Relax Inc Take Home Challenge Final Report

Data Overview

The dataset contained a total of 12,000 users, of which only 8823 were active (had logged in after creating an account) and only 1656 are 'adopted users.'

Feature Engineering

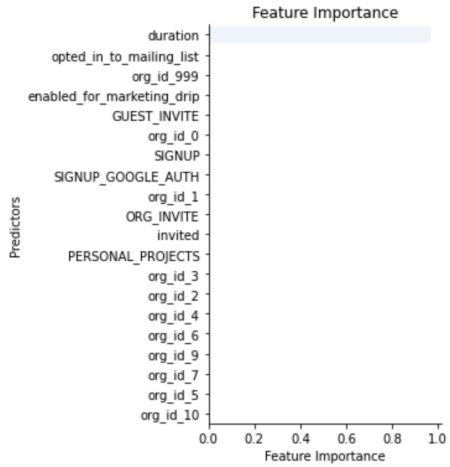
From the creation_time and last_session_creation_time, I created a new feature 'duration' which tracks the number of days between creation time and last session time. This tells us how long after signing up, the user is still using the app.

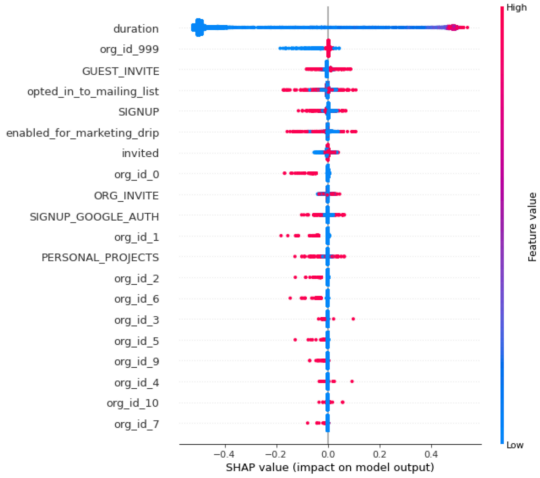
Data Modelling

I built a random forest classifier to predict adopted users, the model achieved an overall accuracy of 95%.

Feature Importance & SHAP

The relative importance of each feature in the model is shown in fig 1. We can see that the duration feature explains for over 96% of the model. From fig 2, we can deduct that users who received an invite are more likely to be an adopted user. Those who signed up on the website are less likely to be adopted user. Those who enabled for marketing drip are also less likely to be adopted user.





Recommendations

Duration is by far the most important feature, this is expected as a user is likely to be adopted if they are still using the app after a long period of time. In order to increase duration, I would recommend giving incentives for users to use the app often e.g. daily login bonuses. Users who enabled for marketing drip are actually less likely to be an adopted user. I would recommend reviewing the marketing drip and improving it's content to keep users engage. A/B testing could be implemented for future marketing drip content to see what works best with users.

Users who signed up through an invite are more likely to become an adopted user, I would recommend encouraging current users to send guest invites e.g. increasing the quota for guest invites (if there are currently any restrictions) or adding incentives for sending guest invites e.g. in-app credit to use for each invited guest who signs up. On the contrary, users who signed up via the website are less likely to become an adopted users. I would recommend decreasing advertising budget that generates traffic to the website's sign-up page and focus efforts on encouraging internal invites instead.

Future Scope

It would be interesting to collect user demographic data to see if a particular demographic is more likely to become an adopted user. In terms of the current data, given more time I would like to try other classification models (logistic regression, xgboost etc.) and feature selection techniques.