

Relax Inc Take Home Challenge Findings

By Cameron Hicks

Through proven data science methods, I analyzed the user data and engagement data provided by Relax Inc to identify factors that can predict future user adoption. The user data contained 12,000 users, 1,656 of whom were defined as 'Adopted Users' based on logging in 3 or more times in a 7 day period, equating to 13.8% of current users were adopted users.

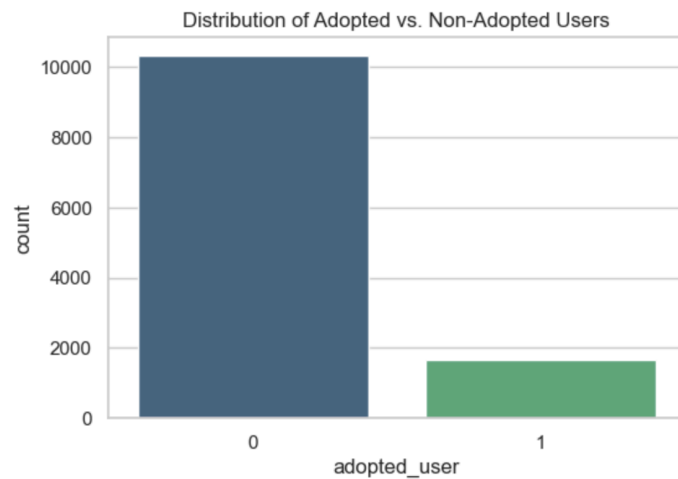
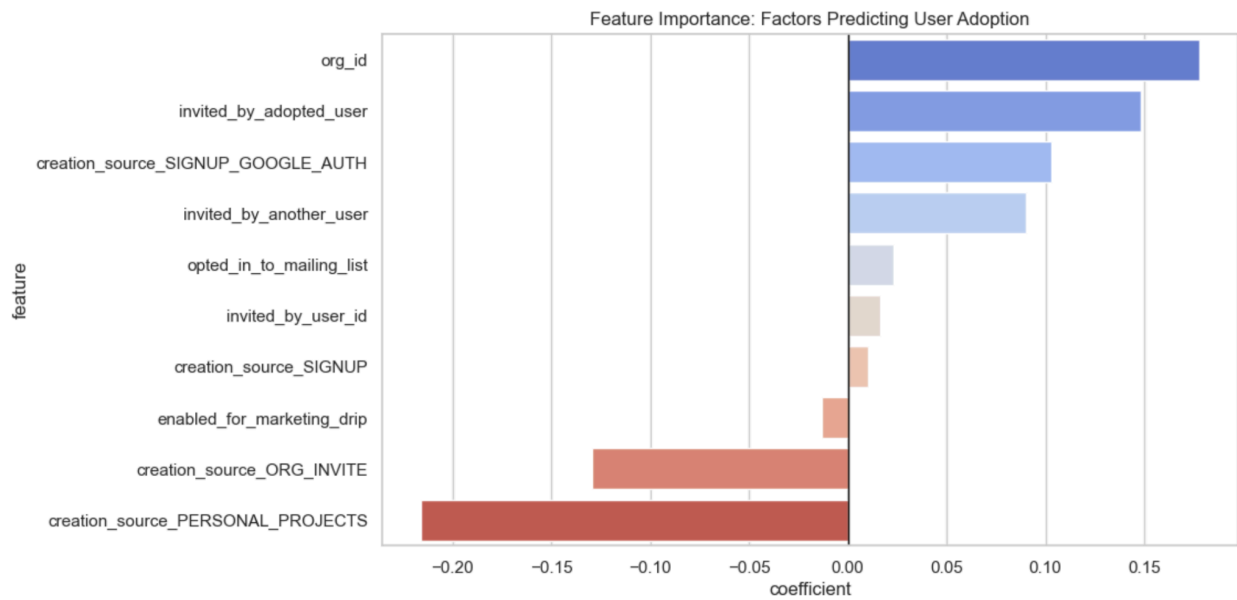
Initial Exploratory Data Analysis showed that being invited by an adopted user is significantly more impactful than simply being invited by another user. Additionally, Organization and Creation Source both showed that specific values had higher adoption rates than others. There is an opportunity to target these specific sources to increase overall adoption rates.

As part of the Exploratory Data Analysis only, I compared account age, total logins, and user activity by month and week. Total logins showed a high correlation to user adoption, indicating opportunities to convert users with high login rates that do not quite meet adoption requirements. Account Age also showed a high correlation to adopted user indicating there may be opportunity to target users with longer tenure who do not meet the adopted requirement to convert. Interestingly, weekly and monthly login averages did not show a high correlation to adoption rates. All of these features were removed prior to training the model to prevent data leakage and increase true accuracy.

To determine the top features, I trained a Logistic Regression model for its ease of interpretation and linear feature modeling. From the model, we can determine that the top 3 features that can predict future user adoption are **Organization ID, being invited by an Adopted User, and creation source via Google Auth.**

	feature	coefficient
2	org_id	0.177860
5	invited_by_adopted_user	0.148341
9	creation_source_SIGNUP_GOOGLE_AUTH	0.103219
4	invited_by_another_user	0.090195
0	opted_in_to_mailing_list	0.023350
3	invited_by_user_id	0.016314
8	creation_source_SIGNUP	0.010281
1	enabled_for_marketing_drip	-0.013302
6	creation_source_ORG_INVITE	-0.129485
7	creation_source_PERSONAL_PROJECTS	-0.215968

Visualizations



There are 1656 adopted users and 10344 non-adopted users of the 12000 in the dataset
13.8% of users are adopted users

