Sentiment Analysis using Conventional rule-based approach vs Hugging Face Pretrained Transformers Model RoBERTa

Gift Idama
Towson University
gidama1@students.towson.edu

***Abstract*— Sentiment analysis of news headlines is concerned with identifying positive and negative news in order to use it in decision support systems for stock trend predictions. This study investigates the usage of the traditional sentiment analysis approach on a parsed dataset and then compares it to a more complex pretrained model that focuses more on the context of the sentiment.**

Keywords*— *news headlines, web scraping, machine learning, exploratory data analysis, sentiment analysis, parsed dataset

I. INTRODUCTION

Breaking news will cause stock prices to change. This is known as price discovery. Sometimes, Investors will evaluate the potential effects of the new knowledge on stock prices. Additionally, you'll see how the news affects prices. The kind of reaction we see depends on whether the news is positive or negative news. It's also critical to determine whether the news will materially affect investors' equity. The amount of information accessible makes things potentially quite confusing. Depending on the source and veracity, investors may occasionally be unsure of whether to act or not. Investors must determine whether the news will have a material impact on their equity positions. Classifying this numerous news into good or bad sentiments is very essential in determining the type of reaction to expect from investors and traders.

Sentiment analysis is the process of determining if a piece of text expresses positive or negative sentiment. Businesses typically use it to analyze sentiment in social media data, assess brand reputation, and understand clients.

The Finviz news section is divided into blog entries and news. Both originate from well-chosen sources, which guarantees a certain level of quality. You can view news for a specific company by performing a search, even though news for any firm or organization is available in this section. You can browse news on just Apple, if you search for a stock like Apple. This is great since it enables you to keep tabs on the performance of the stocks you follow.

The Collins dictionary defines "bad news" as "anything that will give you worry or issues." When you say something is good news, you indicate it will be advantageous to you. We will categorize different stock news headlines as positive and negative for this project using an adopted ML model known as “RoBERTa”

II. BUILDING THE DATASET

In this project, we decided to build our own dataset and convert it into a csv file. We parsed various news material into a desirable and well-structured dataset utilizing a web scraping script written in a Python.

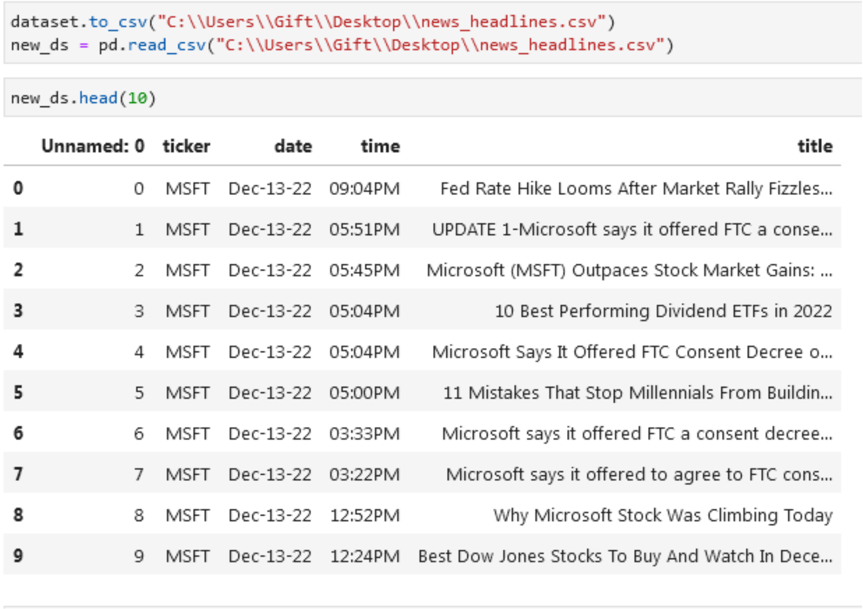
Going into the extraction process in details, we imported the urlopen library and the request library from the python package urrllib, we also imported the librarcy called beautiful soup from the python package bs4.

- The urlopen was integrated to open the contents of the website url, the request library takes this url and returns the html data of the url and parses the html content. We saved this html content into a response variable and then feed this response variable into a beautiful soup object to get the html source code of the website. The source code is stored in a news table dictionary.
- Using a for loop, we iterate through the news table dictionary and extract the title, date and time from their individual tags. The date and time were concatenated so we had to use a "split method" to separate them individually. The date, time, title and ticker are appended into a list and then convert said list into a pandas dataframe. The pandas dataframe was converted into a csv file format for readability, accessibility, and suitability. The name of this well-structured pandas dataset was called "news headlines".

Our dataset contains 4 columns and about 500 observations (we limited it to 500 for better visualization of results. The dataset can be expanded to millions of observations). The web scraping script is flexible and scalable.

The four categories that make up the news headlines dataset are listed below.

1. time: This contains the specified time a news was posted.
2. date: This tells you what day, month, or year a news item was published or disseminated.
3. title: The content in this category relates to news headlines. The most important category for our project is this one. We'll concentrate our sentiment analysis on this category. The news will be classified as either positive, bad, or neutral based on the headlines.
4. ticker: The company's ticker (AAPL – apple, AMZN – amazon).



```
dataset.to_csv("C:\\Users\\Gift\\Desktop\\news_headlines.csv")
new_ds = pd.read_csv("C:\\Users\\Gift\\Desktop\\news_headlines.csv")
```

```
new_ds.head(10)
```

	Unnamed: 0	ticker	date	time	title
0	0	MSFT	Dec-13-22	09:04PM	Fed Rate Hike Looms After Market Rally Fizzles...
1	1	MSFT	Dec-13-22	05:51PM	UPDATE 1-Microsoft says it offered FTC a conse...
2	2	MSFT	Dec-13-22	05:45PM	Microsoft (MSFT) Outpaces Stock Market Gains: ...
3	3	MSFT	Dec-13-22	05:04PM	10 Best Performing Dividend ETFs in 2022
4	4	MSFT	Dec-13-22	05:04PM	Microsoft Says It Offered FTC Consent Decree o...
5	5	MSFT	Dec-13-22	05:00PM	11 Mistakes That Stop Millennials From Buildin...
6	6	MSFT	Dec-13-22	03:33PM	Microsoft says it offered FTC a consent decree...
7	7	MSFT	Dec-13-22	03:22PM	Microsoft says it offered to agree to FTC cons...
8	8	MSFT	Dec-13-22	12:52PM	Why Microsoft Stock Was Climbing Today
9	9	MSFT	Dec-13-22	12:24PM	Best Dow Jones Stocks To Buy And Watch In Dece...

Fig 1: The parsed dataset after web scraping

III. RELATED WORK

Real-time data collection makes streaming data a great source for data analysis. Such data's accessibility and availability, which are its main qualities, aid in accurate analysis and forecasting. Various sentiment analysis projects exist, many ground-breaking proposals have been proposed in the past in areas of different models such as BERT, Flair, SBert, Vader and so much more. These works play a huge role in inspiring this project.

During the outbreak, Wang et al. (2020) offer a public sentiment analysis that can provide illuminating information for formulating effective public health measures. They study posts on the well-known Chinese social media platform Sina Weibo, classifying the sentiment into three groups (positive, neutral, and negative) and summarizing the themes of posts using the TF-IDF (term frequency-inverse document frequency) model. Analyzing social media posts with unfavorable emotion can help us comprehend the experiences and provide models for other nations. The findings shed light on how social sentiment has changed over time as well as the subject matters that have been linked to unfavorable sentiment on social media platforms. Results from the TF-IDF topic extraction model and the BERT classification model were supplied with high precision.

With the assumption that news articles have an impact on the stock market, Kalyani et al. (2016)'s approach uses data from financial news articles about a firm to forecast its future stock movement. This is an effort to investigate the connection between news and stock trend. They employed a dictionary-based strategy for this. Words that convey general and financial-specific sentiments were used to develop the dictionaries for positive and negative words. They constructed categorization models based on this data. The findings demonstrate that Support Vector Machine (SVM) and Random Forest (RF) exhibit strong performance across all tests.

IV. METHODOLOGY

4.1 Problem Definition:

Sentiment Analysis is used to determine the feelings or sentiments expressed in texts. It is an NLP technique that allows you to understand the nuances in human texts and to determine if data is either positive, negative or neutral. There are currently several conventional rule-based approaches to sentiment analysis and in this project, we have identified the well-known VADER model (Valence aware dictionary and sentiment reasoner). VADER model is a very efficient rule-based sentiment analyzer that employs a "bag of words" approach, but it has a drawback in the sense that VADER neglects to take into consideration the relationships between words and so it ignores the context behind words (with context being crucial in human speech). This brings up the project's main point of interest. We utilize a pretrained model called RoBERTa to classify our data and return the sentiment of the fed data. This model, which is a sophisticated transformer model uses deep learning in the form of pretraining and examines the context behind human speech or sentences. The dataset will first undergo basic exploratory data analysis (EDA) in order to get insights and findings from it. To better explain the method, we will use a sample news headline and perform some simple analysis on the parsed information using the Python Natural Language Toolkit (NLTK).

Our project's main objective is to compare the sentiment prediction of the conventional rule-based approach to the adopted pretrained model.

Here are some of the main tools we utilized for this project is.

NLTK (natural language processing toolkit) – For sentiment analysis and ML algorithms

BeautifulSoup – For web scraping and parsing the news headlines

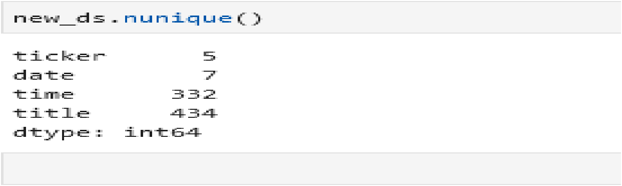
Pandas – for creating our dataframe

Matplotlib – For visualization of our results and findings

Hugging Face transformers.

4.2 Exploratory Data Analysis:

In order to have a better understanding of our data, let’s explore it a little.



```
new_ds.nunique()

ticker        5
date          7
time        332
title       434
dtype: int64
```

Fig 2: number of unique, ticker, date, time, title.

Our dataset type is of integer and there are 5 tickers/companies in our dataset. This dataset takes into consideration news from the last 7 days hence we have a unique date of 7 values. There are 332 unique times various news headline were posted, several news headlines from different sources might have been posted at the same time. For the title column, it is quite interesting, unique value should be around 500 but instead we have 434 unique news headlines. This could mean that maybe various news outlets posted the same news or the website repeated news headlines. This repeated has no effect on our analysis so it can be ignored.

```
new_ds.shape
```

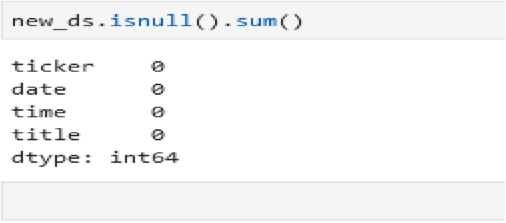
```
(500, 4)
```

Fig 3: number of observations in the dataset

As mentioned earlier, we have about 500 observations and 4 categories/columns.

4.3 Checking for missing values:

Presence of missing values will affect our sentiment analysis, as no sentiment would be returned. So, we need to remove columns with missing values.



```
new_ds.isnull().sum()
```

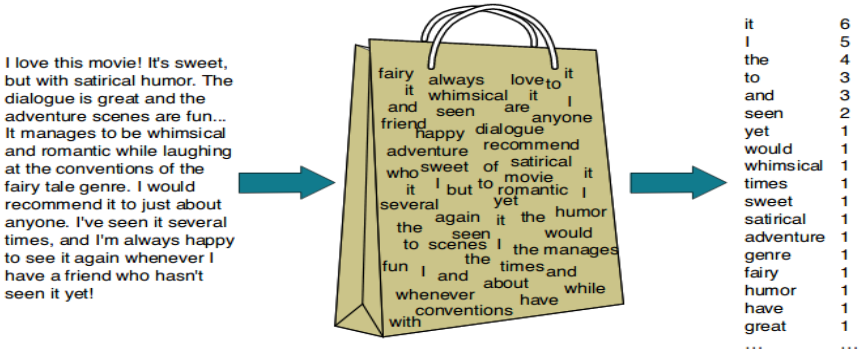
```
ticker    0
date      0
time      0
title     0
dtype: int64
```

Fig 4: Checking for missing values

Our dataset currently has no missing values. So that’s all about it for our exploratory data analysis. Next, we are going to understand how he rule-based approach and how the pretrained model works.

4.4 The rule-based approach: Vader.

VADER is a rule-based sentiment analyzer that employs a “bag of words” approach to sentiment analysis. Vader breaks down a sentence into individual words called tokens, checks a lexicon to see if the words exists, if it does, it returns the polarity score of each individual word.



I love this movie! It's sweet, but with satirical humor. The dialogue is great and the adventure scenes are fun... It manages to be whimsical and romantic while laughing at the conventions of the fairy tale genre. I would recommend it to just about anyone. I've seen it several times, and I'm always happy to see it again whenever I have a friend who hasn't seen it yet!

it	6
I	5
the	4
to	3
and	3
seen	2
yet	1
would	1
whimsical	1
times	1
sweet	1
satirical	1
adventure	1
genre	1
fairy	1
humor	1
have	1
great	1
...	...

Fig 5: Bag of words approach

Vader builds upon two widely known text analysis libraries. LIWC (Linguistic Inquiry and word count) and ANEW (Affective Norms for English words). The LIWC library looks at the polarity of different words and try to figure out if a word is positive or negative. The creators of Vader have categorized about 900 words into these two categories

The ANEW library looks at the intensity of different words. How positive or how negative is a word. It usually has a range of 1-9 with 1 being negative, 5 being neutral and 9 being positive. Below we came up with a pictorial description of how the Vader model works.

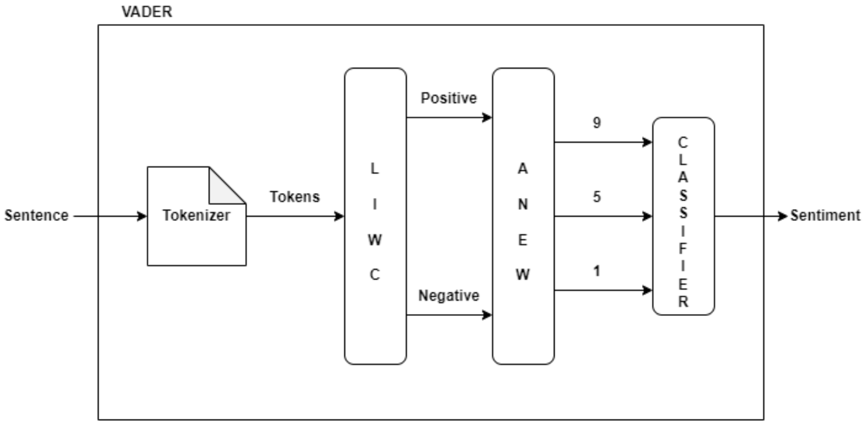
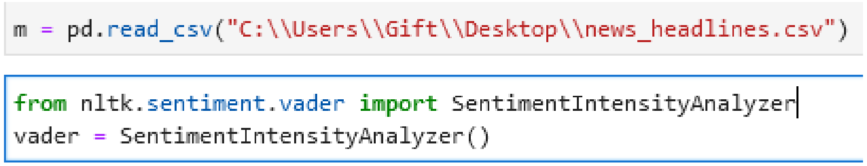


Fig 6: Pictorial description of the Vader model

4.5 Vader implementation on the dataset.

After reading the csv file, we implement the Vader model by calling the Sentiment Intensity Analyzer from the NLTK module.

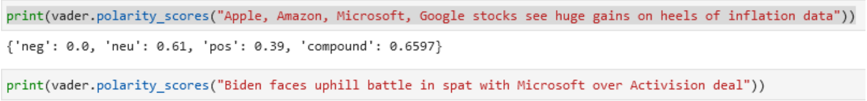


```
m = pd.read_csv("C:\\Users\\Gift\\Desktop\\news_headlines.csv")
```

```
from nltk.sentiment.vader import SentimentIntensityAnalyzer
vader = SentimentIntensityAnalyzer()
```

Fig 7: Calling the vader model.

We test the model if it is implemented by running a sample sentence into the model to check its polarity score and see how the model performs.



```
print(vader.polarity_scores("Apple, Amazon, Microsoft, Google stocks see huge gains on heels of inflation data"))
{'neg': 0.0, 'neu': 0.61, 'pos': 0.39, 'compound': 0.6597}

print(vader.polarity_scores("Biden faces uphill battle in spat with Microsoft over Activision deal"))
```

Fig 8: A sample representation

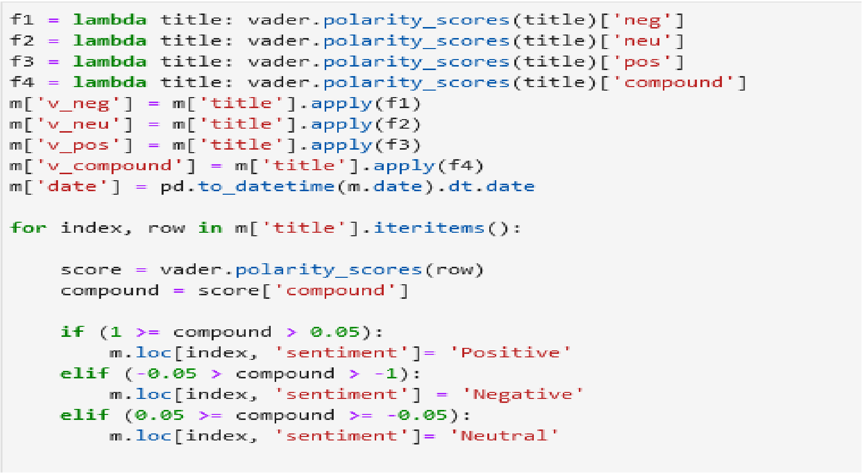
Analyzing the sample representation, the polarity score of each sentiment were shown, negative being 0, neutral has a polarity score of 0.61 and positive sentiment has a polarity score of 0.39. Looking at the polarity scores, one could imply that the sample sentence is a neutral sentence whereas based on human intuition the sample sentence clearly shows a positive sentiment. This bring us to a unique feature in the Vader model. The Vader does not really pay attention to the polarity scores of each sentiment but rather focuses on the compound score. Vader uses the compound score to predict the sentiment of a sentence. The compound score is arrived at by simply taking the scores of each individual word, add them up and normalize it. We then arrive at a final score between the range of -1 and 1. That final score is known as the compound score.

The closer it is to 1 shows a positive sentiment and the closer it is to -1 shows a negative sentiment. The ranges are shown as follow.

[-1, -0.05] negative, [-0.05, 0.05] neutral, [0.05, 1] positive.

For our sample sentence, 0.65 lie in the positive range, implying that the sample sentence is a positive sentiment.

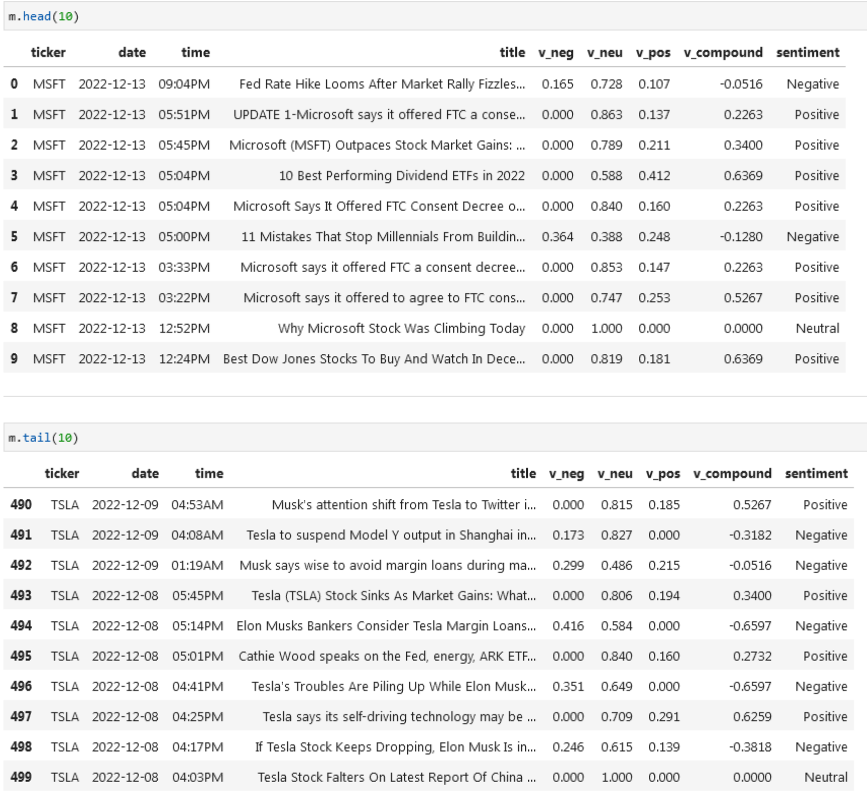
We understand how the Vader model works using a sample sentence, let us run it on the entirety of the dataset using a python for loop we wrote.



```
f1 = lambda title: vader.polarity_scores(title)['neg']
f2 = lambda title: vader.polarity_scores(title)['neu']
f3 = lambda title: vader.polarity_scores(title)['pos']
f4 = lambda title: vader.polarity_scores(title)['compound']
m['v_neg'] = m['title'].apply(f1)
m['v_neu'] = m['title'].apply(f2)
m['v_pos'] = m['title'].apply(f3)
m['v_compound'] = m['title'].apply(f4)
m['date'] = pd.to_datetime(m.date).dt.date

for index, row in m['title'].iteritems():

    score = vader.polarity_scores(row)
    compound = score['compound']

    if (1 >= compound > 0.05):
        m.loc[index, 'sentiment']= 'Positive'
    elif (-0.05 > compound > -1):
        m.loc[index, 'sentiment'] = 'Negative'
    elif (0.05 >= compound >= -0.05):
        m.loc[index, 'sentiment']= 'Neutral'
```

Fig 9: Iteration of the vader model on the title row

Breaking down the python script, we implemented the lambda function to focus on the polarity score of each sentiment. The we created new columns in the dataset to take in the polarity scores of each iteration. Iterating through the title column, we pick out the compound score of each row and put it in an if-else condition with the ranges of the normalized compound score. If compound score lies in the negative range, return negative sentiment else return positive sentiment if compound score lies in the positive range else return neutral sentiment if the compound score lies in the neutral range. We get a new dataset with the appended columns showing the polarity scores of each iteration.



```
m.head(10)
```

	ticker	date	time	title	v_neg	v_neu	v_pos	v_compound	sentiment
0	MSFT	2022-12-13	09:04PM	Fed Rate Hike Looms After Market Rally Fizzles...	0.165	0.728	0.107	-0.0516	Negative
1	MSFT	2022-12-13	05:51PM	UPDATE 1-Microsoft says it offered FTC a conse...	0.000	0.863	0.137	0.2263	Positive
2	MSFT	2022-12-13	05:45PM	Microsoft (MSFT) Outpaces Stock Market Gains: ...	0.000	0.789	0.211	0.3400	Positive
3	MSFT	2022-12-13	05:04PM	10 Best Performing Dividend ETFs in 2022	0.000	0.588	0.412	0.6369	Positive
4	MSFT	2022-12-13	05:04PM	Microsoft Says It Offered FTC Consent Decree o...	0.000	0.840	0.160	0.2263	Positive
5	MSFT	2022-12-13	05:00PM	11 Mistakes That Stop Millennials From Buildin...	0.364	0.388	0.248	-0.1280	Negative
6	MSFT	2022-12-13	03:33PM	Microsoft says it offered FTC a consent decree...	0.000	0.853	0.147	0.2263	Positive
7	MSFT	2022-12-13	03:22PM	Microsoft says it offered to agree to FTC cons...	0.000	0.747	0.253	0.5267	Positive
8	MSFT	2022-12-13	12:52PM	Why Microsoft Stock Was Climbing Today	0.000	1.000	0.000	0.0000	Neutral
9	MSFT	2022-12-13	12:24PM	Best Dow Jones Stocks To Buy And Watch In Dece...	0.000	0.819	0.181	0.6369	Positive

```
m.tail(10)
```

	ticker	date	time	title	v_neg	v_neu	v_pos	v_compound	sentiment
490	TSLA	2022-12-09	04:53AM	Musk's attention shift from Tesla to Twitter i...	0.000	0.815	0.185	0.5267	Positive
491	TSLA	2022-12-09	04:08AM	Tesla to suspend Model Y output in Shanghai in...	0.173	0.827	0.000	-0.3182	Negative
492	TSLA	2022-12-09	01:19AM	Musk says wise to avoid margin loans during ma...	0.299	0.486	0.215	-0.0516	Negative
493	TSLA	2022-12-08	05:45PM	Tesla (TSLA) Stock Sinks As Market Gains: What...	0.000	0.806	0.194	0.3400	Positive
494	TSLA	2022-12-08	05:14PM	Elon Musks Bankers Consider Tesla Margin Loans...	0.416	0.584	0.000	-0.6597	Negative
495	TSLA	2022-12-08	05:01PM	Cathie Wood speaks on the Fed, energy, ARK ETF...	0.000	0.840	0.160	0.2732	Positive
496	TSLA	2022-12-08	04:41PM	Tesla's Troubles Are Piling Up While Elon Musk...	0.351	0.649	0.000	-0.6597	Negative
497	TSLA	2022-12-08	04:25PM	Tesla says its self-driving technology may be ...	0.000	0.709	0.291	0.6259	Positive
498	TSLA	2022-12-08	04:17PM	If Tesla Stock Keeps Dropping, Elon Musk Is in...	0.246	0.615	0.139	-0.3818	Negative
499	TSLA	2022-12-08	04:03PM	Tesla Stock Falters On Latest Report Of China ...	0.000	1.000	0.000	0.0000	Neutral

Fig 10 & 11: Results showing the sentiment of the vader model.

Analyzing the results of the Vader model and taking into consideration the main drawback of the Vader model, using row 8 and row 490 as a case study, the model returns a neutral sentiment for row 8 and a positive sentiment for row 490 when clearly analyzing based off human intuition, row 8 implies a positive sentiment and row 490 implies a negative sentiment. This backs up the claim about the main drawback of the Vader model.

Visualizing the Vader model, we can deduce the number of negative news and positive news about a company in a day.

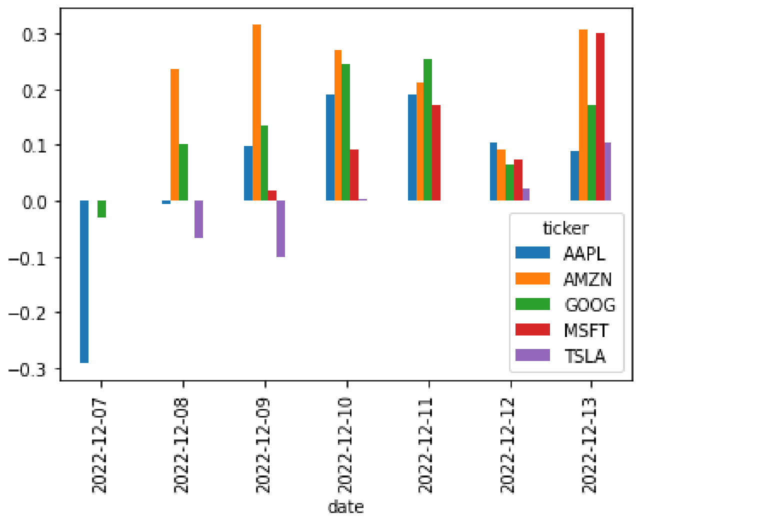


Fig 12: Visualization of the result of the Vader model.

Breaking down this visualization, we can see that on the 7th of December, Majority of the news about Apple and Google were negative news, comparing it to the 13th of December, all companies had a majority of positive news on that day.

4.6 Hugging Face Pretrained Model: RoBERTa

This major drawback to Vader was addressed in this project using Hugging Face pretrained model RoBERTa. RoBERTa is a variant of BERT. It is uses pretraining approach in calculating the polarity scores of a sentence. Unlike BERT, it uses dynamic masking to mask various aspects of a sentence and the passes it through an algorithm, the algorithm uses each old knowledge it learns to make better predictions. It uses a transfer of knowledge approach to predict the polarity scores of a sentence. RoBERTa mask different tokens in the input sentences and then epoch trained. Hence the input sentence doesn't remain constant and doesn't remain the same for all epochs.

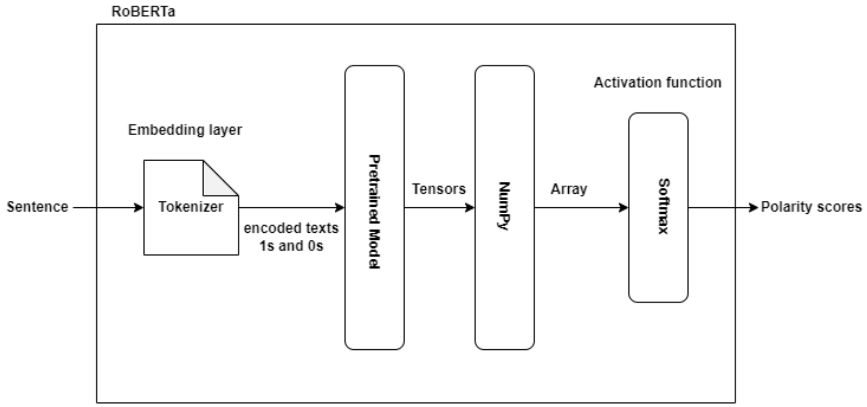
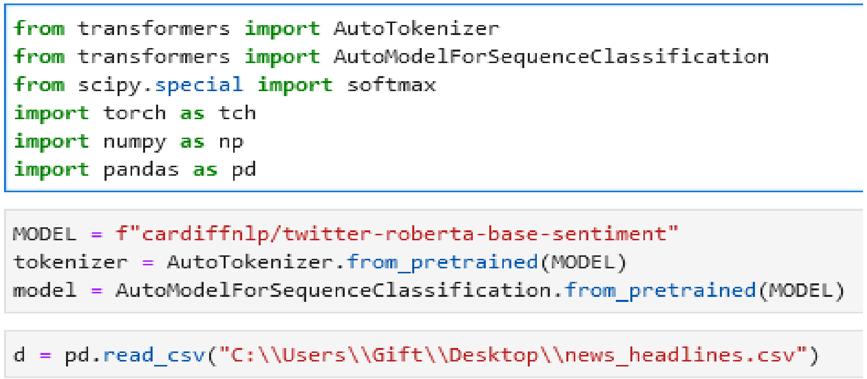


Fig 13: Pictorial description of the RoBERTa model.

4.7 Implementing RoBERTa on our dataset.

I implemented the Auto Tokenizer to tokenizes the sentence by breaking or splitting it into individual words. Then, we are gonna pull in a model provided by Hugging Face that has been trained on a bunch of data When we run the Auto Tokenizer and the Sequence classification methods and load it from a pretrained model, it will pull down the model weight that has been stored. This is very great, because it is basically performing transfer learning. The model was trained on a bunch of twitter comments, and we don't have to retrain the model at all, we can simply use the trained weights and apply it to our dataset.



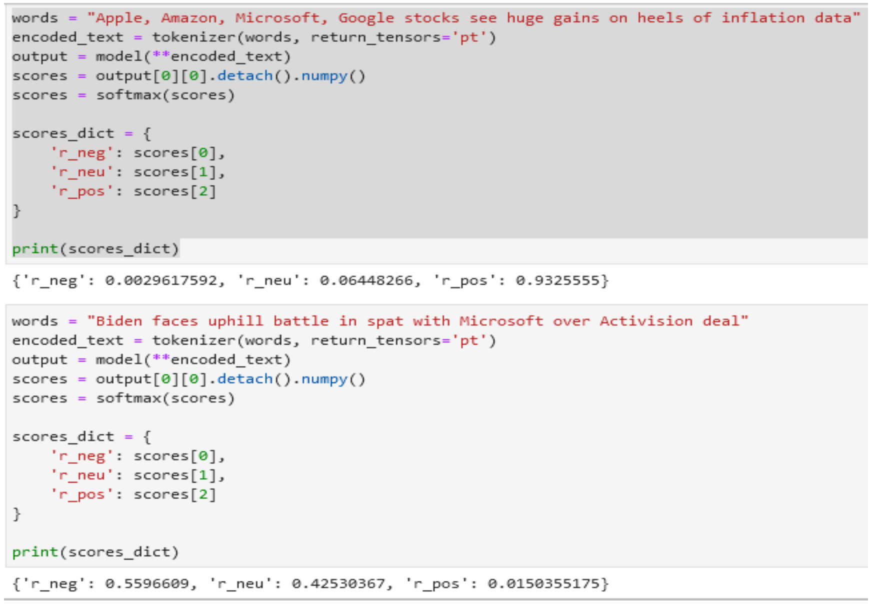
```
from transformers import AutoTokenizer
from transformers import AutoModelForSequenceClassification
from scipy.special import softmax
import torch as tch
import numpy as np
import pandas as pd
```

```
MODEL = f"cardiffnlp/twitter-roberta-base-sentiment"
tokenizer = AutoTokenizer.from_pretrained(MODEL)
model = AutoModelForSequenceClassification.from_pretrained(MODEL)
```

```
d = pd.read_csv("C:\\Users\\Gift\\Desktop\\news_headlines.csv")
```

Fig 14: Extracting the weights from the pretrained model

After implementing our model, we test it out on a sample sentence to check the polarity score and see how it performs.



```
words = "Apple, Amazon, Microsoft, Google stocks see huge gains on heels of inflation data"
encoded_text = tokenizer(words, return_tensors='pt')
output = model(**encoded_text)
scores = output[0][0].detach().numpy()
scores = softmax(scores)

scores_dict = {
    'r_neg': scores[0],
    'r_neu': scores[1],
    'r_pos': scores[2]
}

print(scores_dict)
```

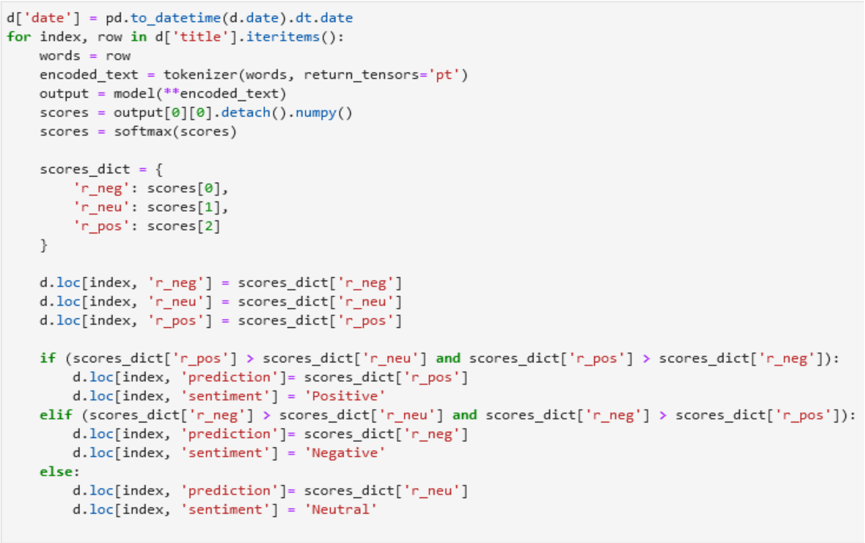
```
{'r_neg': 0.0029617592, 'r_neu': 0.06448266, 'r_pos': 0.9325555}
```

```
words = "Biden faces uphill battle in spat with Microsoft over Activision deal"
encoded_text = tokenizer(words, return_tensors='pt')
output = model(**encoded_text)
scores = output[0][0].detach().numpy()
scores = softmax(scores)

scores_dict = {
    'r_neg': scores[0],
    'r_neu': scores[1],
    'r_pos': scores[2]
}

print(scores_dict)
```

```
{'r_neg': 0.5596609, 'r_neu': 0.42530367, 'r_pos': 0.0150355175}
```

Fig 15: Sample representation of the RoBERTa model.

Analyzing the sample representation 1, the polarity score of each sentiment were shown, negative being 0.029, neutral has a polarity score of 0.06 and positive sentiment has a polarity score of 0.93. Looking at the polarity scores, one could imply that the sample sentence is a positive sentence and based on human intuition the sample sentence clearly shows a positive sentiment. We have a brief understanding of the RoBERTa model, now we are going to implement it on our dataset by using a for loop to iterate through the "title" column to return the polarity scores.



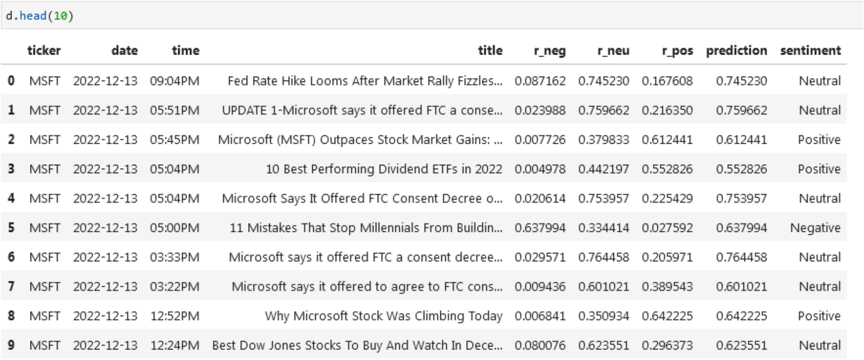
```
d['date'] = pd.to_datetime(d.date).dt.date
for index, row in d['title'].iteritems():
    words = row
    encoded_text = tokenizer(words, return_tensors='pt')
    output = model(**encoded_text)
    scores = output[0][0].detach().numpy()
    scores = softmax(scores)

    scores_dict = {
        'r_neg': scores[0],
        'r_neu': scores[1],
        'r_pos': scores[2]
    }

    d.loc[index, 'r_neg'] = scores_dict['r_neg']
    d.loc[index, 'r_neu'] = scores_dict['r_neu']
    d.loc[index, 'r_pos'] = scores_dict['r_pos']

    if (scores_dict['r_pos'] > scores_dict['r_neu'] and scores_dict['r_pos'] > scores_dict['r_neg']):
        d.loc[index, 'prediction']= scores_dict['r_pos']
        d.loc[index, 'sentiment'] = 'Positive'
    elif (scores_dict['r_neg'] > scores_dict['r_neu'] and scores_dict['r_neg'] > scores_dict['r_pos']):
        d.loc[index, 'prediction']= scores_dict['r_neg']
        d.loc[index, 'sentiment'] = 'Negative'
    else:
        d.loc[index, 'prediction']= scores_dict['r_neu']
        d.loc[index, 'sentiment'] = 'Neutral'
```

Fig 16: Iteration of RoBERTa model on the title row

Breaking down the python script. the sentences are stored in a variable called "words". Because the model does not understand words but rather 1s and 0s, the sentence is passed into a tokenizer which returns an embedded text of 1s and 0s as the output. The embedded text is then passed into the model, the model analyzes it and returns a tensor as output. The tensors are converted into a NumPy array that can be stored locally. This array is then passed through a SoftMax activation function and returns an array of three different individual scores.

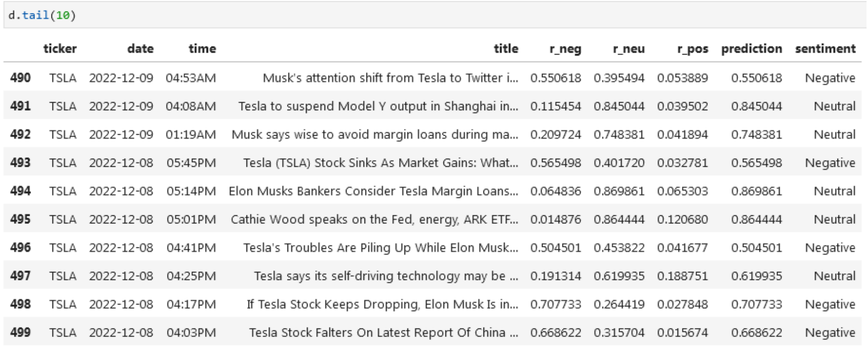
If-else statement is used to specify each polarity score (positive, neutral and negative).

We ran the script and we arrived at the following scores and sentiments in our dataset.



```
d.head(10)
```

	ticker	date	time	title	r_neg	r_neu	r_pos	prediction	sentiment
0	MSFT	2022-12-13	09:04PM	Fed Rate Hike Looms After Market Rally Fizzles...	0.087162	0.745230	0.167608	0.745230	Neutral
1	MSFT	2022-12-13	05:51PM	UPDATE 1-Microsoft says it offered FTC a conse...	0.023988	0.759662	0.216350	0.759662	Neutral
2	MSFT	2022-12-13	05:45PM	Microsoft (MSFT) Outpaces Stock Market Gains: ...	0.007726	0.379833	0.612441	0.612441	Positive
3	MSFT	2022-12-13	05:04PM	10 Best Performing Dividend ETFs in 2022	0.004978	0.442197	0.552826	0.552826	Positive
4	MSFT	2022-12-13	05:04PM	Microsoft Says It Offered FTC Consent Decree o...	0.020614	0.753957	0.225429	0.753957	Neutral
5	MSFT	2022-12-13	05:00PM	11 Mistakes That Stop Millennials From Buildin...	0.637994	0.334414	0.027592	0.637994	Negative
6	MSFT	2022-12-13	03:33PM	Microsoft says it offered FTC a consent decree...	0.029571	0.764458	0.205971	0.764458	Neutral
7	MSFT	2022-12-13	03:22PM	Microsoft says it offered to agree to FTC cons...	0.009436	0.601021	0.389543	0.601021	Neutral
8	MSFT	2022-12-13	12:52PM	Why Microsoft Stock Was Climbing Today	0.006841	0.350934	0.642225	0.642225	Positive
9	MSFT	2022-12-13	12:24PM	Best Dow Jones Stocks To Buy And Watch In Dece...	0.080076	0.623551	0.296373	0.623551	Neutral



```
d.tail(10)
```

	ticker	date	time	title	r_neg	r_neu	r_pos	prediction	sentiment
490	TSLA	2022-12-09	04:53AM	Musk's attention shift from Tesla to Twitter i...	0.550618	0.395494	0.053889	0.550618	Negative
491	TSLA	2022-12-09	04:08AM	Tesla to suspend Model Y output in Shanghai in...	0.115454	0.845044	0.039502	0.845044	Neutral
492	TSLA	2022-12-09	01:19AM	Musk says wise to avoid margin loans during ma...	0.209724	0.748381	0.041894	0.748381	Neutral
493	TSLA	2022-12-08	05:45PM	Tesla (TSLA) Stock Sinks As Market Gains: What...	0.565498	0.401720	0.032781	0.565498	Negative
494	TSLA	2022-12-08	05:14PM	Elon Musks Bankers Consider Tesla Margin Loans...	0.064836	0.869861	0.065303	0.869861	Neutral
495	TSLA	2022-12-08	05:01PM	Cathie Wood speaks on the Fed, energy, ARK ETF...	0.014876	0.864444	0.120680	0.864444	Neutral
496	TSLA	2022-12-08	04:41PM	Tesla's Troubles Are Piling Up While Elon Musk...	0.504501	0.453822	0.041677	0.504501	Negative
497	TSLA	2022-12-08	04:25PM	Tesla says its self-driving technology may be ...	0.191314	0.619935	0.188751	0.619935	Neutral
498	TSLA	2022-12-08	04:17PM	If Tesla Stock Keeps Dropping, Elon Musk Is in...	0.707733	0.264419	0.027848	0.707733	Negative
499	TSLA	2022-12-08	04:03PM	Tesla Stock Falters On Latest Report Of China ...	0.668622	0.315704	0.015674	0.668622	Negative

Fig 17&18: Results of the RoBERTa model

Analyzing the results of the RoBERTa model and taking into consideration the main drawback of the Vader model here as well, using the same rows used above (row 8 and row 490), the RoBERTa model returns the actual sentiment of the sentence if we use a baseline of human intuition. This shows that the RoBERTa model performs better where the Vader model falls short. In the next section, we will compare both results and highlight areas where the RoBERTa model performed better than Vader.

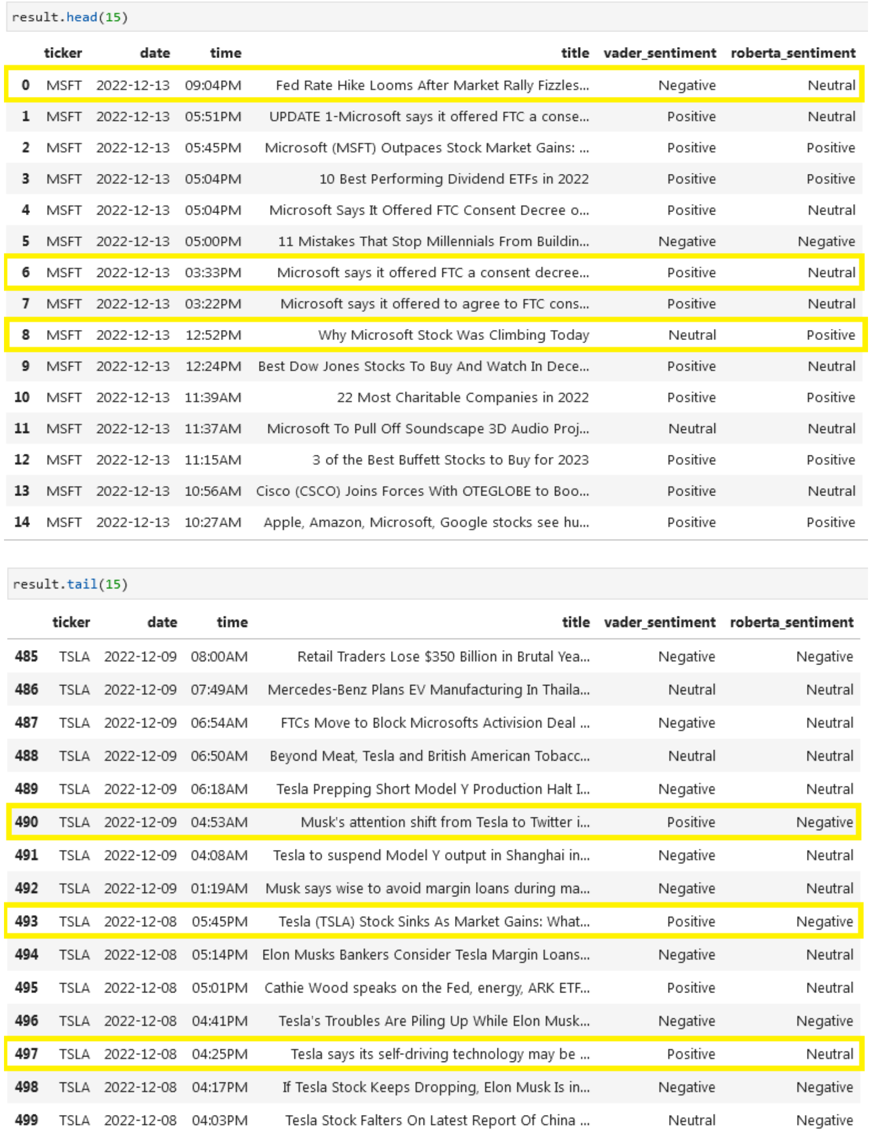
V. RESULT & DISCUSSIONS

5.1 Testing both models' performance:

Most sentiment analysis algorithm would categorize the data into positive/negative/neutral. For testing the accuracy, we use the rule of thumb to measure performance of each model in accordance with the intuition of the user. So, to test the performance of each model, we ask a group of students to rate each title on the dataset based on their own intuition and we found out that the RoBERTa model outdid the Vader model in terms of accuracy of sentiment. About 90% of student's sentiment correlated with the RoBERTa prediction compared to 65% of correlation with the Vader model.

5.2 Comparing both models' results:

Below is an analysis and comparison of both models with some parts highlighted to show how the RoBERTa model gave a better sentiment than the Vader model.



```
result.head(15)
```

	ticker	date	time	title	vader_sentiment	roberta_sentiment
0	MSFT	2022-12-13	09:04PM	Fed Rate Hike Looms After Market Rally Fizzles...	Negative	Neutral
1	MSFT	2022-12-13	05:51PM	UPDATE 1-Microsoft says it offered FTC a conse...	Positive	Neutral
2	MSFT	2022-12-13	05:45PM	Microsoft (MSFT) Outpaces Stock Market Gains: ...	Positive	Positive
3	MSFT	2022-12-13	05:04PM	10 Best Performing Dividend ETFs in 2022	Positive	Positive
4	MSFT	2022-12-13	05:04PM	Microsoft Says It Offered FTC Consent Decree o...	Positive	Neutral
5	MSFT	2022-12-13	05:00PM	11 Mistakes That Stop Millennials From Buildin...	Negative	Negative
6	MSFT	2022-12-13	03:33PM	Microsoft says it offered FTC a consent decree...	Positive	Neutral
7	MSFT	2022-12-13	03:22PM	Microsoft says it offered to agree to FTC cons...	Positive	Neutral
8	MSFT	2022-12-13	12:52PM	Why Microsoft Stock Was Climbing Today	Neutral	Positive
9	MSFT	2022-12-13	12:24PM	Best Dow Jones Stocks To Buy And Watch In Dece...	Positive	Neutral
10	MSFT	2022-12-13	11:39AM	22 Most Charitable Companies in 2022	Positive	Positive
11	MSFT	2022-12-13	11:37AM	Microsoft To Pull Off Soundscape 3D Audio Proj...	Neutral	Neutral
12	MSFT	2022-12-13	11:15AM	3 of the Best Buffett Stocks to Buy for 2023	Positive	Positive
13	MSFT	2022-12-13	10:56AM	Cisco (CSCO) Joins Forces With OTEGLOBE to Boo...	Positive	Neutral
14	MSFT	2022-12-13	10:27AM	Apple, Amazon, Microsoft, Google stocks see hu...	Positive	Positive

```
result.tail(15)
```

	ticker	date	time	title	vader_sentiment	roberta_sentiment
485	TSLA	2022-12-09	08:00AM	Retail Traders Lose $350 Billion in Brutal Yea...	Negative	Negative
486	TSLA	2022-12-09	07:49AM	Mercedes-Benz Plans EV Manufacturing In Thaila...	Neutral	Neutral
487	TSLA	2022-12-09	06:54AM	FTCs Move to Block Microsofts Activision Deal ...	Negative	Neutral
488	TSLA	2022-12-09	06:50AM	Beyond Meat, Tesla and British American Tobacc...	Neutral	Neutral
489	TSLA	2022-12-09	06:18AM	Tesla Prepping Short Model Y Production Halt I...	Negative	Neutral
490	TSLA	2022-12-09	04:53AM	Musk's attention shift from Tesla to Twitter i...	Positive	Negative
491	TSLA	2022-12-09	04:08AM	Tesla to suspend Model Y output in Shanghai in...	Negative	Neutral
492	TSLA	2022-12-09	01:19AM	Musk says wise to avoid margin loans during ma...	Negative	Neutral
493	TSLA	2022-12-08	05:45PM	Tesla (TSLA) Stock Sinks As Market Gains: What...	Positive	Negative
494	TSLA	2022-12-08	05:14PM	Elon Musks Bankers Consider Tesla Margin Loans...	Negative	Neutral
495	TSLA	2022-12-08	05:01PM	Cathie Wood speaks on the Fed, energy, ARK ETF...	Positive	Neutral
496	TSLA	2022-12-08	04:41PM	Tesla's Troubles Are Piling Up While Elon Musk...	Negative	Negative
497	TSLA	2022-12-08	04:25PM	Tesla says its self-driving technology may be ...	Positive	Neutral
498	TSLA	2022-12-08	04:17PM	If Tesla Stock Keeps Dropping, Elon Musk Is in...	Negative	Negative
499	TSLA	2022-12-08	04:03PM	Tesla Stock Falters On Latest Report Of China ...	Neutral	Negative

Fig 19: Results showing the comparison between the two models.

VI. CONCLUSION

Rule based approach are pre-developed manually and includes a lexicon method for making sentiment analysis. It has a drawback which makes the result of its accuracy not-so-effective. It ignores the context behind a sentence. One thing w noticed about rule-based approach is that it is prone to human bias. If the people preparing the lexicon or bags of words does not have enough domain knowledge, the model results may be inaccurate.

A pretrained model on the other end is a far better approach to implement in sentiment analysis. It takes into consideration the context behind a sentence or text. Due to the AI pretraining, it keeps learning in order to become more efficient than it was. It can also be trained to detect sarcasm, irony and negation which the rule-based approach cannot do. One major drawback or critique of the RoBERTa is the lack of its compound score which makes it difficult to properly visualize its results.

REFERENCES

1. *BMO - Personal Banking, Credit Cards, Loans & Investing*. (n.d.). https://www.bmo.com/main/personal
2. FINVIZ.com - Stock Screener. (n.d.-b). https://finviz.com/.
3. Text Analytics. (n.d.). MonkeyLearn. https://monkeylearn.com
4. Nemes, L., & Kiss, A. (2021). Prediction of stock values changes using sentiment analysis of stock news headlines. Journal of Information and Telecommunication, 5(3), 375–394. https://doi.org/10.1080/24751839.2021.1874252