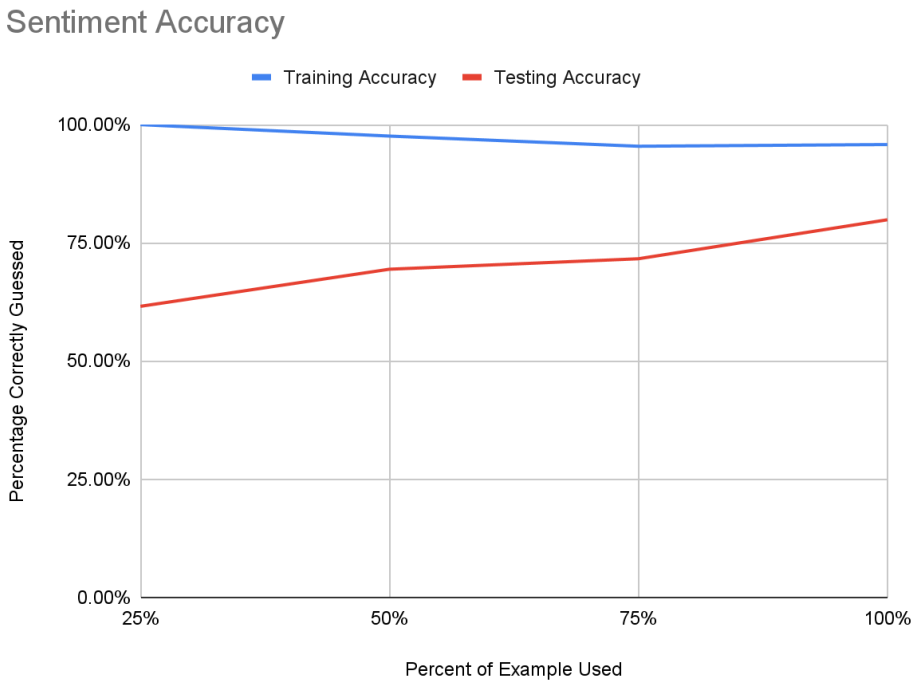
James Smith
Joshua Tumlinson

Sentiment Analysis -- CS 331H



The main difficulties that we encountered were trying to visualize the math in the form of code as well as debugging. It took me a few minutes to tens of minutes to look at the equations and realize how I should implement them with the data that we had. This wasn't necessarily hard, but took some time and a few tries to get right. A second difficulty encountered was debugging. For the longest time, we had testing accuracies of >98%, which we knew couldn't be right. After much debugging which included finding small errors and some larger mathematical errors, we realized that we were including the labels in the vectors. When the program sees that 100% of positive reviews contain the label, it's clear to see why it would have such a high probability of success. After removing the labels, the accuracy came down to a much more understandable percentage. Seeing it slowly climb over time as we add more examples makes sense. The high training data success and its eventual decline also makes sense. With so few examples, it makes sense that it is most likely matching the reviews to their training reviews. As we add more examples, it becomes harder to match these reviews as there are a greater number of reviews using those words.