



The Good, the Bad, and the Neutral

Using classification models to assess
character alignment in comics

Noah Jaffe

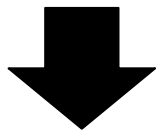
Which attributes are associated with good and evil?

- Inferential: looking for trends in the data



Features and Target

- Secret identity?
- Sex
- Hair and eye color
- GSM?
- Alive?



Good

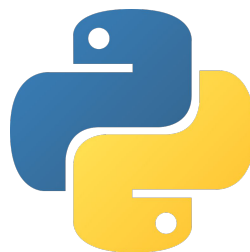
Neutral

Bad



Data cleaning

kaggle



Pandas

Numpy

Matplotlib

Seaborn

Images: <https://pngimage.net/database-logo-png-4/>,
<https://commons.wikimedia.org/wiki/File:Python-logo-not-ext.svg>, <https://en.wikipedia.org/wiki/Kaggle>

Models

- Logistic regression
- KNN
- SVM
- Naive Bayes
- Decision Tree
- Random Forest
- XGBoost

XGB performed best, but...



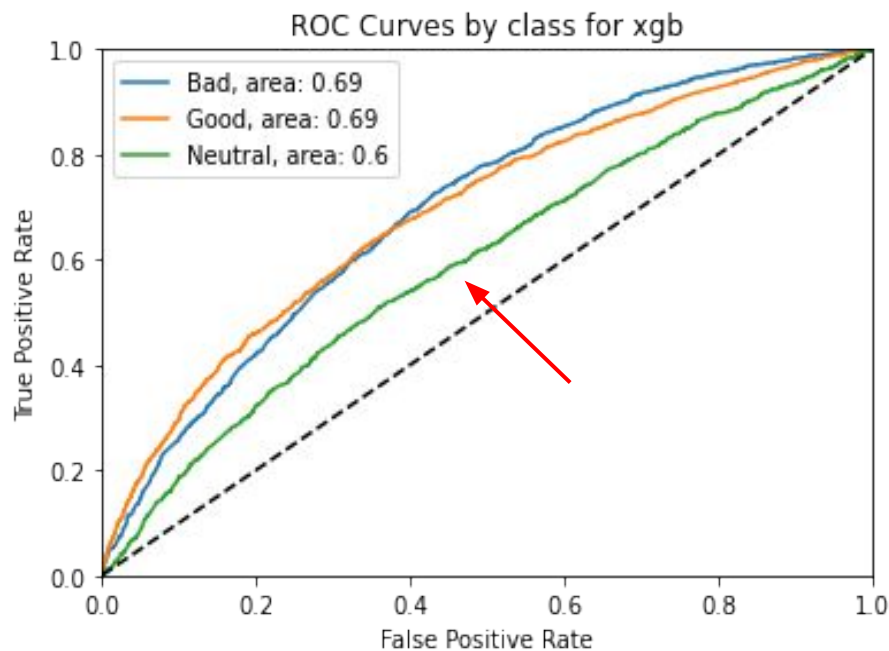
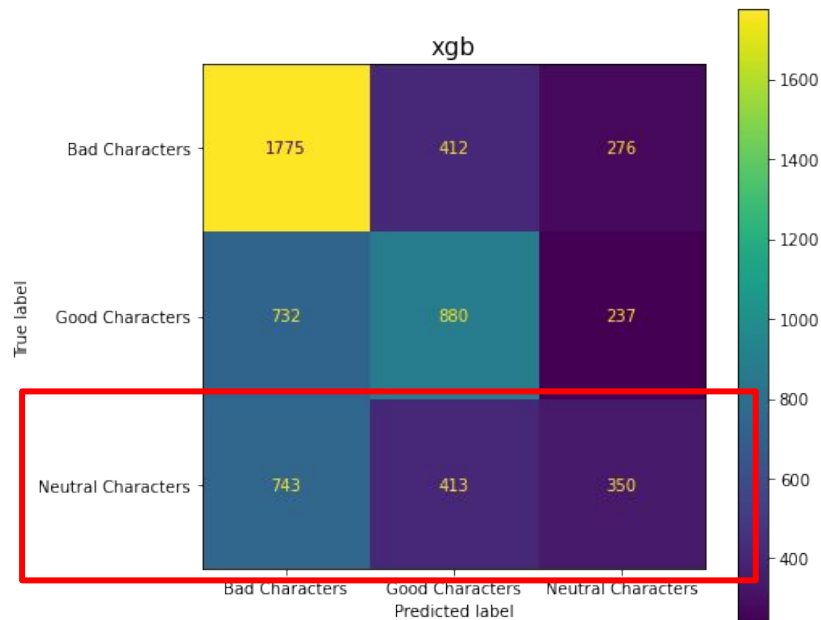
Neutral category poorly predicted

precision recall f1-score support

0 0.58 0.46 0.51 2419

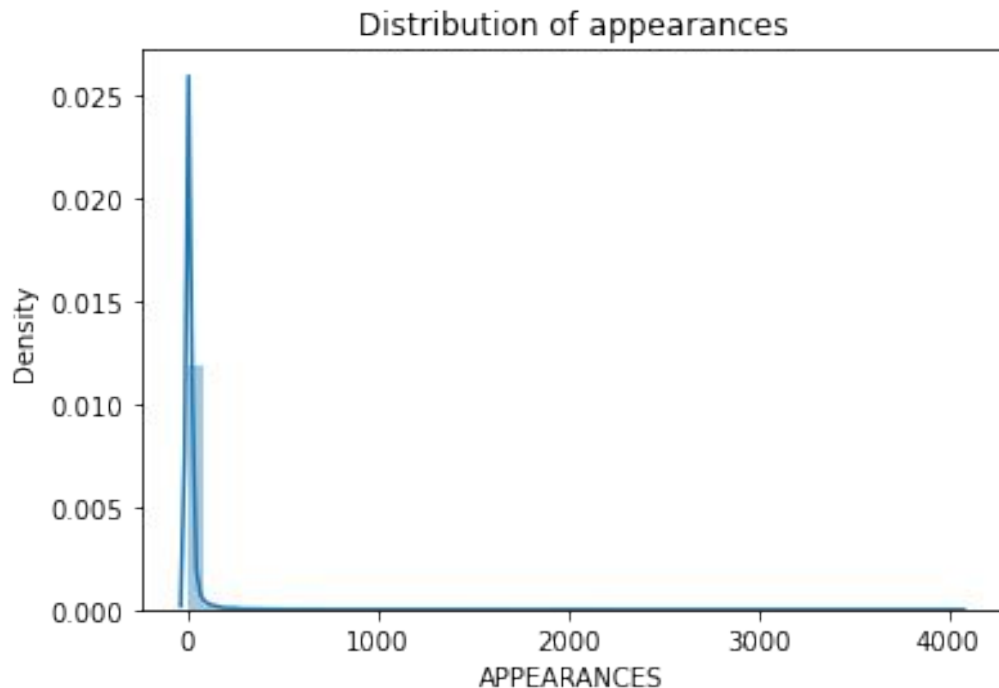
1 0.60 0.29 0.39 1916

2 0.40 0.08 0.13 1483



Why is predicting alignment difficult?

- Dearth of well-developed characters
- Neutral characters have less extreme traits
- Lacked *personality* data
 - NLP/sentiment analysis



Takeaways

- Character tropes and trends are fluid
- What **actually** determines characters' morality?
- How have characters changed?



Gritty!

Images: <https://www.themarysue.com/celebrate-superman-month-with-a-historical-timeline-part-1-1930s-to-1960s/>
<https://www.popsugar.com/entertainment/First-Picture-Henry-Cavill-Batman-v-Superman-35141922>



<https://ew.com/movies/2019/10/14/deadpool-writers-marvel-future/>
<https://screenrant.com/venom-cast-character-guide/>, https://batman.fandom.com/wiki/The_Dark_Knight_Rises,
<https://www.empireonline.com/tv/reviews/the-boys-season-2/>,
<https://www.amazon.com/SHAZAM-POSTER-ORIGINAL-Advance-ZACHARY/dp/B07MBP4LMG>

Demo!



Image: <https://www.insider.com/captain-america-civil-war-concept-art-2016-9>

Questions?



Important features



<https://metro.co.uk/2019/10/10/brie-larson-confirms-female-mcu-stars-asked-women-marvel-movie-10893821/>

