1. Names of all group members and at the top of your report

Names: Elliot Selveira, Cayden Blake, Taylor Farnsworth

1. A brief description of the problem

Our project is focused on predicting the emotion used in different texts. For instance, we are using different learning models that will receive a tweet and can correctly identify if the emotion used in it will be happiness, sadness, ect.

1. How and from where is the data being gathered

The data that we have gathered has been collected from Kaggle.com which has 40000 rows of data with a tweet from twitter and the emotion expressed in the text. We downloaded the file and uploaded it to GitHub where we are able to access it at any time.

1. A description of the data set including:

* Actual example instances, including a reasonable representation (continuous, nominal, etc.) and values for each feature
* How many instances and features you plan to have in your final data set

The dataset has 40000 rows of information to train and test on. We have decided to break down the number of emotions to identify into 9 different groups: sadness, fun, neutral, worry, surprise, love, anger, happiness, and relief. We have broken down the tweets into categories of the number of each type of punctuation, Uppercase letters to lowercase letters ratio, text length, as well as a bag of words with different key words. Below is an example of a few rows of our data.

1. What machine learning models are you initially trying to learn with

We decided that we will try the machine learning models of Decision Trees, Back Propogation, and k-Nearest Neighbors as our initial approaches to the problem. We will see the accuracy of each of these and see if a different model would be better suited for the problem at hand.

1. Brief discussion of plans and schedule to finish the project