Relax Technologies Take Home Writeup

For the completed Jupyter Notebook with step-by-step, click [here]

Executive Summary

A Random Forest Classifier model was the final model considered with an accuracy rating of 86.6% since this is an appropriate model for a binary classification problem.

In terms of features the Relax can actually influence (listed below), we found statistically significant but not particularly substantial predictive power features to make predictions on:

- Did the user create a least one session? (*session_created)*
- Did the user enable marketing drip (*enabled_for_marketing_drip)*
- Did the user opt into the mailing list (*opted_in_to_mailing_list)*
- Was the user invited by another user (*invited)*

However, we did find that certain organizations (identified by organizational ID) do have higher rates of adopted users which identifies one possible avenue for Relax to consider. Relax may benefit from focusing on organization with higher adoption rates regarded targeted marketing. This may allow them to better utilize their marketing budget for higher returns on generating adopted users.

Notes of Concern & Recommendations

Two main concerns arose while conducting this analysis that would be potentially useful to consider. First, the predictive model generally tried to stay to features that Relax could actually act upon, make logical sense to consider, or were dense enough to warrant. Therefore, features like *creation_time* or *invited_by_user_id* were either not used or modified to be better suited for modeling. This did not leave many features to work with and may be a consideration for Relax in that they don't have many "levers" as it currently stands according to the data to try and influence whether a user becomes an adopted user. As such, Relax may want to consider what further methods they can employ to make more users adopted users.

Second, the low adopted user rate out of the total users provided (only 13.8% of all users are adopted users based on the metric[1] provided by Relax) does suggest Relax may benefit from a "deeper dive" into its current adopted users. Currently, there are not many adopted users to really explore to begin with from a raw data standpoint. This may mean that it would be beneficial for Relax to do some more qualitative analyses (e.g., focus groups, customer interviews, etc.) to garner more information from these uses to understand more about what may make them likely to become adopted users, which in turn may provide useful leads regarding identifying new features that can be tracked (or generated) to help develop more predictive power for models.

---

[1] Defining an "adopted user" as a user who has logged into the product on three separate days in at least one seven day period

Appendix – Tables and Figures

Table 1: Account Creation Pathway Versus % of Adopted Users

Account Creation Source Versus Adopted User Percentage			
	Total Users	Adopted Users	% of Adopted Users
Guest Invite	2,163	295	13.64%
Organizational Invite	4,254	592	13.92%
Personal Project	2,111	299	14.16%
Signup (General)	2,087	284	13.61%
Signup with Google	1,385	186	13.43%

We can see that how the user creates their account does not seem to have much predictive power regarding whether a user will eventually become an adopted user or note

Table 2: Opt Into Mailing List Versus % of Adopted Users

Opt Into Mailing List Versus Adopted User Percentage			
	Total Users	Adopted Users	% of Adopted Users
No	9,006	1,237	13.74%
Yes	2,994	419	13.99%

We can see here that there doesn’t appear to be an easily discernable difference between whether a user opts into the mailing list and whether they ultimately become an adopted user

Table 3: Enabled For Marketing Drip Versus % of Adopted Users

Enabled For Marketing Drip Versus Adopted User Percentage			
	Total Users	Adopted Users	% of Adopted Users
No	10,208	1,403	13.74%
Yes	1,792	253	14.12%

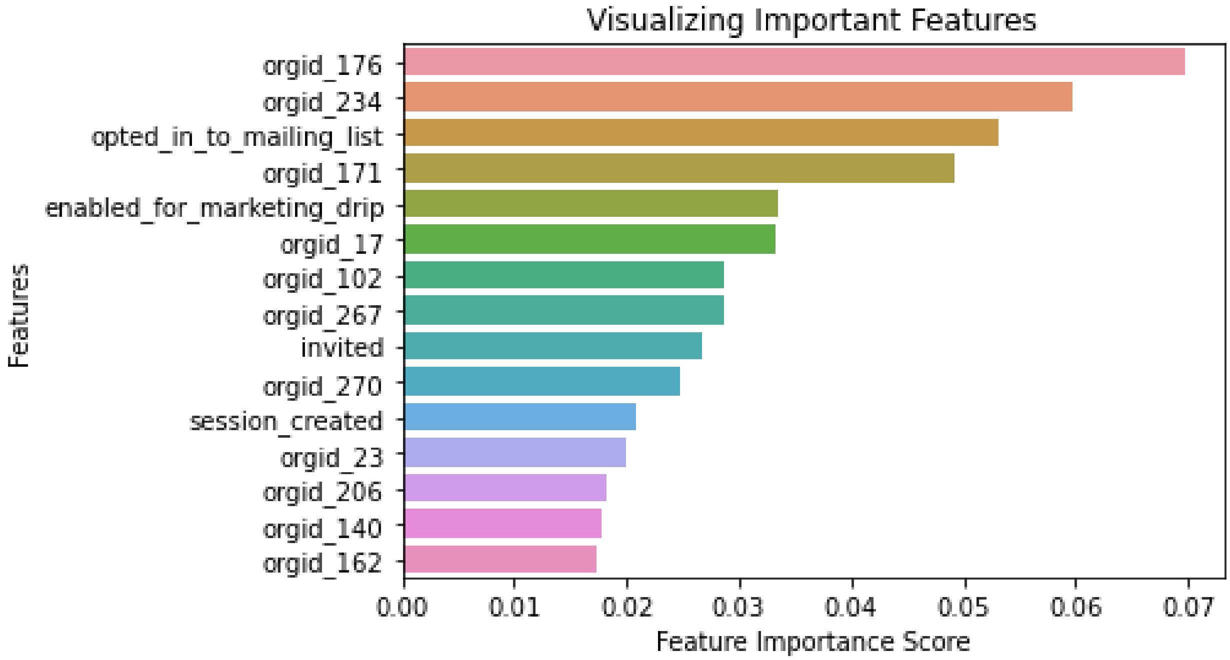
Once again we can see not to strong of an indicator regarding a user becoming adopted but it is at least slightly more predictive than some of the other features.

Table 4: Invited By Another User Versus % of Adopted Users

Invited By Another User Versus Adopted User Percentage			
	Total Users	Adopted Users	% of Adopted Users
No	5,583	769	13.77%
Yes	6,417	887	13.82%

We can see hardly any difference between someone being invited by another user and whether they become an adopted user

Figure 1: Feature Importance From Random Classifier Model



This figure shares the importance of the top 15 features from the Random Classifier Model. You can see that the four features we highlighted are on the figure. However, one important observation is that many of the organizational IDs make it into the top 15 (some above the other features) which may mean that it can make sense for Relax to initial target those organization regarding getting higher adopted user rates. For example, *Organization 176* has the strongest predictive power of all the other features so it may make sense for Relax to initially explore who that organization has a higher adopted user rate than most other organizations.