PREDICTING WHO THE ASSHOLES ARE

By Robert Malka
Springboard Capstone #2
Dec 2020

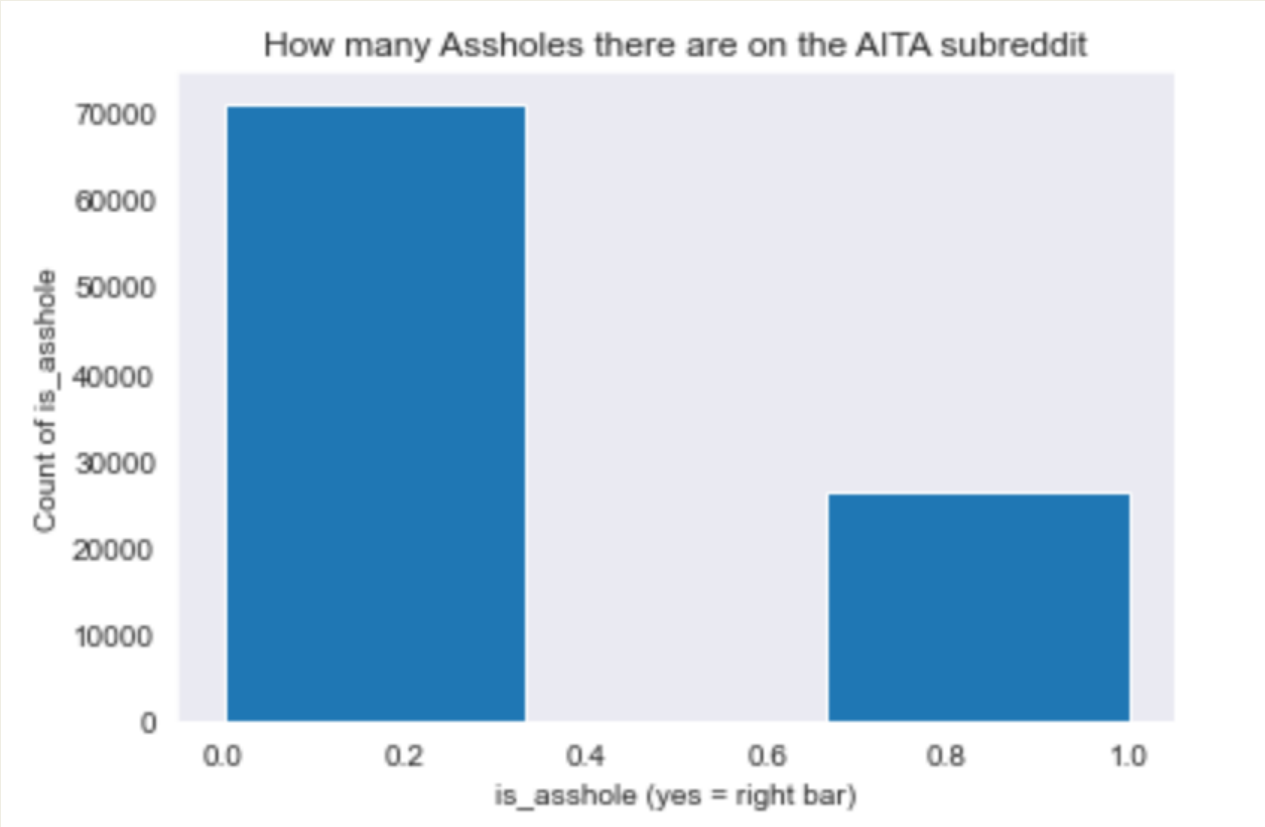
With thanks to mentors Benjamin Bell and Kenneth Gil-Pasquel!

The Business Problem

- What features best help predict whether someone will see a controversial situation we're communicating positively or negatively?
  - *Applicable to businesses pitching, the field of self-help, criminal defendants before a judge & jury.*
- Explicitly: What features help predict whether or not a community judges someone to be an asshole?

The Data

- Scraped from the AITA (Am I The Asshole) Subreddit on Reddit.com
- >25% considered assholes
- ~50% about family;
- ~40% about relationships;
- ~10% about work.



The Features

- Whether the post was edited
- Frequency of pronouns (I, he/she/they)
- Length of post
- Question mark at the end of the post
- Reading level of post
- “Would I be the asshole” vs “Am I the asshole”
- Use of certain nouns, verbs, adjectives
- Categorizing posts according to family, work, or relationships
- And others!

The Wordcloud

The Notable Findings

- Question Mark at end of the post is slightly correlated with NTA
- The valence of a post's tone (positive/negative) is *not* correlated
- Higher likelihood of being an asshole if the issue is about sex (33% vs global avg ~26%)
- Younger the poster relative to his subject (e.g. talking about parents), the less likely he is to be an asshole. The converse also applies.
- Reading level of the post makes no difference
- Use of pronouns have minimal correlation with community judgment
- If post was edited, correlative with YTA.

The Modeling

- Type: Supervised Learning
- Binary Classification: 1 for YTA (You’re the asshole) and 0 for NTA (Not the asshole)
- Imbalanced data – 25% YTA
- Tools: Scikit Learn
- Scoring = “roc_auc”
- Data splitting into train/test sets (50%, 50%)
- Weighted data to take care of imbalance problem
- Used cross-validation for hyperparameter tuning (5-fold cv)

The Modeling Continued

- F-Score = ideal measurement for model effectiveness.
- Classification Algorithms Used:
  - *Decision Tree*
  - *Random Forest*
  - *Logistic Regression*
  - *Gradient Boosting*

Model Metrics + Comparison

Model	Gridsearch Parameters	Recall	Precision	Accuracy	F1
Decision Tree Classifier	Criterion = Entropy, max depth = 4	0.577	0.555	0.551	0.669
Random Forest Classifier	Criterion = Gini, n_estimators =97, max depth = none (chose 12 from prev. gridsearches)	0.573	0.580	0.577	0.577
Logistic Regression	C = 0.01, penalty = 12	0.551	0.571	0.566	0.561
Gradient Boosting	Criterion = mae, learning rate = 0.1, loss = deviance, max depth = 3	0.549	0.559	0.555	0.669

Winner: Decision Tree!? (Random Forest + Gradient Boosting deserve further investigation.)

Future Directions

- Group words by parts of speech
- Examine the tone of each post in more nuanced ways and compare
- Weigh comments based on scores and awards (put YTA/NTA on a spectrum).
- Grouping titles in more nuanced groupings.

Thank you!