



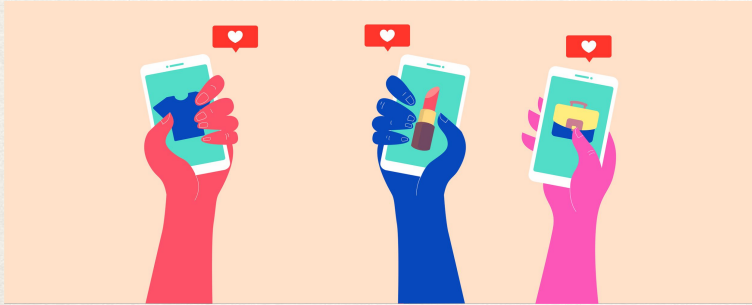
Instagram Account Type Predictor

- Galvanize DSI Capstone Project

Liying (Lily) Li

Dec 2018

Business Application

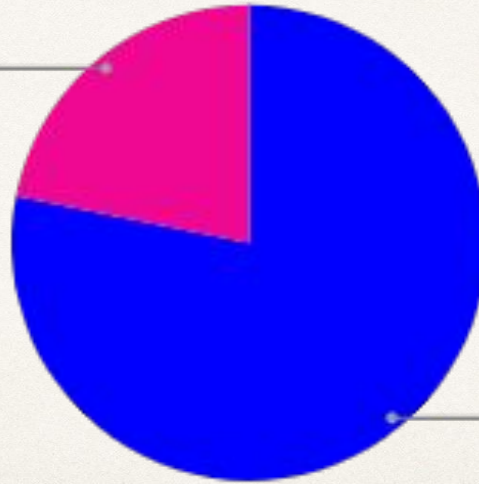


- Discover potential marketing power of Instagram personal accounts

Data Exploration

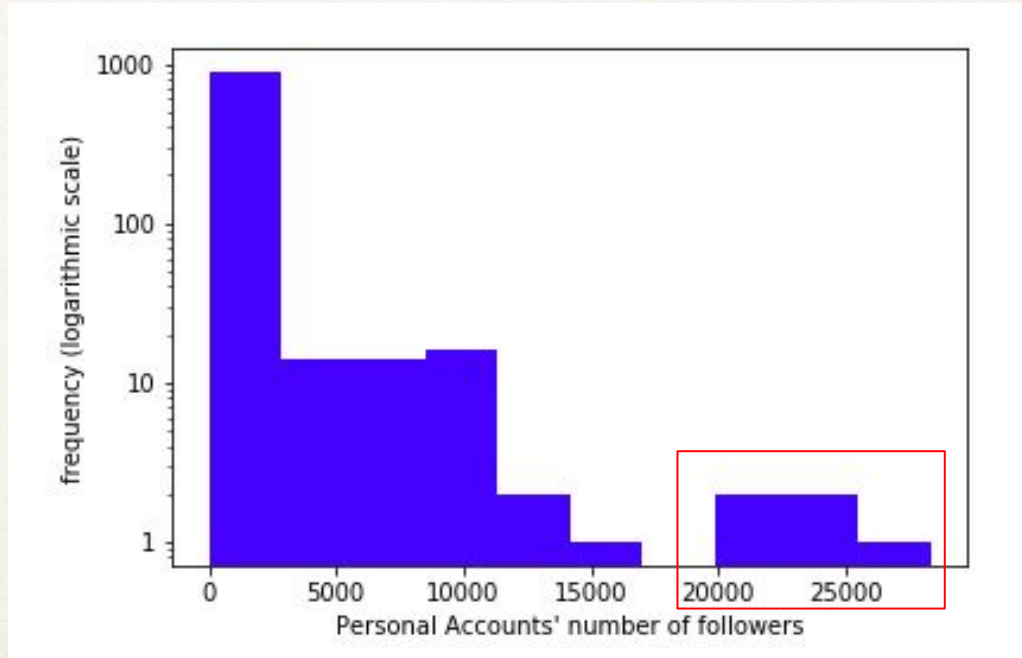
Account Type

Business Accounts
21.8%



Personal Accounts
78.2%

Data Exploration



Data Preparation

- Collected users data from Instagram's official page

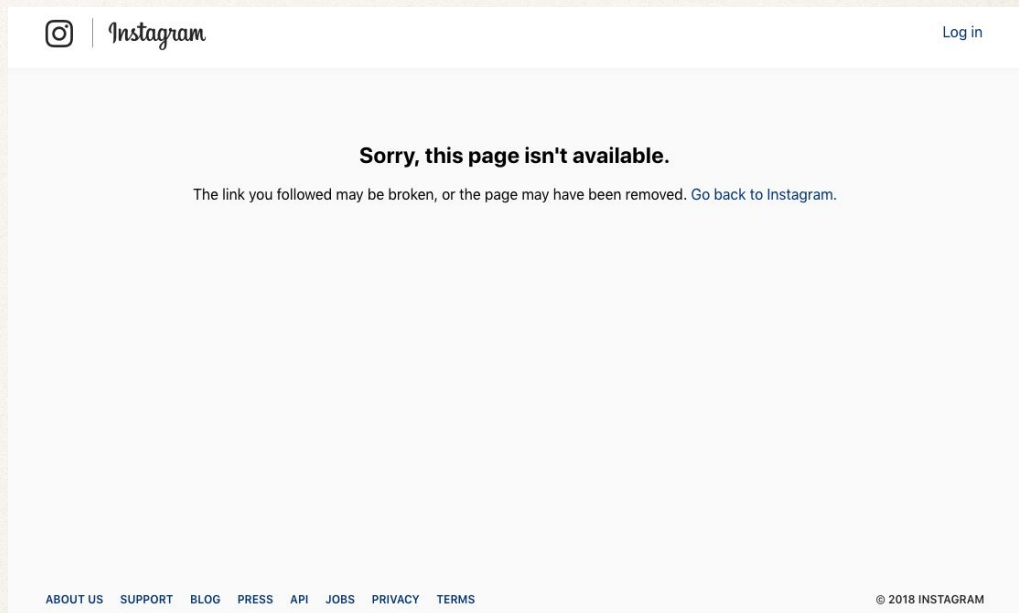


'Target' (labeled data)

- Account type information ('is_business': True/False) is available and can be scraped from any user's Instagram page.

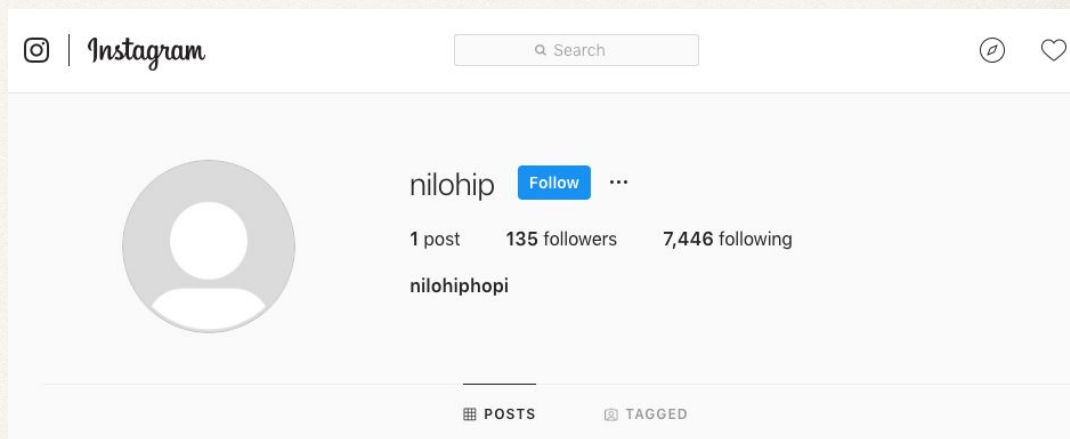
Feature Engineering

- Missing Data
- Added feature
“avg_likes_five_recent_posts”



Feature Engineering

- Missing Data
- Added feature
“avg_likes_five_recent_posts”

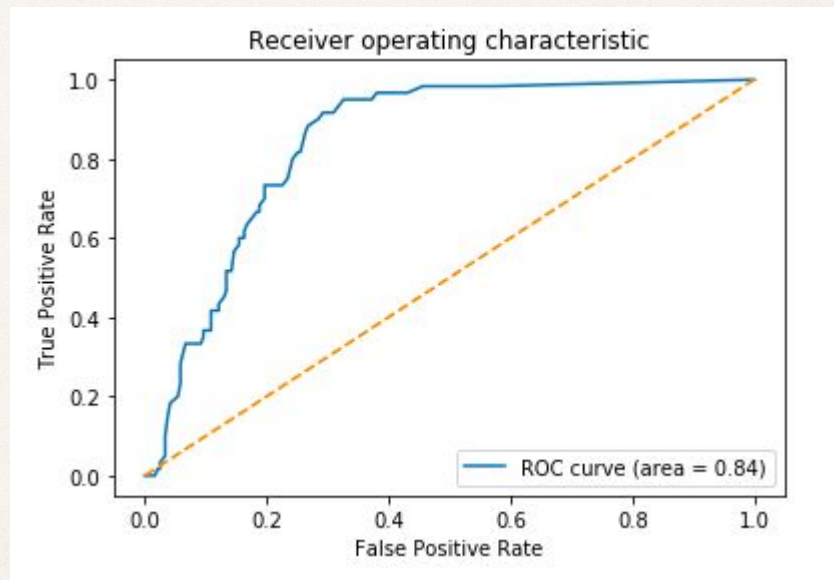


Feature Engineering

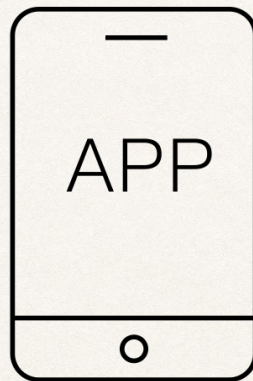
- Missing Data
- Added feature
“avg_likes_five_recent_posts”

Model: Random forest

- Train-test split
- Logistic Regression
- Best model in random forest



Demo



Further Steps

- Website domain name registration
- Some NLP on Biography.

Thank you!

Liying (Lily) Li
Dec 2018



https://github.com/liyingli0818/instagram_account_type_predictor



<https://www.linkedin.com/in/liyingli0818/>



lilyli.0818@gmail.com