

## Report

When we perform the bootstrapping, we use  $b=1000$ . We chose this value instead of  $b=10000$  because we did not have the time or patience to wait for the program to run with this larger value. From our results, it seems that  $b=1000$  is a sufficiently large number to support robust statistical analysis.

From our experiments, we observe that one minus p-value is positively correlated to the effect size. This makes sense. As the effect size increases, the difference between the two systems is less likely to be caused by random noise. Therefore, the p-value would decrease, and 1 minus the p-value would increase. Our plots of 1 - p-value vs effect size demonstrate that the relationship between these two variables follows a logistic trend. This was not something that we initially anticipated would happen. There is probably a mathematical reason for this that we would be interested in learning more about.

Both of our plots display this relationship between the values. Our plot acknowledges when systemA and systemB use the same feature representation vs when they use a different feature representation does not display any meaningful differences from the basic plot. This indicates that there is no statistically significant difference between the binary and the count feature representations. It is interesting to see how the mere presence of words (binary) can be just as meaningful and powerful of a feature representation as the counts of the words.

## Reflection

What was easy about this assignment?

- T1-T3 were the easiest parts of this assignment, as most of this code was derived from our previous project. Additionally, T7 and T9 were relatively straightforward as well.

What was challenging about this assignment, or parts of this assignment that you couldn't get working correctly?

- T8 was the most challenging aspect of this assignment as it took the most time to code, as well as a good chunk of time to run the code as well.

What did you like about this assignment?

- It was interesting to put everything together that we've learned over this semester into one final project.

What did you dislike about this assignment?

- The timing of this assignment was a bit tough, as we had slightly less time than previous assignments, as well as many other computer science courses have assignments due at the same time.

How did your team function? Include details regarding what each team member contributed, how the team communicated with each other, and how team software development & design is as accomplished.

- Our team functioned relatively well under the circumstances. We created a group chat to communicate for the assignment. Luke started us off with T1-T3, as well as T7 and T9.

Together, Luke, Humza, and Chris worked on T4-T5, and Geethika worked on T6 and T8. All group members worked T10.

What did you learn from this assignment?

- This assignment gave our group hands-on experience with Naive Bayes classification and Bootstrapping, two foundational processes of Natural language processing.