



# Sentiment Analysis

## With Logistic Regression

By Nechama Borisute

# Business Problem

Challenge: Lack of Real-time Insight into Public Sentiment on Social Media Platforms



- Customer Perception
- Emerging Trends
- Marketing Strategy



Solution: Real time tweet sentiment analysis

# Data Source

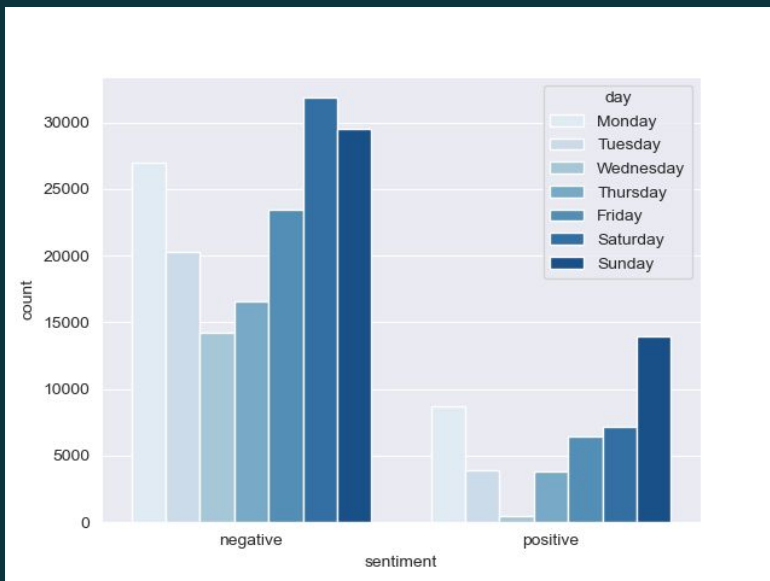
The dataset is from **kaggle**

There are **1,600,000** tweets in the dataset.



The data is from the year **2009** in the months April and May.

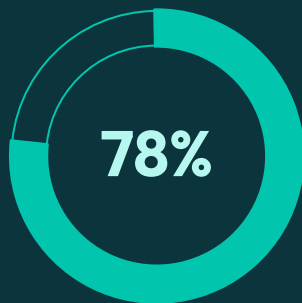
# Tweet Trends



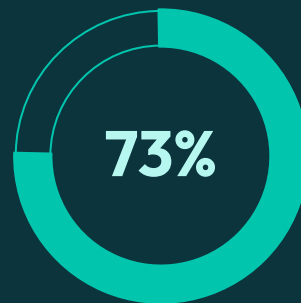
- Tweets are more negative in general.
- Negative tweets hit the roof on Saturday.
- Positive tweets are at the highest on Sundays
- On Wednesday people are DONE with positivity

# Model Flow

Metric = Accuracy

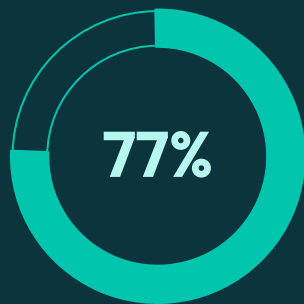


MultinomialNB

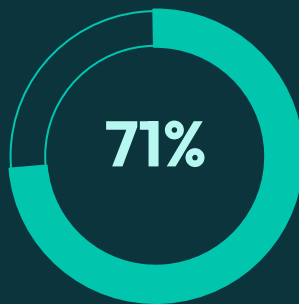


Random Forests

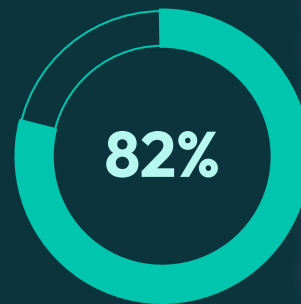
Logistic Regression



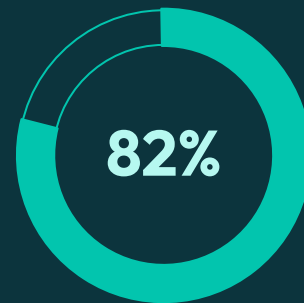
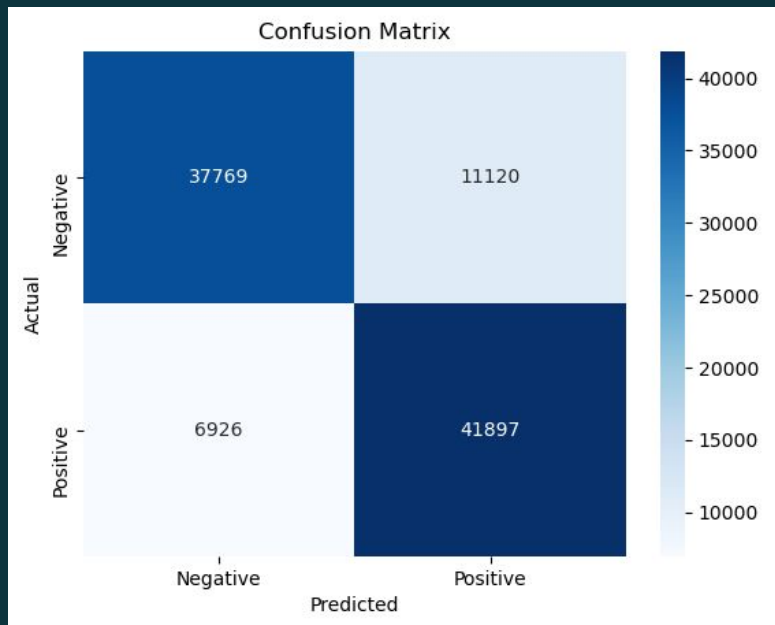
BernoulliNB



XGBoost



# Logistic Regression Model



- Best performing model with 82% accuracy.

# Recommendations

Use this model to:

- Detect emerging trends
- Analyze public opinion towards a brand/product
  - Understand customer perception
  - Inform marketing strategies

# Next Steps

**Develop  
Deployment  
plan**

**Implement  
APIs and  
Integration**

**Finalize  
Documentation**



The slide features a dark teal background. In each of the four corners, there is a vertical column of small, light teal squares. The top-left column has 10 squares, the top-right has 7, the bottom-left has 15, and the bottom-right has 14.

# Thank You

ANY QUESTIONS?