Identifying Differences In Political Subreddits

By Ricky Everest

Banned And Controversial Subreddits

Banned

- **r/Incels - promoting violence -misogyny**

- **r//ChapoTrapHouse - promoting violence**

- **r/GenderCritical - transphobia -promoting hate**

r/Incels has been banned from Reddit

This subreddit was banned due to a violation of our content policy, specifically, our sitewide rules regarding violent content.

Controversial

- **r/FemaleDatingStrategy - homophobia, transphobia, misandry**

- **r/KotakuInAction -harassment -brigading**

- **r/TheRedPill -antifeminism, rape culture, hegemonic masculinity**

Goals

Identify any major and/or potentially harmful shifts in comments

Create a model that can accurately categorize comments from subs with similar but differing politics

Based on findings formulate ways to generalize detection of political shifts in subreddits

Destiny

- **Identifies as left wing, audience identifies as center left**

- **Debates against far right and left figures**

- **Banned for Inciting violence and Hate speech on twitch**

- **Some complaints of harassment from fans**

About Community

Hub for all things Destiny (the streamer, not the game).

Created Oct 27, 2008

173k
Daliban

● **3.2k**
Online

Top 1%
Ranked by Size

Vaush

- **Far left /socialist**

- **Banned on twitch for racism and calls to violence**

- **Debates against far right and left figures**

- **Fans have doxed political opponents**

About Community ...

Subreddit for the socialist streamer & media critic Vaush

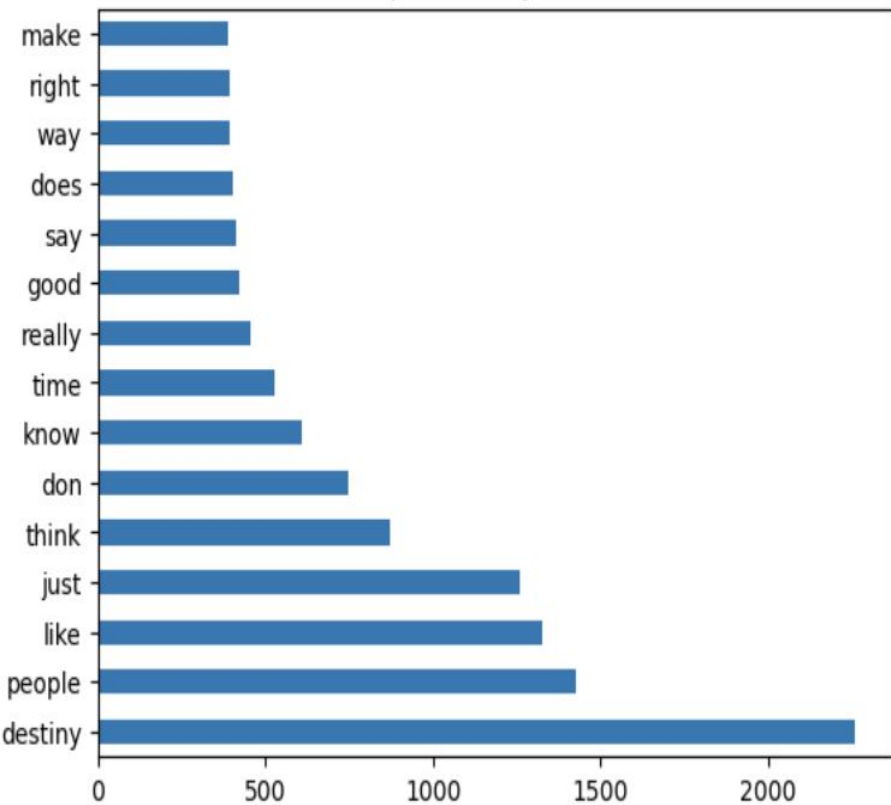Created May 3, 2019

57.7k
Comrades

● **861**
Antifa Supersoldiers

Data

- 9934 from Nov 3rd to Jan 12th combined post title and text combined
- 5000 from r/Destiny
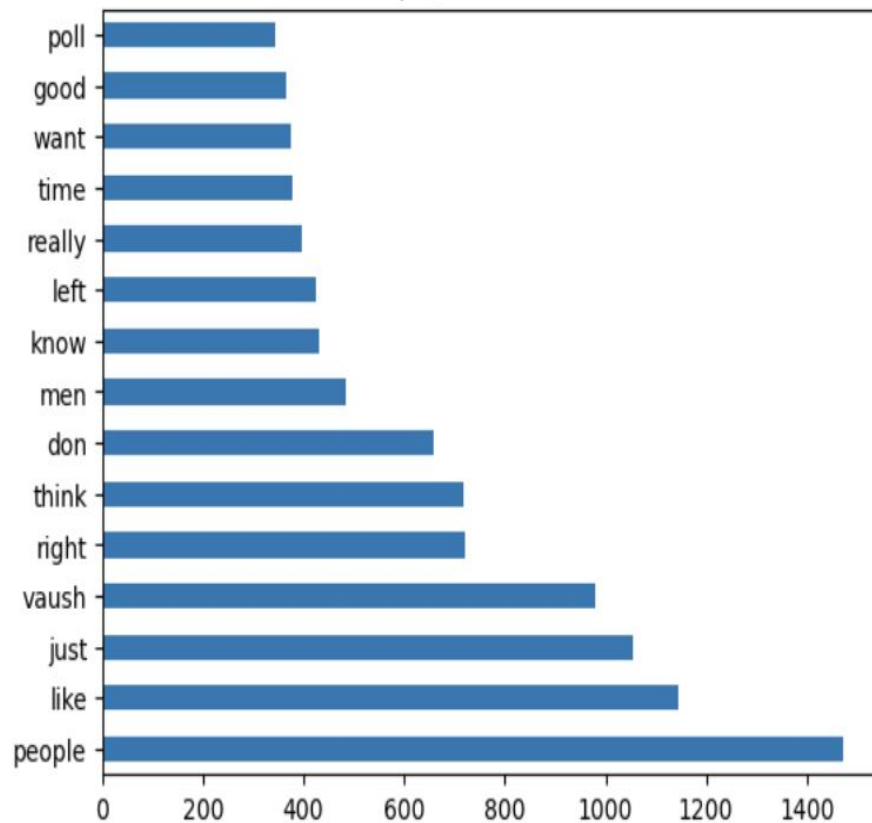- 4934 from r/VaushV

Initial Analysis

- CountVecorizing Text
- Quantifying occurrences of words and phrases
- Sentiment Analysis
- Word Count

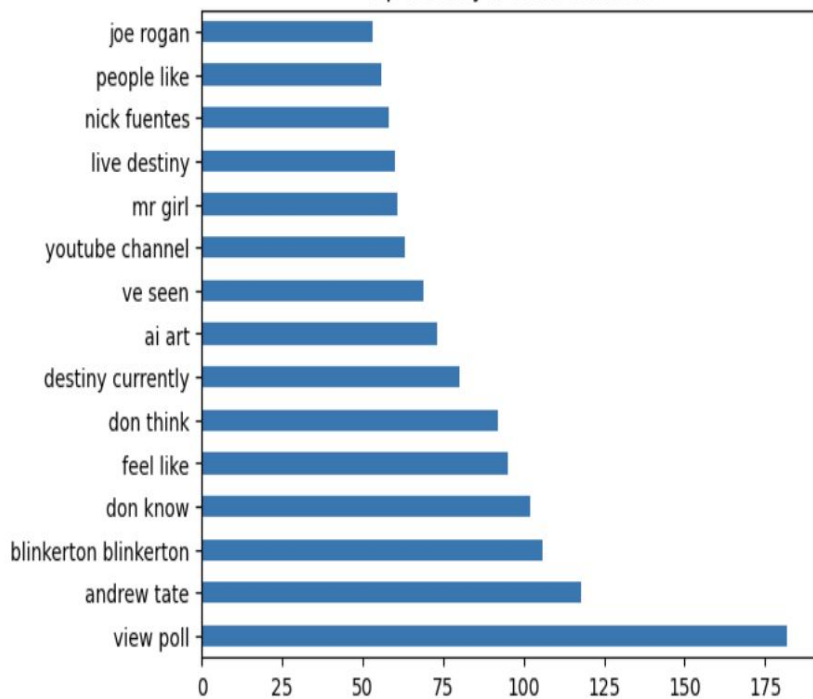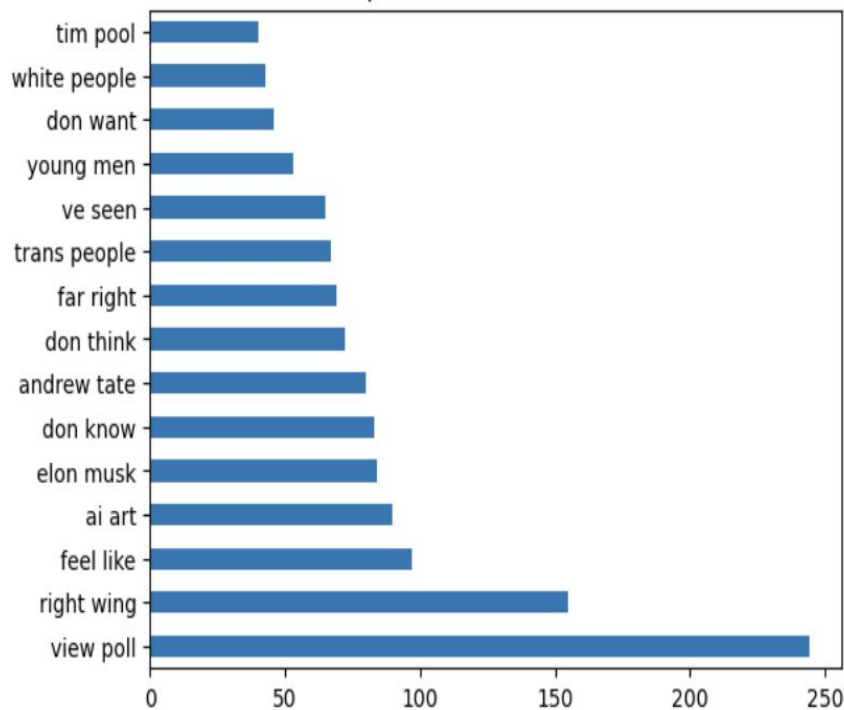# Single Words Don't Reveal Much

Top r/Destiny Words

Top r/Vaush Words

Some Community Differences With 2 Word Combos

Top Destiny 2 Word Combos

Top Vaush 2 Word Combos

Word counts

Vaush

Word Count max 2598

Post word count mean 48

Post word count med. 13

Destiny

Word Count max 4182
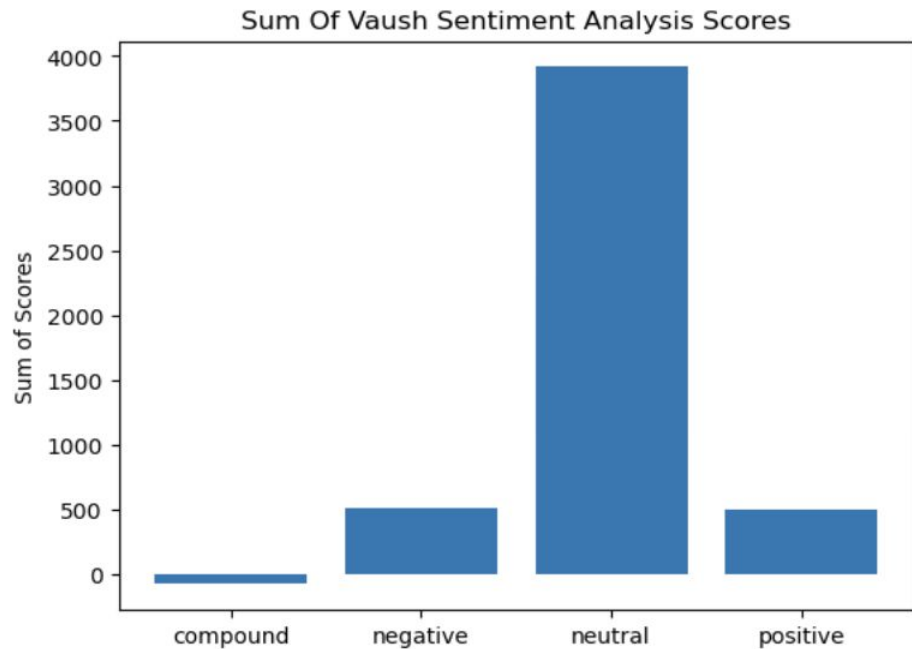
Post word count mean 56

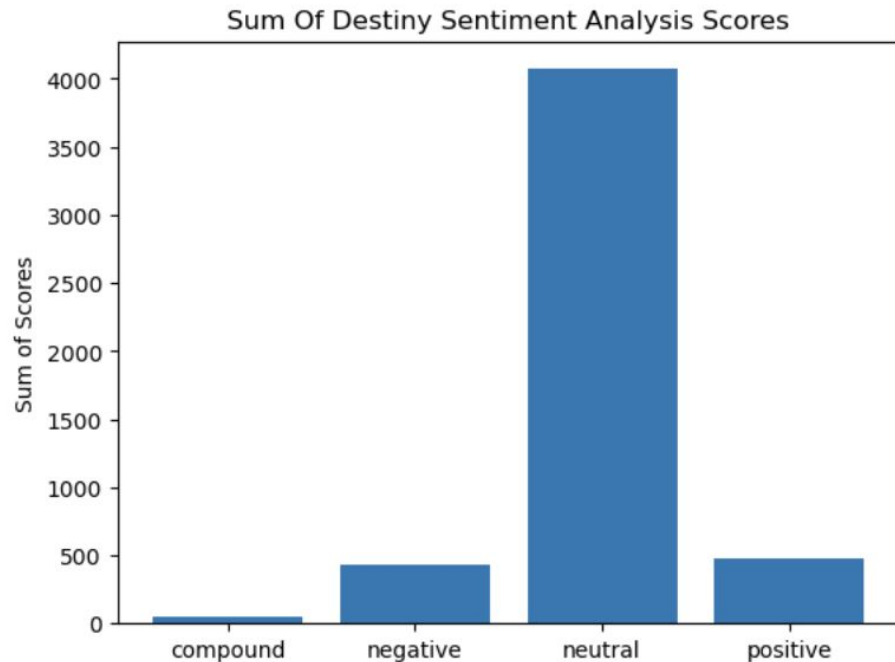**Post word count med. 13**

Vaush
Compound score   -77.9468

most sane south park fan: .3 positive

Sum Of Vaush Sentiment Analysis Scores

Destiny
compound score   44.275

Fine day isn't it fellow factorio enjoyers: .5 positive

Sum Of Destiny Sentiment Analysis Scores
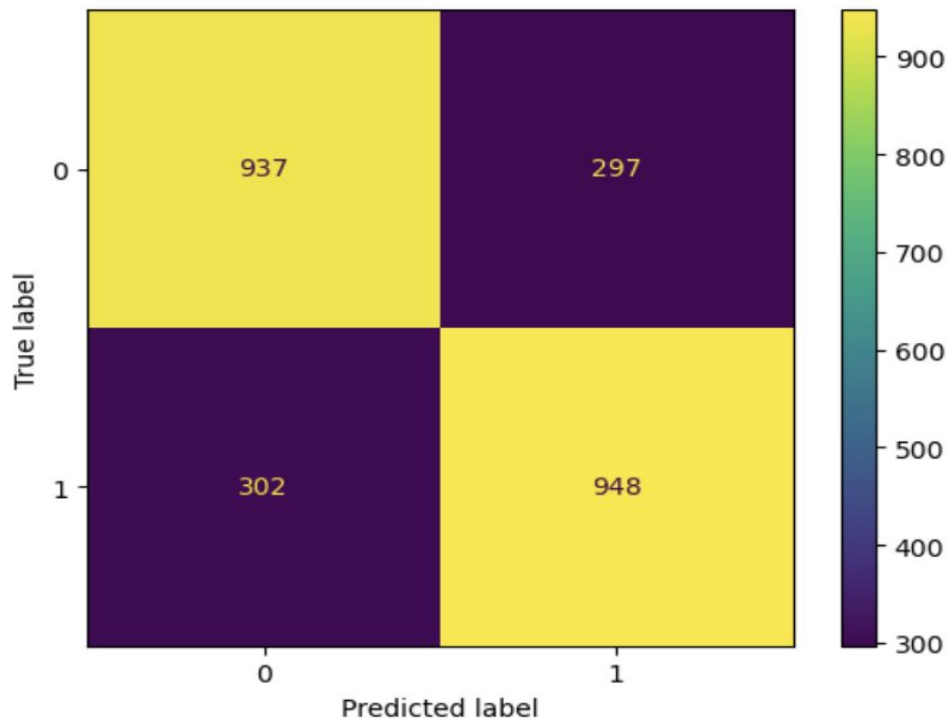
Tested Models

- XGBOOST, ADABOOST, NAIVE BAYES,LOGISTIC REGRESSION, RANDOM FOREST
- All Count Vectorized (no tf idf)
- A/B tested numerical features
- A/B tested lemmatization
- All single estimator models grid searched various hyperparameters
- 6 Variations of stacked models

My 2nd best performing model stacked the lvl one estimators XGBOOST, ADABOOST, NAIVE BAYES, AND RANDOM FOREST
The level 2 estimator was a logistic regression

Baseline Accuracy
Destiny   0.503356
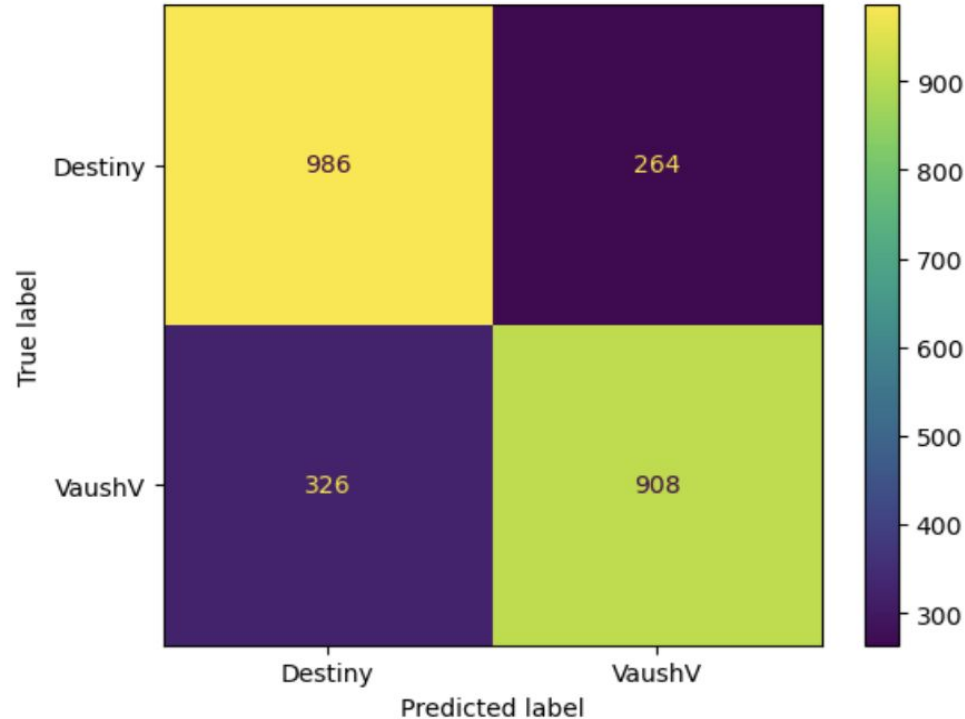**VaushV   0.496644**

Test Accuracy .75

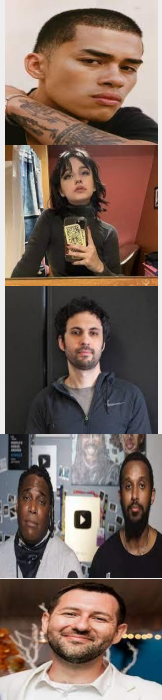# My Best model just used the The multinomial Naive Bayes classifier

Baseline Accuracy
Destiny   0.503356
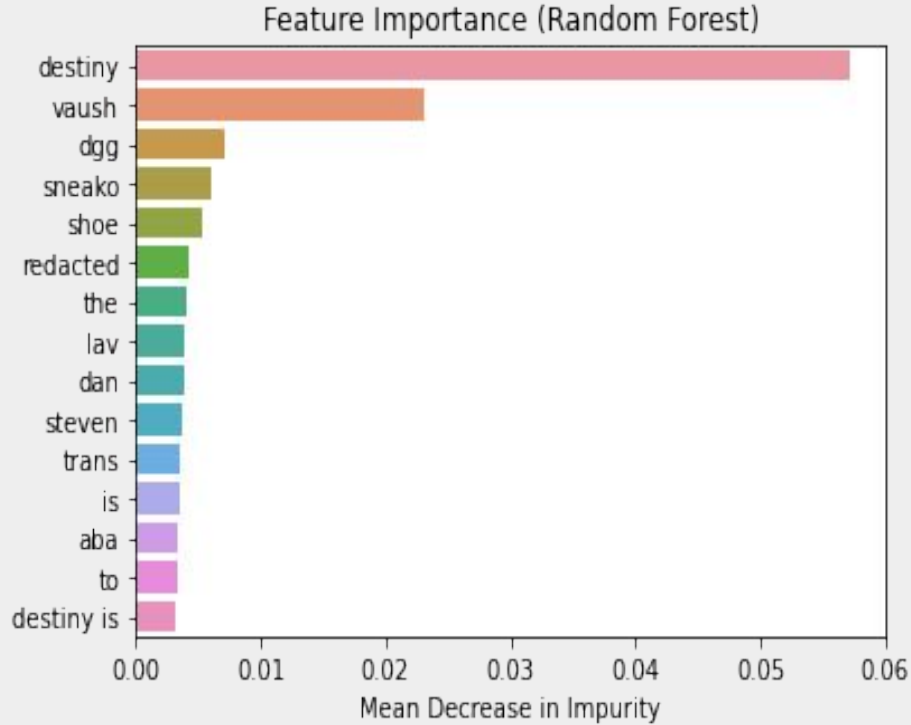**VaushV   0.496644**
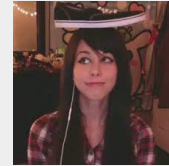
Test Accuracy .76

Important Words In Random Forest

Destiny

Vaush

Feature Importance (Random Forest)

Conclusions/Recommendations

There were some major differences in rhetoric but they mostly revolved around memes and community members rather than politics.

In another attempt I would pick a timeframe via the reddit api when both communities were debating politics with each other.

Shifting political opinions may be easier to detect via sentiment analysis on words/phrases like trans, left or right wing and the total overall usage of the certain words and phrases.

While the final model was decent at differentiating communities it had a lot of help from using Destiny's name.
I'd want to see how the models perform without the creators' names. (Destiny is a banned word on r/Vaush)