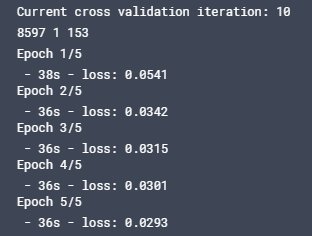
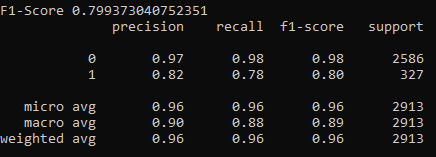
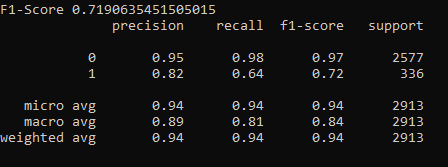
Training using CuDNNLSTM was slow for some reason

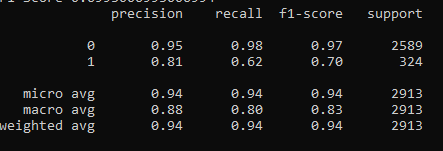


**GloVe MLP Neural Network Classifier**

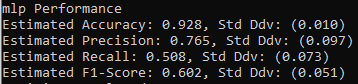
Comments filtered, w/o cross validation



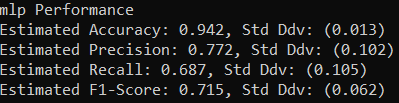




With Cross Validation (5 epochs)

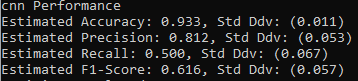


With Cross Validation (30 Epochs)

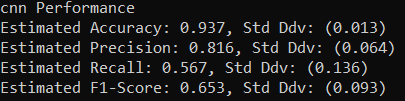


**GloVe CNN Classifier**

With Cross Validation (5 epochs)



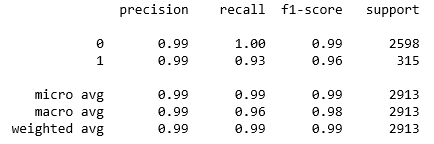
With Cross Validation (30 epochs)



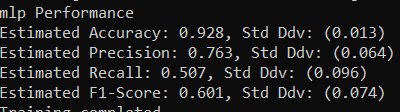
**Word2Vec (skipgram) MLP Neural Network Classifier**

518 features, including bag of words w/ stop words (variables and functions) removed:

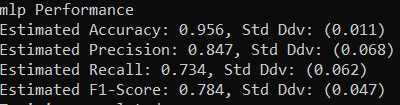
Accuracy: 0.9909028492962582

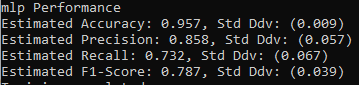


150 word2vec features, comments filtered

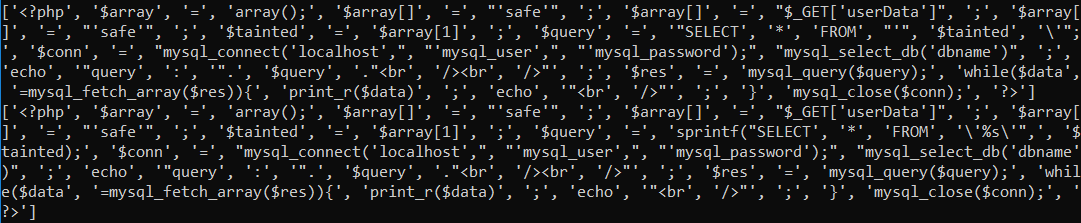


Word2vec (30 epochs), comments filtered, getCorpus function uses text.split() instead of gensim.utils.simple\_preprocess(text)



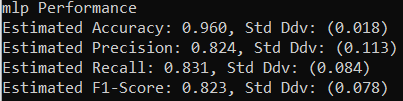


This may have heavy overfitting though. This is an example of what the getCorpus function outputs under the wordlist



Word2Vec (30 epochs), comments filtered, getCorpus function uses PHPParser(text).tokens() instead of gensim.utils.simple\_preprocess(text)

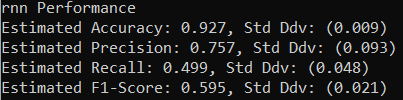
Before editing StopWordsList to include variables with $



After editing StopWordsList to include variables with $

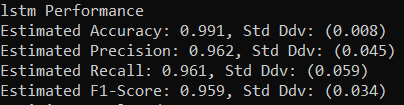
**Word2Vec (skipgram) RNN Classifier**

150 word2vec features, comments filtered

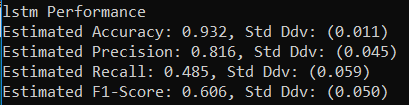


**Word2Vec (skipgram) LSTM Classifier**

150 word2vec features

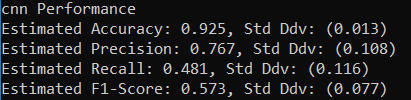


150 word2vec features, comments filtered



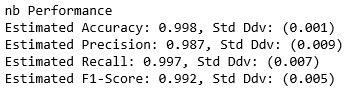
**Word2Vec (skipgram) CNN Classifier**

150 word2vec features, comments filtered

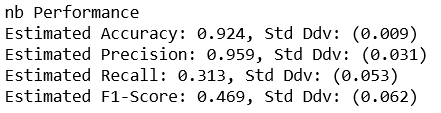


**Gaussian Naïve Bayes Classifier After Cross Validation**

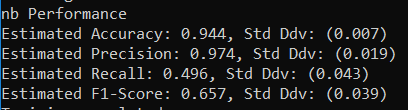
518 features, including bag of words w/ stop words (variables and functions) removed:



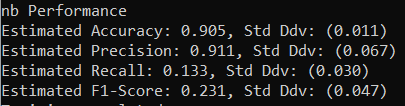
~500 features, did feature selection, including bag of words w/ stop words (variables and functions) removed:



~500 features, did feature selection, including bag of words w/ stop words (variables) removed. Filtered comments:

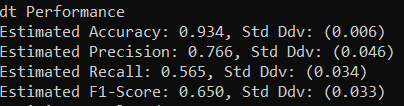


139 features, including bag of words (features only) w/ stop words (variables) removed:



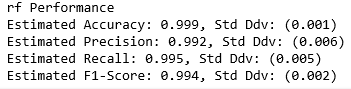
**Decision Tree Classifier After Cross Validation**

~500 features, did feature selection, including bag of words w/ stop words (variables) removed. Filtered comments:

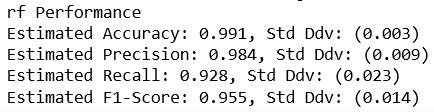


**Random Forest Classifier After Cross Validation**

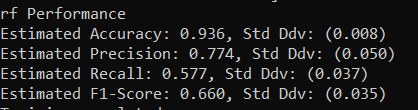
518 features, including bag of words w/ stop words (variables and functions) removed:



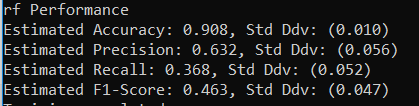
~500 features, did feature selection, including bag of words w/ stop words (variables and functions) removed:



~500 features, did feature selection, including bag of words w/ stop words (variables) removed. Filtered comments:

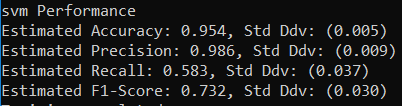


139 features, including bag of words (features only) w/ stop words (variables) removed:



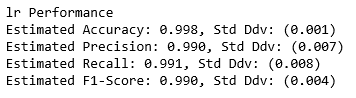
**SVM Classifier After Cross Validation**

~500 features, did feature selection, including bag of words w/ stop words (variables) removed. Filtered comments:

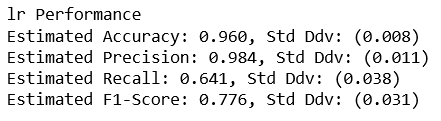


**Logistic Regression Classifier After Cross Validation**

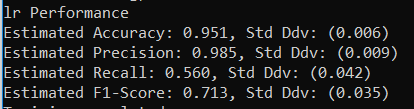
518 features, including bag of words w/ stop words (variables and functions) removed:



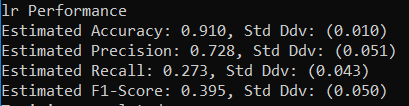
~500 features, did feature selection, including bag of words w/ stop words (variables and functions) removed:



~500 features, did feature selection, including bag of words w/ stop words (variables) removed. Filtered comments:

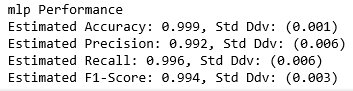


139 features, including bag of words (features only) w/ stop words (variables) removed:

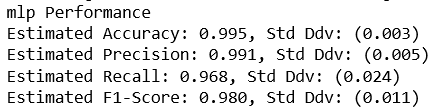


**MLP Classifier After Cross Validation**

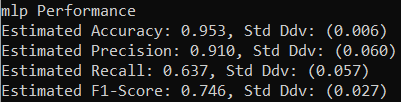
518 features, including bag of words w/ stop words (variables and functions) removed:



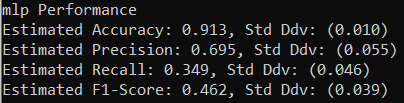
~500 features, did feature selection, including bag of words w/ stop words (variables and functions) removed:



~500 features, did feature selection, including bag of words w/ stop words (variables) removed. Filtered comments:

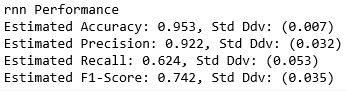


139 features, including bag of words (features only) w/ stop words (variables) removed:



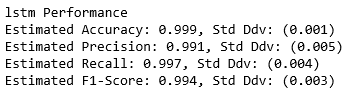
**RNN Classifier After Cross Validation**

~500 features, did feature selection, including bag of words w/ stop words (variables) removed. Filtered comments:

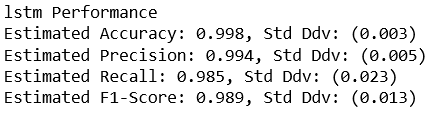


**LSTM Classifier After Cross Validation**

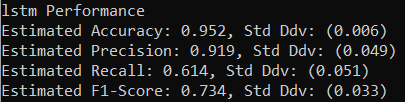
518 features, including bag of words w/ stop words (variables and functions) removed:



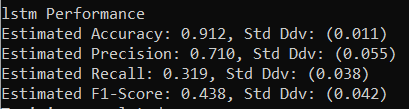
~500 features, did feature selection, including bag of words w/ stop words (variables and functions) removed:



~500 features, did feature selection, including bag of words w/ stop words (variables) removed. Filtered comments:

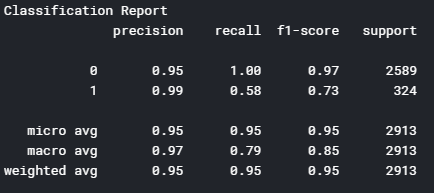


139 features, including bag of words (features only) w/ stop words (variables) removed:



**CNN (No Cross Validation Yet)**

518 features, bag of words, comments filtered



**CNN After Cross Validation**

518 features, bag of words, comments filtered

