

Introduction & Problem Statement

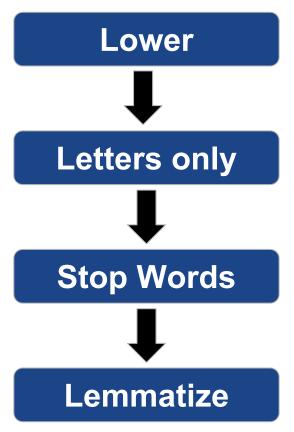
Someone at Reddit accidentaly deleted the conspiracy and climate skeptics subreddits.

Stakeholders:

- Reddit
- 839k r/conspiracy subscribers
- 11k r/climateskeptics subscribers



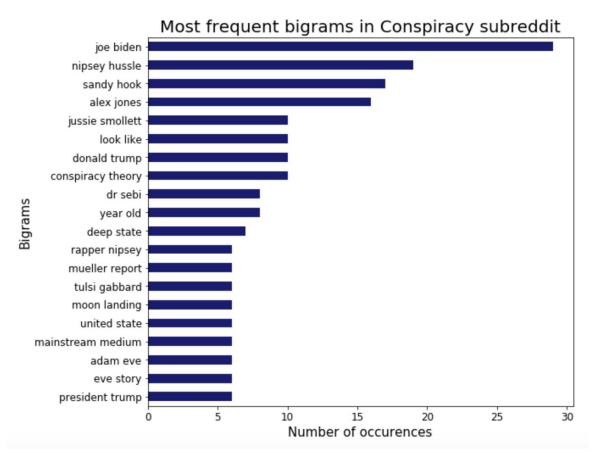
Cleaning and EDA



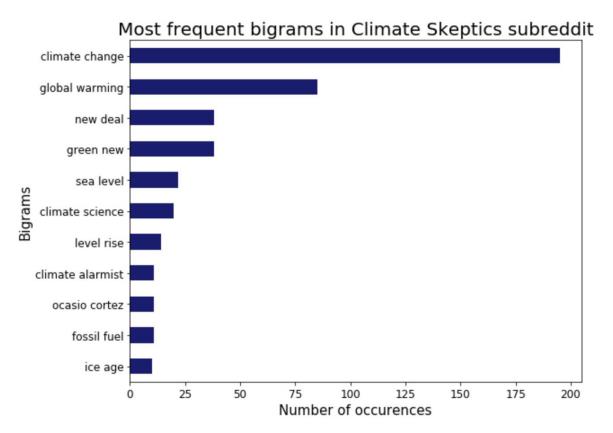
Word Count

Conspiracy	3517
Climate Skeptics	2809
Common	1153

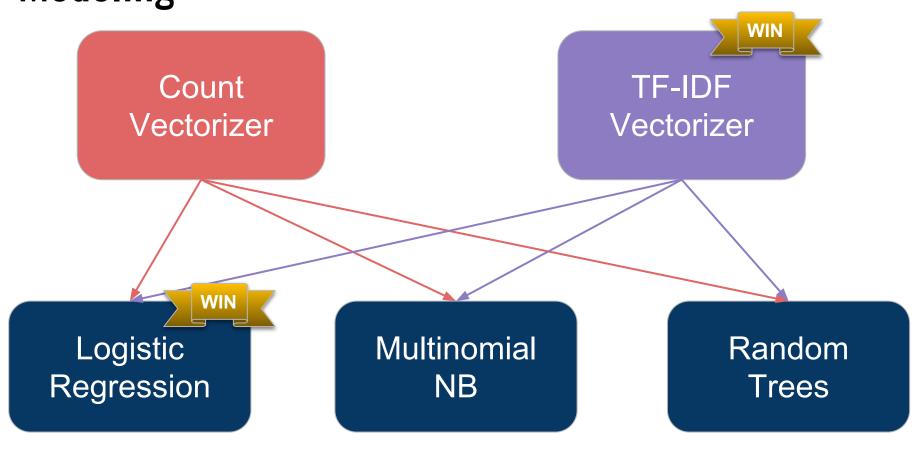
Cleaning and EDA

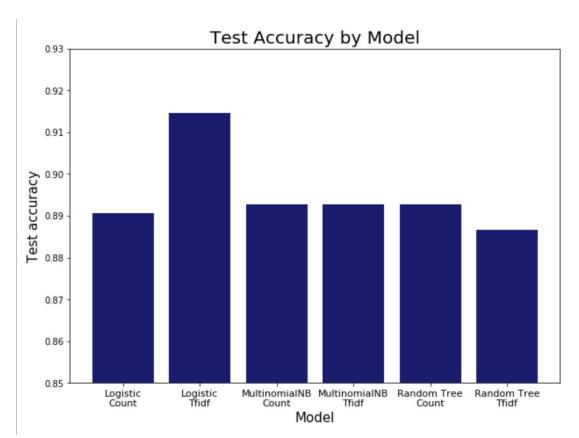


Cleaning and EDA



Modeling





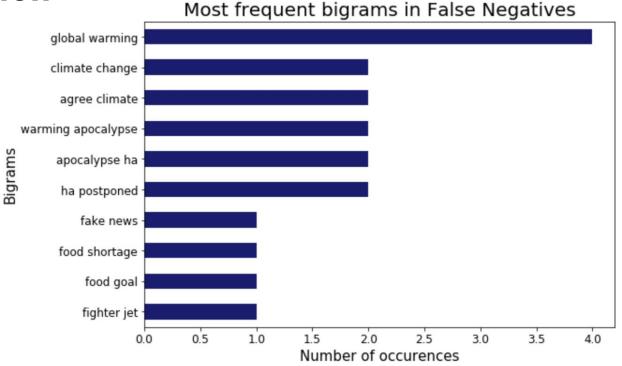
Baseline
Accuracy =
50.7%
By predicting
majority

Best Model: Logistic Regression with TF-IDF Vectorizer

Accuracy	91.5%
Recall	95.6%
Specificity	87.5%
Precision	88.1%

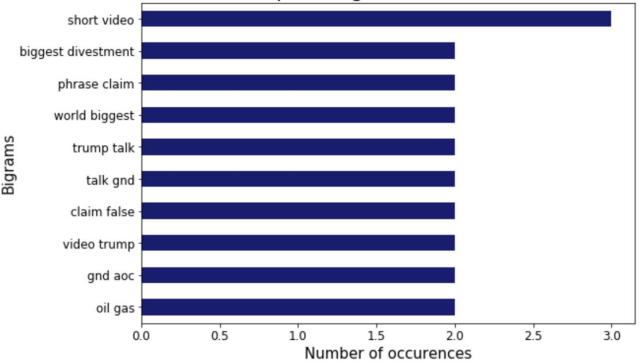
Confusion Matrix

	Actual Conspiracy	Actual Climate Skeptics
Predicted Conspiracy	206	28
Predicted Climate Skeptics	15	218



Posts that were predicted as Climate Change but are actually from Conspiracy





Posts that were predicted as Conspiracy but are actually from Climate Skeptics

Conclusion



r/climateskeptics

10.7k

75

Subscribers

Online



839k

2.5k

Subscribers

Online

We managed to successfully rebuild both subreddits with 91% accuracy!