**EXTRACREDIT REPORT**

I have applied bert base uncase model on olid data set and fine tuned it to make predictions on test dataset. Here, also I have preprocessed the tweets using clear tweets function in my code and then tokenized it with bert tokenizer, then have trained the training set and did fine tuning using following parameters and evaluated the model on testing data and calculated the probabilities..

Parameters used for fine tuning using torch:

**num\_train\_epochs=1,              # total number of training epochs**

**per\_device\_train\_batch\_size=64,  # batch size per device during training**

**per\_device\_eval\_batch\_size=64,   # batch size for evaluation**

**warmup\_steps=100,                # number of warmup steps for learning rate scheduler**

**weight\_decay=0.01,               # strength of weight decay**

**logging\_dir='./logs',            # directory for storing logs**

**logging\_steps=40,**

Your tasks for the EC are:

1. Using transformers and pytorch, fine-tune a bert-based-uncased model on the subtask A of the OLID dataset (OFF or NOT).

done

1. Produce predictions for the test portion of OLID.

File attached in the zip folder for predictions

1. Calculate accuracy.

Accuracy on test data:

{'accuracy': 0.8383720930232558, 'f1': 0.6759906759906759}

1. In a **SEPARATE** report, report the results and the previous “simpler” models’ results from previous homeworks.

**All results have been on test data :**

**With bert base uncase : 83 % accuracy and f1 score : .67**

**Logistic Regression : 80 %**

**Naïve Baye’s : 74 %**

**MLP with pretrained glove embedding : 0 accuracy on test data**

**MLP with three hidden layer with fined tuned embeddings : 67%**

1. Compare and contrast the results for the models and provide explicit discussion.

I have preprocessed the tweets then created bag of words using count vectorizer and I have passed it to logistic regression for training the same vocab is used for testing data as well. I am getting accuracy of **80 percentage** in case of logistic regression and for naïve baye’s I have used Bernoulli method which is giving an accuracy of **74 percentage** on test data.

However using bert transformer, I am getting the good accuracy of 83%.

However, MLP had given zero accuracy on test data.

With increasing layers accuracy has been resulted **75 percentage**  using three hidden layer on train test split on entire data but when tested on actual test file given I am getting zero accuracy perhaps model has gone overfitted. I have used pretrained model glove for embeddings.

However multilayer perceptron with three layers with word2vec fine tuned embeddings has resulted in average accuracy of **75 % on train data and 67 % on test data**.

6. Hand in your code, report, and prediction files. (The prediction files may be in the same format that previous homeworks had.)

Code, report and prediction file all attached in zip