

EXECUTIVE SUMMARY

Relax Inc. User Adoption Predictors

Background: Relax Inc. wants to increase its ROI by understanding its user adoption rate and finding factors that might predict future adoption. This analysis is based on user logins for a two-year period (May 31, 2012 to June 6, 2014) and information on the 8,823 users corresponding to those logins.

Question 1: User adoption rate. The user adoption rate (defined as three or more logins in any 7-day period) was 18.8% (1,656 out of 8,823).

Question 2: Predictors of user adoption.

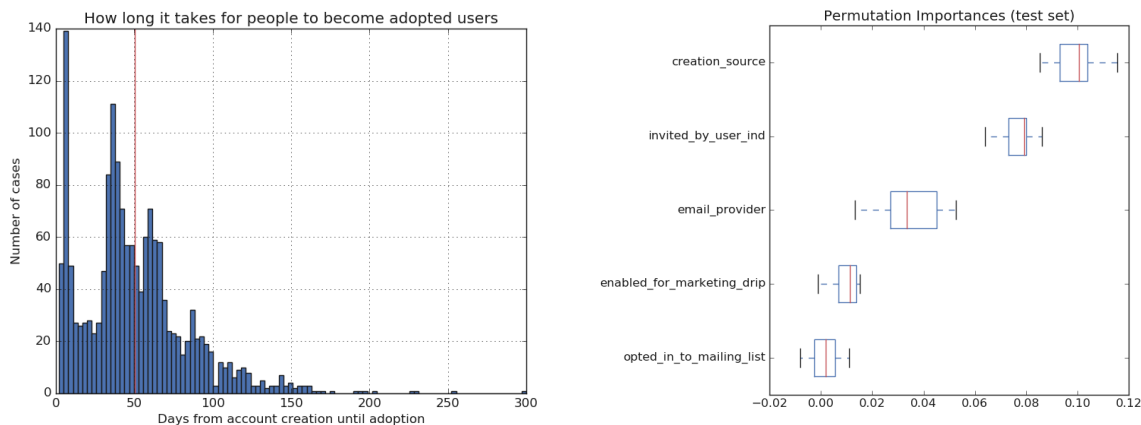
We considered five possible predictors of future user adoption: creation source, email provider (a proxy for geographic location?), mailing list opt-in, whether the user was invited by another user, and marketing drip. A random forest algorithm was used, with cross-validation.

Finding 1: Test accuracy was reasonable but not stellar (64%), given this relatively small set of predictors, with 40% of the adopted users recalled by these predictors.

Finding 2: The strongest predictors were creation source (*hobby users* and *by invitation*), having an invitation from another user, and email provider (*hotmail* users most likely).

Conclusion: We found some moderate predictors of user status but they are of questionable actionable status. More predictors are needed to refine the model.

Future work: Future work might investigate cyclical effects and also the factors influencing early vs late adoption (the average adoption period is 50 days but there is some variation, see below).



Other finer analytic points:

1. Censored data ($n = 932$) were removed, leaving final sample of 7,891 users.
2. Date-based variables were excluded to avoid data leakage.
3. 5-fold cross-validation was used, final hyper parameters (estimators = 5, features = 3)
4. Permutation test used for feature importance as less biased
5. Overfitting was avoided (typical of random forests)