

Fake Account Filter

Project Deliverable 3, MAIS 202

1. Final Training Results

I stuck with the SVM model, however I switched to NuSVC instead of LinearSVC. The final model has 90% accuracy on the training set and 89% accuracy on the test set. Compared to the preliminary results, the test set performance saw an increase by 2% and the training set performance saw a decrease by 1%. Despite the decrease in the training performance, I think this new model performs more consistently than the first one. In other words, it generalizes better to unseen data, which is preferable.

Here is the correlation matrix:

0 – Real account 1 – Fake account

Predicted	0.0	1.0
Actual		
0.0	56	4
1.0	9	51

2. Final Demonstration Proposal

I want my final application to use web scraping to get data from Instagram accounts given an inputted username. Using that user data, after standardizing it, my model should predict whether it thinks that it's a fake account or not. I have close to no experience in any web development, so I'm just looking to build something more functional than aesthetically pleasing. I believe I'll use Flask to create the webapp and host it on my laptop for the presentation. With the help of my exec buddy and of the internet, I hope I'll come up with something good.