

Team Members:

Zoher Ghadyali
Kiki Chandra

Data Set:

We are going to be working with the Rotten Tomatoes dataset.

Project Learning Goals:

- Zoher
 - In Software Design, we used someone else's tool for sentiment analysis and I didn't really know how it worked but it worked well for us. Gaining any sort of insight into sentiment analysis and maybe even potentially improving the analysis done on the text for the Oscar Predictor would be an awesome thing to learn how to do in this project.
 - This project poses a really cool visualization problem. I don't really know how to visualize text-based data which you don't normally show temporally or spatially so it would be awesome to explore some visualizations to see what is best for text.
 - Since this project has us implement a sentiment analysis algorithm, I will be gaining insight into how sentiment analysis works and how I can play with my own on the Oscar tweets dataset. For this project, our final deliverable will be an iPython notebook(s) that walks through our project and what we accomplished and we plan to include relevant visualizations. We will have to determine what kind of visualization will work for us and what won't to be clear and concise.
- Kiki
 - I would like to delve more into natural language processing, because I am taking a class on this topic at Wellesley and it's an area of computer science that is very interesting. My focus would definitely be on figuring out different ways to improve predictions, best practices, and understanding how to develop algorithms for sentiment analysis in more depth.
 - The purpose of evaluating this dataset is to classify sentiment displayed in Rotten Tomatoes reviews. In order to do this, we will have to learn more about how sentiment analysis algorithms work, and how they can be trained to improve accuracy. Once we do that, we also explore how to create better visualizations that can concisely demonstrate information, which is always a useful skill to have.