**CS 196 - Sentiment Analysis Research (Prem and Aditya)**

**Research:**

* We will need to use Jupyter Notebook to code this so we can dissect the code in multiple parts

**Inputs and Outputs:**

* Input(01 - neutral) -> collection of movies ranging from sad, funny, and happy
* Input(02 - calm) ->
* Input(03 - happy) ->
* Input(04 - sad) ->
* Input(05 - angry) ->
* Input(06 - fearful) ->
* Input(07 - disgust) ->
* Input(08 - surprised) ->

**Training Dataset:**

* The dataset we will use is from this link: <https://zenodo.org/record/1188976#.YWjf5ZNucq1>
* We will need to implement supervised learning because we are given data with a known output of what emotion.
* We are given 1404 files of data which should be enough to get an accuracy of over 90%.
* Measure of success will be accuracy of how the computer perceives the emotion of an input.
* We will also need a validation set to fine tune the results of the prediction of the album.
* I do not think we will need feature scaling as all the data comes from the same dataset and moreover, the data is all controlled.
* The features we will most likely use is emotion and emotional intensity.