# Measuring Risk Tolerance - Crypto or Stocks?

**Seung Woo Choi**

# Disclaimer

# An Introduction to Two Subreddits



r/wallstreetbets

r/SatoshiStreetBets
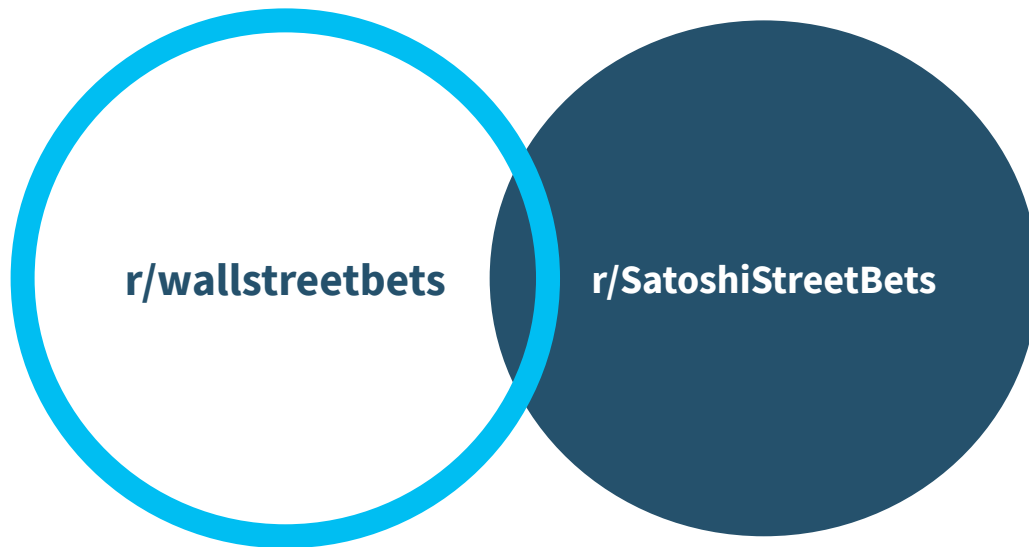
# Problem Statement

1. **Build a classification model with accuracy of 80%+**
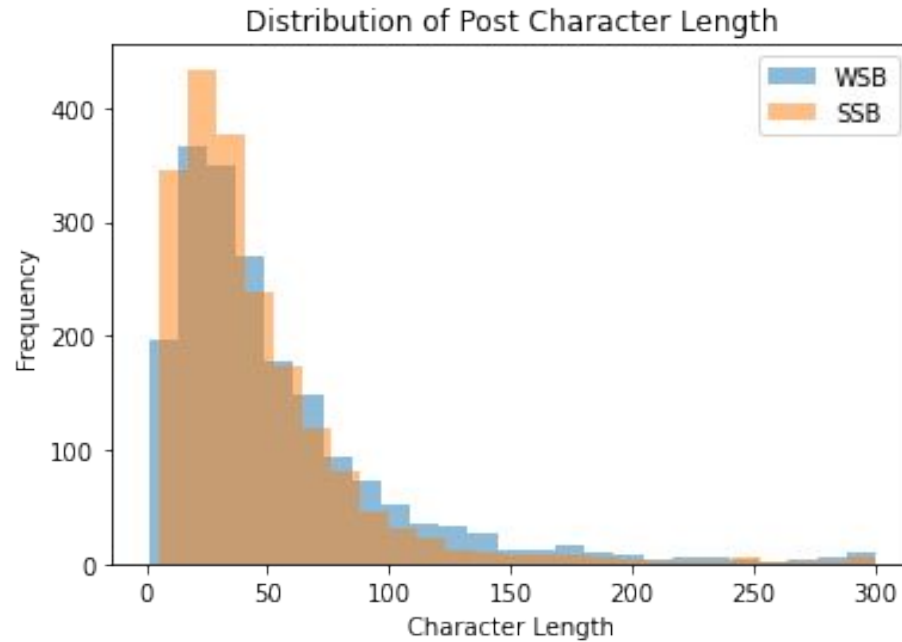2. **Determine words unique to two subreddits to understand risk profile**
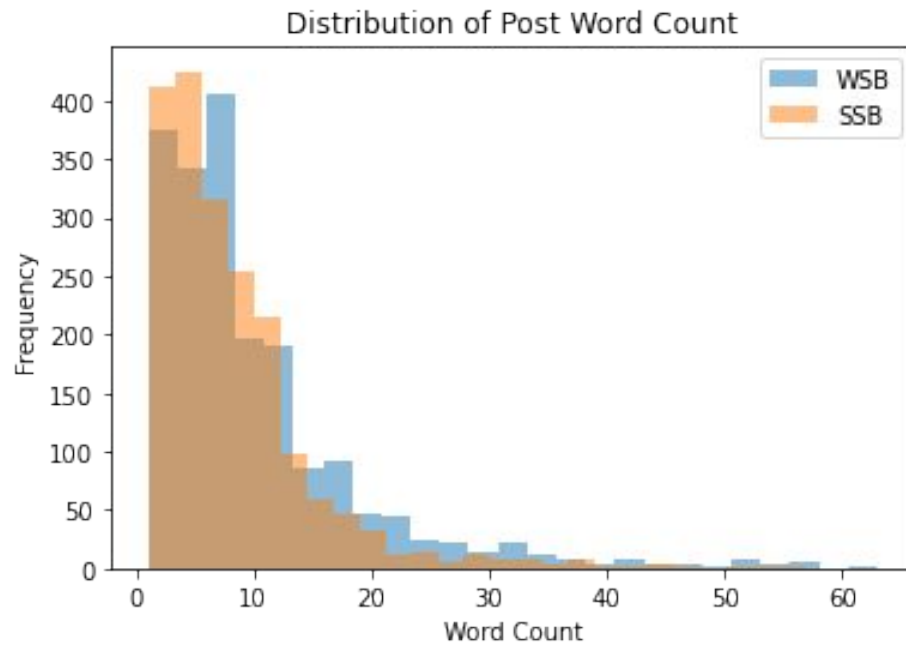
# Data Source

**Pushshift's web API**

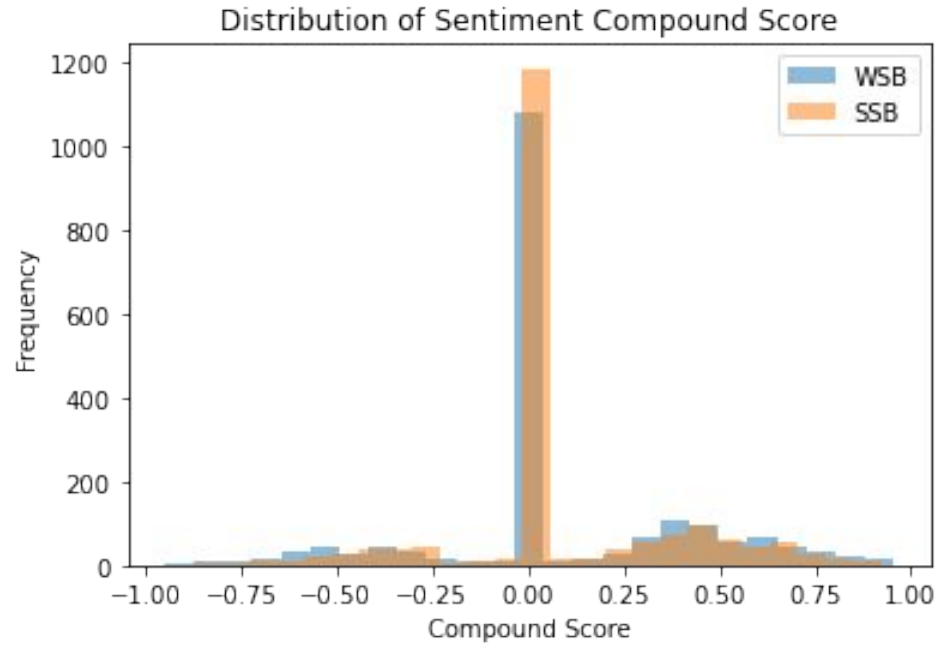**Monday, March 1, 2021
9:06:40 PM**

# Exploratory Data Analysis
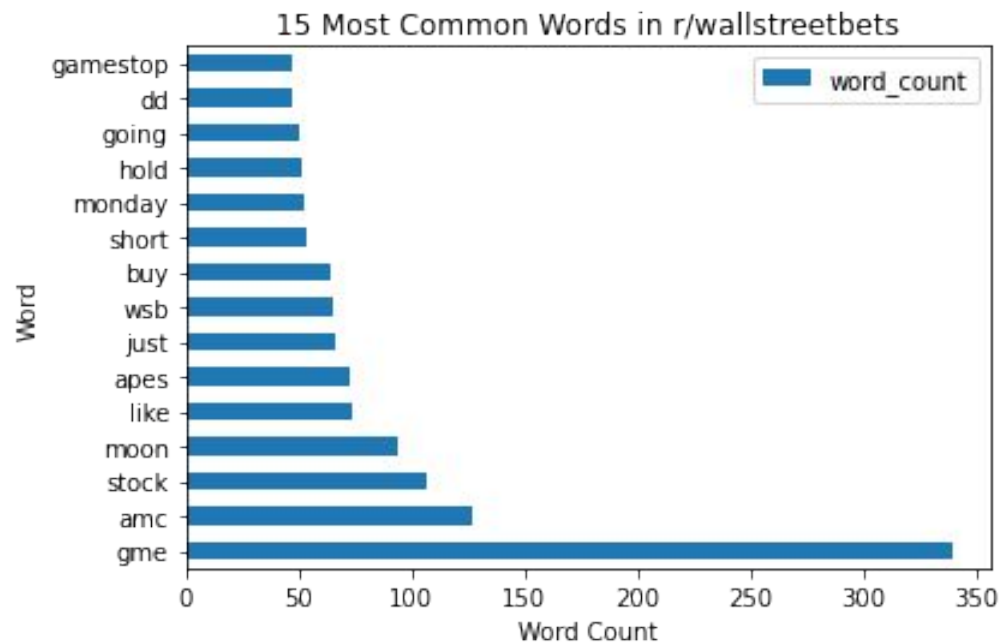
**Character Length of Posts**

**Word Count of Posts**

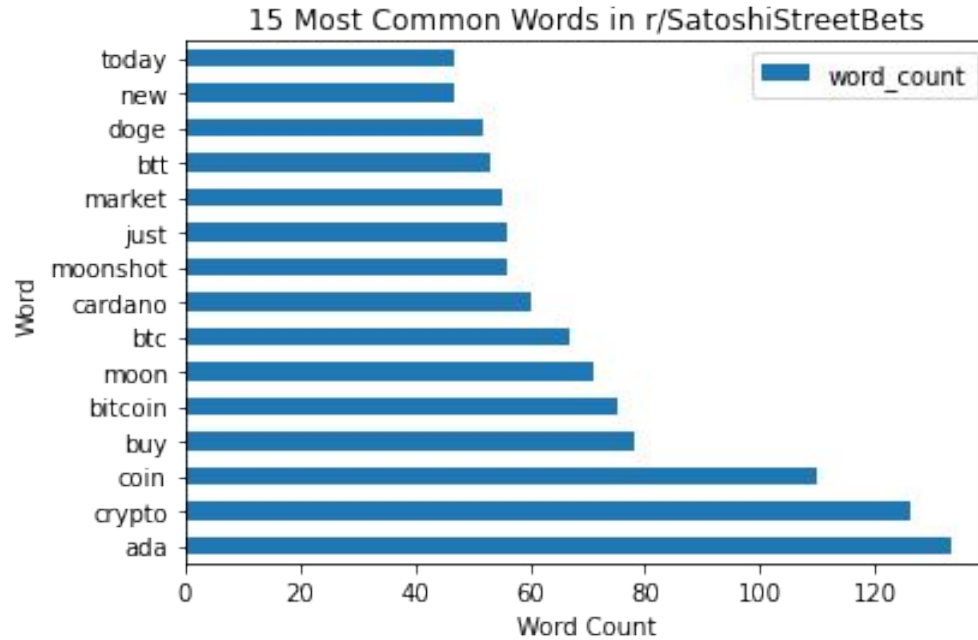Distribution of Sentiment Compound Scores

Boxplot grouped by subreddit
Sentiment Compound Scores
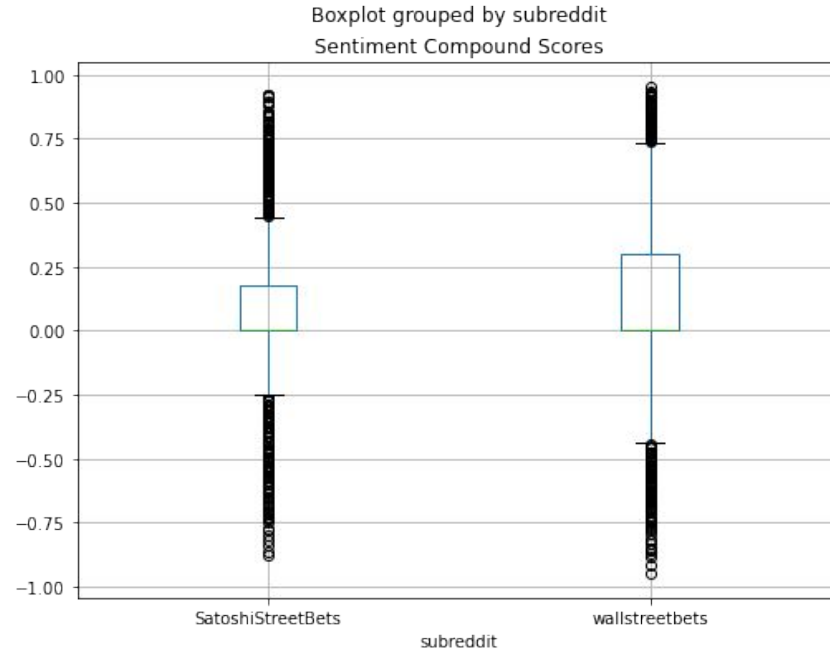
**Ranges and Outliers in Sentiment Compound Scores**

15 Most Common Words in r/wallstreetbets

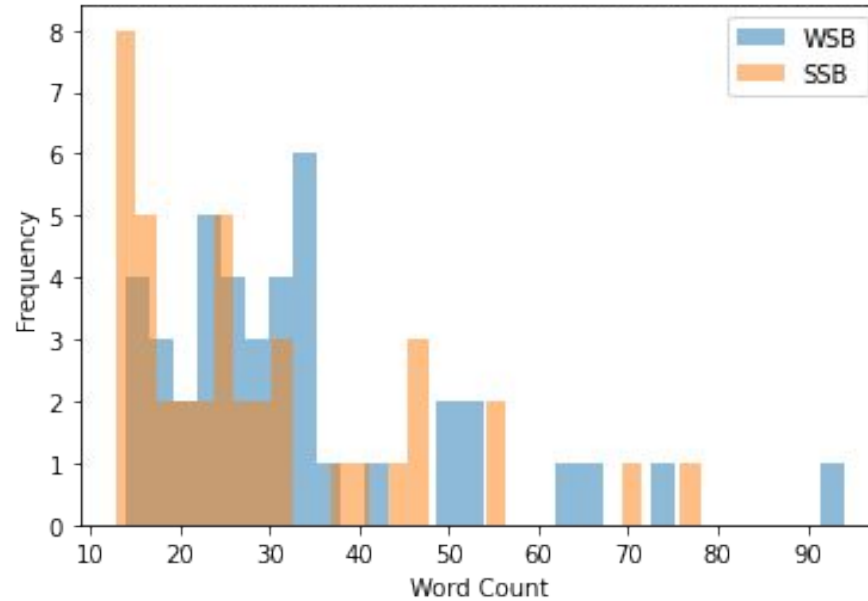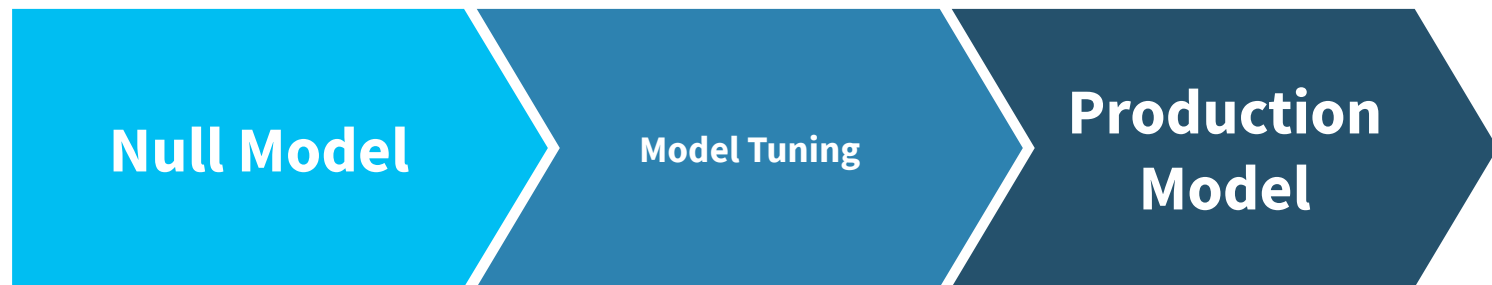15 Most Common Words in r/SatoshiStreetBets

**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/wallstreetbets**

**59 Unique, Most Common Words for r/SatoshiStreetBets**

# Modeling

**Null Model** → Model Tuning → **Production Model**

# Conclusion

# r/wallstreetbets
stocks

# r/SatoshiStreetBets
cryptocurrencies

# 81.69%
accuracy on unseen data

**Thank you!**