

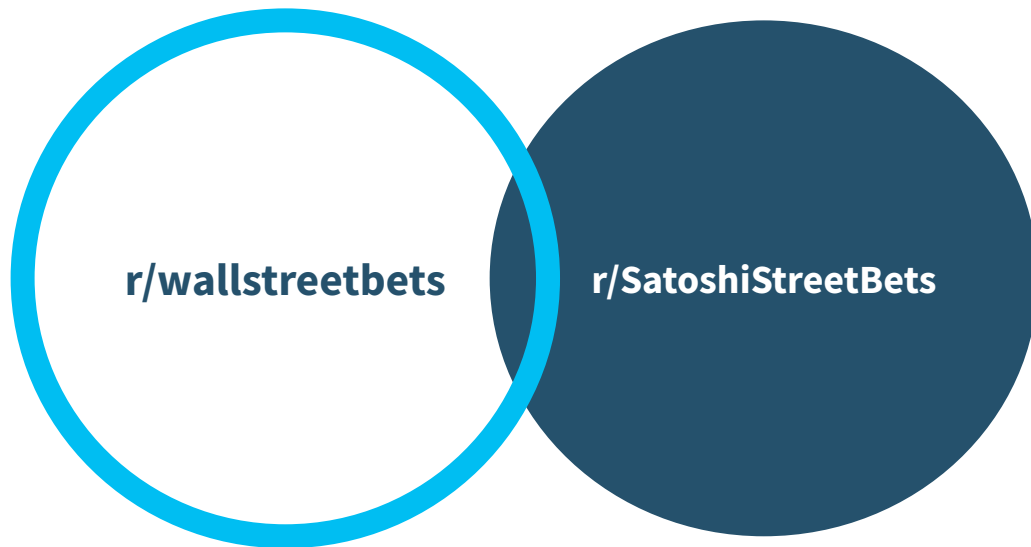
r/wallstreetbets or r/SatoshiStreetBets? Predicting the subreddit of a post using NLP

Seung Woo Choi



Disclaimer

An Introduction to Two Subreddits



Problem Statement

1. Build a classification model with accuracy of 80%+
2. Determine words unique to two subreddits



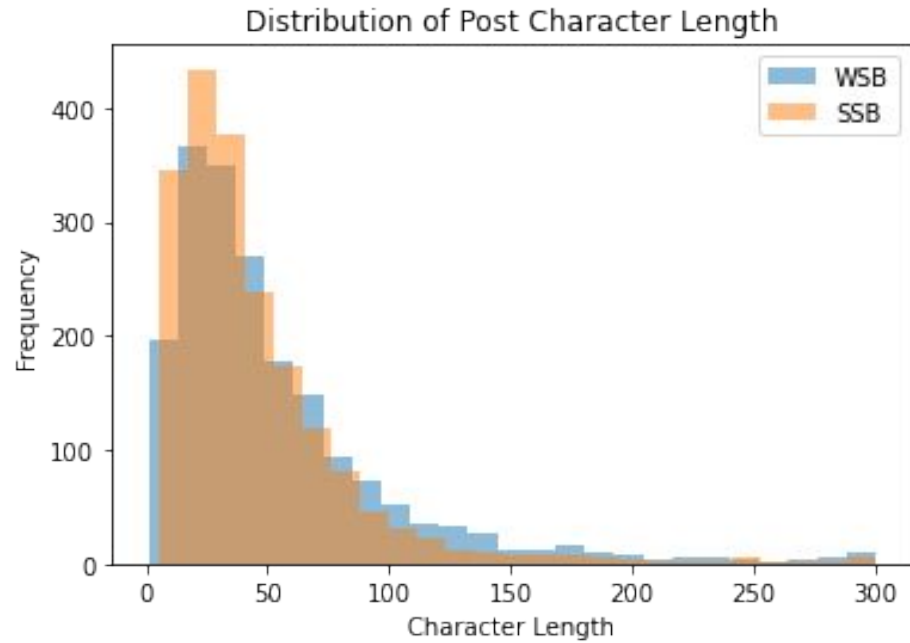
Data Source

Pushshift's web API

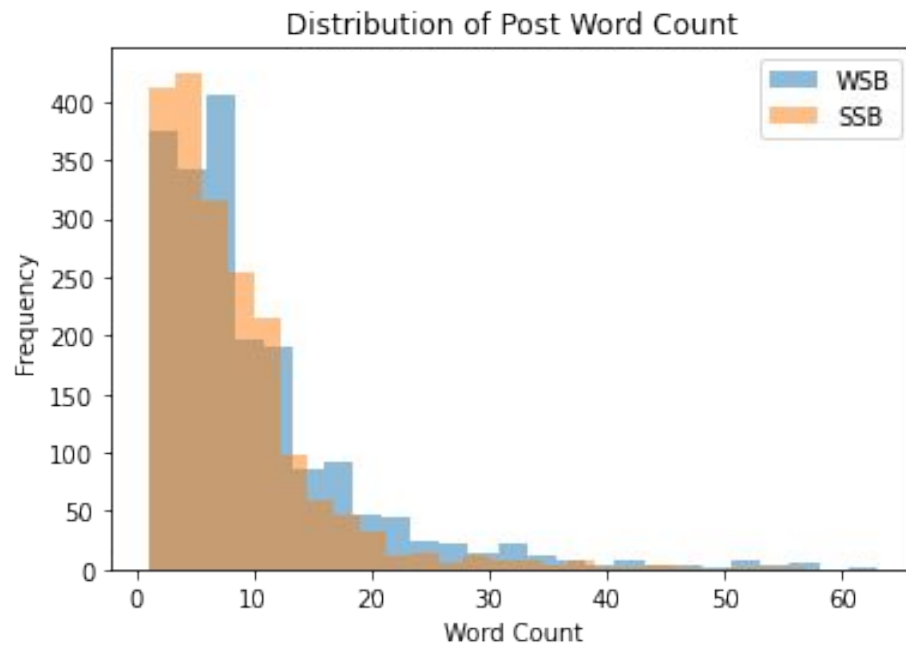
**Monday, March 1, 2021
9:06:40 PM**



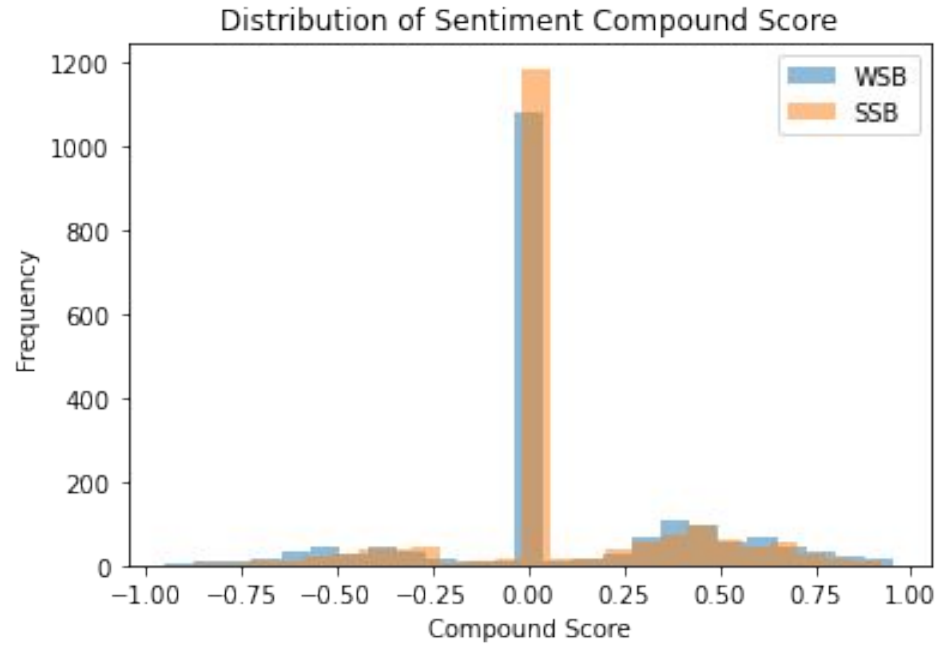
Exploratory Data Analysis



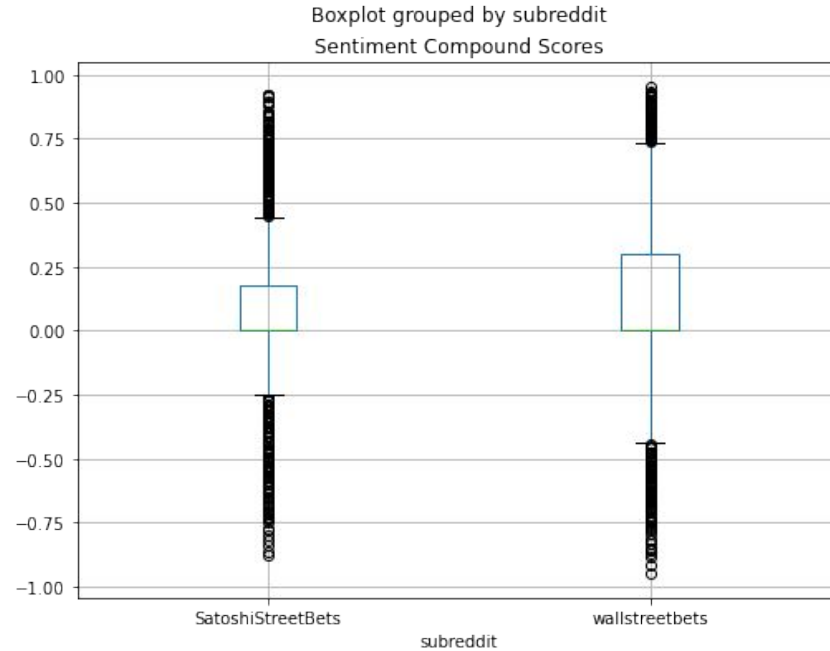
Character Length of Posts



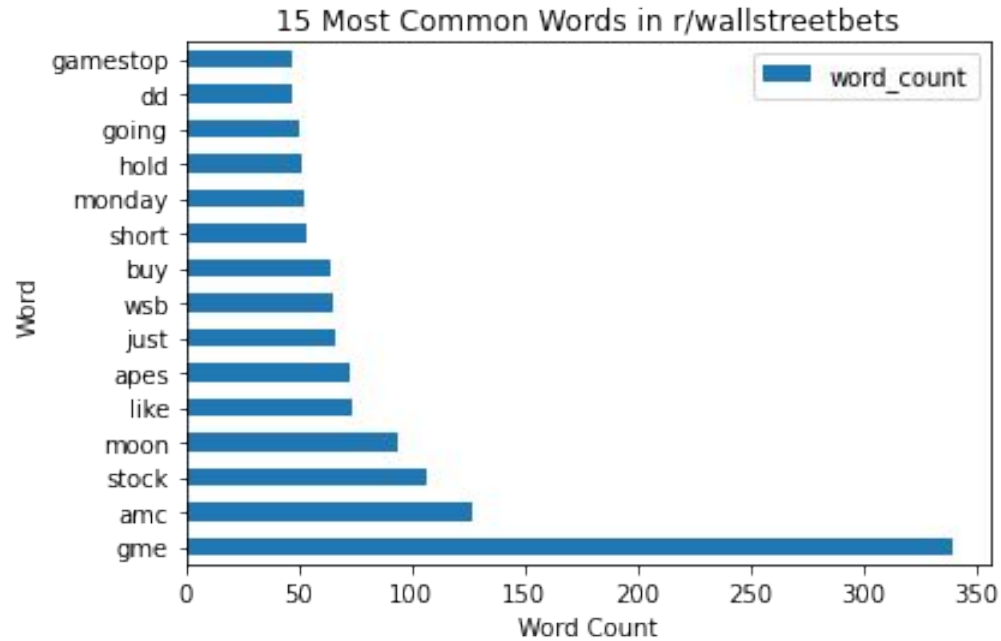
Word Count of Posts



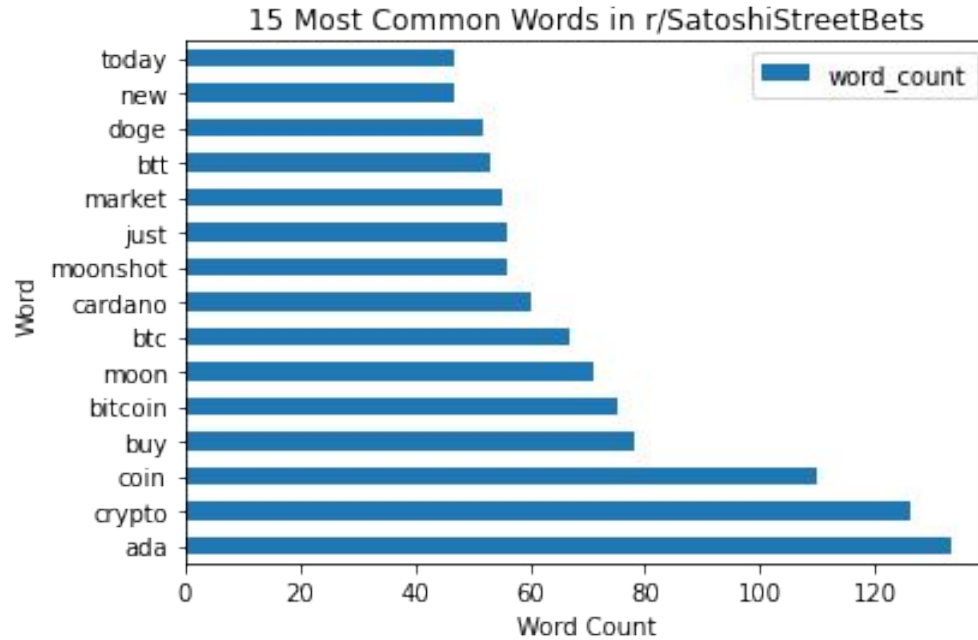
Distribution of Sentiment Compound Scores



Ranges and Outliers in Sentiment Compound Scores



15 Most Common Words in r/wallstreetbets

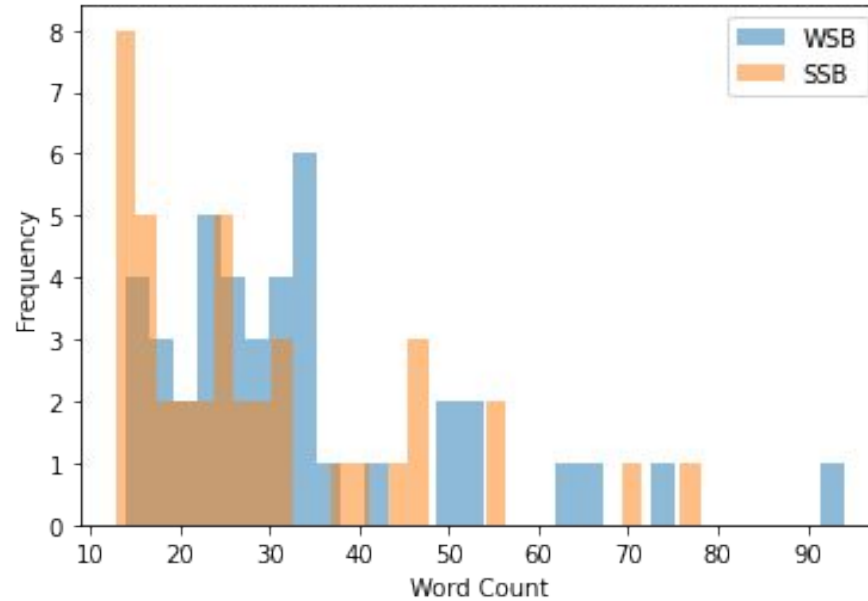


15 Most Common Words in r/SatoshiStreetBets

A word cloud of 41 common words shared between WSB and SSB. The words are arranged in a roughly circular pattern, with some words being larger and more prominent than others. The words include: day, help, 10, long, right, let, amp, know, bought, 2021, dip, short, like, today, think, big, time, yolo, buy, just, going, moon, monday, market, new, got, trading, make, hold, buying, money, did, don, good, people, price, guys, holding, way, ve, and news. The words are in various colors and sizes, with 'moon' and 'buy' being the largest.

41 Most Common Words Shared between WSB and SSB

Distribution of Most Common Words Shared between WSB and SSB



Most Common Words Shared between WSB and SSB



59 Unique, Most Common Words for r/SatoshiStreetBets

Modeling



Conclusion

r/wallstreetbets

stocks

r/SatoshiStreetBets

cryptocurrencies

80.85%

accuracy on unseen data

Thank you!