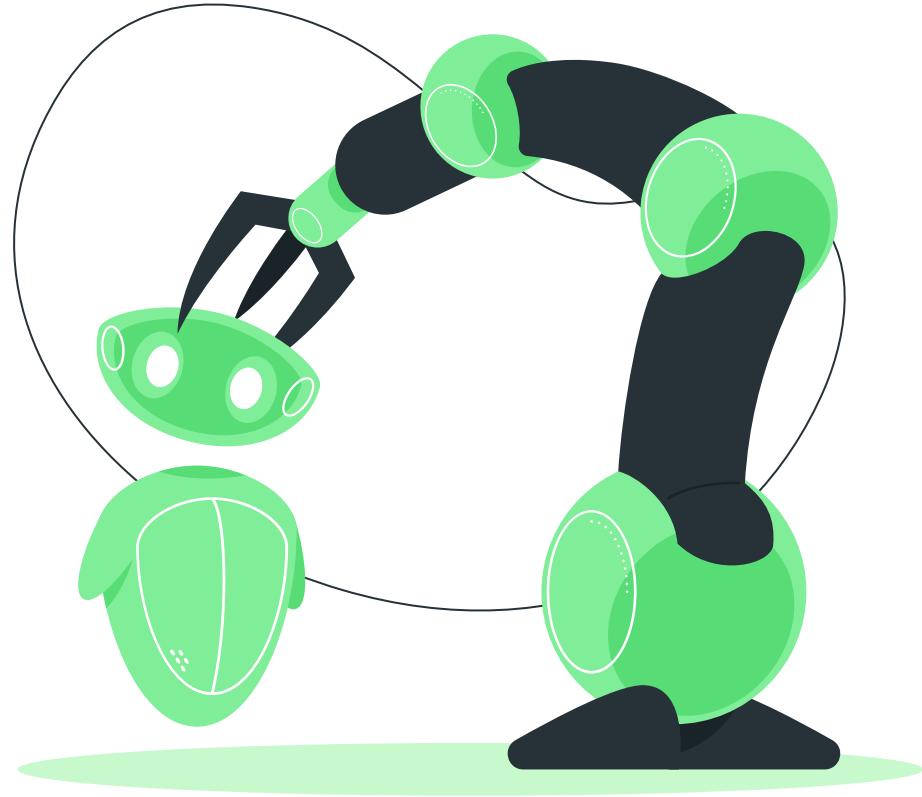


Team

UNDEFINED

Meenakshi Balgobin
Ashish Venkiah
Nikhil Ken Ramessur
Yuvraj Seegolam

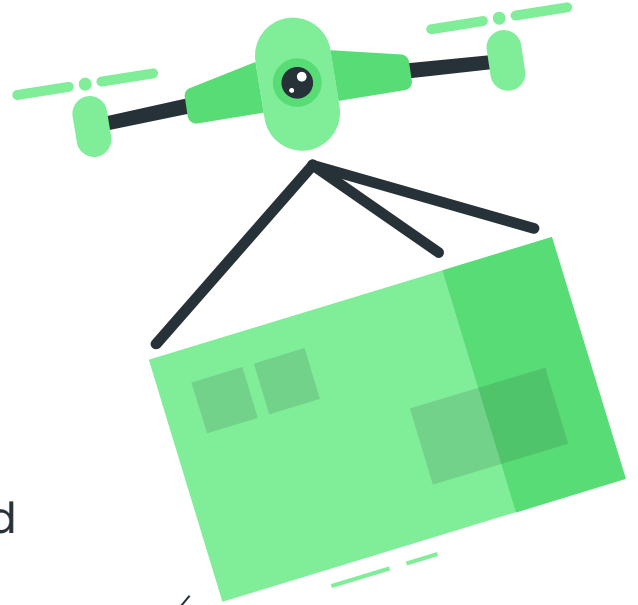


Topic

On the the 17th of June 2022, the U.S Food & Drug Administration authorized use of moderna & Pfizer vaccine to children down to 6 months.



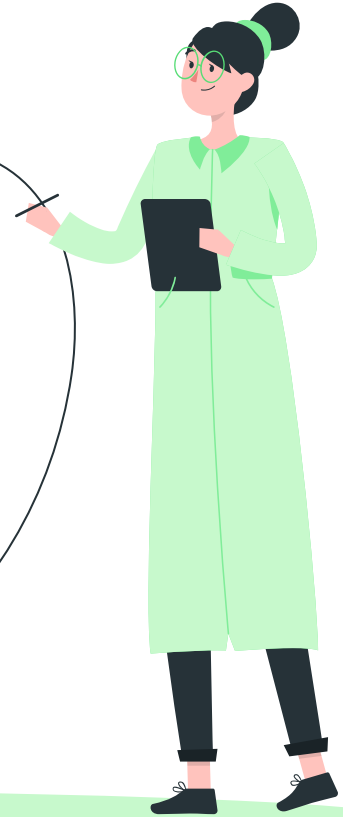
How do people feel about this?



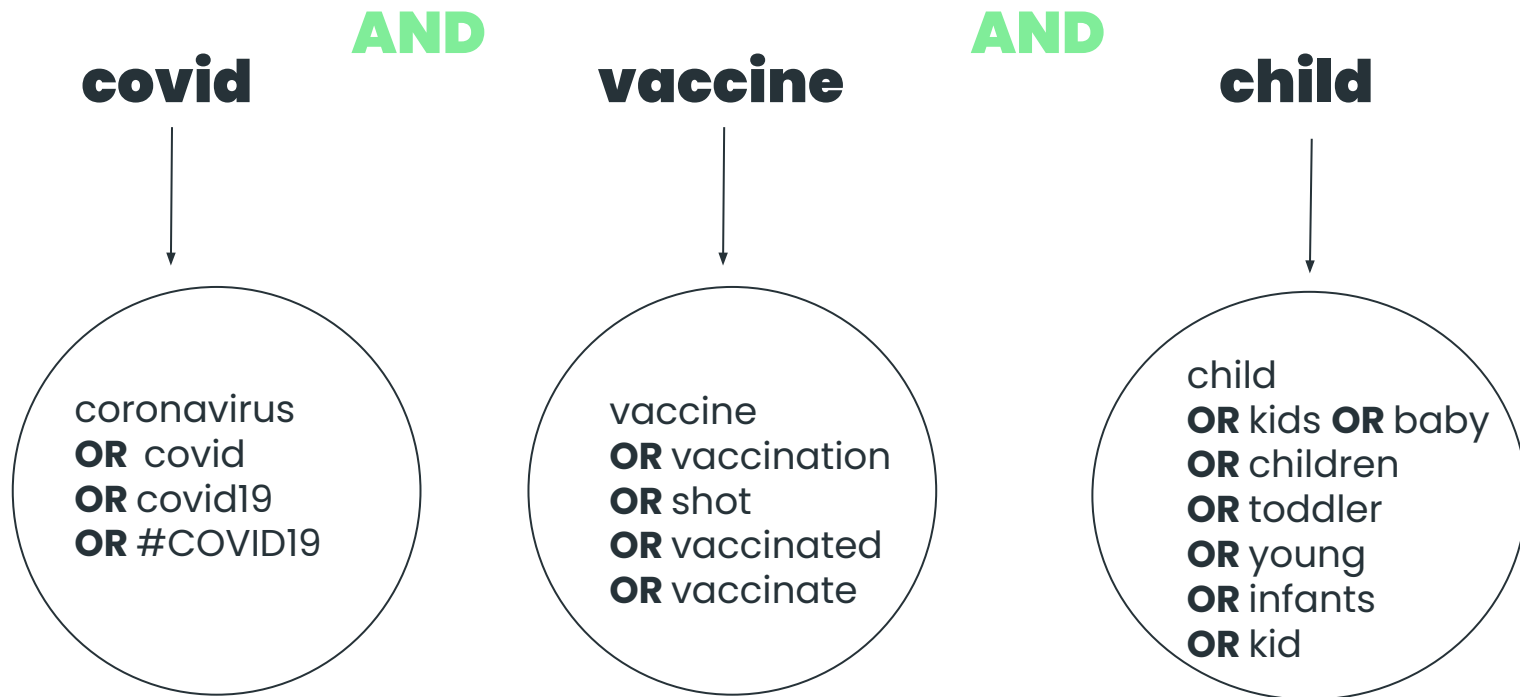


Why

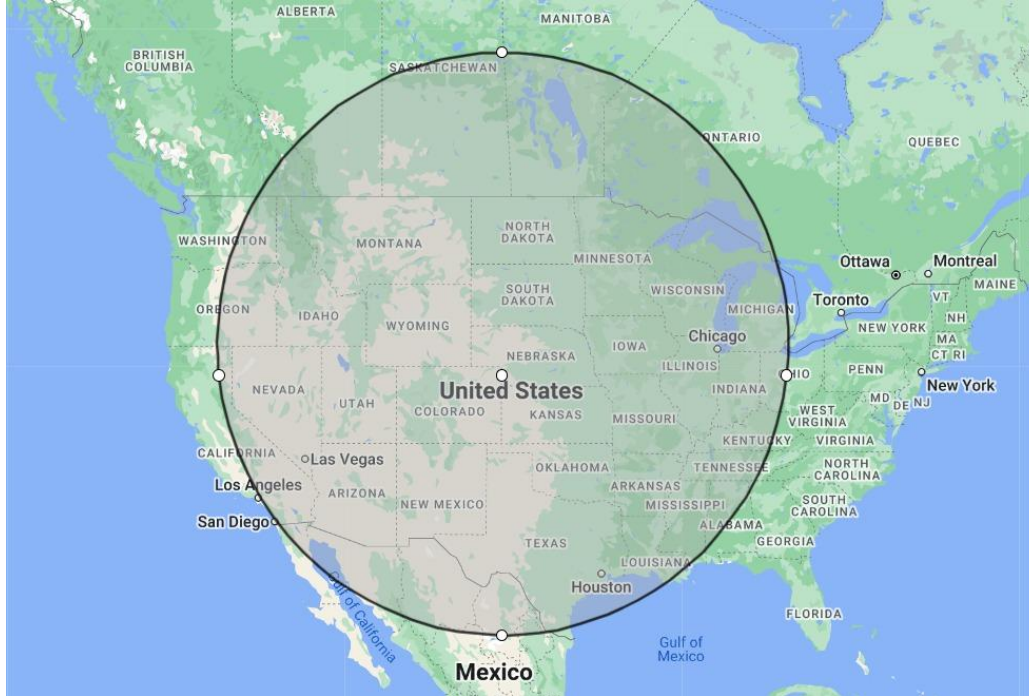
1. To analyse US citizens opinions
2. For other countries to analyse the impact of this decision



Query Visualization



Target Locations



1600 km
United States

Latitude: 39.036356,
Longitude: -101.575207

Data Exploration

- 1000 tweets
- 7% tweets have emojis
- 114 duplicates found
- 189 characters in average
- 88% tweets have uppercase words

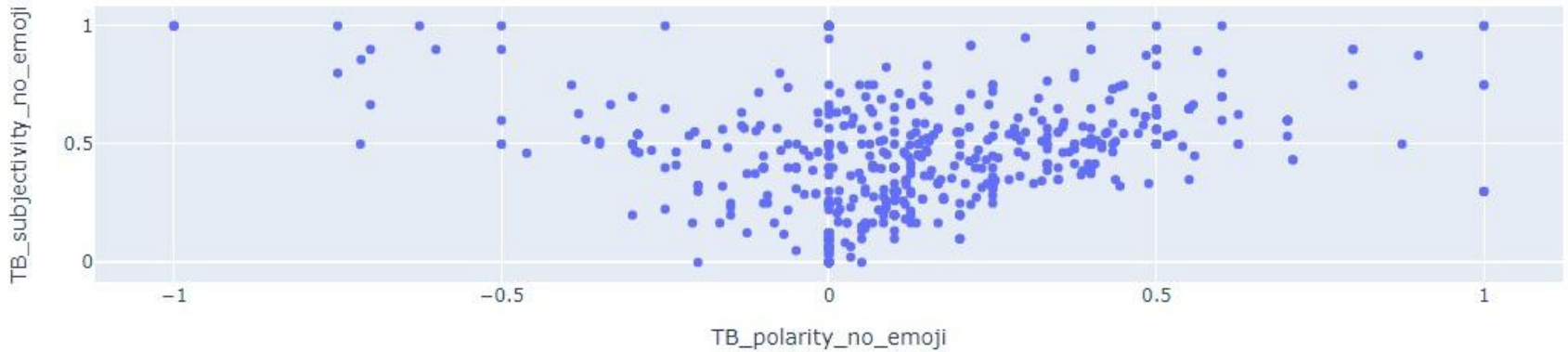


Data Pre-Processing

- Remove links, mentions
- Remove 114 duplicates
- Remove stopwords for TextBlob



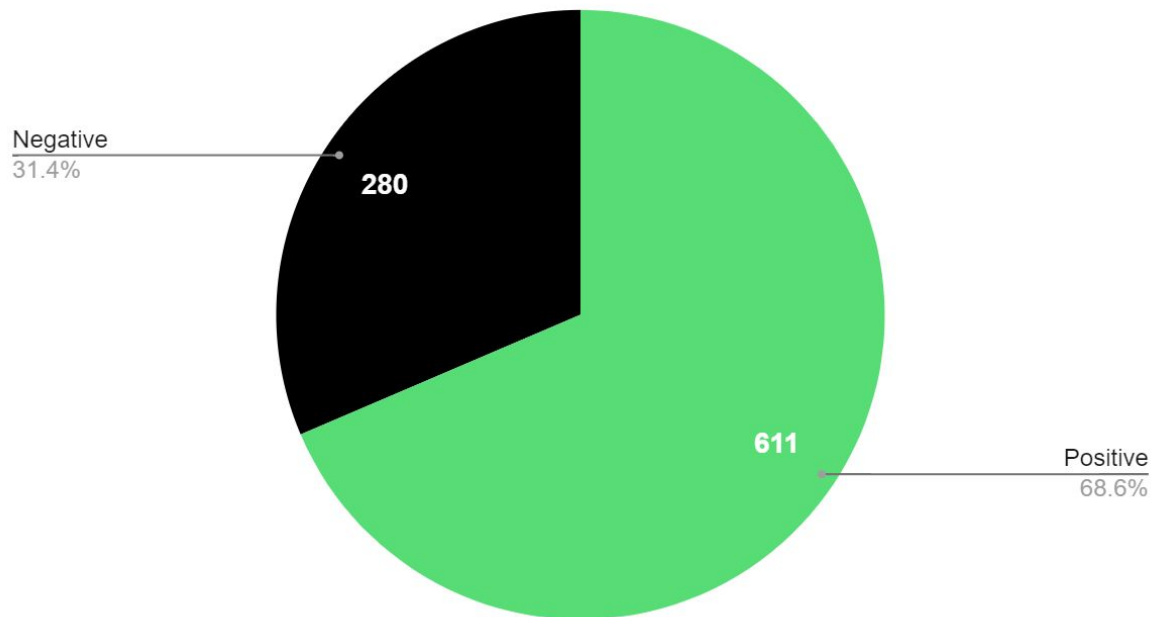
Sentiment Analysis – TextBlob



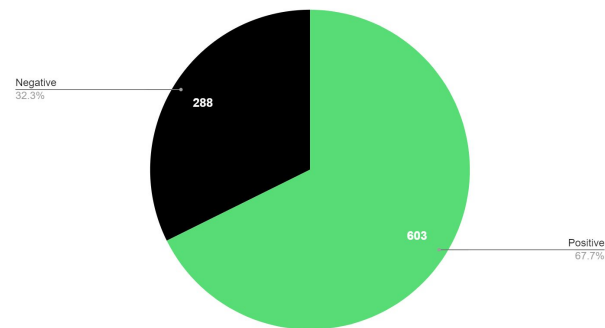
- Most of the tweets are **Neutral** or **Positive**
- Most of the tweets are **Factual**

Sentiment Analysis - CT-BERT

CT-BERT

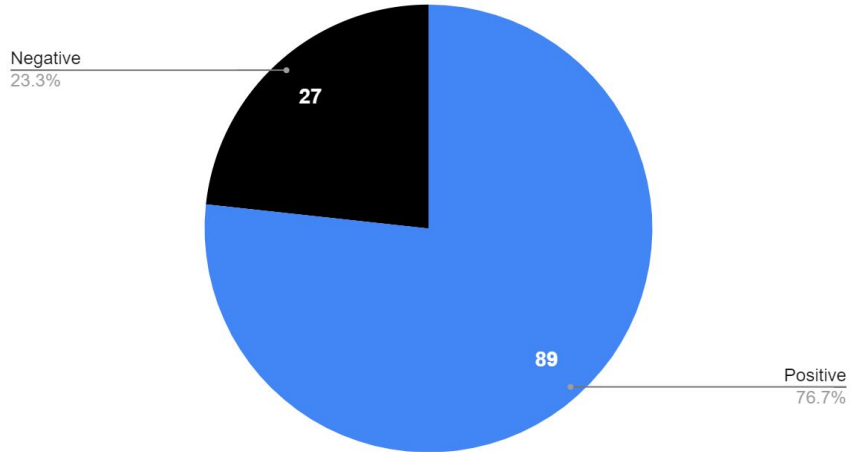


BERT



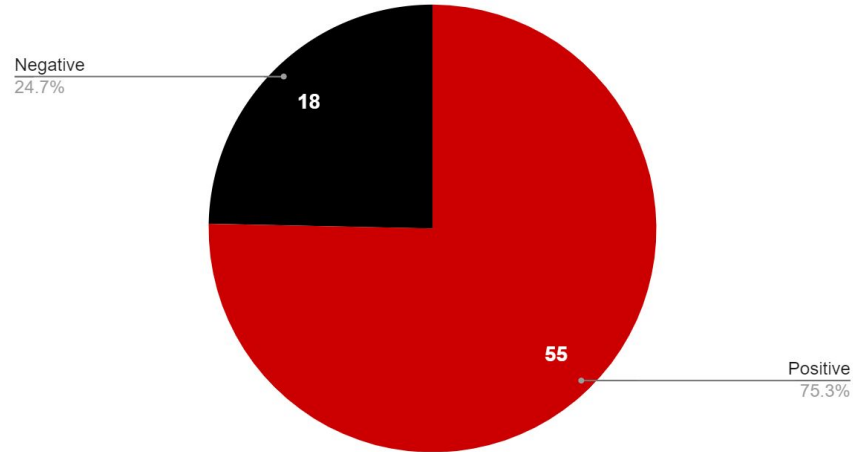
Impact of Politics

California



Blue State

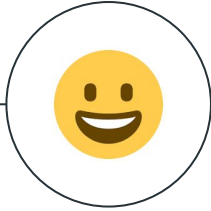
Texas



Red State

Side Quest – Emoji Processing

1. Extract emojis in tweets



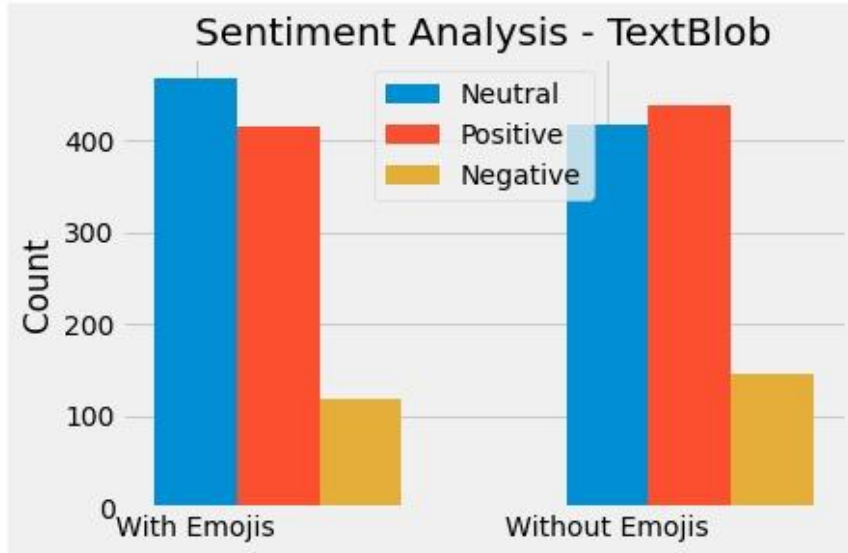
Happy
face

2. Convert to Text

Existing text
+
Emoji text

3. Append to tweet text

Side Quest – Impact of Emojis



Emojis increase **Neutrality** of tweets

Additional challenges

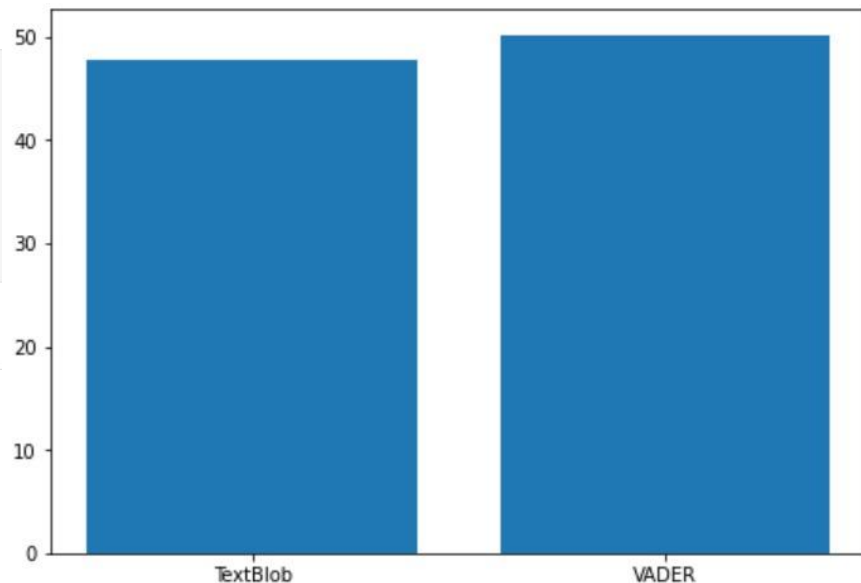
Q1: For what percentage of tweets did the TextBlob & VADER tools predict correctly the sentiment? Which model performs better?

```
TBtotal = len(df[df['TBCorrect'] == True])
VADERtotal = len(df[df['VADERCorrect'] == True])

TBpercentage = (TBtotal / len(df['TBSentiment'])) * 100
VADERpercentage = (VADERtotal / len(df['VADERSentiment'])) * 100
print(TBpercentage)
print(VADERpercentage)
```

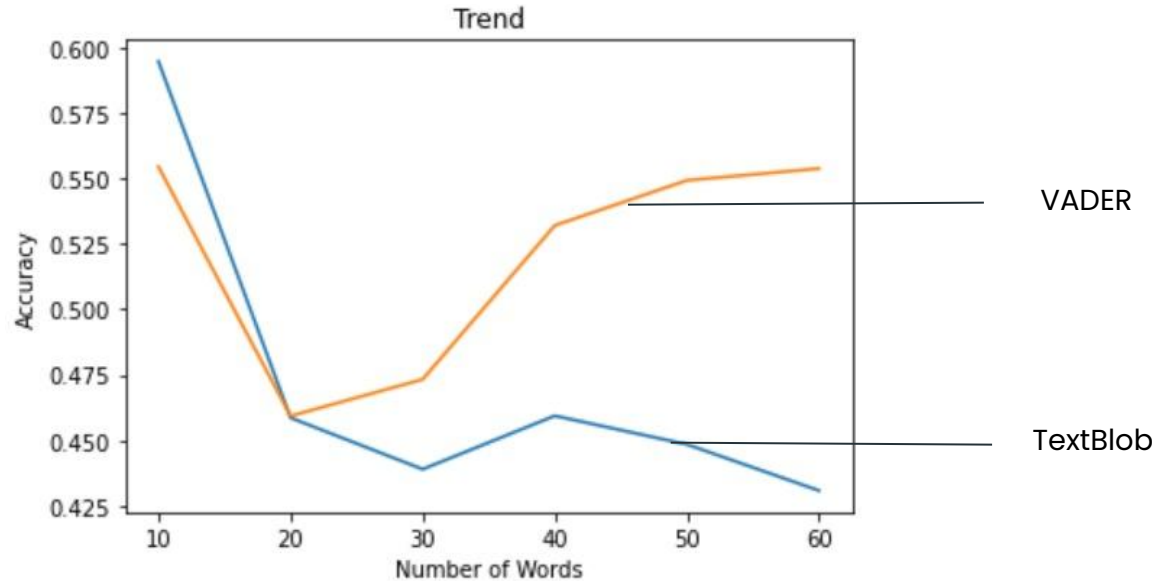
```
47.8363679356822
50.224639394655945
```

**VADER performs slightly better
(~ 2%)**



Additional challenges

Q2: For each models, does the number of words (or chars) in a tweet affect the accuracy of the assigned sentiment



Conclusion!



Team UNDEFINED

Any Questions ?