

Data Science SF 13

Facebook Recruiting IV: Human or Robot?

Tucker Silva

General Assembly

The Kaggle competition, *Facebook Recruiting IV: Human or Robot?*, is a currently active competition with all required data sorted and provided. The goal is to identify online auction bids that are placed by "robots", helping the site owners easily flag these users for removal from their site to prevent unfair auction activity. In the train data set, only 5% of the bidders are identified as robots. This extreme imbalance makes it difficult to predict which bidders will be humans and which will actually be robots.

The information in the data set provided included bidder I.D., payment account, address, outcome (robot or human), bid I.D., auction, merchandise, device, time, country, ip, and url. Most information, such as bidder I.D. and payment account, is obfuscated to protect the privacy of bidders. In order to view all the information together, I merged the bidder data set and the bids data set on the bidder I.D. so that my data frame would show all the bids with the bidder's information.

In order to run cross validation and linear regression, I needed to change all the strings of information to integers. I used LabelEncoder to change all the string variables. I then used all the features to try to predict the outcome. The first time running logistic regression I did not specify a class weight. The accuracy score was 87%, which it turns out, is not very good. Just by going through the whole data set and guessing "not spam" every time, the accuracy score would be 95% because of the class imbalance. Time is shown to be the only predictor of which class the outcome will be. Looking at the classification report, it can be seen that "not spam" is being predicted 87% of the time and "spam" is not being

predicted correctly at all. Next, I ran Logistic Regression with the same feature and set the class weight to "auto". This time the accuracy score was 13%. Time is again the only significant predictor of the classes. Although the accuracy is a lot worse than the previous model, the classification report shows that the 13% was all correctly predicted spam.

While this is where I ended my work on the project for this class, my next step will be to look at how the other features relate to the time feature and create new columns that show their relation. I will then test my model on those features to see if the outcome can be better predicted. Additionally, I will try another method entirely, such as Random Forests, to see if it fares better.