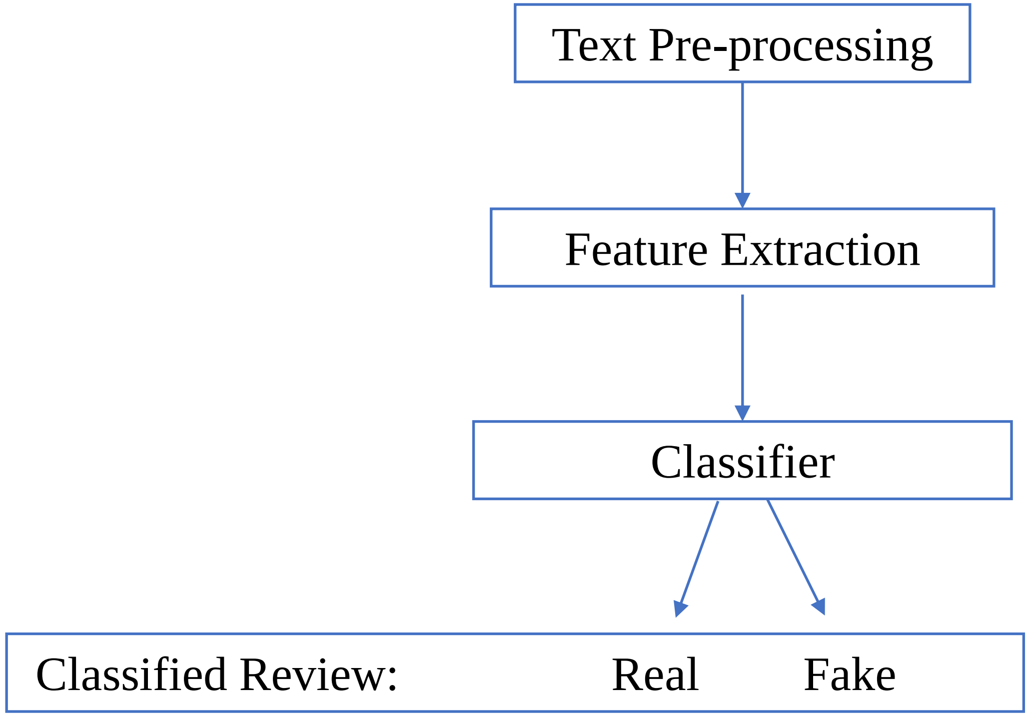
Fake Consumer Review Detection

C. Somers

Problem Overview

- Ever increasing importance of e-commerce on our daily lives forces us to trust opinions of others.
- 93% of consumers say online reviews impact their purchasing decisions (Zhong-Gang et al., 2015).
- This importance results in sellers falsifying reviews to try increase sales leading consumers into believing in dishonest product features or quality.
- Humans are poor at discerning these false reviews however machines can do a much better job.

Methodology: Supervised Learning



Data

- 21,000 Amazon reviews labelled real or fake.
- Balanced dataset (10,500 real vs 10,500 fake).
- Balanced across 19 different categories.
- Rating counts are varied.
- Features:
  - Raw review text
  - Rating
  - Purchase verification
  - Category

Data cont.

Columns keys

`class` = class label the sentence (**this is our target label**).

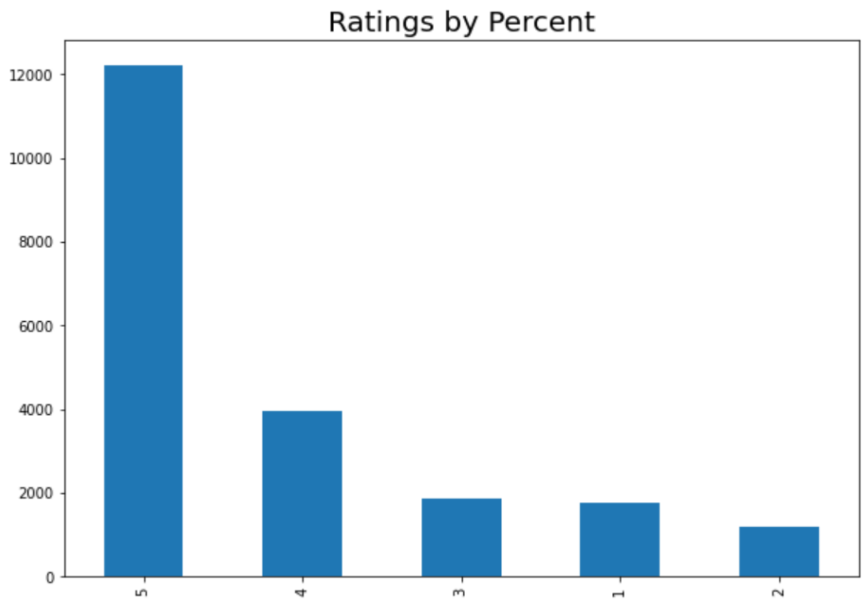
```
0 - Fake
1 - Real
```

`RATING` = rating by customer (range 1 - 5)

`VERIFIED_PURCHASE` = binary label confirming if the product was truly purchased or not

`PRODUCT_CATEGORY` = numerical categories (range 1 - 19)

`raw_sentence` = raw sentence text



Exploratory Data Analysis

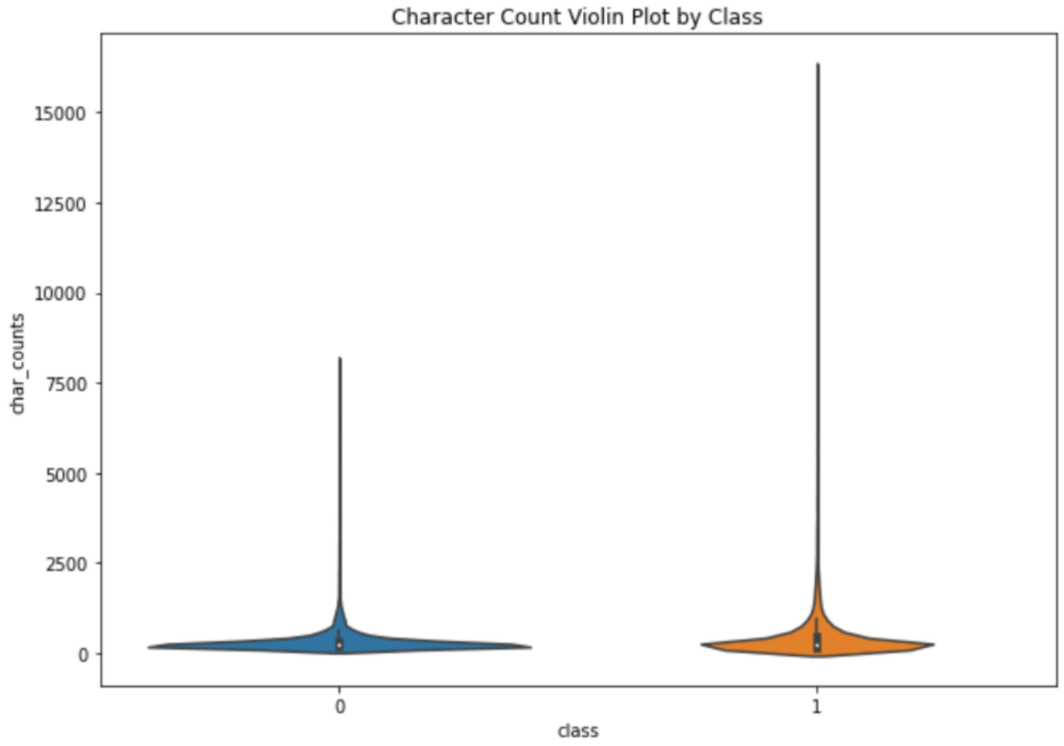
Real Review Word Cloud

Fake Review Word Cloud

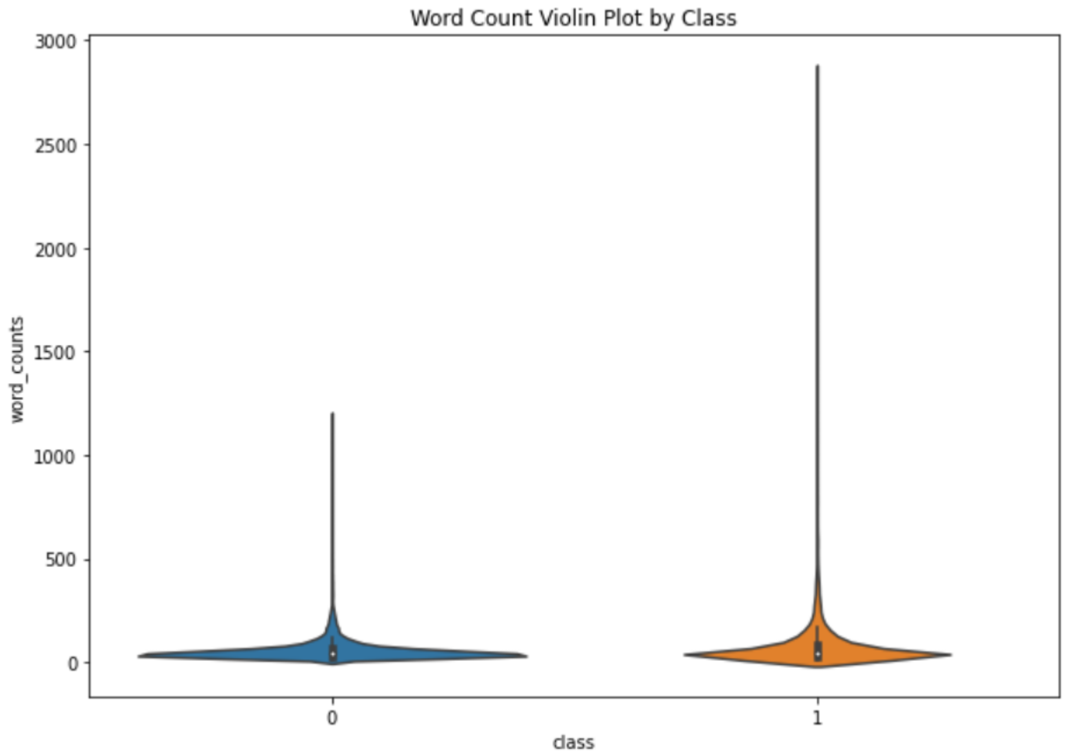
Note - Content of text shares overwhelming similarities to the human eye

Exploratory Data Analysis

```
REAL CHAR_COUNT MEAN: 428.18
FALSE CHAR_COUNT MEAN: 316.61
```




```
REAL WORD_COUNT MEAN: 79.09
FALSE WORD_COUNT MEAN: 59.29
```



Note – Clear discrepancy between text structures of real and fake reviews

Feature Selection

- Clear structural differences allowed me to employ structural NLP features in tandem with given features
- Features for each review:
  - Category (Categorical)
  - Rating (Discrete)
  - Verified Purchase (Binary)
  - Word Count
  - Character Count
  - Unigrams
  - Bigrams

Model Evaluation

- Scored model based on F1-Score as it balances the problem with missing fake reviews (customers mislead/unhappy) and falsely classifying fake reviews (accidental customer banning).
- This, I believe, is the best evaluation metric for the problem.
- Due to strong base lightGBM scoring, hyperparametric optimization resulted in only a minor improvement in scoring (~81% F1).

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

Model Evaluation Cont.

-----------------CLASSIFICATION REPORT-----------------

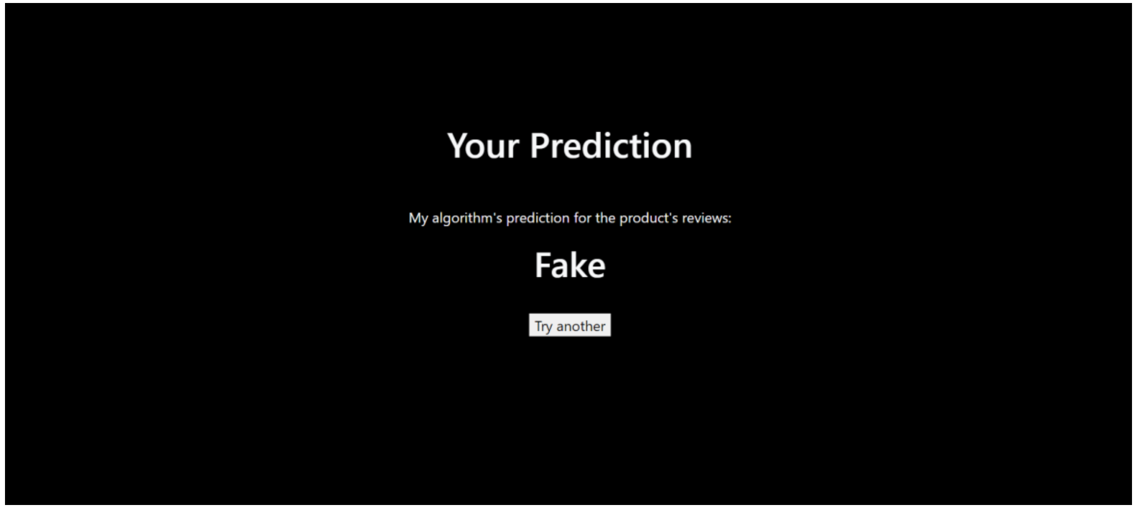
	precision	recall	f1-score	support
0	0.83	0.74	0.79	1062
1	0.76	0.85	0.80	1038
accuracy			0.79	2100
macro avg	0.80	0.80	0.79	2100
weighted avg	0.80	0.79	0.79	2100

Final Out-of-Sample F1-Score: 0.803

Out of Sample (Tuned) lightGBM Evaluation

Deployment

- Deployed using Flask.
- Scrapes data from first-N pages of Amazon reviews.
- Averages N-model estimates to determine the use of false reviews.



5/5 stars

Thank you!
I hope this helps you uncover reviews you can trust.