This was a very strange challenge because the features in the data set were not very comprehensive at all. There were really only a handful of variables to use in order to predict. We decided to use a Random Forest classifier to predict whether or not a particular user was a part of the adopted users, which I am not sure what that means in relation to the organization. However, the strongest feature that predicted whether or not a user was a part of this class was the organizational id. All other variables performed poorly, but the organization ID singly handedly provided an accuracy score of nearly 85%. I don’t believe that the problem is coherently defined, nor do I feel like I have enough background information to go about clarifying the problem or adding additional information. Our model clearly shows that given organizations utilize the platform more, and therefore organizations should be target and sought to implement this application(?). From there, we can the predict the users of that given organization are more likely to be in this adopted class.

The other approach we were working on towards was counting the number of times in a given week this user accesses the application on three separate days. From this perspective, we could turn the problem into a regression problem, however, a handful of dummy variables are not enough to be able to truly predict anything. I would greatly appreciate any feedback because I am still confused as to what this problem is actually looking for.