

## Relax Take-Home Challenge

The user data and user data engagement data were imported into a Jupyter notebook to answer the question what factors predict future user adoption. The following analysis steps were taken:

1. Read both the csv files to analyze the columns
2. Using the adopted user criteria filtered the unique user ids and added an adoption column to the users file.
3. Cleaned the dataset and dropped the irrelevant columns
4. Applied one hot encoder and random forest and performed a Grid search of parameters before fitting the training dataset and calculated the accuracy score
5. Predicted the test model accuracy from the test data
6. Identified which important features are to be considered for future prediction

Weight: 0.103 | Feature: invited\_by\_user

Weight: 0.086 | Feature: opted\_in\_to\_mailing\_list

Weight: 0.084 | Feature: creation\_source\_GUEST\_INVITE

Weight: 0.080 | Feature: enabled\_for\_marketing\_drip

Weight: 0.066 | Feature: creation\_source\_SIGNUP

Weight: 0.049 | Feature: creation\_source\_SIGNUP\_GOOGLE\_AUTH

Weight: 0.020 | Feature: creation\_source\_PERSONAL\_PROJECTS

Weight: 0.015 | Feature: creation\_source\_ORG\_INVITE

Above listed are the feature names and weights listed in order of importance. The top four features that would help determine future predictions are:

invited\_by\_user - if a user was referred by another user

opted\_in\_to\_mailing\_list - whether user has opted into receiving marketing emails

creation\_source\_GUEST\_INVITE - how the account was created (by a guest invitation)

enabled\_for\_marketing\_drip - whether they are on the regular marketing email drip

We can see that people are serious users when they are personally referred by another user or get a guest invite from a known person. Apart from that opting into receiving regular marketing emails can be helpful in retaining the users.