26.2.3_Relax Inc. Take home exercise

Ross Brinkerhoff
11.5.24

The task for this exercise is as follows:
Defining an "adopted user" as a user who has logged into the product on three separate days in at least one seven day period, identify which factors predict future user adoption.

To answer this question performed EDA on the two provided data sets and built a predictive model for user adoption. The findings from these efforts are summarized here. More detail ia available in the accompanying notebook.

EDA did not indicate a substantial difference in adopted and non-adopted users in the major provided features. For each feature listed below, which was most features in the data, the two groups are almost identical (see visualizations in notebook), or

• invited_by_user_id
• opted_in_to_mailing_list
• enabled_for_marketing_drip
• creation_source
• org_id

To better utilize the remaining relevant features (creation_time and last_session_creation_time) I created a new feature, which was the number of days between the account creation and the last session creation. This can be thought of generally time on the platform.This new feature, was the most important feature in the predictive model as shown below. 150 has an increasingly positive correlation with a user becoming an adopted user. This relationship remains high after about 150 days. So given the data available, the factor that best predicts future user adoption is how long they have been using the platform.

Feature Importances from XGB model