

Project Report: Example Take-Home Challenge: Relax Inc.

The “adopted_user” column was computed via the following steps:

1. Combine the user and engagement files
2. Sort by user_id and time_stamp
3. Loop inside each user_id to check if there were 3 consecutive entries for it within a 7day time period.

1656 out of 12000 users were found to be adoptive users, which is roughly just 13%.

For machine learning modeling, RandomForestClassifier, GradientBoostingClassifier, RandomForestRegressor and LogisticRegression were applied to fit and train.

They all came back with 98% accuracy and performance.

Number of visits turned out to be the most important feature in determining the value of adopted_user, followed by org_id.