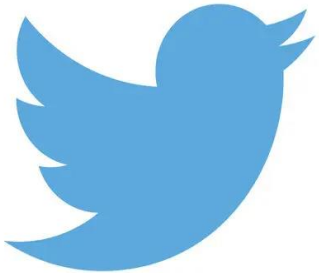
CS226 Project:

# Fast Food Brand Appeal Using Twitter and Sentiment Analysis

Anish Sekar, Brandon Herrera, Justin Schopick, Yogesh Singh

# The Goal




TACO
BELL


Wendy's


BURGER
KING


SUBWAY


KFC


IN·N·OUT
BURGER

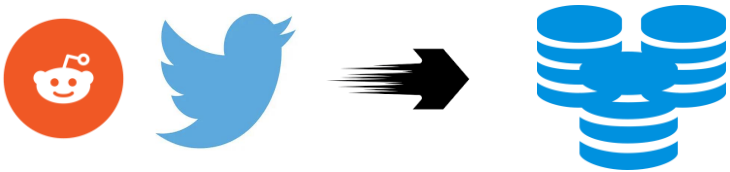

Pizza
Hut

# Motivation

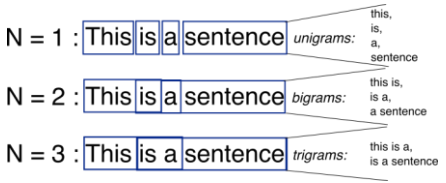
- ~ 4 Terabytes of data generated daily
- Unmoderated data, true opinions
- Hashtags and keyword based filtering makes it easier to separate required data
- Free public API
- Most popular microblogging platform

# Sentiment/Sarcasm model training

- To predict the sentiment of the tweets we trained a model from an existing dataset called Sentiment140.
- The dataset contains 1.6 million tweets that are rated positive or negative
- Sarcasm prediction is trained on a reddit dataset



- We used existing Spark machine learning libraries in Python to train the model.
- We used N-grams and term frequency inverse document frequency (tf-idf) for features.
- The tokenizer removed punctuation and other symbols.

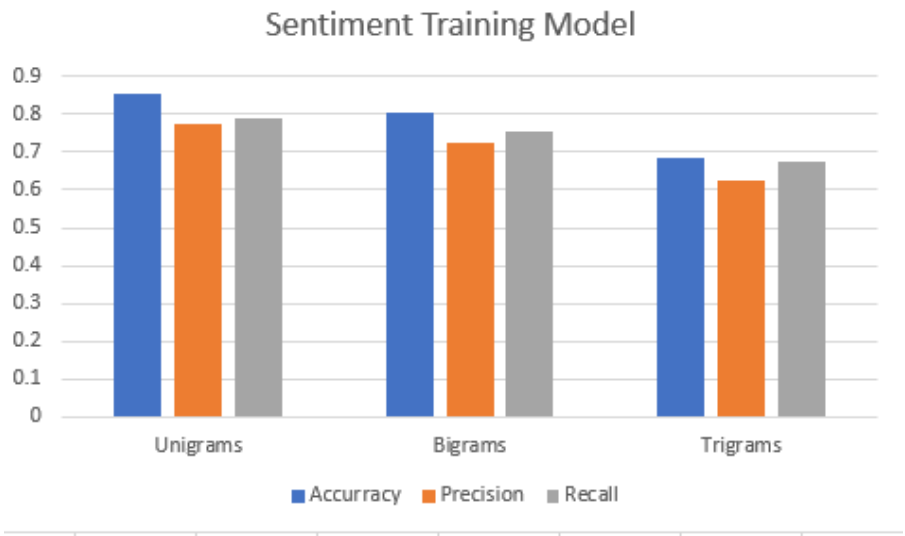
# Datasets – Sentiment Prediction

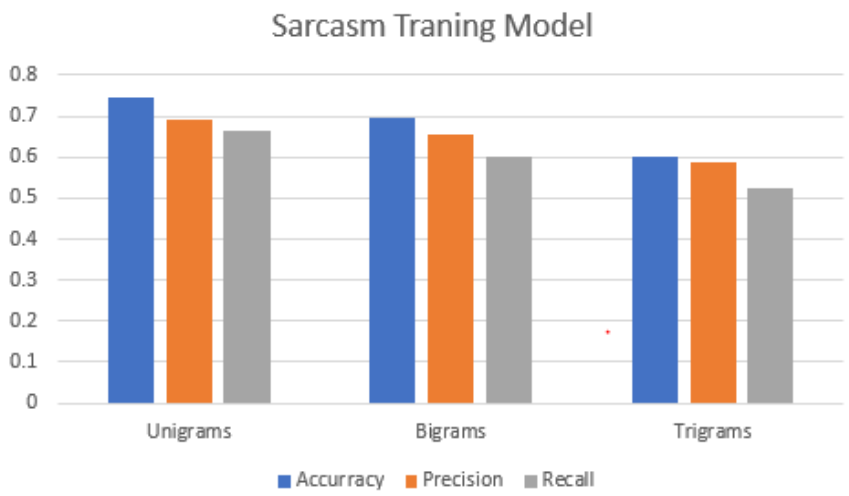
- We gathered around 60GB of tweets over a span of 2 weeks
- We used different combinations of keyword filters
- We use geolocation filters to restrict tweets originating from the united states

# Challenges



# Datasets


Sentiment Training Model
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Unigrams
Bigrams
Trigrams
Accurracy
Precision
Recall


Sarcasm Traning Model
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Unigrams
Bigrams
Trigrams
Accurracy
Precision
Recall

# Visualization