Prem Ananda
Springboard DSC
4/26/2021

Relax Inc. Challenge

Problem Statement: **Which variables best predict future user adoption?**

According to my analysis, the variables that best predict future user adoption are:
1. Age of the user's account,
2. Creation Source, specifically, if the user used a guest invite to sign up.

1. The age of a user's account was computed by subtracting the account creation time from the user's last login. We could potentially target marketing towards users with older accounts. Users with short account ages could mean that they signed up and never used the account, or used it for a short period of time and didn't use it again, Or, they could be newer users.

2. Creation Source: Guest Invites had the strongest feature importance (see figure) with adopted users among creation sources. This is an actionable insight as we could potentially create incentives to encourage referrals/ invites.
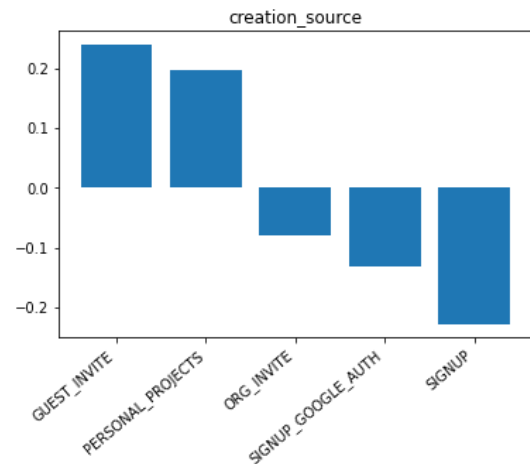
METHOD

In order to analyze the dataset, we needed to isolate the variables, create numeric features from categories (one-hot encoding) for the 'creation sources', create a binary 'adopted_user' feature and calculate relative feature importances. We used logistic regression, since 'adopted users' is a binary feature with the only possible values being 'adopted' or 'not adopted'.

FEATURE	'account_age'	'guest_invite'	'personal_projects'
FEATURE IMPORTANCE	0.41569	0.24026	0.19852

FURTHER RESEARCH: Communicating with the marketing team would be valuable to learn what experiments have been conducted and if we can access the data. We would like to know exactly which inputs marketing can influence so that we could model experiments.