

## Relax Inc. Take Home Challenge Report

The dataset I worked with consisted of two csv files regarding user engagement data for 8,823 users. We sought to find which features of the datasets were most predictive of user adoption, defined as *having logged into the product on three separate days in at least one seven-day period*. The relevant features used were as follows: opted\_in\_to\_mailing\_list; enabled\_for\_marketing\_drip; creation\_source (encoded into 5 columns); act\_history. act\_history was a feature engineered column consisting of the span between the creation date and the last usage date. After dropping users who had only signed up in the previous 2 weeks, and thus had little chance to be considered an adopted user, we ended up with 1,653 adopted users out of 8,429 users. We used a random forest classifier and the accuracy on the test set was 0.9482. The confusion matrix and AUC and ROC curve can be seen in Figure 1.

Figure 1 shows the confusion matrix and AUC and ROC Curve of the results.

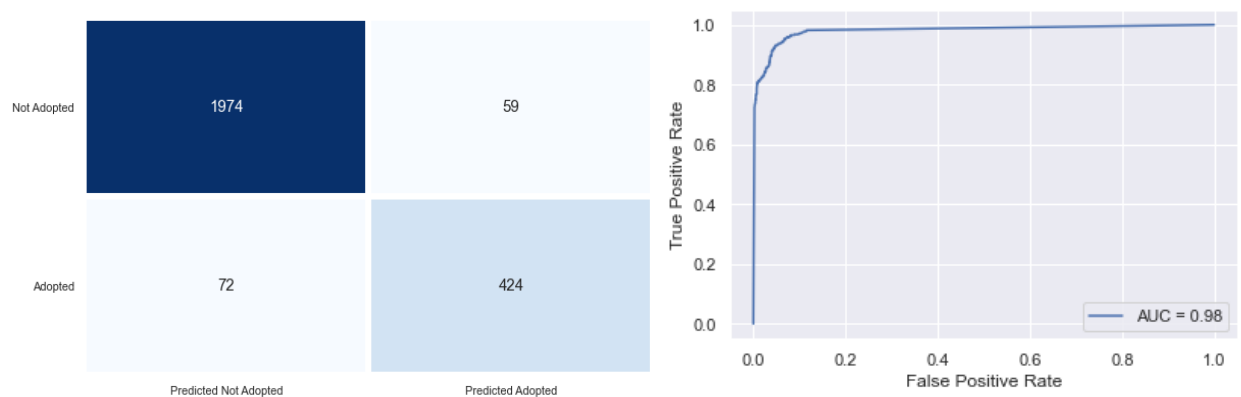
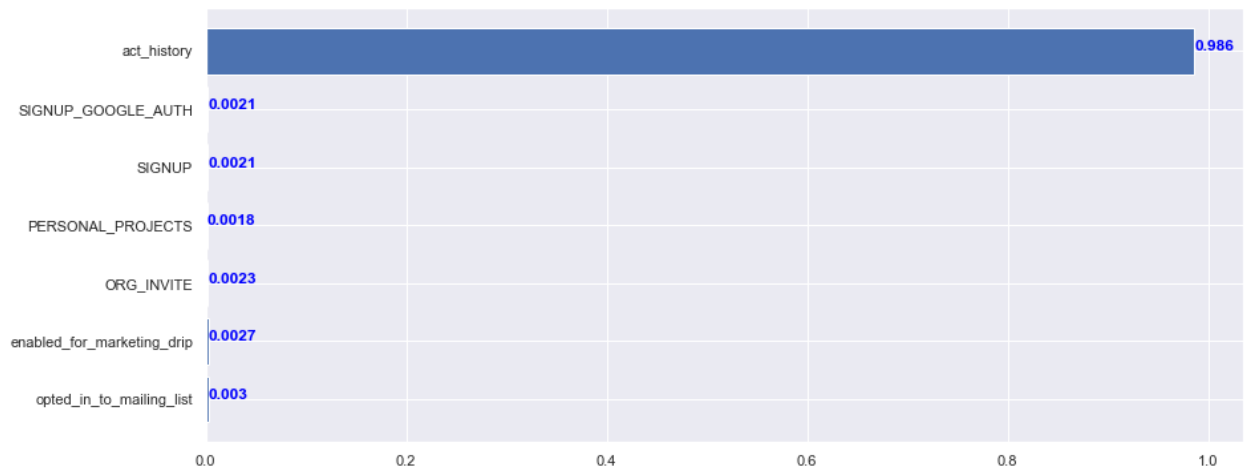
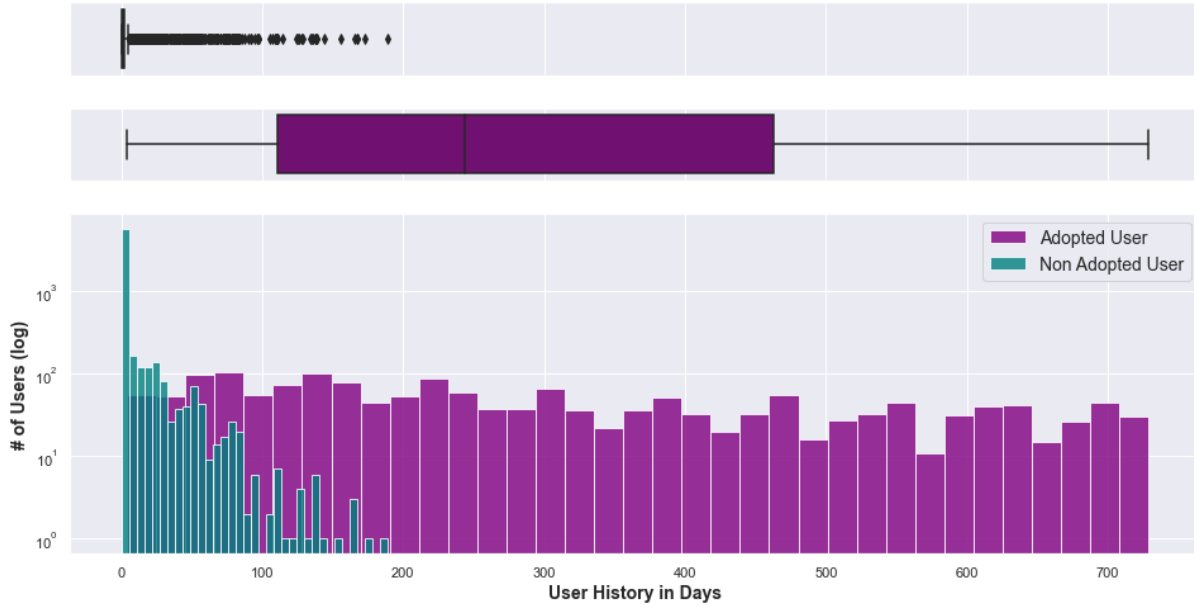


Figure 2. Feature Importances



By far the most important feature in predicting adopted status was act\_history.

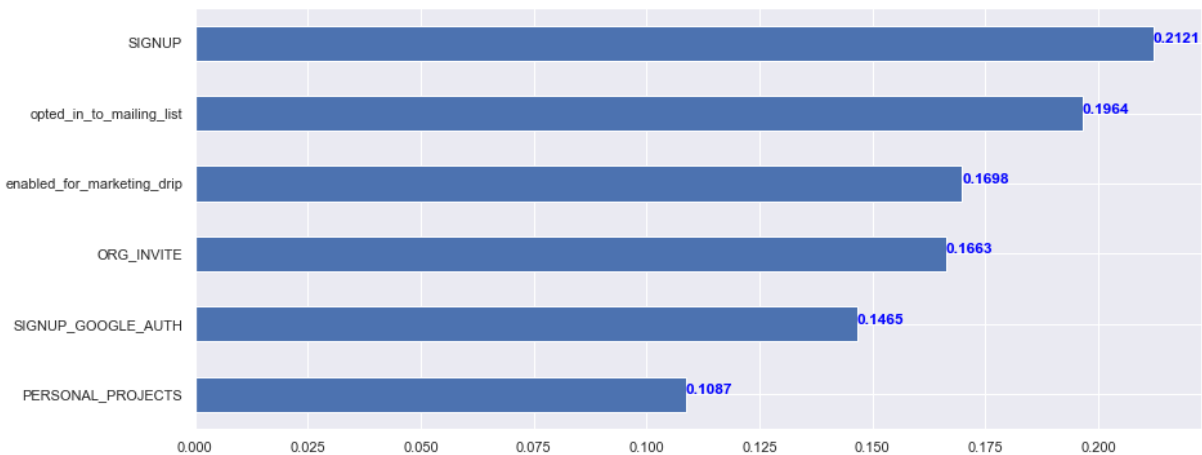
Figure 3. Distribution of users based on account history



In Figure 3 we can see a drop off after ~100 days and a no adopted users at all after 200 days. Thus, our takeaway is that one users stop logging in after the first 100 days or so, we can expect a low chance of becoming an adopted user.

Although account history was by far the biggest predictor, it may be informative to consider which features were most important without the added act\_history feature.

Figure 4. Feature importances without act\_history



We can see in Figure 4 that signup (of the creation\_source feature) was most influential on adopted status, followed by opted\_in\_to\_mailing\_list and enabled\_for\_marketing\_drip.

Based on my findings, I would recommend finding ways to really encourage new users (i.e. those with less than ~100 days from startup) to keep logging in and keep active. Signing up was the second most influential feature, but it has a negative correlation. Opted into the mailing list and were enabled for marketing drip were also predictive. So, for those who signed up, those who did not opt into the mailing list and those who are not enabled for marketing drip we should find ways to keep them more engaged in the first 100 days.