Executive Voices

By:

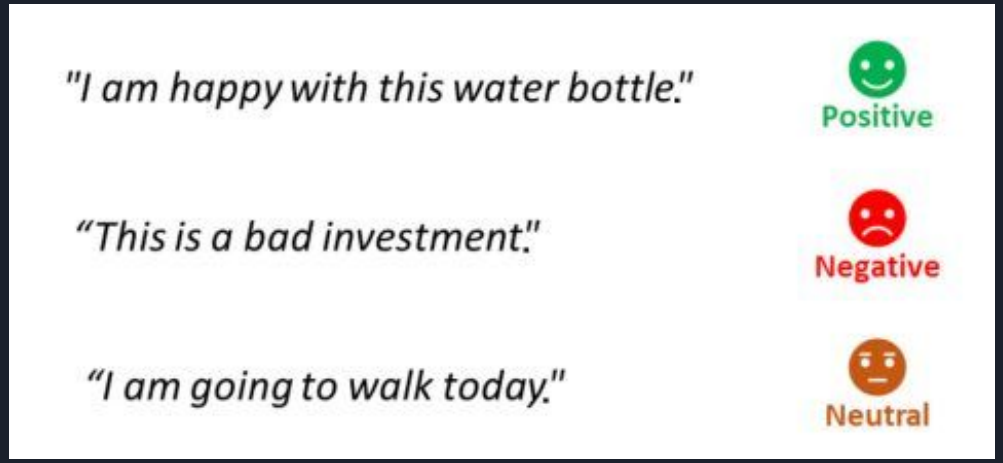
Harry Torres
Blake Gregory
Taylor Tucker

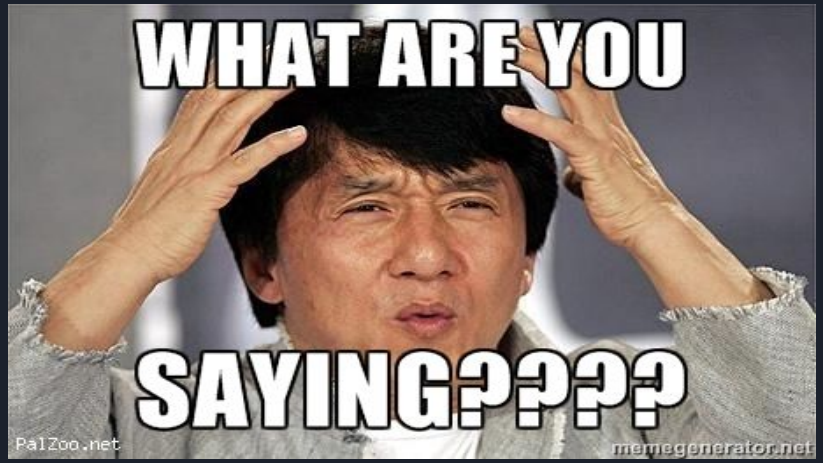
Motivation

We wanted to demonstrate two different methods that are used to perform sentiment analysis, and compare their efficacy when applied to comments made by executives during an earnings call.

Summary

Sentiment Analysis is the most common text classification tool that analyses an incoming message and tells whether the underlying sentiment is positive, negative or neutral.





Approach

We decided to use the CEO comments of Intel's Q4 2020 earnings call as a benchmark for our sentiment classification.

Data Cleanup & Model Training

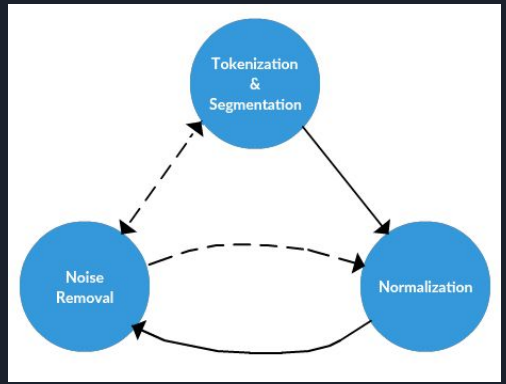
Describe the exploration and cleanup process

Scraped the SEC website for Intel's 10-k's

Reverted to single statement from single 10-k

Cleaned the words in the executive statements:

- Sentence segmentation
- Lower case all words
- Remove digits
- Tokenize words
- Lemmatize words
- Remove stopwords



Method 1:

1. **VADER-Valence Aware Dictionary for sEntiment Reasoning**
    - sentiment reasoning
    - classified as either positive or negative
    - lexicon-based approach
    - calculates intensity sentiment of showing words

Method 2:

2. **LSTM model (Long Short-Term Memory)**
    - effective for sequential data like texts
    - relate the context of a sentence very well
    - supports Long-term dependency
    - Machine learning approach

Model Evaluation

BLEU - an algorithm for evaluating the quality of text which has been machine-translated from one natural language to another.

- We had a Score of 1

f1 Score - The f1 Score is the mean of Recall and Precision, with a higher score as a better model.

Comparison Of Model Accuracy

VADER	73.48%
LSTM	89.89%

Model Evaluation

LSTM

Classification Report for the RNN LSTM Model

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	10
accuracy			1.00	11
macro avg	1.00	1.00	1.00	11
weighted avg	1.00	1.00	1.00	11

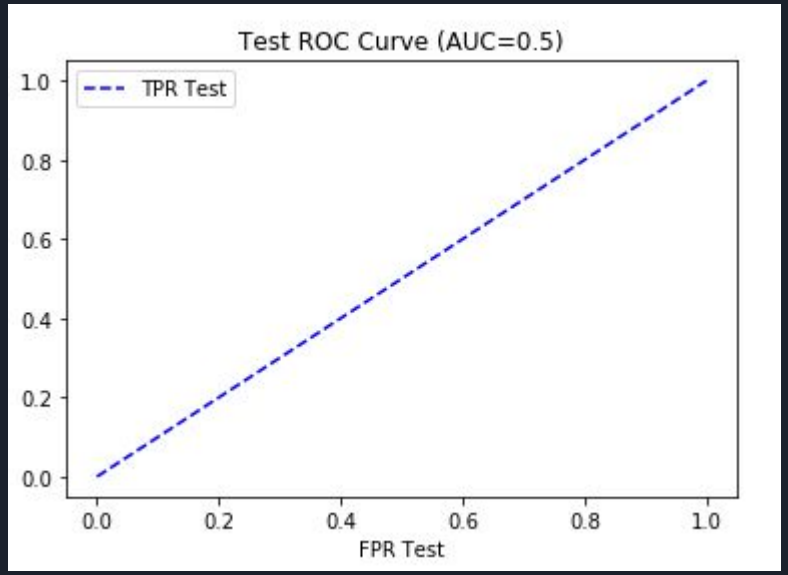
Vader

Classification Report for Vader

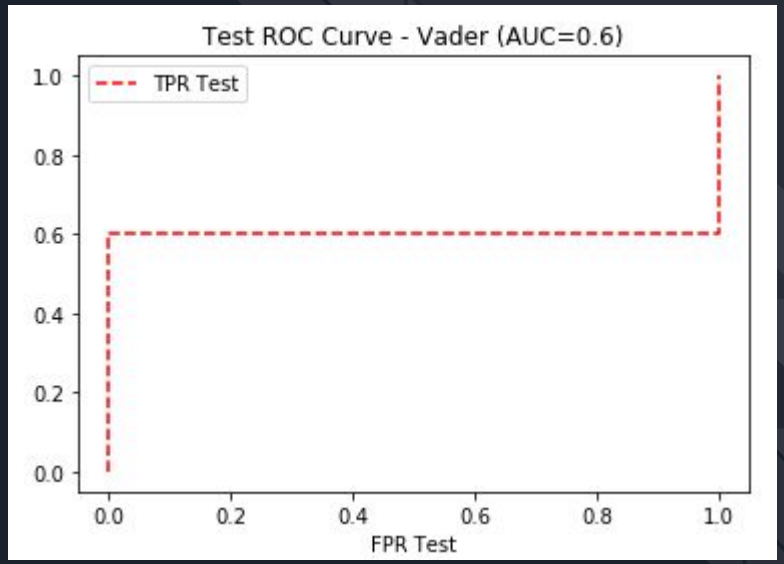
	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	10
accuracy			1.00	11
macro avg	1.00	1.00	1.00	11
weighted avg	1.00	1.00	1.00	11

Model Evaluation

LSTM Loses



Vader Wins



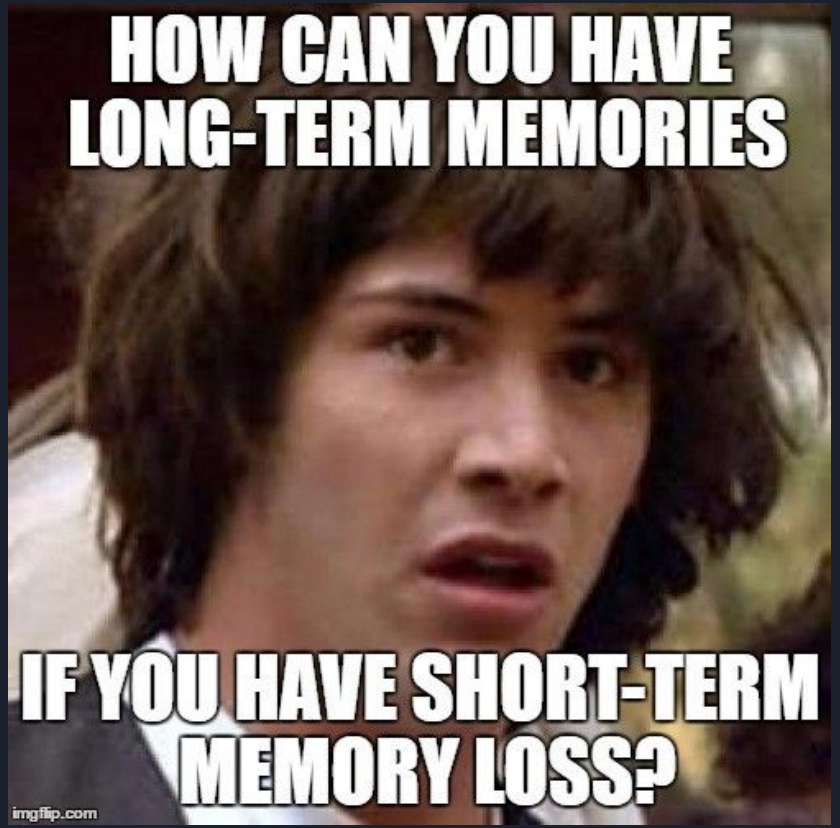
Postmortem

Discuss any difficulties that arose, and how you dealt with them.

- SEC API
- Data binary
- Dataframe

Discuss any additional questions or problems that came up but you didn't have time to answer: What would you research next if you had two more weeks?

Questions


HOW CAN YOU HAVE LONG-TERM MEMORIES
IF YOU HAVE SHORT-TERM MEMORY LOSS?
imgflip.com