

Instructions: Take the two datasets, and try to find what variables help predict if a user will be an adopted user ( a user that opens the app 3 times within a 7 day period)

I was able to locate the adopted users by grouping by a rolling window of 3 time stamps per user. Then I took only the values that were 7 days or less to leave us with only the adopted users. Since we were limited on time (1-2 hours) I wanted to be realistic with my approach. I created dummy variables for our categorical variables and split it into a 70, 30 split. I ran a simple logistic regression and then did a gridsearch to find the best parameters. After running the gridsearch I found the logistic regression coefficients to find what our model weighs most heavily. I found that if someone made a personal project it was a strong indicator that they would not be an adopted user. Users who signed up through google authentication or guest invite showed a strong indication that the user could become an adopted user. Other strong indicators were if the users were invited by another or signed up for the marketing drip.

