**Relax Inc Take Home Challenge:**

In the Relax Inc dataset, there were 12000 users that we analyzed. Of those users, there were 1,358 who were adopted based on the criteria of logging in at least 3 times over any 7-day period. Through feature engineering on the users dataset (recent logins, whether the user was referred, and onehot encoding the creation source), I was able to train a random forest classifier that achieved a respectable prediction of user adoption. The confusion matrix in Figure 1 shows very strong performance predicting the negative class and mediocre performance predicting the positive class.

Figure 1:

Chart, treemap chart

Description automatically generated

Although the model only predicted the positive class correctly 71% of the time, that is far better than other methods when you try to predict something that happens just over 10% of the time. The greatest predictor of user adoption in this case was whether a user had logged in within the last month and that is supported by Figure 2 which shows feature importance in the random forest classifier model below:

Figure 2:

To achieve my results, I tried some other methods to achieve better results like hyperparameter tuning as well as upsampling the minority class in order to generate more accurate predictions. These provided minimal upgrades to the model so the tradeoff of more computational complexity was not worth it. An interesting way to further explore this topic could be to implement a time series model to project logins for the future but for the purposes of this exercise, it was out of scope.