

Relax Take Home Challenge

First, I converted the engagement data to datetime format for easier computation. After determining which users qualified as adopted users, I looked for trends in all reported features of their data, but found nothing obvious. Then, I looked for trends in the non-adopted users' group to see if I could find any recommendations to omit users, but again found nothing obvious, other than perhaps the higher amount of user origins in personal projects. Finally, I built a random forest classifier, hoping for a relatively well performing model which I could pull the feature importances from. The following is the accuracy of the model and the feature importances it calculated:

Classifier score : 0.8186666666666667

Feature Importances :

	Estimated Coefficients	Features
2	0.715587	org_id
0	0.091676	opted_in_to_mailing_list
1	0.069204	enabled_for_marketing_drip
6	0.033891	creation_source_SIGNUP
3	0.027980	creation_source_GUEST_INVITE
4	0.027416	creation_source_ORG_INVITE
7	0.024024	creation_source_SIGNUP_GOOGLE_AUTH
5	0.010223	creation_source_PERSONAL_PROJECTS

That classifier score is not bad, and the feature importances suggest the only feature that really matters is which org the user belongs to. However, looking back at the lists of the users' orgs, the single digit orgs appear a lot in both the adopted and non-adopted users, so studying this will require more time.