

Relax Inc. Take-Home Challenge

To identify which factors predict future user adoption based on the data provided, we need to define what "adopted user" and then perform an analysis of the factors in the "takehome_users" and "takehome_user_engagement" tables. An "adopted user" is defined as a user who has logged in on three separate days within at least one seven-day period. This means we need to:

- Identify users who have logged in for three or more unique days.
- Identify if those days are within a window of seven consecutive days.

We created an adopted column using these criteria, this column will define our target variable.

The dataset is highly imbalanced.

Using the "takehome_user_engagement" table, we generated new features:

- Total number of logins.
 - Average days number between logins.
 - Number of logins during weekends.
- Adopted users logged in more often than non-adopted users.
 - Many of the non-adopted users logged in only once, and a significant number of them had long intervals between logins.
 - Adopted users have more logins during weekends than non-adopted users.

After training a Random Forest model, we inspected features' importance to identify the factors that most predict adoption.

total_logins	0.458160
weekend_logins	0.315354
avg_days_between_logins	0.187043
org_id	0.024829
creation_source_ORG_INVITE	0.002832
opted_in_to_mailing_list	0.002744
creation_source_PERSONAL_PROJECTS	0.002449
enabled_for_marketing_drip	0.002217
creation_source_SIGNUP_GOOGLE_AUTH	0.002200
creation_source_SIGNUP	0.002173

The total number of logins, weekend logins, and average number of days between logins are the most important factors for adoption prediction.

