
Anu Adeboje

Oakland, CA 94619

Metis Online Flex

Beyond Meat & KFC: Tweet Impressions

Abstract

The goal of this project was to create a process or application that could quickly describe the major topics or themes and summarize sentiments in tweets reacting to the launch of Beyond Meat & KFC's plant-based chicken nuggets. This type of real time real world feedback informs stakeholders with the early opinions and judgements of the new product that might be difficult to get in a market research setting.

Tweepy was used initially with the API v1 call methods with variable success, later refinement of the access code yielded better results. Unfortunately the volume of the tweets had greatly diminished by the time of retrieval resulting in a corpus of roughly 3,000 tweets (launch related) before elimination of duplicates. A reference corpus of tweets retrieved with the Beyond Meat as the search terms resulted in a corpus of 2,812 tweets.

Initial sentiment analysis with Vader on the corpus revealed that news outlets and individual social media influencer types dominated the "twitter-sphere" with neutral content, at roughly 40% of the sampled tweets. Visualizing Vader analysis with WordClouds gave some insights, however using scapy's parts of speech tag identified the top ten most common adjectives used in positive, negative, and neutral tweets. This level of analysis revealed the frequent mention of competitor Impossible Foods, along with incorrectly tagged tweets by Vader due to mischaracterizations of words like fire and slap as negative or inedible as neutral.

A larger corpus of tweets might have resulted in more distinct topics, or made possible a topic analysis on tweets by sentiment. After several iterations and hyperparameter tuning the topics found through modeling with highest explained variance ratio mirrored the sentiment analysis in their themes which included positive and negative preferences or news reporting. Ultimately topic modeling via Corex on the Beyond Meat keyword tweets did yield stratified topics, one of which being the new product launch.

Retrieving more tweet data, the utilization of newer NLP's that recognize slang (Equilid), and topic modeling on subsets (achieved by sentiment analysis scores) of the corpus are all improvements that I would make to the pipeline. Wrapping the process in an application would enable a brand or any interested party with the ability to have a high level perspective in response to an event or product launch and respond accordingly.

- Tweet queries were built iteratively to insure only relevant tweets were included, order of words were preserved when necessary, see example queries below:
 - "Beyond+Meat+kfc -is:retweet" ,
 - "Beyond+Meat+chicken -filter:retweets",
 - beyond+fried+chicken OR beyond+plant-based OR beyond+nuggets -filter:retweets"
 - "Beyond Meat"* -filter:retweets"
- Tweets were compiled starting January 4th, 2022
- Analysis aims to included roughly 5,000 tweets, unfortunately rate max and code errors resulted in a smaller than desired corpus

Tools:

- **Tweepy** along with the **Twitter developer API** is used to gather tweets
- **NLTK Vader**, for sentiment analysis
- **Scapy**
 - **Pipeline**
 - **Parts of Speech tagging**
- **LSA, NMF, vs Corex**, for topic modeling

Communication:



Word Cloud from all tweets related to launch



WordCloud Negative Tweets

