Recommendations for Relax

After cleaning and wrangling the data and processing through a model, which happened to have a relatively high accuracy, I was able to determine of the data given, what was most important in determining if a user would be adopted.

I realize that by far, the amount of time was on the platform is the factor that determines whether or not the user is adopted. The organization seems to have an effect, as well as who they are invited by. Whoever the email provider seems to be a close 4th in terms of feature importance, but I am wondering if this is due to the popularity of certain email providers, and can be explained away by imbalance in the data set rather than how important this feature actually is in determining the adoption of the user.

The score that I got on the test set is actually worrying high, and I’m not sure if it is because it somehow got overfit due to the data, or that the small number of features allowed for a more accurate prediction.

For further research, I would want to actually implement more data. There were only about 7 features that we were able to test on, and I assume that the point of this is to eventually get more users adopted, but since we are limited in the information that we are given, it is hard to pinpoint what exactly could be done other than playing with the game of numbers and get many people on the site and hope that a percentage of them stay adopted. With better features, we could be able to do marketing or work on recruiting out of a certain market to increase the amount of adopted users, but with the current dataset, it is hard to say.