

Sentiment Analysis of Tweets

Unlocking the Power of Public Opinion Through Advanced ML

BAX 493B Topics in Business Analytics - Implementing Machine Learning on the Cloud
Team Members: Jodie Chen, Blessy Chinthapalli, Mrimon (Nemo) Guha , Yajur Sehra, Jie Zhu

Agenda

1	Introduction	0.5 min
2	Data Source	1 min
3	Model Methodology	1 min
4	Model Accuracy Demonstration	2 min
5	Competitive Edge	1 min
6	Revenue Model	1 min
7	Market Opportunity and Investor Appeal	1 min
8	Conclusion	0.5 min

Introduction

What is our company's product / service?

Challenge

In the digital age, understanding real-time public sentiment is crucial yet challenging due to the sheer volume of data. Businesses, politicians, and creators need automated, precise tools to gauge public opinion instantly.

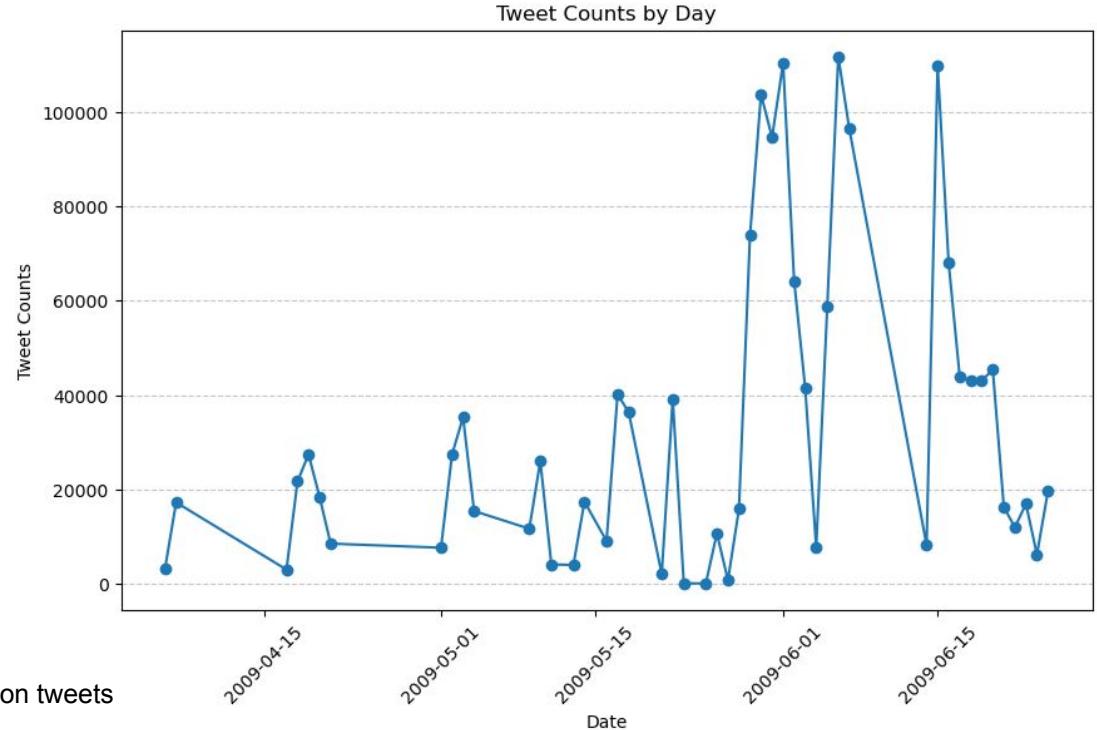
Our Solution

- Overview - Utilizing ML to analyze 1.6 million tweets, providing real-time sentiment analysis to identify positive, negative, and neutral sentiments.
- Technology Stack - Leveraging NLTK for NLP tasks, VADER lexicon for sentiment scoring, and Google Cloud Platform for scalable, secure data handling.
- Features - Real-time insights, high accuracy models, and easy API integration.

Data Source

What ML model and data was used for insights? Is there an alternative to ML?

We analyzed tweets distribution over time, focusing on high-activity periods and predominant sentiments.



* Data source (from Kaggle): Sentiment140 dataset with 1.6 million tweets

* <https://www.kaggle.com/datasets/kazanov/sentiment140>

Model Methodology



What ML model and data was used for insights? Is there an alternative to ML?



Technique: We use NLTK for NLP tasks, VADER lexicon for sentiment scoring, and Google Cloud Platform for scalable, secure data handling.



Edge: ML is the best way to handle such large scale data and GCP helps us easily store and analyze the data without worrying about resourcing or scaling existing hardware resources of any client.



Actionable Insights: Our product will generate example insights like "**80% positive sentiment about upcoming weather**" or "**10% dislike new tax rule,**" showcasing the model's potential impact.

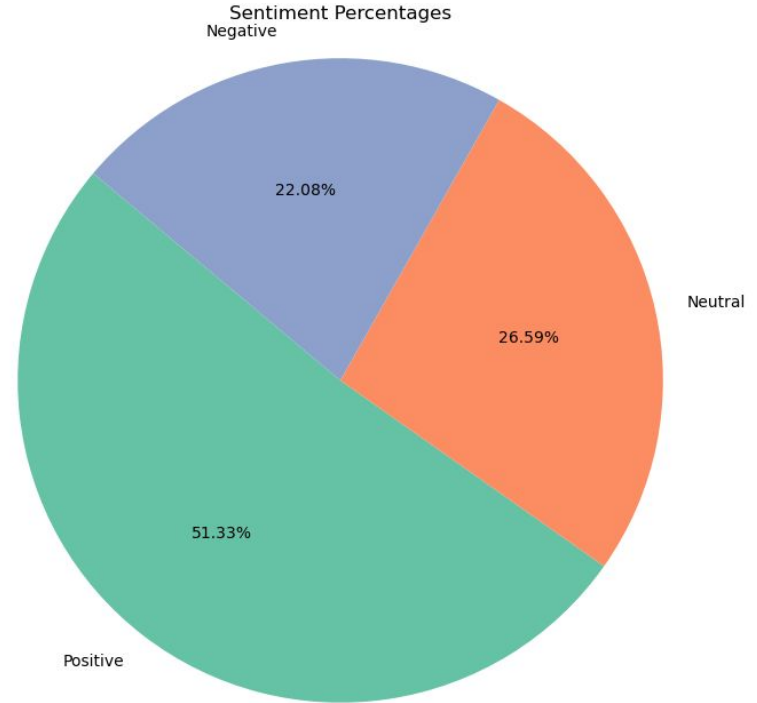
Model Accuracy Demonstration - 1

What ML model and data was used for insights? Is there an alternative to ML?

Among all the sentiments:

51.33% positive / 26.59% neutral / 22.08% negative

These percentages indicate the distribution of sentiment classifications. Positive sentiments appear to be the most common, followed by neutral sentiments, with negative sentiments being the least common among the data.



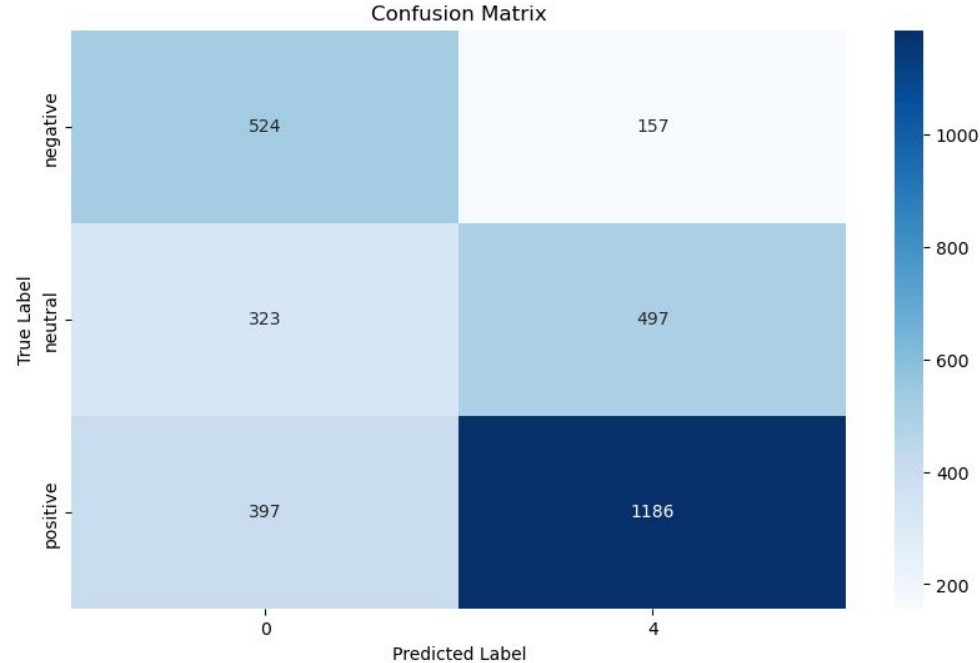
Model Accuracy Demonstration - 2

What ML model and data was used for insights? Is there an alternative to ML?

For the target class 0, instances are:
524 negative / 323 neutral / 397 positive

For the target class 4, instances are:
157 negative / 497 neutral / 1186 positive

The confusion matrix shows how well the sentiment analysis model performed for each sentiment class across different target classes. This indicates that our model is particularly effective at identifying positive sentiments.



Competitive Edge



What sets your company apart from the competition?

- **Ease of use:** Storing the code and using Google Colab to connect with Cloud helps anyone in the team to just run the notebook as desired
- **Data security:** we have used the data in the most secure way by using Access Control in Google Cloud (IAM)
- **Ability to handle large scale data:** Our company can handle large scale data without any delays or latency issues.
- **Continuous improvements:** Our company would improve the model frequently by rebalancing the data, conducting error analysis and implementing advanced techniques to make the model more effective. Thus, it could increase the customer satisfactions overall.

Revenue Model

How the product will earn revenue and from whom? What is the path to profitability?

- **Primary Customers:** Product-based companies (e.g., skincare, smartphones) seeking to understand market sentiments.
- **Revenue Model:** Subscription-based API access, premium features for advanced analytics.
- **Path to Profitability:** Focus on service excellence, continuous innovation, expanding client base, and establishing long-term partnerships.



Market Opportunity and Investor Appeal

What is the opportunity for an investor?

- High growth potential in a booming market
- Scalable model with cloud-based technology
- Diverse client base reduces risk and facilitates cross-selling
- Data security and compliance as key differentiators



Conclusions

- **Advanced Sentiment Analysis:** Leverages machine learning and cloud technologies to analyze millions of tweets for deep user sentiment understanding.
- **High Accuracy Model:** Successfully built to provide real-time, actionable insights with scalable deployment options.
- **Stable Revenue Source:** Subscription-based API access tailored for product-based companies, ensuring continuous revenue.
- **Data Security and Compliance:** Prioritizes creating a reliable and ethical ecosystem for cloud-based solutions, establishing trust and integrity.
- **Trusted Partner:** Positions our company as a dependable provider in sentiment analysis, empowering clients to make informed, data-driven decisions.



Thank You !