# DAYANANDA SAGAR UNIVERSITY

**KUDLU GATE, BANGALORE – 560068**

**Bachelor of Technology**
**in**
**COMPUTER SCIENCE AND ENGINEERING**

## Major Project Phase-II Report

## USING SENTIMENT ANALYSIS TO EXAMINE STOCKS

By

**Ananya Jaikumar (ENG18CS0035)**

**Anmol Shreyas (ENG18CS0040)**

**Gauthami M (ENG18CS0104)**

**Shayiq Shafi Lone (ENG18CS0258)**

**Under the supervision of**
**Dr. Girisha G S**
**Professor & Chairman**
**Department of CSE**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING,**

**SCHOOL OF ENGINEERING**

**DAYANANDA SAGAR UNIVERSITY,**

**(2021-2022)**

DAYANANDA SAGAR UNIVERSITY

## School of Engineering
## Department of Computer Science & Engineering

Kudlu Gate, Bangalore – 560068
Karnataka, India

# CERTIFICATE

This is to certify that the Phase-II project work titled **"USING SENTIMENT ANALYSIS TO EXAMINE STOCKS"** is carried out by **Ananya Jaikumar (ENG18CS0035), Anmol Shreyas (ENG18CS0040), Gauthami M (ENG18CS0104), Shayiq Shafi Lone (ENG18CS0258),** bonafide students of Bachelor of Technology in Computer Science and Engineering at the School of Engineering, Dayananda Sagar University, Bangalore in partial fulfillment for the award of degree in Bachelor of Technology in Computer Science and Engineering, during the year **2021-2022**.

| | | |
|---|---|---|
| **Dr Girisha G S** | **Dr Girisha G S** | **Dr. A Srinivas** |
| Professor & Chairman CSE | Professor & Chairman CSE | Dean |
| School of Engineering | School of Engineering | School of Engineering |
| Dayananda Sagar University | Dayananda Sagar University | Dayananda Sagar University |
| Date: | Date: | Date: |

**Name of the Examiner** **Signature of Examiner**

1.

2.

# DECLARATION

We, **Ananya Jaikumar (ENG18CS0035), Anmol Shreyas (ENG18CS0040), Gauthami M (ENG18CS0104), Shayiq Shafi Lone (ENG18CS0258),** are students of the seventh semester B.Tech in **Computer Science and Engineering**, at School of Engineering, **Dayananda Sagar University**, hereby declare that the phase-II project titled **"USING SENTIMENT ANALYSIS TO EXAMINE STOCKS"** has been carried out by us and submitted in partial fulfillment for the award of degree in **Bachelor of Technology in Computer Science and Engineering** during the academic year **2021-2022**.

Student Signature

**Ananya Jaikumar:**

**ENG18CS0035**

**Anmol Shreyas:**

**ENG18CS0040**

**Gauthami M:**

**ENG18CS0104**

**Shayiq Shafi Lone:**

**ENG18CS0258**

# ACKNOWLEDGEMENT

It is a great pleasure for us to acknowledge the assistance and support of many individuals who have been responsible for the successful completion of this project work.

First, we take this opportunity to express our sincere gratitude to School of Engineering & Technology, Dayananda Sagar University for providing us with a great opportunity to pursue our Bachelor's degree in this institution.

We would like to thank **Dr. A Srinivas. Dean**, **School of Engineering & Technology**, **Dayananda Sagar University** for his constant encouragement and expert advice. It is a matter of immense pleasure to express our sincere thanks to **Dr. Girisha G S, Department Chairman**, **Computer Science, and Engineering**, **Dayananda Sagar University,** for providing the right academic guidance that made our task possible.

We would like to thank our guide **Dr. Girisha G S, Professor & Department Chairman**, **Dept. of Computer Science and Engineering**, **Dayananda Sagar University**, for sparing his/her valuable time to extend help in every step of our project work, which paved the way for smooth progress and the fruitful culmination of the project.

We would like to thank our Project Coordinators Dr. Meenakshi Malhotra and Dr. Bharanidharan N, and all the staff members of Computer Science and Engineering for their support.

We are also grateful to our family and friends who provided us with every requirement throughout the course. We would like to thank one and all who directly or indirectly helped us in the Project work.

# TABLE OF CONTENTS

Page

# NOMENCLATURE USED

| | |
|---|---|
| HTML | Hypertext Markup Language |
| Bs4 | Beautiful Soup |
| NLTK | Natural Language Toolkit |
| VADER | Valence Aware Dictionary and Sentiment Reasoning |

# LIST OF FIGURES

# LIST OF TABLES

# ABSTRACT

The stock market is one of the most sensitive fields, where sentiments of the people can change the trend of the entire market. Actually, there are many factors, affect the movement of the stock market and, the sentiments of the traders are also one of them drive the market.

In present times, stock market investment plays an inevitable role in the finance sector as high stock market value is considered the parameter of high economies. The volatile nature of the stock market has equal chances for earning money and losing money as well. But if the situation can be predicted, investors can make a profit or minimize their losses.

Hence, AI companies are now using sentiment analysis in the stock market to predict the market trend or movement of a particular stock. The aim of the project is to predict if the stock price of a company will increase or decrease based on top news headline

# CHAPTER 1

# INTRODUCTION

# CHAPTER 1 INTRODUCTION

## 1.1 Purpose

Stock sentiment analysis can be used to determine investors' opinions of a specific stock or asset. Sentiment may at times hint at future price action. There are various factors that influence stock sentiment, which include news (economic, political and industry related) and social media. These factors help influence stock sentiment as they impact stock market volatility, trading volume and company earnings.

## 1.2 Intended audience

The intended audience that this project will be targeting would be:

- Investors
- Stock market traders
- General Audience
- Investor firms
- Beginner Investors

## 1.3 Intended Use

This project mainly aims to target the market by analyzing the sentiment of the stocks/shares from the given news headlines, data of previous prices of shares and predict the end sentiment and the closing price of the stocks efficiently.

## 1.4 Scope or Product Scope

The project manly focuses on:

- Sentiment Analysis
- Descriptive Analysis of the Stock
- Machine Learning

# CHAPTER 2

# PROBLEM DEFINITION

# CHAPTER 2 PROBLEM DEFINITION

- "Stock Price Prediction Using Sentiment Analysis" method is used in this project for predicting stock prices using news articles. The changes in stock prices of a company, the rises and falls, are correlated with the public opinions being expressed by the people in various platforms about that company.
- In this project we mainly implement how to utilize new headlines to assess market sentiment and predict the behavior of stock of certain company, market and stock indexes to identify an opportunity for trading
- A prediction model for finding and analyzing correlation between contents of public opinion and stock prices and then making predictions for future prices can be developed by using machine learning.

# CHAPTER 3

# LITERATURE REVIEW

# CHAPTER 3 LITERATURE SURVEY

| Authors, Title, Year of Publication | Methodology | Outcomes | Limitations |
|---|---|---|---|
| **Authors**: Shri Bharathi and Angelina Geetha **Year of Publication**: 2017 **Title**: Sentiment Analysis for Effective Stock Market Prediction | The main aim of this research paper is to provide a detailed analysis using different methodologies of stock market forecasting to predict future trends and stock returns. | Two types of results are achieved in the proposed system. One is without sentiment the moving average value is calculated and second one is with sentiment the moving average is calculated. Both the results will show the improvements in the stock market forecasting. | More refinements are needed for the current processing steps, such as using the sentiment analysis in the news pre-processing to construct higher level features. |
| **Authors:** Isaac Kofi Nti1, Adebayo Felix Adekoya, Benjamin Asubam Weyori **Year of Publication:** 2020 **Title:** Predicting Stock Market Price Movement Using Sentiment Analysis: Evidence From Ghana | Using web news, financial tweets posted on Twitter, Google trends and forum discussions, the current study examines the association between public sentiments and the predictability of future stock price movement using Artificial Neural Network (ANN). | The experimental setup with stock data (January 2010 to September 2019) of three (3) companies listed on the Ghana stock exchange shows that the stock market is predictable using public sentiments. | The limited size of the dataset obtained from Twitter shows that investors in developing countries such as Ghana hardly share their views on SNSs concerning market trends. Hence, it makes it insufficient to wholly depend on public sentiment from a single source to predict stock market movement in such countries. |

# CHAPTER 4

# PROJECT DESCRIPTION

# CHAPTER 4 PROJECT DESCRIPTION

## 4.1. Proposed Design:

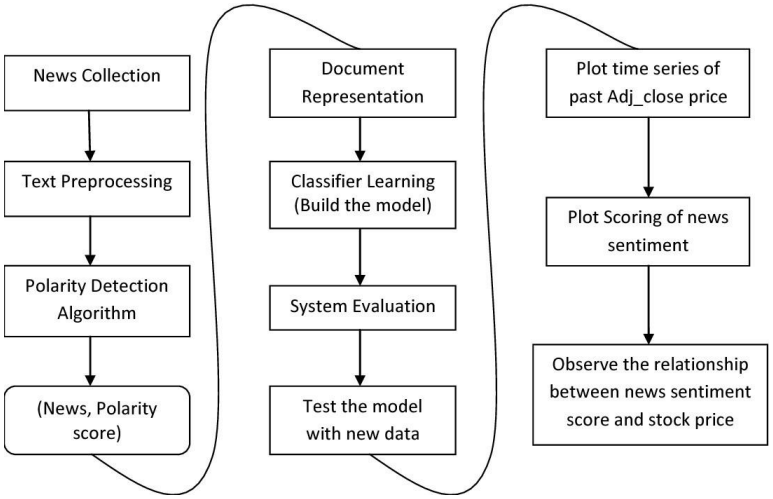



Figure 1 : Proposed Block Diagram

1.Polarity Detection Algorithm – VADER uses a combination of A sentiment lexicon is a list of lexical features which are generally labeled according to their semantic orientation as either positive or negative.

2.We compare the predicted closing prices and actual closing prices to find the accuracy of the model.

## 4.2. Assumptions and Dependencies:

- Assume the user knows the names of the stocks in the ticker form.

# CHAPTER 5

# REQUIREMENTS

# CHAPTER 5 REQUIREMENTS

## 5.1 Functional Requirements:

1. Vader Sentiment Analyzer – It specifies the positivity and negativityscore and also talks about how positive or negative a sentiment is.
2. Beautiful Soup - To scrape headlines from Finviz website.
3. Pandas (Python Data Analysis Library) – to analyze and run sentimentanalysis on the article headlines.
4. Plotly – for visualization of our results
5. Flask – For Web Application

## 5.2 Non-Functional Requirements:

1. Correctness – the ability to predict the sentiment of theheadlines
2. Availability – it should be available for general audience to use
3. Ease of use – the project should be easy to use even for beginners

## 5.3 Hardware Requirements:

1. Processor - Intel core i3 and above series.
2. RAM - 4GB and above.

## 5.4 Software Requirements:

1. Python Interpreter
2. ML modules like pandas, ploty, vadersentiment, textblob, regular expression and flask

# CHAPTER 6

# METHODOLOGY

# CHAPTER 6 METHODOLOGY

This project is a beginner-friendly Python and Data Science application focused on building a script to analyze the sentiment of news articles of stocks on FinViz. Use the Requests module in Python, parse the HTML returned in BeautifulSoup and get the Article Data, apply Sentiment Analysis on the data with NLTK and finally visualize the results of our data frame in Pandas with MatPlotLib

- Yse Beautifulsoup in Python to scrape article headlines from FinViz
- FinViz is a free website that makes stock data easily accessible to traders and investors. Gather the stock data from FinViz for a specific stock ticker.
- Then, use Pandas (Python Data Analysis Library) to analyze and runsentiment analysis on the article headlines
- Applying sentiment analysis on the titles is actually the easiest part of the entire project. With NLTK (Natural Language Toolkit) comes a beautiful submodule called VADER that allows us to pass in a string into its function andget back a funky looking result.
- Finally, use Plotly for visualization of the results
- To visualize this data frame in Plotly to see how the Stocks fared every day and every hour from public perception in news articles.
- Build a dashboard web app showing the stock sentiment scores using Python Flask.

# CHAPTER 7

# EXPERIMENTATION

# CHAPTER 7 EXPERIMENTATION

**Web Scrapping Algorithm**

```
req = Request(url=url, headers={'user-agent': 'my-app'})
response = urlopen(req)
html = BeautifulSoup(response, features='html.parser')
news_table = html.find(id='news-table') news_tables[ticker] =
news_table
```

**Sentiment Analysis Algorithm**

```
vader = SentimentIntensityAnalyzer()
f = lambda title: vader.polarity_scores(title)['compound']df['compound']
= df['title'].apply(f)
```

**Creating a Flask App**

```
app = Flask(__name__)
```

**get_news()**

The get_news() function takes in a ticker name, uses the urllib3 Python library to obtain relevant news headlines from the FinViz website, it outputs the raw html code.

**parse_news()**

The parse_news() function takes in the raw HTML code from above and uses the BeautifulSoup library to parse the headlines and corresponding dates/times into a Pandas DataFrame.

**score_news()**

The score_news() function takes in the DataFrame from above and gives sentiment score for each headline using the NLTK Vader library. It outputs the sentiment scores (negative, neutral, positive, aggregated score) as additional columns in the DataFrame.

**plot_hourly_sentiment() and plot_daily_sentiment()**

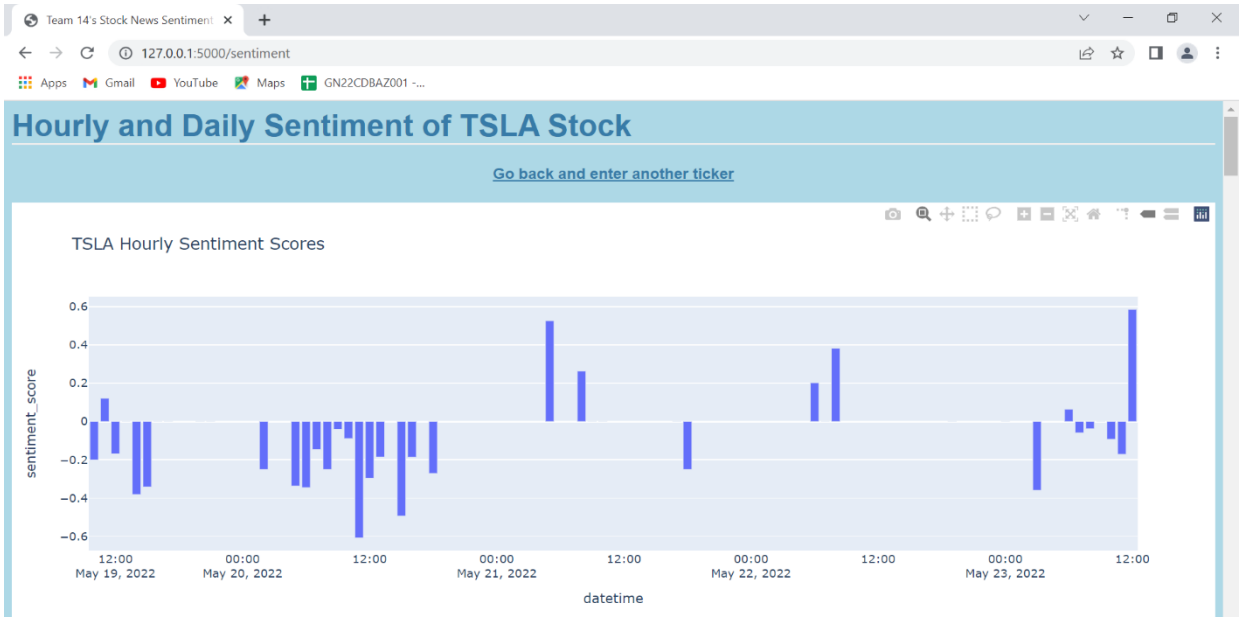
Finally the plot_hourly_sentiment() and plot_daily_sentiment() functions resamples the respective hourly and daily sentiment scores in the DataFrame above, and plot it out in an interactive chart using the Plotly package.

# CHAPTER 8

# TESTING AND RESULTS

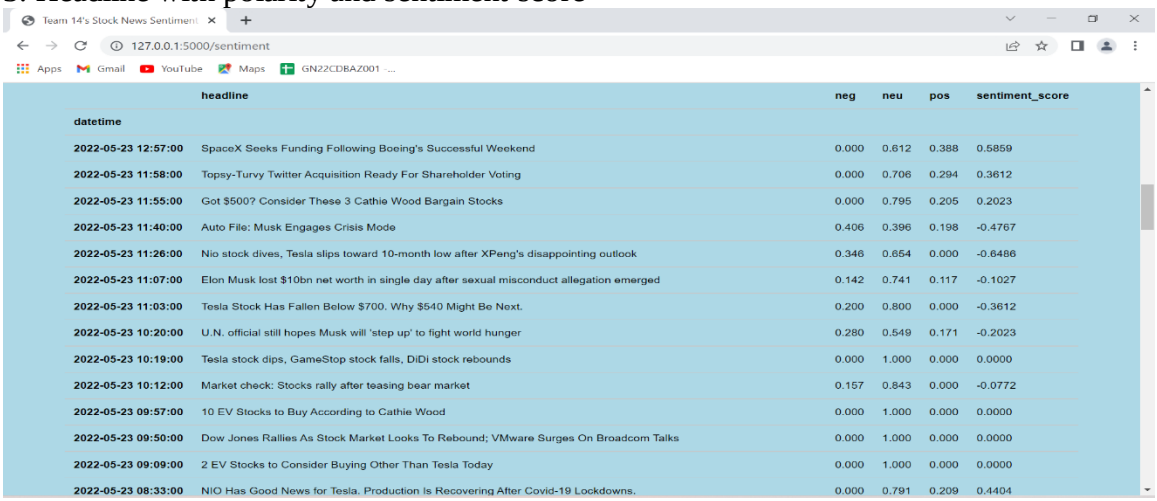
# CHAPTER 8 TESTING AND RESULTS

1. Hourly Sentiment Scores




2. Daily Sentiment Scores

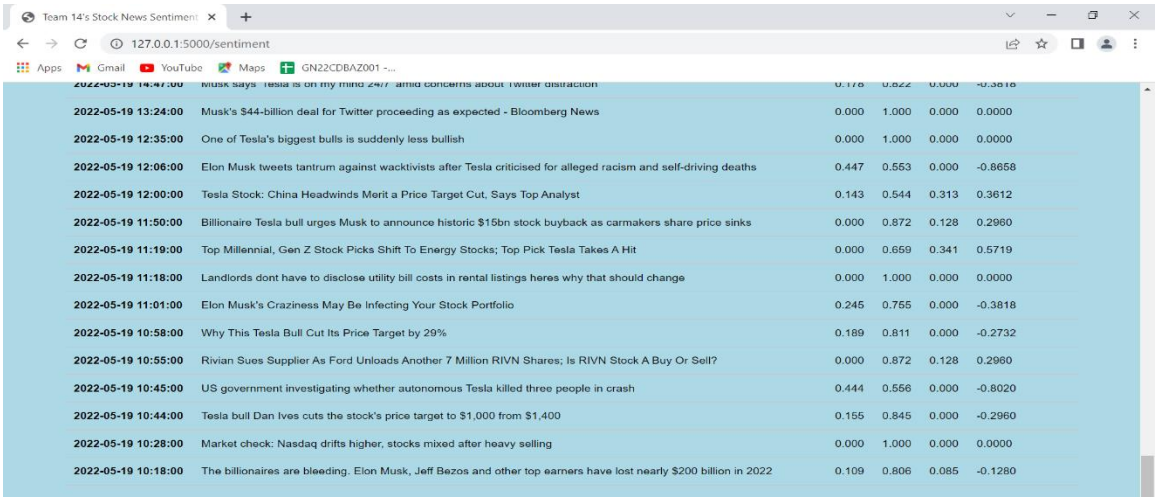
## 3. Headline with polarity and sentiment score


Team 14's Stock News Sentiment

127.0.0.1:5000/sentiment

Apps Gmail YouTube Maps GN22CDBAZ001 -...

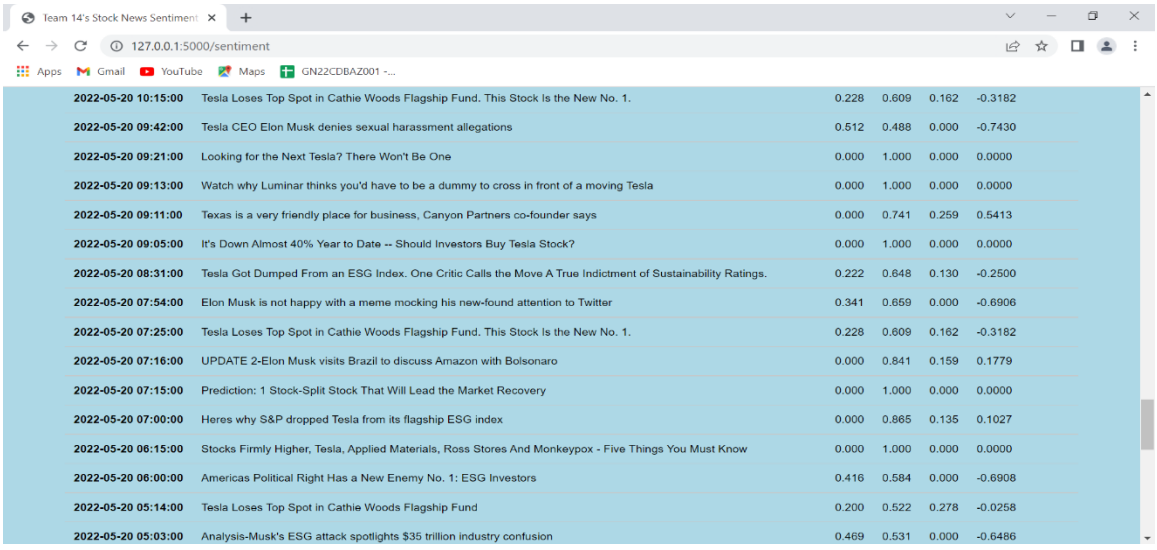
| datetime | headline | neg | neu | pos | sentiment_score |
|---|---|---|---|---|---|
| 2022-05-23 12:57:00 | SpaceX Seeks Funding Following Boeing's Successful Weekend | 0.000 | 0.612 | 0.388 | 0.5859 |
| 2022-05-23 11:58:00 | Topsy-Turvy Twitter Acquisition Ready For Shareholder Voting | 0.000 | 0.706 | 0.294 | 0.3612 |
| 2022-05-23 11:55:00 | Got $500? Consider These 3 Cathie Wood Bargain Stocks | 0.000 | 0.795 | 0.205 | 0.2023 |
| 2022-05-23 11:40:00 | Auto File: Musk Engages Crisis Mode | 0.406 | 0.396 | 0.198 | -0.4767 |
| 2022-05-23 11:26:00 | Nio stock dives, Tesla slips toward 10-month low after XPeng's disappointing outlook | 0.346 | 0.654 | 0.000 | -0.6486 |
| 2022-05-23 11:07:00 | Elon Musk lost $10bn net worth in single day after sexual misconduct allegation emerged | 0.142 | 0.741 | 0.117 | -0.1027 |
| 2022-05-23 11:03:00 | Tesla Stock Has Fallen Below $700. Why $540 Might Be Next. | 0.200 | 0.800 | 0.000 | -0.3612 |
| 2022-05-23 10:20:00 | U.N. official still hopes Musk will 'step up' to fight world hunger | 0.280 | 0.549 | 0.171 | -0.2023 |
| 2022-05-23 10:19:00 | Tesla stock dips, GameStop stock falls, DiDi stock rebounds | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-23 10:12:00 | Market check: Stocks rally after teasing bear market | 0.157 | 0.843 | 0.000 | -0.0772 |
| 2022-05-23 09:57:00 | 10 EV Stocks to Buy According to Cathie Wood | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-23 09:50:00 | Dow Jones Rallies As Stock Market Looks To Rebound; VMware Surges On Broadcom Talks | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-23 09:09:00 | 2 EV Stocks to Consider Buying Other Than Tesla Today | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-23 08:33:00 | NIO Has Good News for Tesla. Production Is Recovering After Covid-19 Lockdowns. | 0.000 | 0.791 | 0.209 | 0.4404 |


Team 14's Stock News Sentiment

127.0.0.1:5000/sentiment

Apps Gmail YouTube Maps GN22CDBAZ001 -...

| | | | | | |
|---|---|---|---|---|---|
| 2022-05-19 13:24:00 | Musk's $44-billion deal for Twitter proceeding as expected - Bloomberg News | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-19 12:35:00 | One of Tesla's biggest bulls is suddenly less bullish | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-19 12:06:00 | Elon Musk tweets tantrum against wacktivists after Tesla criticised for alleged racism and self-driving deaths | 0.447 | 0.553 | 0.000 | -0.8658 |
| 2022-05-19 12:00:00 | Tesla Stock: China Headwinds Merit a Price Target Cut, Says Top Analyst | 0.143 | 0.544 | 0.313 | 0.3612 |
| 2022-05-19 11:50:00 | Billionaire Tesla bull urges Musk to announce historic $15bn stock buyback as carmakers share price sinks | 0.000 | 0.872 | 0.128 | 0.2960 |
| 2022-05-19 11:19:00 | Top Millennial, Gen Z Stock Picks Shift To Energy Stocks; Top Pick Tesla Takes A Hit | 0.000 | 0.659 | 0.341 | 0.5719 |
| 2022-05-19 11:18:00 | Landlords dont have to disclose utility bill costs in rental listings heres why that should change | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-19 11:01:00 | Elon Musk's Craziness May Be Infecting Your Stock Portfolio | 0.245 | 0.755 | 0.000 | -0.3818 |
| 2022-05-19 10:58:00 | Why This Tesla Bull Cut Its Price Target by 29% | 0.189 | 0.811 | 0.000 | -0.2732 |
| 2022-05-19 10:55:00 | Rivian Sues Supplier As Ford Unloads Another 7 Million RIVN Shares; Is RIVN Stock A Buy Or Sell? | 0.000 | 0.872 | 0.128 | 0.2960 |
| 2022-05-19 10:45:00 | US government investigating whether autonomous Tesla killed three people in crash | 0.444 | 0.556 | 0.000 | -0.8020 |
| 2022-05-19 10:44:00 | Tesla bull Dan Ives cuts the stock's price target to $1,000 from $1,400 | 0.155 | 0.845 | 0.000 | -0.2960 |
| 2022-05-19 10:28:00 | Market check: Nasdaq drifts higher, stocks mixed after heavy selling | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-19 10:18:00 | The billionaires are bleeding. Elon Musk, Jeff Bezos and other top earners have lost nearly $200 billion in 2022 | 0.109 | 0.806 | 0.085 | -0.1280 |


Team 14's Stock News Sentiment

127.0.0.1:5000/sentiment

Apps Gmail YouTube Maps GN22CDBAZ001 -...

| | | | | | |
|---|---|---|---|---|---|
| 2022-05-20 10:15:00 | Tesla Loses Top Spot in Cathie Woods Flagship Fund. This Stock Is the New No. 1. | 0.228 | 0.609 | 0.162 | -0.3182 |
| 2022-05-20 09:42:00 | Tesla CEO Elon Musk denies sexual harassment allegations | 0.512 | 0.488 | 0.000 | -0.7430 |
| 2022-05-20 09:21:00 | Looking for the Next Tesla? There Won't Be One | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-20 09:13:00 | Watch why Luminar thinks you'd have to be a dummy to cross in front of a moving Tesla | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-20 09:11:00 | Texas is a very friendly place for business, Canyon Partners co-founder says | 0.000 | 0.741 | 0.259 | 0.5413 |
| 2022-05-20 09:05:00 | It's Down Almost 40% Year to Date -- Should Investors Buy Tesla Stock? | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-20 08:31:00 | Tesla Got Dumped From an ESG Index. One Critic Calls the Move A True Indictment of Sustainability Ratings. | 0.222 | 0.648 | 0.130 | -0.2500 |
| 2022-05-20 07:54:00 | Elon Musk is not happy with a meme mocking his new-found attention to Twitter | 0.341 | 0.659 | 0.000 | -0.6906 |
| 2022-05-20 07:25:00 | Tesla Loses Top Spot in Cathie Woods Flagship Fund. This Stock Is the New No. 1. | 0.228 | 0.609 | 0.162 | -0.3182 |
| 2022-05-20 07:16:00 | UPDATE 2-Elon Musk visits Brazil to discuss Amazon with Bolsonaro | 0.000 | 0.841 | 0.159 | 0.1779 |
| 2022-05-20 07:15:00 | Prediction: 1 Stock-Split Stock That Will Lead the Market Recovery | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-20 07:00:00 | Heres why S&P dropped Tesla from its flagship ESG index | 0.000 | 0.865 | 0.135 | 0.1027 |
| 2022-05-20 06:15:00 | Stocks Firmly Higher, Tesla, Applied Materials, Ross Stores And Monkeypox - Five Things You Must Know | 0.000 | 1.000 | 0.000 | 0.0000 |
| 2022-05-20 06:00:00 | Americas Political Right Has a New Enemy No. 1: ESG Investors | 0.416 | 0.584 | 0.000 | -0.6908 |
| 2022-05-20 05:14:00 | Tesla Loses Top Spot in Cathie Woods Flagship Fund | 0.200 | 0.522 | 0.278 | -0.0258 |
| 2022-05-20 05:03:00 | Analysis-Musk's ESG attack spotlights $35 trillion industry confusion | 0.469 | 0.531 | 0.000 | -0.6486 |

# REFERENCES


1. Predicting Stock Market Price Movement Using Sentiment Analysis: Evidence From Ghana Isaac Kofi Nti1, 2* , Adebayo Felix Adekoya3 , Benjamin Asubam Weyori4
2. Sentiment Analysis for Effective Stock Market PredictionShri Bharathi1* Angelina Geetha2
3. Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty Aditya Bhardwaja*, Yogendra Narayanb , Vanrajc , Pawana , Maitreyee Duttaa
4. K. Mizumoto, H. Yanagimoto and M. Yoshioka, Sentiment Analysis of Stock Market News with Semi-Supervised Learning,In 2012 IEEE/ACIS11th International Conference on Computer and Information Science (ICIS), pp. 325–328, May (2012).
5. J. Bollen, H. Mao and X. Zeng, Twitter Mood Predicts the Stock Market,Journal of Computational Science, vol. 2, no. 1, pp. 1–8, (2011).

# GITHUB LINK

https://github.com/ananya-jaikumar/Major_Project_Team14