

# Relax Inc Data Analysis Report

## **Problem statement:**

The Relax Inc dataset contains data on 12,000 users who have signed up for a product in the last two years. The goal of this report is to find the ***important factors that predict future user adoption***. *User adoption is determined by whether or not a user has logged into the product on three separate days in a consecutive seven day period.*

## **Dataset:**

The dataset contains a total of 12,000 users out of which only 8,823 users had login timestamps available to determine user adoption. Out of the 8,823 dataset, only 1,656 users were classified as 'adopted users' resulting in a highly imbalanced dataset.

## **Methodology:**

I used the Boruta-Py package to determine the most important features in the model. Boruta-Py trains a random forest classifier on the dataset and calculates the relevance/importance by gathering the Z-scores. Before determining feature importance, data cleaning steps were performed and an additional feature ('length of user history') was determined based on the account creation time stamp.

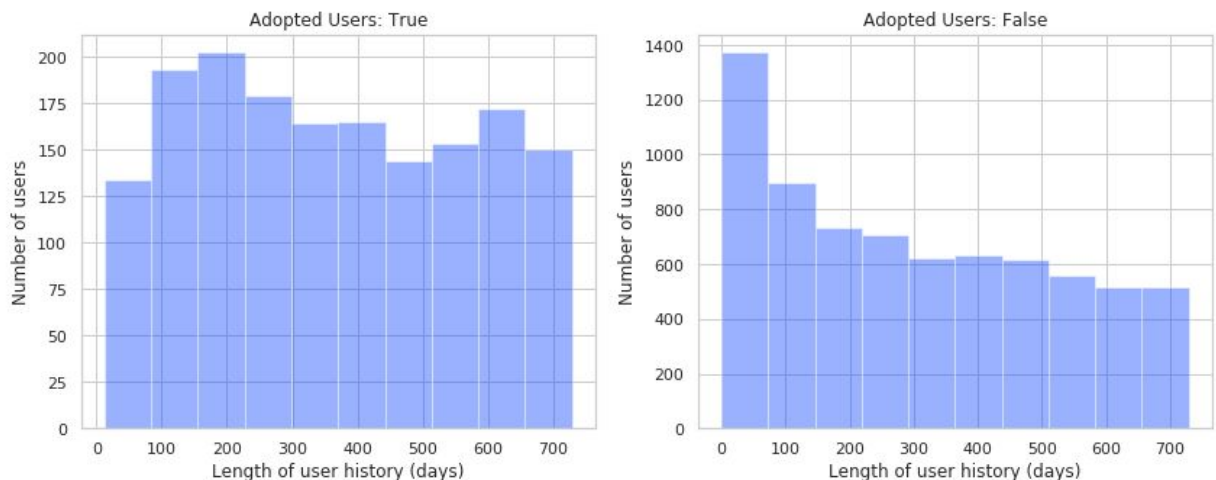
## **Results:**

The most important features for determining user adoption were found to be:

- 1) Organization id
- 2) Length of user history
- 3) Opted to mailing list
- 4) Enabled for marketing drip

Unfortunately, the dataset does not provide any information on the different organizations users belong to except the id, so we can't get more useful information here.

A histogram of the length of user history by 'adopted user type' is shown below. The histogram indicates that a large number of users who are not adopted users have a usage history of < 100 days indicating that users who lose interest in the product within the first 100 days are not expected to be adopted users. On average, user length for an adopted user was found to be 364 days for an adopted user and 305 days for a non-adopted user.



Based on the findings, my recommendation would be to reach out to organizations adopted users belong to and convert non-adopted users by making them enroll in the mailing list and drip marketing (since those are important features), and incentivizing them to use the product longer.