



Problem Statement:

Will FIFA officials benefit from combining soccer and world cup sub reddit into one.

By Jeremy Hart

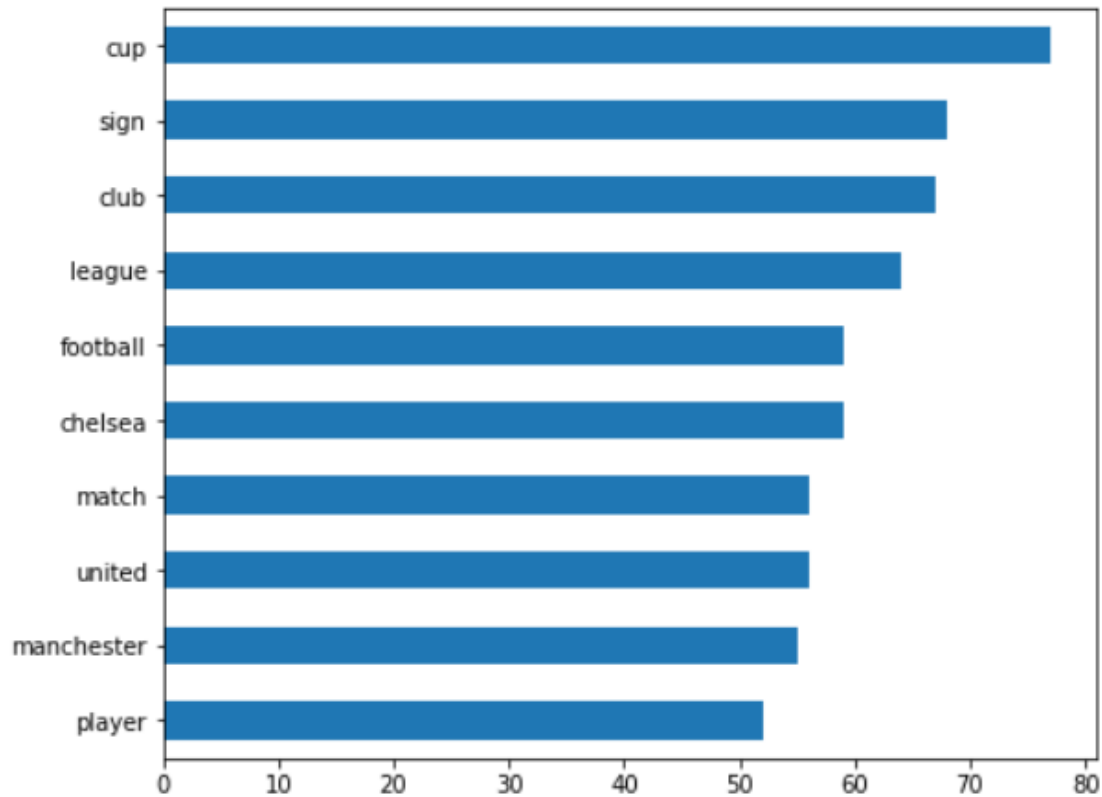
# Soccer Leagues vs World Cup

- Europe leagues have 40 professional leagues
- The 40 leagues have 1000 clubs in Europe
- The US top league the MLS has 28 clubs
- There are 32 teams in every World cup representing 32 different nations
- The World Cup happens every 4 years

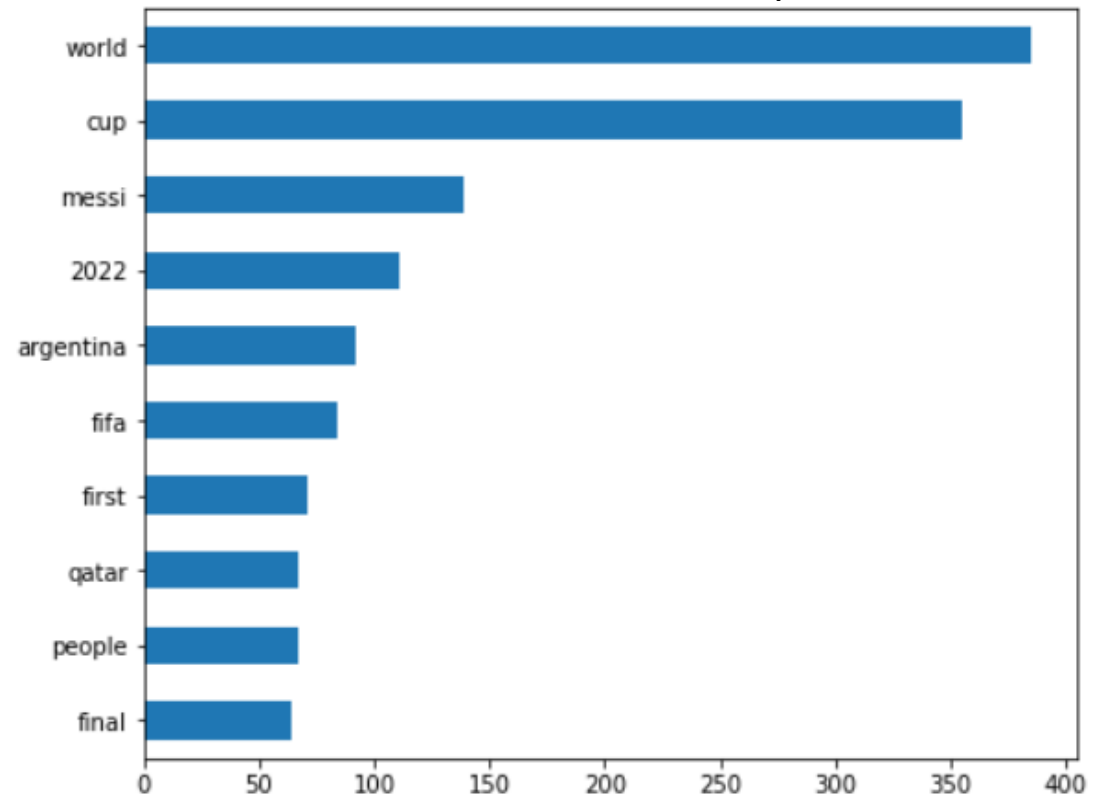


# Soccer and World Cup Top Words

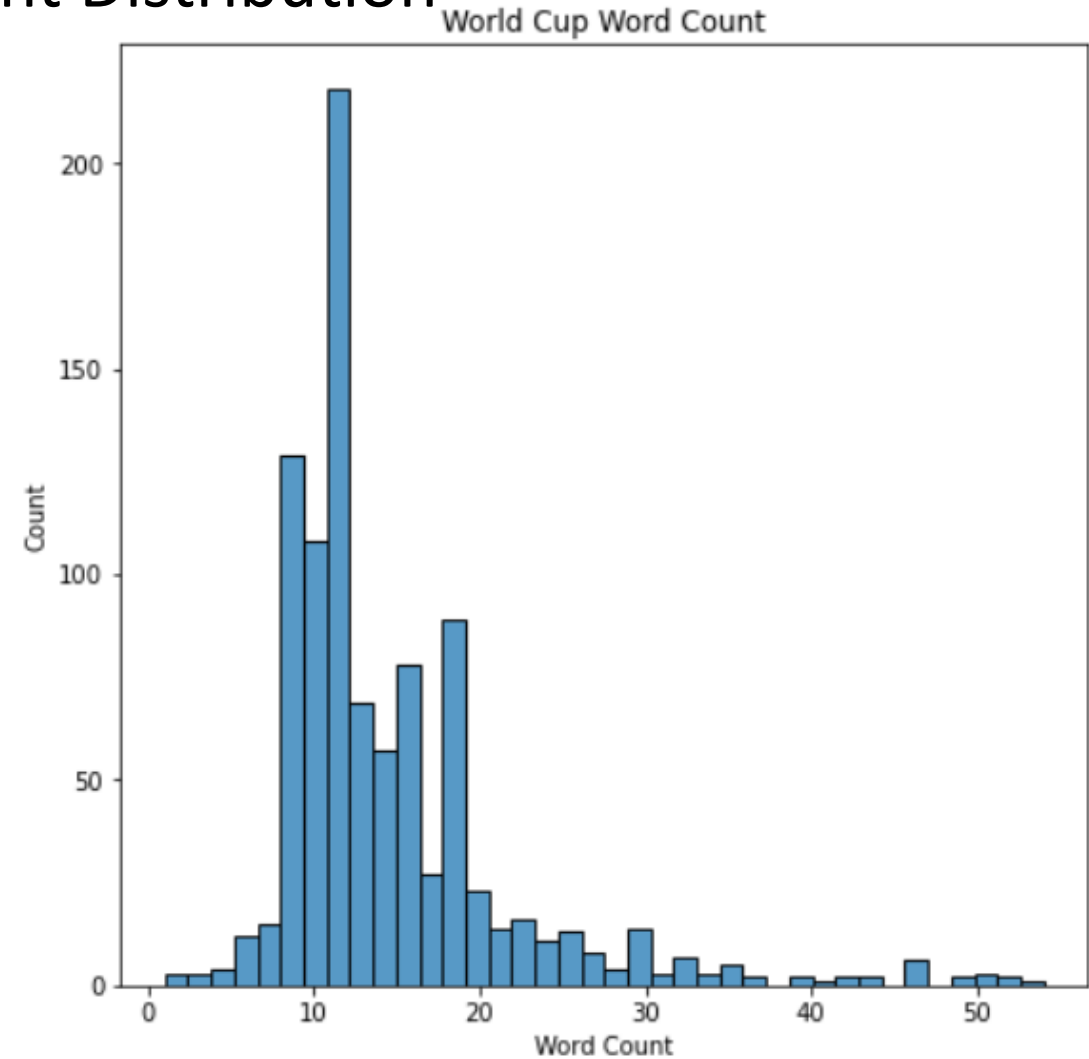
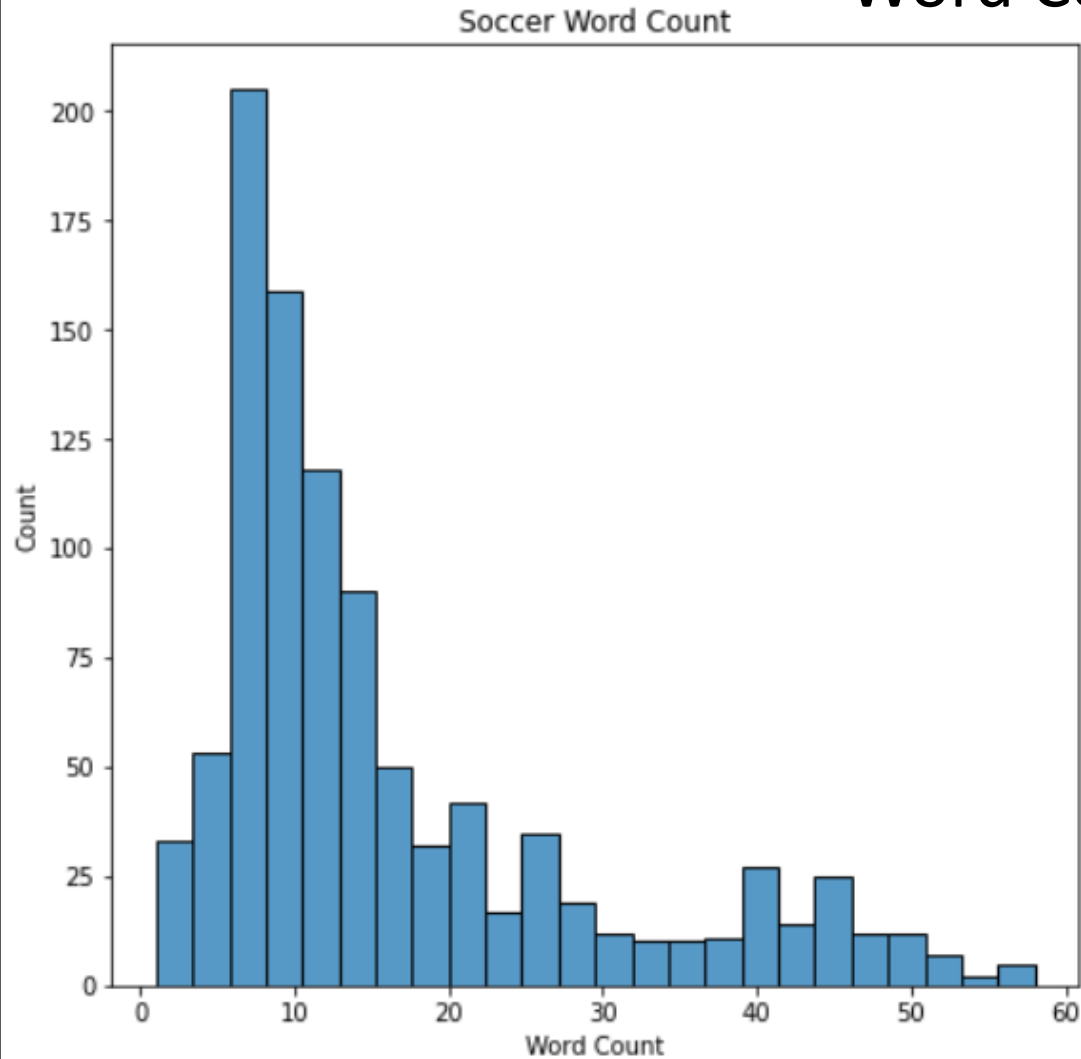
Soccer



World Cup



# Soccer and World Cup Word Count Distribution



# Analysis

---

- Three of the four models had accuracy above 80%
- KNN model performed the worst with a 63% accuracy score
- Random Forest performed the best with the highest accuracy and lowest misclassification rate

Models	Random Forest	Logistic Regression	K-Nearest-Neighbor	AdaBoost
Accuracy Score (Test)	87%	82%	63%	85%
Accuracy Score (Train)	96%	98%	76%	99%
Misclassification Rate (Test)	13%	18%	37%	15%
Misclassification Rate (Train)	3%	1.5%	23%	.4 %

# Conclusions

- The accuracy scores are not low enough to justify combining the subreddits
- FIFA officials could spend some funding towards moderators of the world cup subreddit to reduce clutter and spam
- Possibly making another subreddit exclusively for the world cup that year could make it more manageable
- Leveraging advertisements geared towards those who only tune into soccer during the world cup could be helpful.