Tweets on Google and Apple Products NLP Classifier

By: Konnor Clark

- **Classify Twitter sentiment**

- **Gauging the negative or positive feedback**

- **Categorize positive or negative**?

- **Terms and hashtags** of **positive or negative sentiment**?

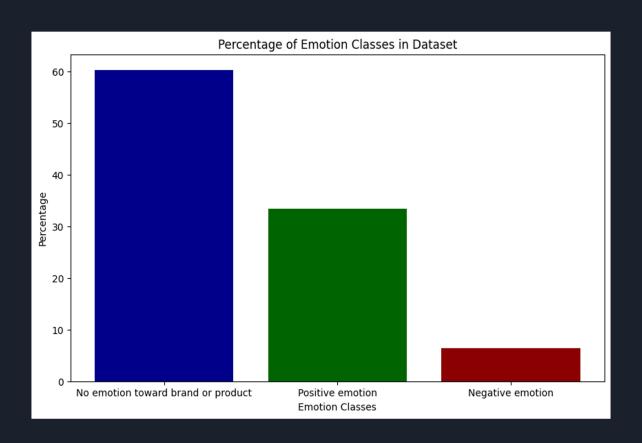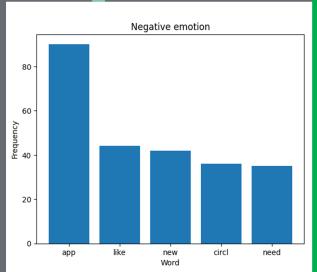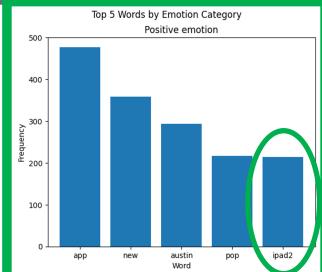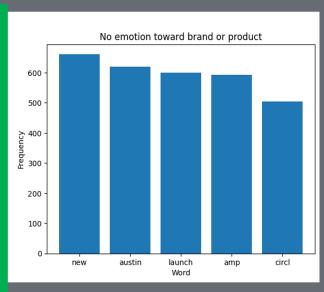# Data Understanding and Analysis: 3 Features

Text

Emotion

Product

Data Understanding and Analysis: Class Imbalance

Percentage of Emotion Classes in Dataset

Data Understanding and Analysis

Top 5 Words by Emotion Category
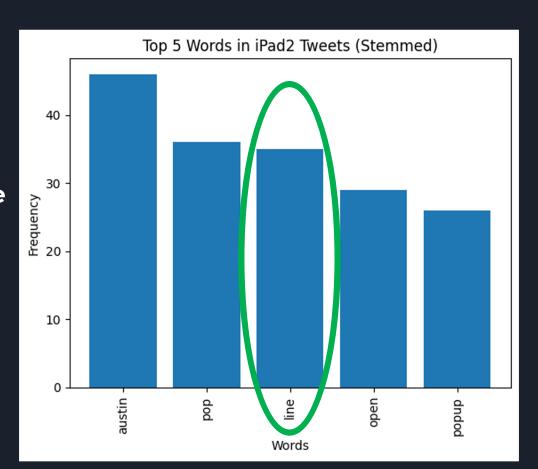
"Line" most common **unique** word in positive iPad2 tweets

Top 5 Words in iPad2 Tweets (Stemmed)

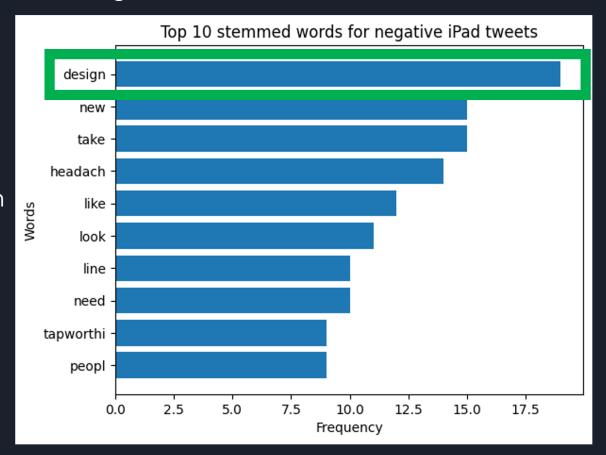Tweets with "Line" and "iPad2"

Lines

Interest and Discussion
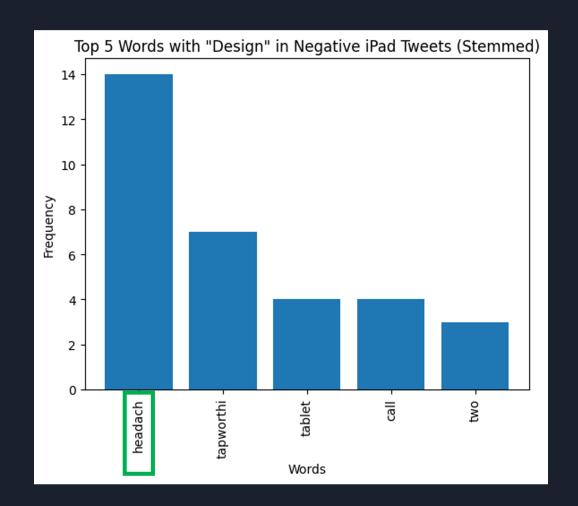
Data Inaccuracies: Tweets Labeled as Positive

- ipad2 takes over #sxsw. the line was already halfway down the block around noon. craziness

- days later there's still a line out the door at the #sxsw apple store for the ipad2.

- the line for the apple popup store at #sxsw stretches around the block...guess i won't be getting the ipad2

Word Cloud for Negative Emotion, iPad Product Tweets
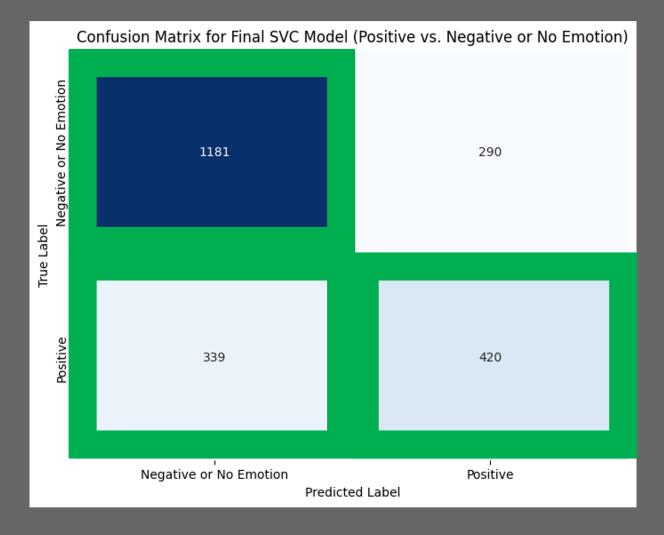

"Design" most common word in negative iPad tweets

"Headache" and "Design" most common pair in negative iPad tweets

Top 5 Words with "Design" in Negative iPad Tweets (Stemmed)

Mislabeled iPad Tweets with Stems "Design" and "Headach"

- attending @mention ipad design headaches #sxsw {link}

- in ipad design headaches: take two tablets, call me in the am panel – excited to hear @mention live! #sxsw

- headed for ipad design headaches (2 tablets, call in the morning) #sxsw {link}
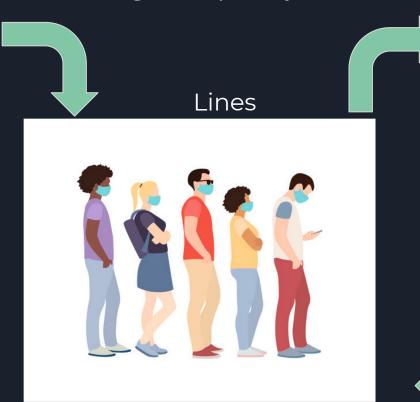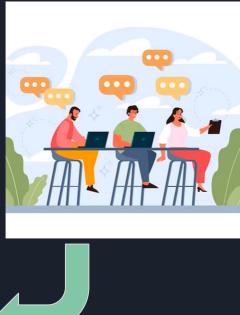
Precision: 55%
Recall: 59%
CV: 65%

Confusion Matrix for Final SVC Model (Positive vs. Negative or No Emotion)

	Negative or No Emotion	Positive
Negative or No Emotion	1181	290
Positive	339	420

True Label

Predicted Label

Recommendations from Data

Follow High Frequency Stems

Common Words

Lines

Interest and Discussion
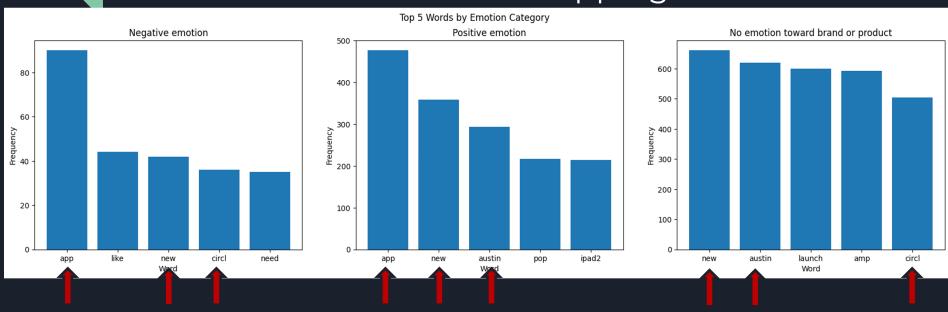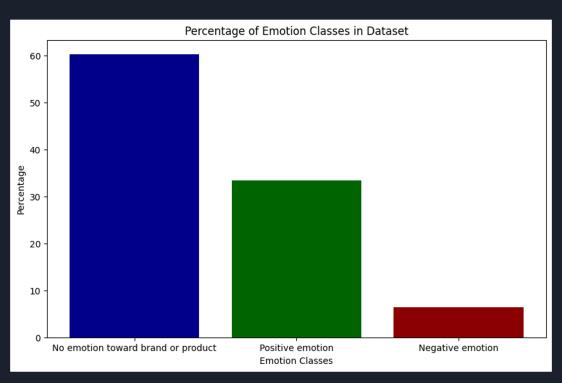
Further removal of overlapping tokens

Top 5 Words by Emotion Category

Additional Data

Percentage of Emotion Classes in Dataset
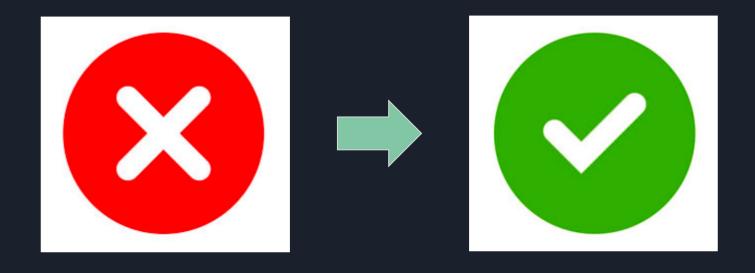
Correct Mislabeled Data

Next Steps

- Testing **new data** can begin after recommendations.

- **New data's words** may prove useful, as demonstrated

- Corresponding **information** from **three classes** of **tweets** will provide Google **keen insights**.

Thank you

I would be happy to answer any questions

You may find my contact information below:

https://www.linkedin.com/in/konnorclark/