# DETECTING CONSUMER REVIEW FRAUD

SI 671 FALL 2020
MCCOY DOHERTY
NICOLAS ORTEGA

An exploration into the viability of conducting fraud-detection of consumer-sourced online product reviews by mining the review text with natural language processing techniques to comparatively analyze differentiating patterns and extract features to train an effective machine-learning fraud-detection system.

## THE PROBLEM

Consumers are more likely to trust platforms that accept and display user reviews, but positive and negative fake reviews alike seem a growing form of malpractice in ecommerce, often to the detriment of consumers and competing honest businesses.

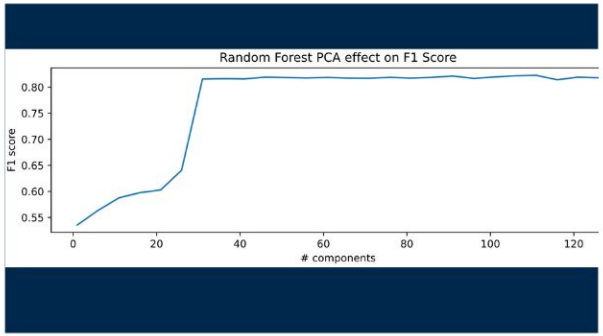
## THE MISSION

Given a dataset of 21K labelled (fraudulent or legitimate) Amazon reviews, we sought out to explore:

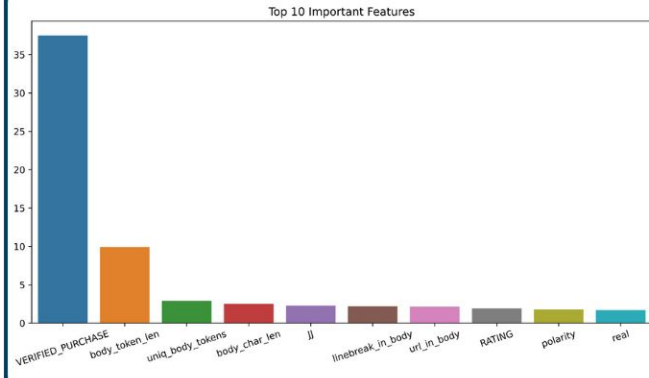
- Can we accurately detect fraudulent reviews using machine learning?
- What features differentiate fraudulent reviews from legitimate ones?
- Is a review being "Amazon verified" substantially indicative of validity?

## THE METHODS

- Feature Engineering using review text
- Binary encoding of human-readable textual features
- 129 non-label features for classification:
  - Random Forest
  - XGBoost
  - Multi-Layer Perceptron
- PCA @ 35 components
- F1 score and Recall as evaluation metrics




## THE RESULTS




- Random Forest (tuned):
  - **F1**: 82.0457, **Recall**: 89.00
- MLP (tuned):
  - **F1**: 81.9454, **Recall**: 87.8571
- XGBoost (tuned):
  - **F1**: 81.4798, **Recall**: 87.0476
- Performance @ PCA=35 components