**Relax Inc. Analysis - prepared by Prerna Saxena**

**Introduction**

The goal of this assignment was to analyze data on users from Relax Inc. to identify which factors predict future user adoption. An “adopted user” is defined as a user who has logged into the product on three separate days in at least one seven-day period.

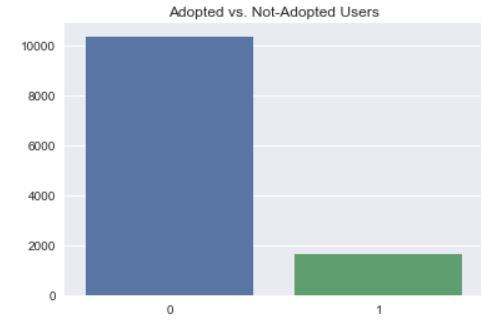
**Analysis**

In order to gauge the predictive strength of the features leading to user adoption I created random forest model as it is known to have yielded better baseline results as compared to other algorithms including logistic regression, Decision Tree, k-nearest neighbors. I further used 5-fold cross validation for an exact estimate of accuracy and fetched 83.17 overall accuracy.

Our most important features turn out to be following:-



**Conclusion and Future direction**

Although, we got a decent accuracy with Random Forest classifier, yet our model suffers from class imbalance heavily. And that is evident from the confusion matrix. This sets the future direction of this analysis. Oversampling/Undersampling/SMOTE techniques should be adopted and algorithms need re-run after that. Feature engineering can be incorporated to come up with a better feature with more power for example Creation Time and Last Creation Session Time could be engineered to find out the duration of usage. Features should be checked for any outliers as well.