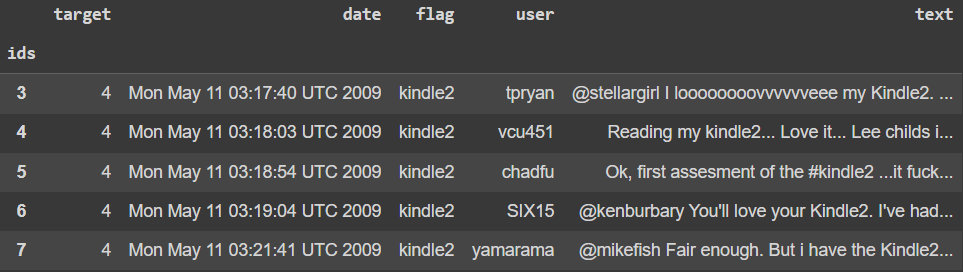
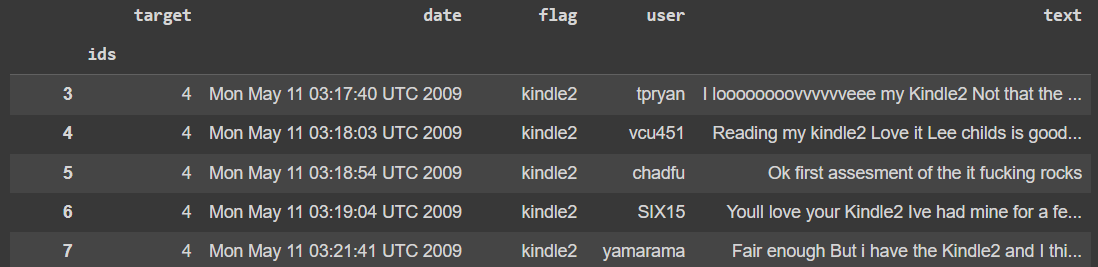
# Explore & Prepare the Dataset

In this section we first applied the lambda function on the text column for removing the links, hash tags and username. Then we use the translator to remove the punctuation.

Before:

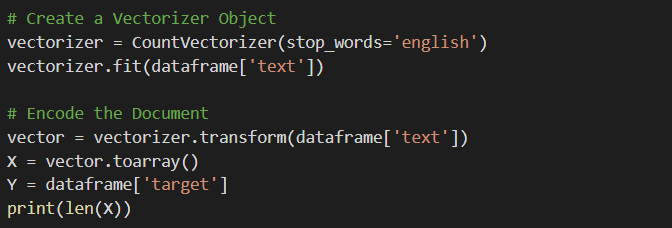


After:



# Build the BOW

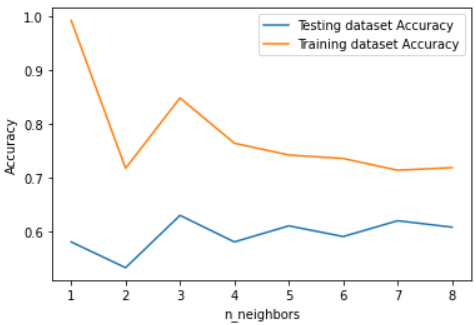
In this section we first change the stop words and then added the vectorize words to build the bags of words for our model.



# Build Models (KNN, DT, SVM)

In this section we trained the KNN modal with different Nearest Neighbors. We choose the neighbors with the best accuracy. In this case we choose the 7 neighbors, because it is giving us best accuracy.

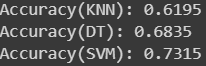
We will again train our model on 7 neighbors for our model evaluation.

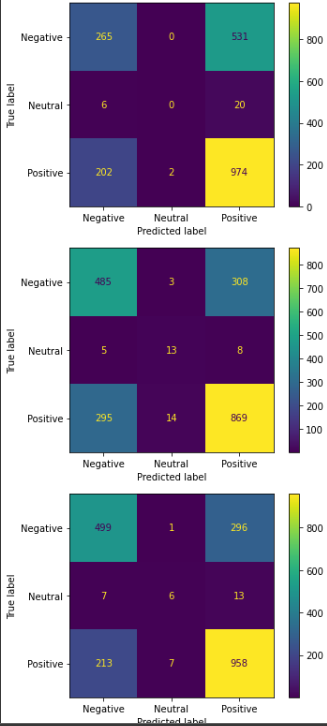
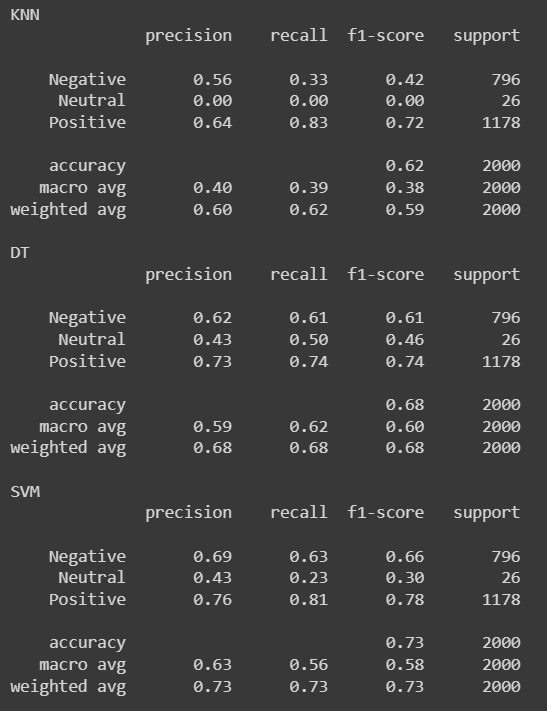


Then we train the Decision Tree (DT) and Support Vector Machine (SVM) on our dataset.

# Model Evaluation

This is section we evaluated our model performance by giving the output of the predicted values.





# Build CNN using word embedding

In this module we first used our dataset we replace the values of target variables from 2 to 1 and 4 to 2. Because we are limited to the output values from 0-2.

1. Means Negative
2. Means Neutral
3. Means Positive

Then we used the bag of words representation for our dataset. After that we split our dataset for training and testing purposes.

For the output we used the SoftMax to get the probabilities with three output categories.

Then we fit our model for training and validation. We used the 10 epochs for our model.