), given that ‘premise’ is true.

Dataset: We are using the Standard Natural language Inference Dataset (snli1.0) dataset for our task.

Sample :



Field	**Description**
gold_label	represents the target label. Entailment means that the two sentences agree with each other, contradiction means that the two sentences contradict each other and neutral means that the two sentences neither entail nor contradict each other.
sentence1	This is the first sentence of the pair. This is also called premise.
sentence2	This is the second sentence in the pair. This is also called hypothesis.

We are using the json versions of the data file present in the snli1.0.zip for our task which contains:

550152	training points
10000	validation points
10000	test points

Preprocessing of Data:

For the task of preprocessing of data we are using the nalutal language toolkit(nltk) library and steps involved in the preprocessing are as follows.

1. Removing of stop words.

2. Conversion to lower case.

3. Stemming.

4. Tokenization.

5. Conversion of labels.

Approaches:

1. Logistic Regression : In this approach we are training a logistic regression model which will classifies the the sentences in one of the three possible outcomes. In this method we are using TF-IDF features using scikit in python.

In this approach of logistic regression for tf-idf we are using n-gram range of (1,5) and 50000 features for our trianing.

Accuracy in Logistic Regression task = 59.36

2. RNN : In this approach we are using deep neural libraries. Particulary we are using RNN which is much better when we have to solve long term dependencies. As we are deciding the class of based on a long sentence so RNN will help us to look into the long term dependencies present in the sentences.

Hyper parameter tuning : In these task we trained many models models tuned upon number of features and layers. We use 100,200 and a500 for feature selection and 1 and for number of layers.

Loss Function : Cross Entropy

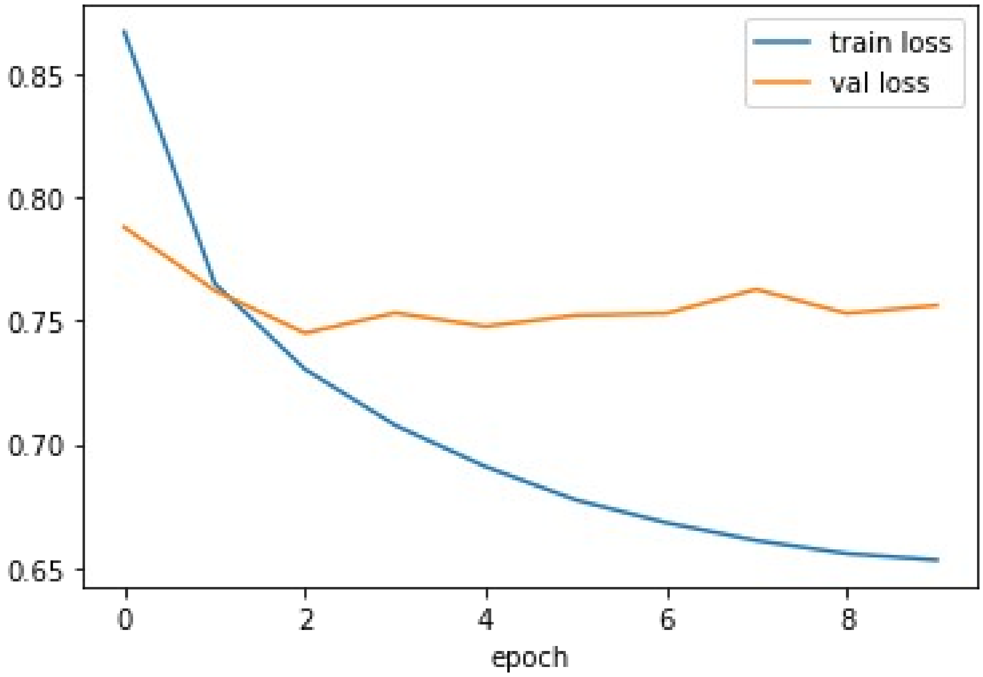
Optimizer : Adam

Sr. No.	Features	Hidden Layers	Accuracy(%)
1.	100	1	66.68
2.	200	1	67.24
3.	500	1	67.55
4.	100	2	66.87
5.	200	2	68.13
6.	500	2	68.66

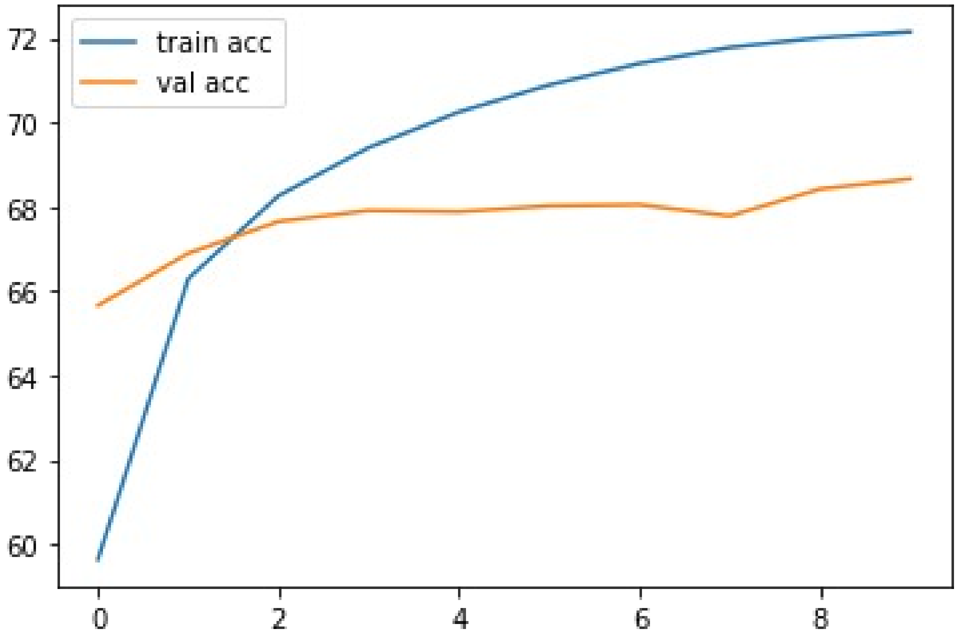
For final reference task we used 500 features and 2 hidden layers. We also tried architecture of LSTM in our task and it slightly same performing as RNN. We didn’t try concept of attention which might be able to increase its accuracy.

Loss:

Below graph (for: 500,2) shows the training loss and validation loss and there are epochs on the x-axis. We trained and tested our models for upto 10 epochs.



Graph(for 500,2) for the training accuracy and validation accuracy, similar epochs on x- axis.



Inference/Output :

1. Output for logistic regression classification is saved in tfidf.txt.

2. Ouput for RNN based approached is saved in deep_model.txt.

RUN: To run the inference task execute main.py