Sentiment Analysis on Twitter

Josh Grimes

Business Problem

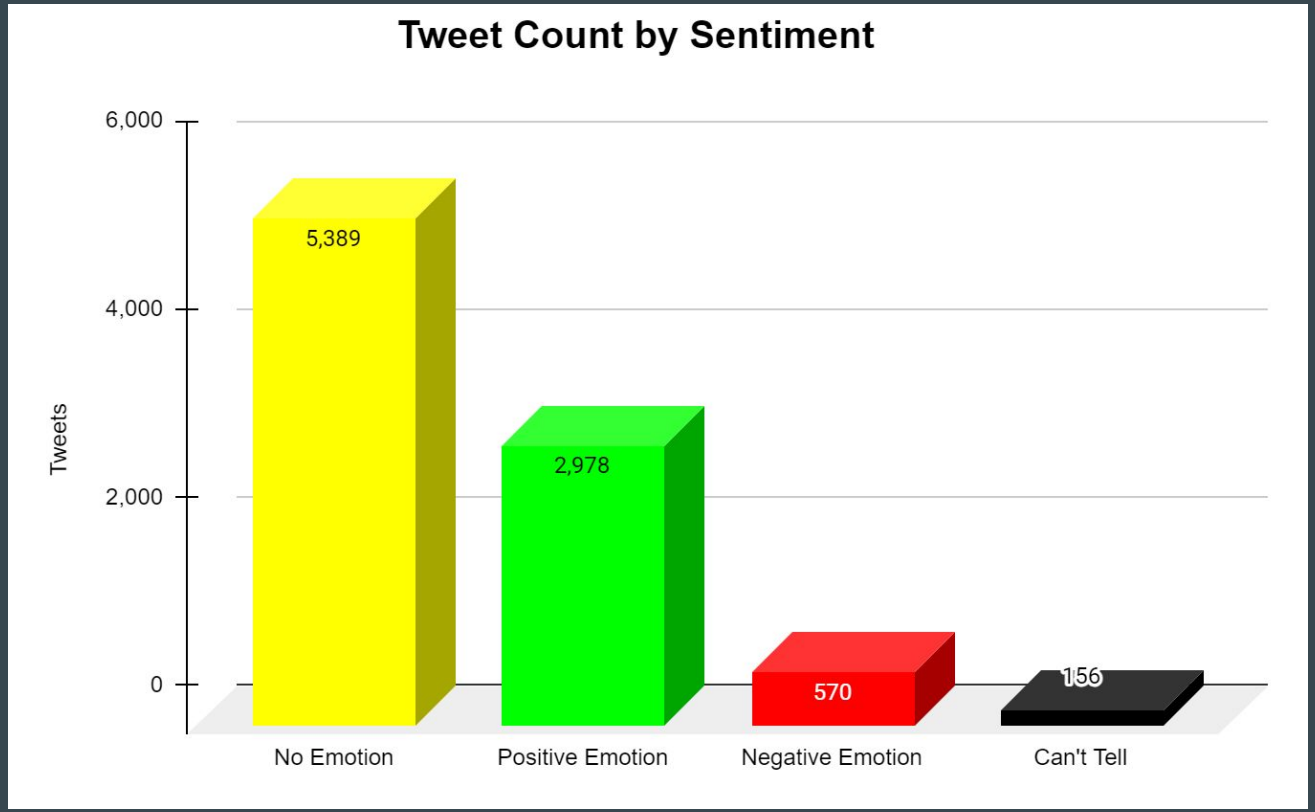
- 500 million tweets per day
- Limited Resources
- Valuable Insights
  - Customer Preferences
  - Needs
  - Expectations
- Customer Service

Goal

- Create a model that can predict a tweet's sentiment without a person having to look at it

Data

- 9,000 tweets
- CrowdFlower
  - Text
  - Product
  - Sentiment



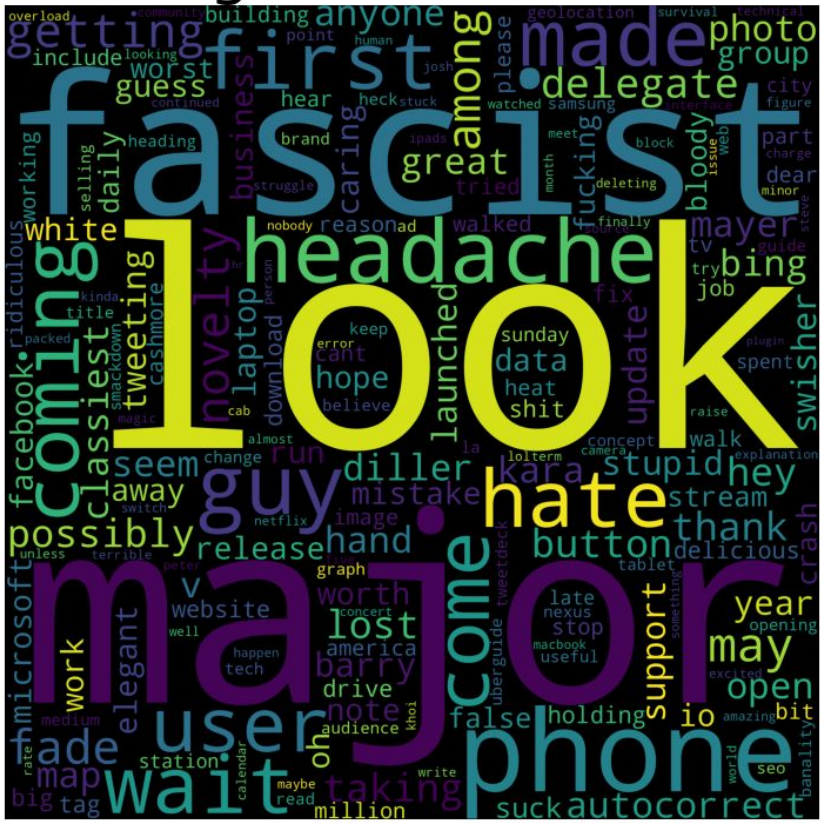
Bag of Words Model

Steps:

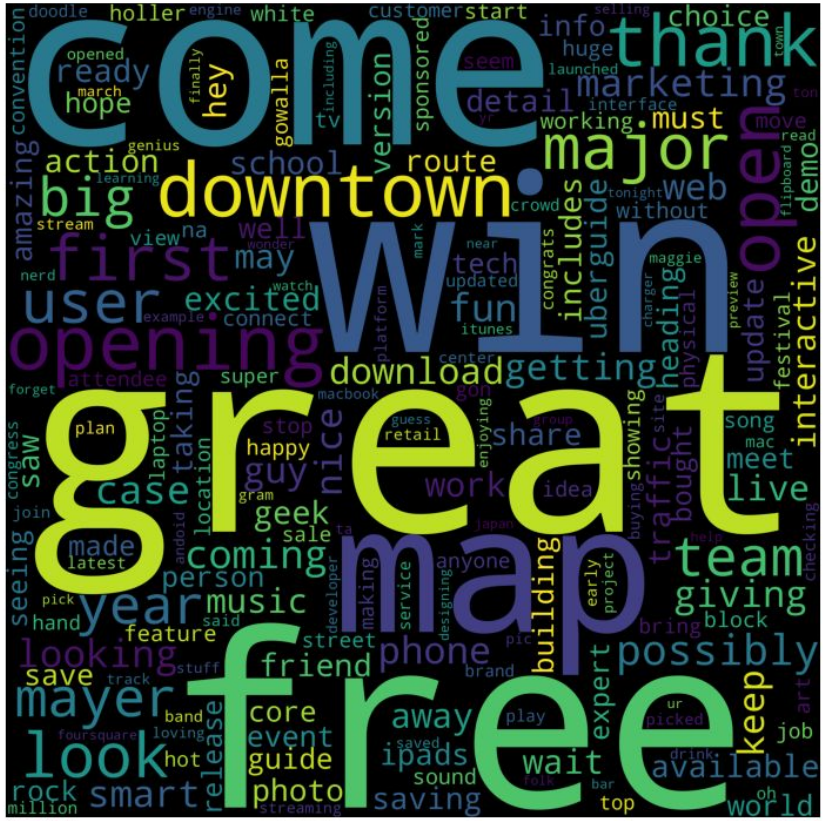
1. Group by sentiment
2. Normalize
3. Find Most Important
4. Turn them into Numbers
5. Put them in the Model

Positive vs. Negative Bags

Negative Emotion



Positive Emotion



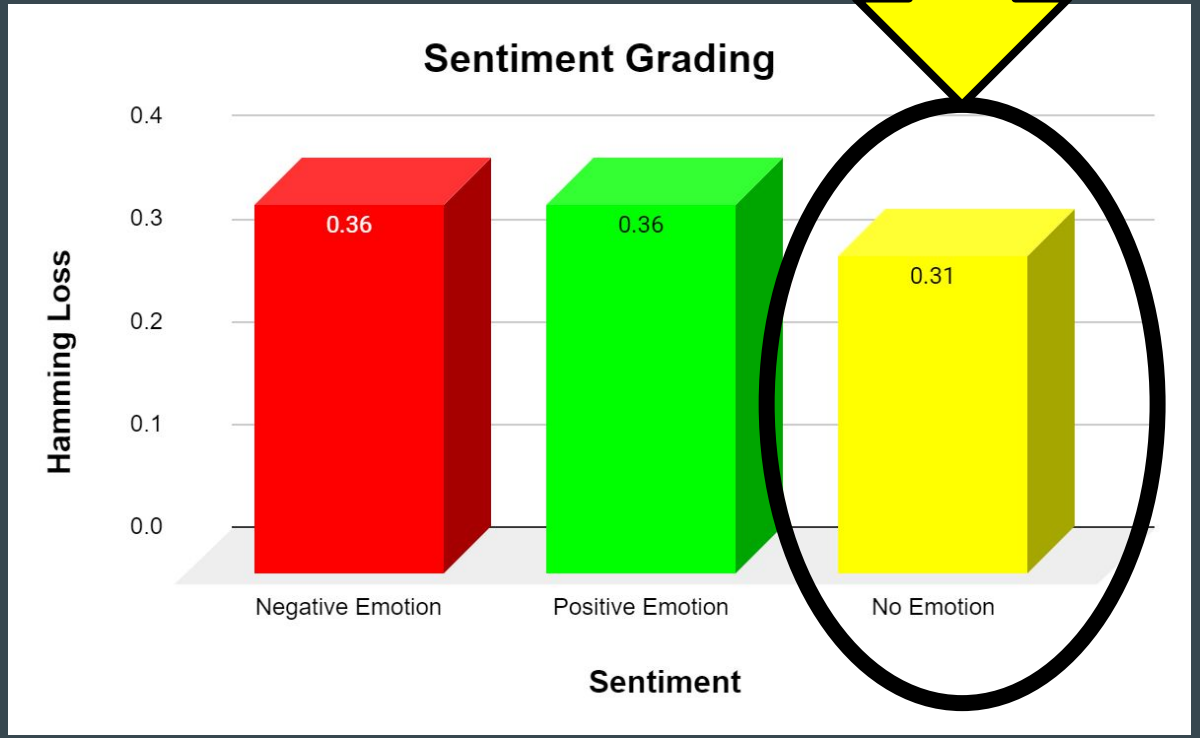
Model

- Turn words to numbers
- Predict the sentiment of tweet based on frequency of words

Tweet	fascist	hate	headache	great	thank	win
tweet1	1	1	1	0	0	0
tweet2	0	0	0	1	1	1

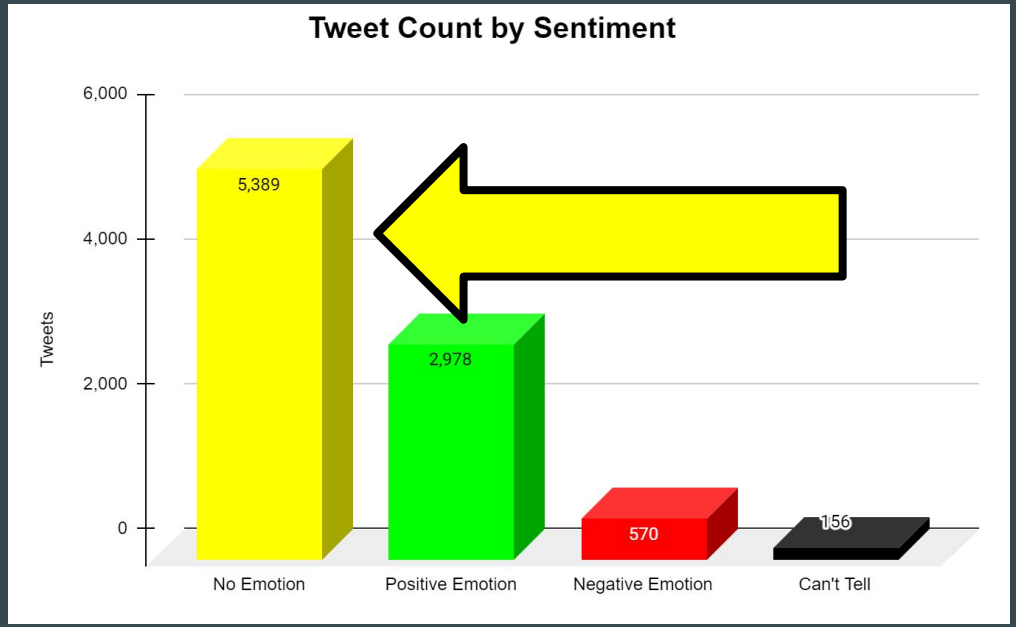
Results

- Best at predicting no emotion
- Lower = Better



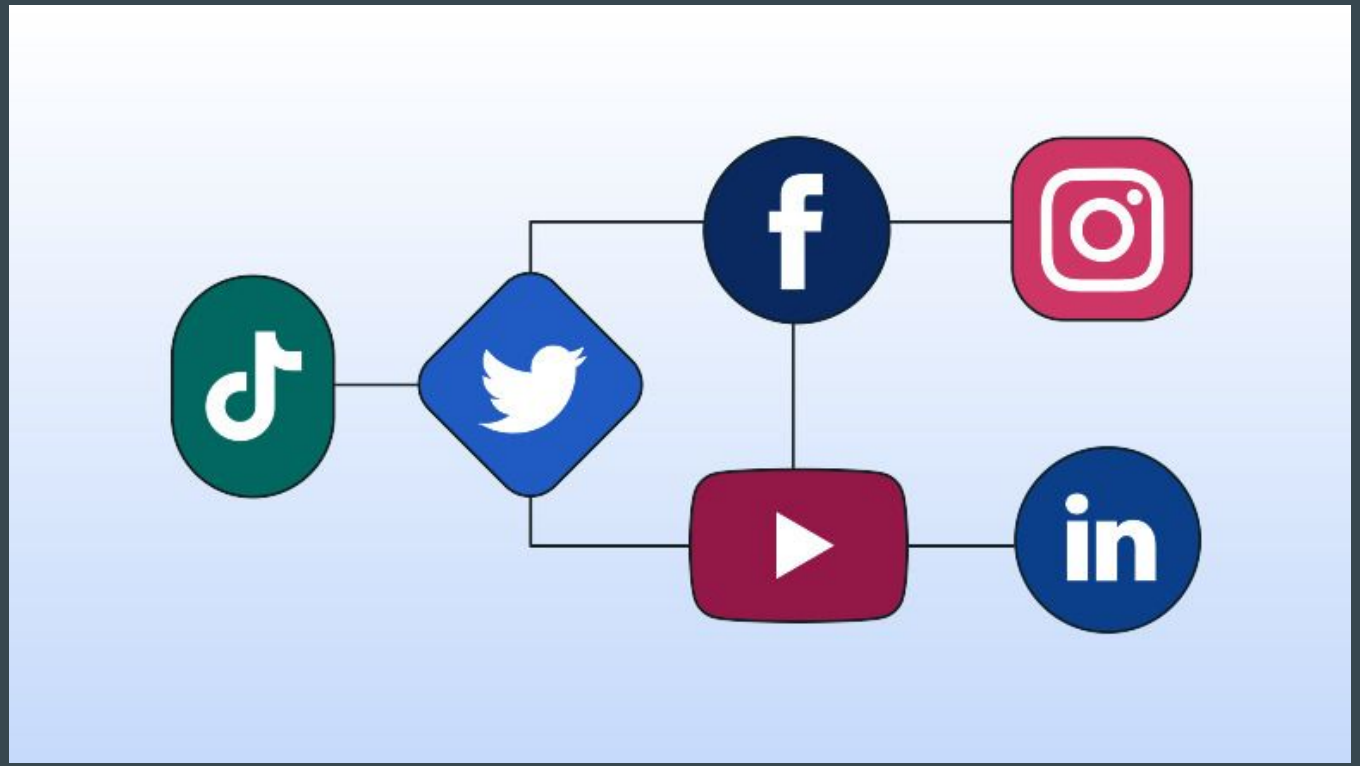
Conclusions

- The emotion behind a tweet is predictable by the text
- Model is best at predicting a lack of emotion
  - Can Ignore these tweets
  - Positive Emotion = Useful Feedback
  - Negative Emotion = Customer Service

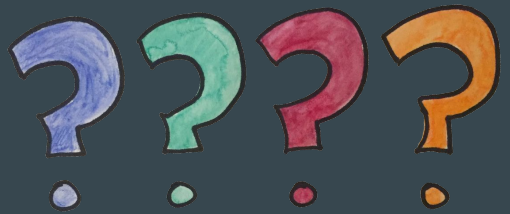


Next Steps

- More Data
- Add Features
- Other Platforms



Thank You




JOSHUA GRIMES
Data Scientist
919-273-1520
josh.g.grimes@gmail.com
Fayetteville, North Carolina
https://www.linkedin.com/in/joshua-g-grimes/