

A Project Report
On
The Buzzing Stocks Using NLP

*Submitted in partial fulfillment of the requirement for
the award of the degree of*

B.Tech AI & ML



(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

Under The Supervision
Of
Dr. A. Suresh Kumar

Designation
Assistant Professor

Submitted By

Vivek Singh

18SCSE1180050

Yashvardhan Singh

18SCSE1180048

**SCHOOL OF COMPUTING SCIENCE AND ENGINEERING DEPARTMENT OF COMPUTER
SCIENCE AND ENGINEERING GALGOTIAS UNIVERSITY, GREATER NOIDA, INDIA**

December,
2021-2022



**SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA**

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the thesis/project/dissertation, entitled “**The Buzzing Stocks Using NLP**” in partial fulfillment of the requirements for the award of the B. tech submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of September 2021 to December 2021, under the supervision of Dr. A. Suresh Kumar, Assistant Professor, Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering , Galgotias University, Greater Noida

The matter presented in the thesis/project/dissertation has not been submitted by me/us for the award of any other degree of this or any other places.

Vivek Singh/ 18SCSE1180050
Yashvardhan Singh/ 18SCSE1180048

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr. A. Suresh Kumar
Assistant Professor

CERTIFICATE

The Final Thesis/Project/ Dissertation Viva-Voce examination of Vivek Singh / 18SCSE1180050 and Yashvardhan Singh / 18SCSE1180048 has been held on _____ and his/her work is recommended for the award of B. tech in Computer Science & Engineering.

Signature of Examiner(s)

Signature of Supervisor(s)

Signature of Project Coordinator

Signature of Dean

Date: December, 2021
Place: Greater Noida

Acknowledgement

We take this opportunity to express our sincere gratitude to **Prof. Dr. A. Suresh Kumar Assistant Professor**, Department of CSE **GALGOTIAS UNIVERSITY**, GREATER NOIDA. Deep Knowledge & keen interest of our supervisor in the field of “*Web Application and Data Science*” to carry out this project. His endless patience, scholarly guidance, strong motivation, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude other faculty member and the staff of CSE department of **GALGOTIAS UNIVERSITY**, GREATER NOIDA to finish our project.

We would like to thank our entire course mate in **GALGOTIAS UNIVERSITY**, GREATER NOIDA, who took part in this discuss while completing the course work.

Abstract

Problem Stated

Stocks? What do you mean by stocks? It is a security which gives you stockholders a small share of ownership in company. Companies issues stocks to get money for paying debts, launching new products, expanding new markets etc. If you are a beginner and wanted to know about stocks but doesn't know how it works due to lack of knowledge about the market you will definitely require something which will save your time and give you accurate picture of today's market

Problem Solution

Our web app will provide End-to-End news updates happening around the world that lets safely invest in the companies you wanted to with visiting thousands of sites just by one click and go. Without any worrying about money. In addition to the protection of your funds. The proposed Application allows the users to communicate and gain knowledge via news and notifications.

This paper aims to develop an online aggregator of news feed from the market finance such as Yahoo Finance etc. The most important factors influencing the stock market are stock issues. This study is an attempt to create a model that predicts how you can use NEL to improve the stock market news supply of online stock exchanges and news polarity that could affect changes in stock price trends. The novelty of our approach is to accurately predict the sensory model in all real-time news that is available in the public domain and getting insights from a raw and unstructured data is very important. Our approach uses emotion analysis in Really Simple Syndication (RSS) news feeds that impact on stock market prices. Therefore RSS news feed data is collected over a period of time. This proposed paper uses an emotion analysis algorithm to generate emotion in RSS news feeds. In our experimental study, the algorithm approach suggests that stock news sentiment can be used to predict a decline in stock prices, either up or down.

KEYWORDS NAMES: Sentiment Analysis, Opinion Mining, RSS Reader, Stock Market News.

Tools and Technology

Technologies or software used for this app are VS Code, Spacy, Jupyter Notebook or Google Colab, Streamlit.

Result

It will be cross-platform application with up-to-date stocks news and all data of the market etc.

KEYWORDS NAMES: Sentiment Analysis, Opinion Mining, RSS Reader, Stock Market News.

Contents

Title	Page No.
Candidates Declaration	I
Acknowledgement	II
Abstract	III
Contents	IV
List of Table	V
List of Figures	VI
Acronyms	VII
Chapter 1 Introduction	10-11
1.1 PROBLEM FORMULATION	12
1.2 PROJECT CATEGORY	12
1.3 TOOLS AND TECHNOLOGY	13-14
1.4 LITERATURE REVIEWS	15-16
1.5 METHODOLOGY	17
1.6 FEASIABILITY	18
Chapter 2 Project Design	
2.1 UML DIAGRAM	19-20
2.2 MVVM	21-22
2.3 DATA RELATED TO STOCKS	23
2.4 DESIGN	24-26
2.5 WORKING	27-29
Chapter 3 TERMS	30-31
3.1 CODE	32-42
Chapter 4 Results and Discussion	43-49
Chapter 5 Conclusion and Future Scope	50
Chapter 6 Reference	51
Chapter 7 Publication/Copyright/Product	52

List of Table

S.No	Enrollment Number	Admission Number	Student Name	Degree / Branch	Sem
1.	18021180049	18SCSE1180050	Vivek Singh	B.Tech/CSE	7
2.	18021180047	18SCSE1180048	Yashvardhan Singh	B.Tech/CSE	7

List of Figures

S.No.	Title	Page No.
1	UML DIAGRAM-1	19
2	UML DIAGRAM-2	20
3	MVVM	22
4	APPLICATION DESIGN	24
5	HEADINGS	25-26
6	RSS link	27
7	TABULAR DATA	28
8	HEADING BAR	29
9	RESULT	43-49

Acronyms

APP	Application
DB	Database
IDE	Integrated development environment
RSS	Really Simple Syndication
NEL	Name Entity Linking
NLP	Natural Language Processing
NER	Name Entity Recognition

1. INTRODUCTION

The most important factors influencing the stock market are stock news. This study is an attempt to create a model that predicts news coherence that could affect changes in stock price trends. The novelty of our approach is to accurately predict the sensory model in all real-time news that is available in the public domain. The Stock Market has more than 2000 stocks and knowing news and financial data of all of them is impossible. It is very important for one to pick a sector of the market and make well informed investments or trading. This project helps one to have a custom feed that allows the user to get specific news with the market rate and data. It picks all the relevant headlines from various platforms and runs it through the market database. Then the user gets to compare or analyze the impact of the trending news and the rates of the stock. We help our traders to have a hassle free research and investing experience.

Getting insights from raw and unstructured data is of vital importance. Uploading a document and getting the important bits of information from it is called information retrieval. Information retrieval has always been a major task and challenge in NLP. And we can use NER (or NEL—Named Entity Linking) in several domains like finance, drug research, e-commerce, and more for information retrieval purposes.

We'll get the textual data from Really Simple Syndication-RSS news feeds on the internet and extract the names of buzzing stocks. We'll then pull their market price data to test the authenticity of the news before taking any position in those stocks.

[1] The application delivers few of its unique features as briefed below:

-
1. It is time efficient basically reduces the time to visit websites by bringing all the information into one single web application. Public can quickly access RSS feeds without having to visit each website;
 2. Easy-to-use, and smooth tabular display. While some existing media editors may offer very complex connectors and advanced settings, it can be very difficult to use them, especially for a non-technical person. The improved connector comes with a solution to this problem, relying on a simple but complete visual interface. At the same time, it provides a responsive web interface, which is available on both large screens and mobile devices;
 3. Spam removal – most of the applications offer spam and ad for articles for profit generation. It does not alter the user experience by adding advertisements;
 4. It has the ability to filter and sort the content according to user-defined criteria.
-

In rundown, we have recognized the accompanying difficulties while picturing information identified with the continuous change in the stock price:

The obstacles in trend analysis: The [2] Stocks and Yahoo Finance sites embraced a novel methodology of relating the all outnumber of pricing with the range of a cloudy circle and setting it at the market area. While this web application can introduce a preview of present status, it doesn't act as a financial adviser. The conventional pattern analysis are regularly not intelligent and show a comparable pattern for some banks in a jumbled manner.

1.1 PROBLEM FORMULATION

To monitor the market pricing and all the data which are influenced by day to day change in the uplift-ment and downfall of the share market price. It is very difficult for an individual to check the accurate pricing rate up to date and screen the development of the influenced funds, share market, and stocks. Along these lines, to track the stocks, we should attempt to assemble a web application that powerfully keeps the record and updates the information of the data regarding the market pricing and aware the individuals in the whole world.

1.2 PROJECT CATEGORY

We have created the buzzing stocks news feed to serve as a tool for artificial employees to get information about the stocks. The project is based on the finance category as well as this project also belongs to the category of educational research. The idea of the project is to build a dynamic web application. The following issues need to be addressed to address this issue: How is it useful? While funding or putting your money in stocks you can get info about the market in that particular area of expertise/stock when you are going to enter the stock market, to avoid such circumstances like losing your funds.

1.3 TOOLS AND TECHNOLOGY

1) STREAMLIT

Streamlit is an open python framework for building Machine Learning and Data Science web applications. We can quickly improve web applications and run them easily using Streamlit. Streamlit lets you write an application in the same way you write a python code. Streamlit makes it easy to work in loops of code interaction and viewing results in a web application. Developed in San Francisco, California and was founded in **2018** by Adrien Treuille, Amanda Kelly, and Thiago Teixeira

2) JUPYTER NOTEBOOK

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. Its uses include data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

3) VS Code Editor

Visual Studio Code is a source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git.

4) GOOGLE COLAB

Colaboratory, or “Colab” for short, is a product from Google Research. Colab **allows anybody to write and execute arbitrary python code through the browser**, and is especially well suited to machine learning, data analysis and education.

5) SPACY

spaCy is a **free, open-source library** for advanced **Natural Language Processing** (NLP) in Python.

If you're working with a lot of text, you'll eventually want to know more about it. For example, what's it about? What do the words mean in context? Who is doing what to whom? What companies and products are mentioned? Which texts are similar to each other?

spaCy is designed specifically for **production use** and helps you build applications that process and "understand" large volumes of text. It can be used to build **information extraction** or **natural language understanding** systems, or to pre-process text for **deep learning**.

1.4 LITERATURE REVIEWS

Stock markets are, without a doubt, the most important and vital component of the world economy. However, the impact of stock markets on a country's economy may differ from the way foreign stock markets affect their economies. This is because the impact of the stock market on the economy depends on a variety of factors such as the structure of the stock market, its relationships with other components of the financial system, the country's governance system etc. All of these things are different in each country; therefore, the impact of the stock market on the national economy is also different. Over the years, the market capitalization of the Indian capital has undergone significant changes in institutions that have led to a reduction in transaction costs, significant improvements in efficiency, visibility and security. All these changes have brought about economic development through the stock market. In the same way, economic growth resulting from changes in technology, products and new jobs is expected to create a high demand for stock market development.

A. **[3]Rakesh H.M (2014), A Study On Individual Investors Behavior In Stock Markets Of India, IJMSS (Vol.02, Issue-02), ISSN:2321-1784:**

This paper proposes to study the behavior of individual investors in the stock market and factors that influence their investment decisions, including awareness level, investment time etc. The research was based on primary data collected in the city of Mysore for 150 respondents, being investors in the stock market. The research paper finds that only 10% of respondents intend to regularly invest in the stock market for more than five years. In other words, the research paper has seen that people do not want to stay committed to the stock market for as long as you can giving a better return. This paper analyzes the annual income and savings annually value by investors, but the level of savings is determined by their level of income. He says that “investors are fully aware of the stock market and feel that market movements are also affects how investors invest in stock markets.” This paper however is remarkably silent on uneducated investors who are absent is familiar with market conditions, market trends and stock price movements. Focuses on savings and information resources in decision-making. Income rate of a person, and determine the investor’s investment pattern. Investor income level determines the type of investment methods the investor chooses to pursue.

B. [4] Reena Rai (2014), Factors Affecting Investors' Decision Making Behavior In The Stock Market: An Analytical Review, Indian Journal of Applied Research (Vol.4, Issue-9), ISSN - 2249-555X:

The study paper aims to study factors that influence investor decision morality on the basis of related subjects. It says various influential factors include a variety of people's factors such as gender, age, education. It is known that such men overconfidence than women. Age plays a role in the human mind as well as the tendency to take risks. It also describes a sometimes, alert attitude and conservatism. On the strongest level of investor decision depends on the capitalization of inflation, politics and media exposure, trend analysis, previous performance of company shares, expected shares and EPS etc. Finally, it concludes that there are various factors that affect the behavior of investors some factors play a small role while others have a significant impact on investor behavior. General features being gender, age, confidence levels, psychological prejudice, risk factors, company performance.

1.5 METHODOLOGY

Have you ever thought of saving money, have you ever thought of investing your money in stocks but afraid due to insufficient knowledge of the market pricing so here we have brought away where you can simply just understand the market pricing with the help of the stock market news. We have used NLP to make a chain between every website to bring them together for the person who really wants to understand the real meaning of stocks by not visiting thousands of websites but by using the BUZZING STOCK NEWS FEED application.

The approach for the problem is very simple using NLP we have created a way to understand the RSS feed and provides you with the main heading or titles on stocks which are trending on the following links mentioned in the RSS feed. This is a small revolution in the field of stocks news where you can find all the details of the market in one go. Now the question arises why BSNF application? We have tried to make things easier for an individual where he/she wanting to invest can invest with authentic knowledge in the domain.

To understand the whole process of the BSNF application all the gathered data is from authentic websites with their XML files. Then linked it with NIFTY500 and yfinance to get the records of the companies and update headlines. And showcase the data in a tabular form for better understanding.

Streamlit is an open source python library used for making the application and the toughest part to be solved as it was a very new tech for us. It took time for us to understand the working of streamlit library and another obstacle was to link the RSS feed using NLP with perfect response as it may give some wrong data which will degrade the whole purpose of the application.

1.6 FEASIBILITY

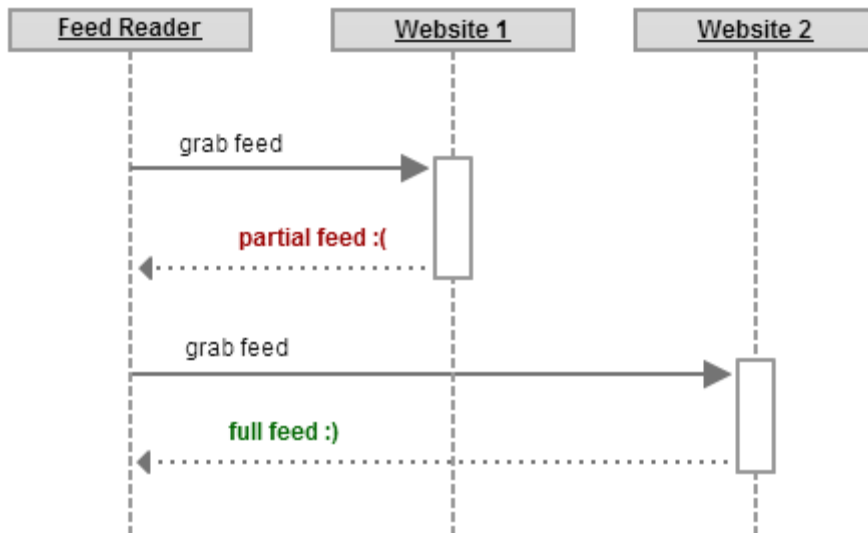
India today is one of the fastest developing countries. We have a population of 1.5 billion people and have one of the highest youth populations in the world. A few shocking facts about the Indian stock market are that only 3 to 4 per cent of the population invest in the stock market. In the recent events of the pandemic, people realised the real meaning of earning and growing money. Previously the country was in a state of earning and saving. This changed in this period because the whole country realised the true importance of money. We have had a massive jump in the numbers of Demat accounts in the last few months. Everyone has adapted to growing their primary income through investing in the Stock market. One of the biggest barriers that stood before and now was that the market timing completely clashes with everyone's professional working hours. Thus, we need to provide them with tools and platforms that help them save time and allow them to invest more wisely. The stock market has two important factors. The first one is the financials of the company that come out quarterly and yearly. The second one is the news. The Indian market is so big and the value of the stock changes due to so many other factors that it gets really time-consuming for people to do the research or even keep up with the news. Hence the use of NLP to fetch important and most recent news on the basis of your choices or the sector in which one would like to invest could be a very effective and great tool to buy. This not only gets you a very good entry point to buy the stock but also help you keep up with the fast moving market. Moreover swing traders have been in the market for ever and have gained a lot of profit out of it. This could be a great tool for them to use and pick the right stocks to invest in intraday trading.

2. PROJECT DESIGN

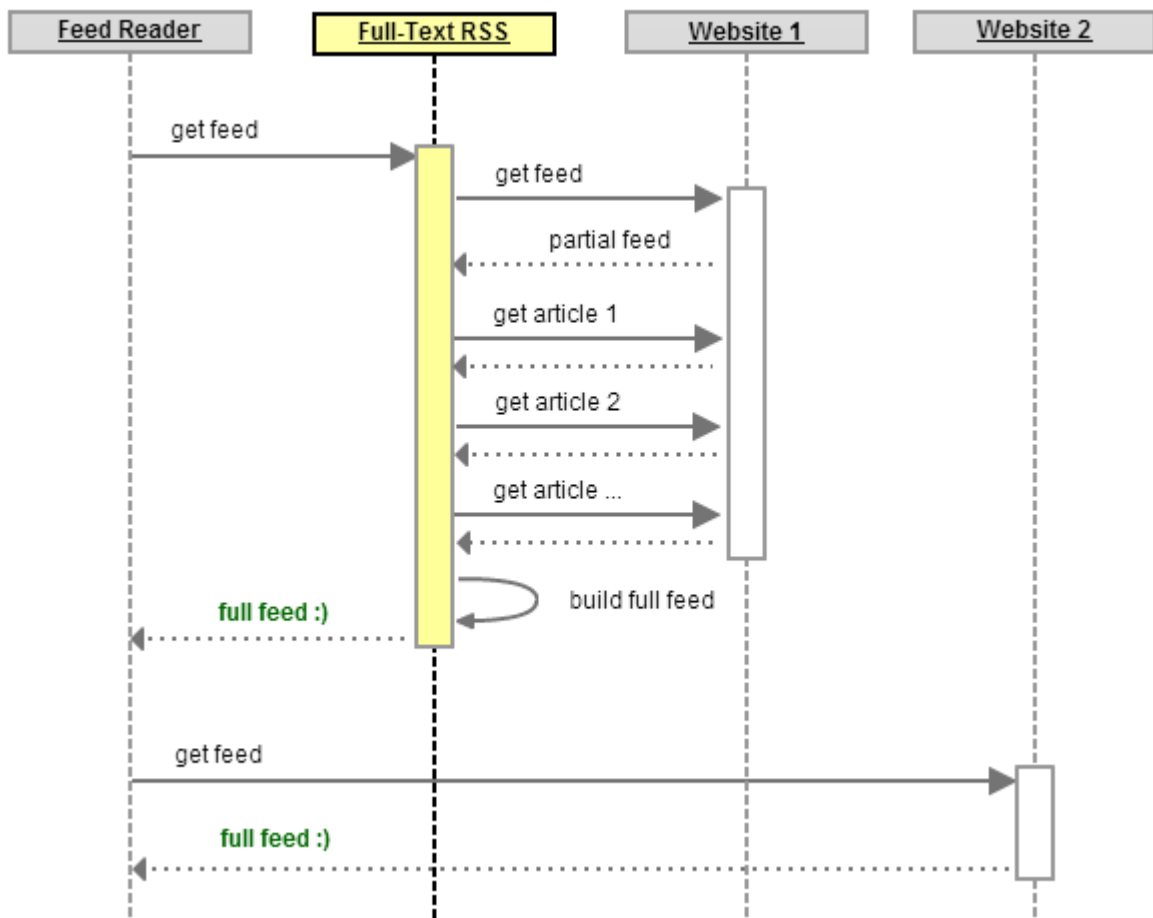
2.1 UML DIAGRAM

We recently updated our RSS Full Text presentation help page with two consecutive diagrams showing how RSS Full Text works with your feed reader.

If you are using a feed reader, you are probably subscribing to the number of web feeds you are interested in. Your server reader then checks those feeds to see new features and pull them for you. Some of these feeds will contain the full content of each article, allowing you to read all of the content in your news reading app. Some feeds will partially contain the content, in anticipation of you visiting the first site to read the full entries. Here is a sequence diagram showing what your server reader will usually do when you sign up for feeds from two websites, Website 1 and Website 2:



In this example, Website 1 returns the partial feed. Instead of subscribing to a feed from Web 1 directly into your feed reader, you can review it to submit the request via RSS Full Text. Showing what happens when you subscribe to a feed this way is shown in the sequence diagram below:



2.2 MODEL-VIEW-VIEW MODEL (MVVM)

MVVM is useful to move business logic from view to ViewModel and Model. ViewModel is the mediator between View and Model which carry all user events and return back the result. The key benefit is allowing true separation between the View and Model and the efficiency that you gain from having that. What it means in real terms is that when your model needs to change, it can be changed easily without the view needing to and vice-versa.

There are three key things that flow out of applying MVVM –

- **Maintainability:** - The presentation layer and the logic are loosely coupled, due to this code is easily maintainable and reusable. As the code base will increase over the course of time, this will help you to distinguish between them.
- **Testability:** - The ViewModel is easier to unit test than code-behind or event-driven code.
- **Extensibility:** - This architecture gives you assurance which enables code to get extensible over the period of time. but keep in mind it's also over job to keep component reusable.

➤ **Model**

The model represents a single source of truth that carries the real-time fetch data or database-related queries.

This layer can contain business logic, code validation, etc. This layer interacts with ViewModel for local data or for real-time. Data are given in response to ViewModel.

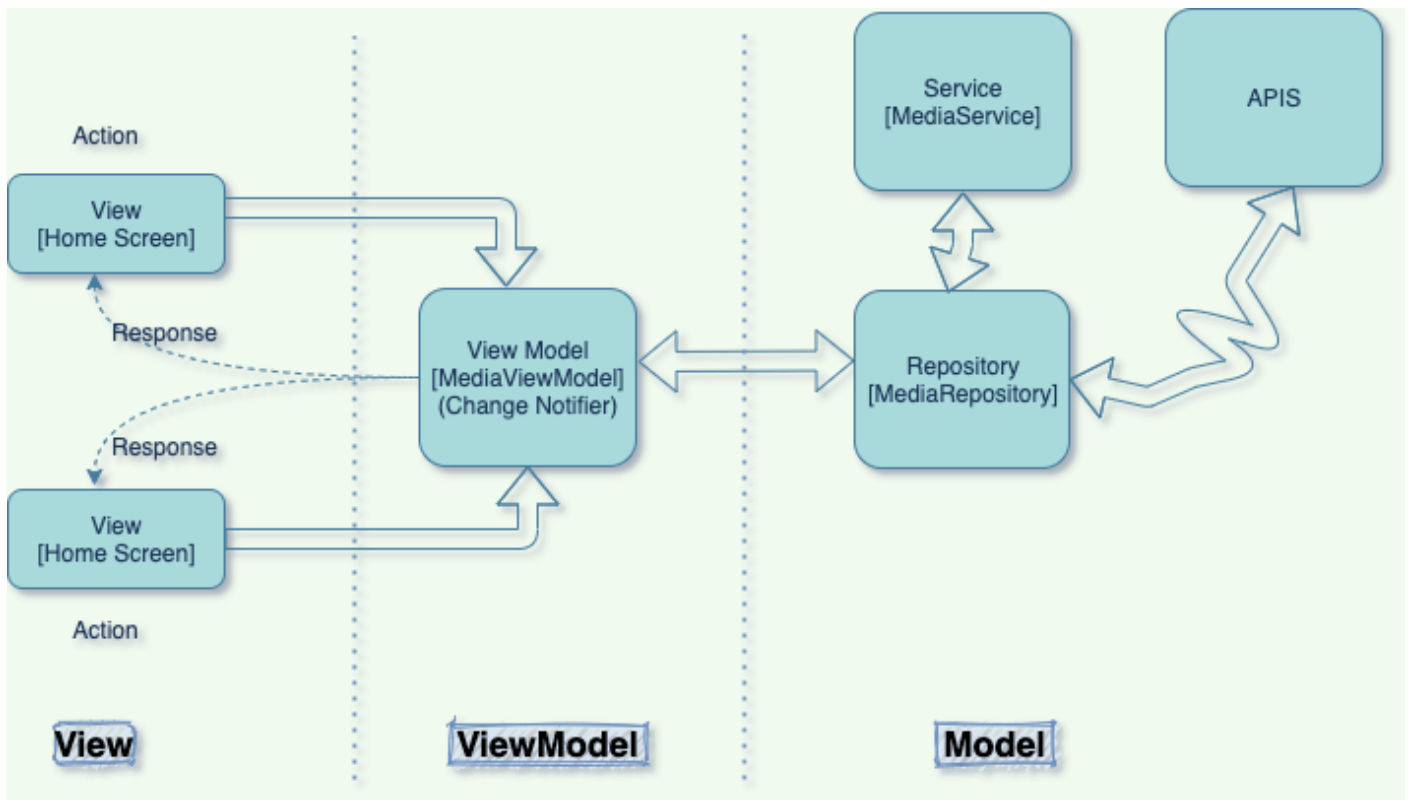
➤ **ViewModel**

ViewModel is the mediator between View and Model, which accept all the user events and request that to Model for data. Once the Model has data then it returns to ViewModel and then ViewModel notify that data to View.

ViewModel can be used by multiple views, which means a single ViewModel can provide data to more than one View.

➤ **View**

The view is where the user is interacting with Widgets that are shown on the screen. These user events request some actions which navigate to ViewModel, and the rest of ViewModel does the job. Once ViewModel has the required data then it updates View.



2.3 DATA RELATED TO STOCK MARKET

For getting the right and authentic stock market news, [5] Economic times, [6] Money Control RSS feeds are been used for the project but you can use any other source maybe your country's RSS feeds or Twitter/Telegram data to make your feed more informative and accurate.

This is a step to apply NEL to build apps from different domains in solving different problems.

A view of a RSS feed:

```
<?xml version="1.0" encoding="UTF-8" ?>
<rss xmlns:atom="http://www.w3.org/2005/Atom" version="2.0">
  <channel>
    <title>Stocks-Markets-Economic Times</title>
    <link>https://economictimes.indiatimes.com/markets/stocks</link>
    <description>Stock Market News: Latest Stock news and updates on The Economic Times. Find Stock Market Live Updates, BSE, NSE Top Gainers, Losers and more.</description>
    <language>en-gb</language>
    <lastBuildDate>2021-09-20T23:00:48+05:30</lastBuildDate>
    <atom:link type="application/rss+xml" rel="self" href="https://economictimes.indiatimes.com/rssfeeds/2146842.cms"/>
    <copyright>Copyright:(C)2021 Bennett Coleman & Co. Ltd, http://info.indiatimes.com/terms/tou.html</copyright>
    <docs>http://syndication.indiatimes.com/</docs>
  </channel>
  <image>
    <title>Economic Times</title>
    <url>https://img.etimg.com/thumb/msid-76939477,width-600,height-60,quality-100/economictimes.jpg</url>
    <link>https://economictimes.indiatimes.com/markets/stocks</link>
  </image>
  <item>
    <title>Ahead of Market: 12 things that will decide stock action on Tuesday</title>
    <description><a href="https://economictimes.indiatimes.com/markets/stocks/ahead-of-market-12-things-that-will-decide-stock-action-on-tuesday/articleshow/86373831.cms"></a>Mazhar Mohammad of Chartviewindia.in said this market may be changing its trajectory from 'buy on dips' to 'sell on rallies' mode.</description>
    <link>https://economictimes.indiatimes.com/markets/stocks/news/ahead-of-market-12-things-that-will-decide-stock-action-on-tuesday/articleshow/86373831.cms</link>
    <image>https://img.etimg.com/thumb/width=1200,imagesize=27553,resize=4,msid=86373831/ahead-of-market-12-things-that-will-decide-stock-action-on-tuesday.jpg</image>
    <guid>Article at EconomicTimes.com with article id : 86373831</guid>
    <pubDate>2021-09-20T21:47:24+05:30</pubDate>
  </item>
</rss>
```

It's quite difficult for a person to understand. Linking each entity with information using a knowledge base [7] Nifty500 companies list. To get daily news update and understand the NLP [8] yfinance library was the major source.

2.4 DESIGN OF THE APPLICATION

- Homepage Screen

Buzzing Stocks ⚡

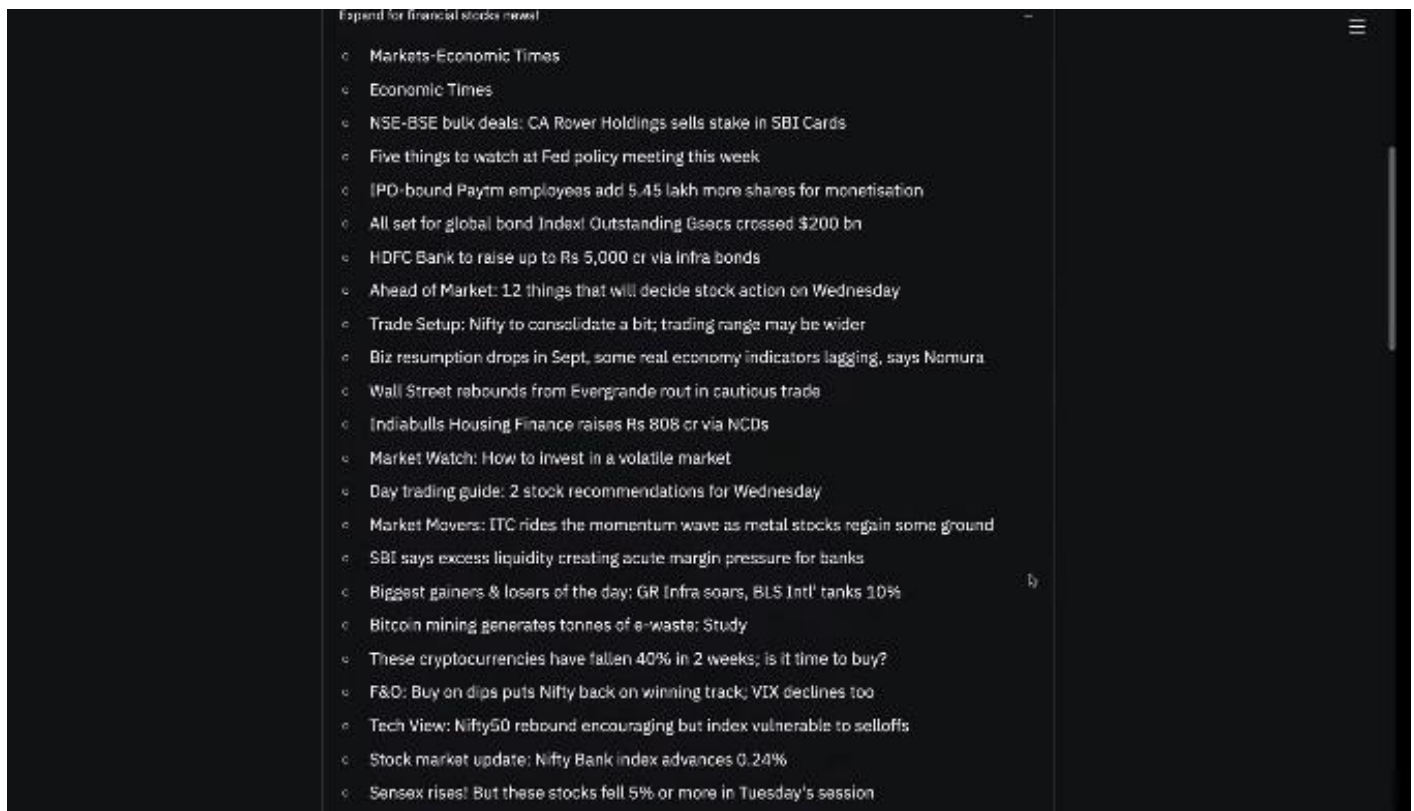
Add your RSS link here!

<https://www.moneycontrol.com/rss/buzzingstocks.xml>

	Org	Symbol	currentPrice	dayHigh	dayLow
0	SBI Cards and Payment ...	SBICARD	1,069.9000	1,079.9000	1,012.0000
1	Federal Bank Ltd.	FEDERALBNK	80.0500	80.5000	78.1500
2	HDFC Bank Ltd.	HDFCBANK	1,551.9500	1,568.6500	1,528.9500
3	Indiabulls Housing Fin...	IBULHSGFIN	208.4500	218.0000	207.0500
4	ITC Ltd.	ITC	241.5000	242.4500	231.7000
5	Adani Green Energy Ltd.	ADANIGREEN	1,120.7000	1,127.0000	1,097.2500
6	Jubilant Foodworks Ltd.	JUBLFOOD	4,135.3500	4,151.8000	3,993.0500
7	Apollo Hospitals Enter...	APOLLOHOSP	4,891.6000	4,918.9500	4,790.1500
8	Marico Ltd.	MARICO	560.7000	570.0000	556.9500
9	Piramal Enterprises Lt...	PEL	2,612.5000	2,624.9500	2,461.0500
10	Hindustan Unilever Ltd.	HINDUNILVR	2,812.4500	2,859.3000	2,796.6500

Expand for financial stocks news! +

- **Headline Screen:**



- Stock market update: Nifty Auto Index falls 0.46%
- Stock market update: Stocks that hit 52-week highs on NSE in today's trade
- Evergrande a Chinese issue; shouldn't affect other markets: Andrew Freris
- Stock market update: Nifty Pharma Index advances 1.29%
- Share market update: Most active stocks in today's market in terms of volume
- Gold marginally down; silver rises by Rs 40
- Stock market update: Stocks that hit 52-week lows on NSE in today's trade
- Stock market update: Nifty IT Index advances 1.94%
- Share market update: Most active stocks of the day in terms of traded value
- Sentiments have turned positive, we are in for a good time, says M&M Fin's Ramesh Iyer
- Sensex reclaims Mt 59K as investors rush to buy beaten-down sectors
- Sensex stages 773 pts recovery, ends 514 pts higher; Nifty50 tops 17,550
- Commodity strategies: Gold, silver, crude, base metals
- Realty stocks in demand after Karnataka cuts stamp duty
- Australian shares recoup losses, end higher on energy boost
- Japan shares end lower on Evergrande woes, bargain-buying limits losses
- FTSE 100 recovers 1% on energy, mining stock boost
- Adani Green shares down 1.33% as Nifty gains
- Jubilant Food shares fall 0.26% as Nifty gains
- Our IPO will provide bountiful rewards to employees: Mobikwik
- Apollo Hospital shares gain 0.1% as Sensex falls
- Debutant stock trades below IPO price; brokerage sees 54% upside
- Marico share price down 0.99 per cent
- Piramal Ent. shares 0.0% as Sensex falls

- Gainers Losers: 10 stocks that moved most on September 21
- NDTV clarifies Adani Group buyout buzz; share price locked at 10% upper circuit
- Indus Towers shares rise 4% as CLSA maintains ₹6,000 buyāC , raises target price
- SBI Cards stock price falls 3% on CA Rover Holdings plan to sell 3.2 crore shares
- Hindustan Unilever share price at 52-week high; brokerages see up to 15% upside
- HCL Tech shares hit new 52-week high on 5-year deal with MKS Instruments Inc
- Cadilla Healthcare share price falls on sale of two brands
- Buzzing Stocks: Anant Raj Global, Kolte - Patil Developers, Visa Steel, and other stocks in news today
- Buzzing Stocks: SBI Cards, Adani Ports, Hindustan Copper and other stocks in news today
- Rakesh Jhunjhunwala offloads additional 98,094 shares in The Mandhana Retail Ventures
- Gainers Losers: 10 stocks that moved most on September 20
- Bigblac Construction share price rises 6% on partnership with SCG International Corporation
- Zee Entertainment shares rise after BlackRock Inc increases stake
- Kotak Mahindra Bank share price gains on acquiring 10% stake in KFin Technologies

2.5 WORKING

App module is divided into these parts:

- 1) RSS link
- 2) PRICING
- 3) HEADLINES

1) RSS Link –

After starting the web app, there should be a screen appear with a search bar showing you an option from which site or RSS feed you want to gather for example I have used ‘Money Control’ here but use can use the others like yahoo finance, Economic times, etc. You just have to copy and paste the link and press enter all the market pricing daily news on stocks shares of any company will appear on your screen.

Add your RSS link here!

<https://www.moneycontrol.com/rss/buzzingstocks.xml>

2) Table Showing the pricing-

Now as you have chosen the best for yourself the RSS feed now a table will appear with all the companies and their symbol of identification which is also unique for every company. This must tell you the current situation of the market whether the stocks are up so that you can invest or whether to not if the market is down.

	Org	Symbol	currentPrice	dayHigh	dayLow	f ↗
0	SBI Cards and Payment ...	SBICARD	1,069.9000	1,079.9000	1,012.0000	
1	Federal Bank Ltd.	FEDERALBNK	80.0500	80.5000	78.1500	
2	HDFC Bank Ltd.	HDFCBANK	1,551.9500	1,568.6500	1,528.9500	
3	Indiabulls Housing Fin...	IBULHSGFIN	208.4500	218.0000	207.0500	
4	ITC Ltd.	ITC	241.5000	242.4500	231.7000	
5	Adani Green Energy Ltd.	ADANIGREEN	1,120.7000	1,127.0000	1,097.2500	
6	Jubilant Foodworks Ltd.	JUBLFOOD	4,135.3500	4,151.8000	3,993.0500	
7	Apollo Hospitals Enter...	APOLLOHOSP	4,891.6000	4,918.9500	4,790.1500	
8	Marico Ltd.	MARICO	560.7000	570.0000	556.9500	
9	Piramal Enterprises Lt...	PEL	2,612.5000	2,624.9500	2,461.0500	
10	Hindustan Unilever Ltd.	HINDUNILVR	2,812.4500	2,859.3000	2,796.6500	

3) Headlines-

Last but not the least at the very bottom of the screen there is another bar which can be expand. It will provide you with the important headlines what all is going in the market which company is expanding more etc. now just by clicking the headline which is attracting you can be completely seen by just a click just click on the headline and you are ready to get some knowledge.

Expand for financial stocks news!



3. TERMS

- **RSS**

RSS formats are preceded by a few web marketing efforts that did not gain widespread popularity. The basic concept of redesigning information about websites dates back to 1995, when Ramanathan V. Guha and others at Apple's Advanced Technology Group developed the Meta Content Framework [9].

The Summary of the RDF Site, the first version of RSS, was created by Dan Libby and Ramanathan V. Guha at Netscape. It was released in March 1999 for use on the My.Netscape.Com portal. This version is known as RSS 0.9. In July 1999, Dan Libby of Netscape produced a new version, RSS 0.91, which simplified the format by removing RDF elements and integrating features from Dave Winer's news marketing format. Libby also renamed the format from RDF to the RSS Rich Site Summary and outlined the continuous improvement of the format "in the text of the future".

- **NLP –**

Indigenous language (NLP) analysis refers to the computer science department — and in particular, the branch of artificial intelligence or AI — which deals with giving computers the ability to understand text and spoken words in the same way that human beings can.

NLP basically combine the computational language rule-based modeling of a human language with statistical, ML, DL models. Together all these technologies enables the computer to process the human language in the form of text or voice data.

Few NLP Tasks are as follow:

- **Speech recognition**, also called speech-to-text, is the function of faithfully converting voice data into text data. Speech recognition is required for any app that follows voice commands or answers spoken questions. What makes speech recognition particularly challenging is how people speak — quickly, consistently, with emphasis and modulation, in a variety of pronunciations, and often using the wrong grammar.
- **Parts of speech tagging** also known grammatical tagging, it is the process of determining the part of speech of a particular word bon its use and context.

- **Natural language generation** is also described as opposite of speech recognition or speech-to-text; it's the task of putting structured information into human language.

- **Anaconda navigator**

Developed and maintained by Anaconda, INC., it was founded by PETER WANG and TRAVIS OLIPHANT in the year 2012.

Anaconda is a distribution of two programming languages known as python and R. It is used for scientific computing which aims to simplify package management and deployment.

- **Visual studio Code**

A code editor for making the app.

3.1 CODE

3.1.1 WORKSPACE

```
<?xml version="1.0" encoding="UTF-8"?>
<project version="4">
  <component name="ChangeListManager">
    <list default="true" id="b5a4c8ec-c1d2-4baf-ba1f-55a8a774fb79" name="Changes"
comment="" />
    <option name="SHOW_DIALOG" value="false" />
    <option name="HIGHLIGHT_CONFLICTS" value="true" />
    <option name="HIGHLIGHT_NON_ACTIVE_CHANGELIST" value="false" />
    <option name="LAST_RESOLUTION" value="IGNORE" />
  </component>
  <component name="MarkdownSettingsMigration">
    <option name="stateVersion" value="1" />
  </component>
  <component name="ProjectId" id="21m4p4mbHpSXhR6Ve1o8bex53Nq" />
  <component name="ProjectViewState">
    <option name="hideEmptyMiddlePackages" value="true" />
    <option name="showLibraryContents" value="true" />
  </component>
  <component name="PropertiesComponent">
    <property name="RunOnceActivity.OpenProjectViewOnStart" value="true" />
    <property name="RunOnceActivity.ShowReadmeOnStart" value="true" />
    <property name="last_opened_file_path" value="$PROJECT_DIR$/../New folder" />
  </component>
  <component name="SpellCheckerSettings" RuntimeDictionaries="0" Folders="0"
CustomDictionaries="0" DefaultDictionary="application-level" UseSingleDictionary="true"
transferred="true" />
  <component name="TaskManager">
    <task active="true" id="Default" summary="Default task">
      <changelist id="b5a4c8ec-c1d2-4baf-ba1f-55a8a774fb79" name="Changes" comment="" />
      <created>1638533109135</created>
      <option name="number" value="Default" />
      <option name="presentableId" value="Default" />
      <updated>1638533109135</updated>
    </task>
  </component>
</project>
```



```
</task>
<servers />
</component>
</project>
```

3.1.2 MODULES

```
<?xml version="1.0" encoding="UTF-8"?>
<project version="4">
  <component name="ProjectModuleManager">
    <modules>
      <module fileurl="file://$PROJECT_DIR$/.idea/NER_News_Feed-master.iml"
filepath="$PROJECT_DIR$/.idea/NER_News_Feed-master.iml" />
    </modules>
  </component>
</project>
```

3.1.3 MISC

```
<?xml version="1.0" encoding="UTF-8"?>
<project version="4">
  <component name="ProjectRootManager" version="2" project-jdk-name="Python
3.10" project-jdk-type="Python SDK" />
</project>
```

3.1.4 PROFILES_SETTINGS

```
<component name="InspectionProjectProfileManager">
  <settings>
    <option name="USE_PROJECT_PROFILE" value="false" />
    <version value="1.0" />
  </settings>
</component>
```

3.1.5 News Feed Buzzing Stock!

1. Import the required libraries - spacy, pandas, requests BeautifulSoup, etc
2. Extract the data from the RSS feed links.
3. NER- spaCy NLP pipeline to process our extracted textual data.
4. NEL - Name Entity Linking
5. Extract the data of these entities(public traded companies) using yahoofinance library.

```
import spacy
import pandas as pd
import requests
from bs4 import BeautifulSoup

resp = requests.get("https://economictimes.indiatimes.com/markets/rssfeeds/1977021501
.cms")
resp

resp.content

soup = BeautifulSoup(resp.content, features='xml')
headlines = soup.findAll('title')

headlines

nlp = spacy.load("en_core_web_sm")

print(headlines[4])
processed_hline = nlp(headlines[4].text)
for token in processed_hline:
    print(token.text, "---", token.pos_)

print(headlines[4])
processed_hline = nlp(headlines[4].text)
for token in processed_hline:
```

```
print(token.text, "---", token.pos_, "---", token.dep_, "---",
      spacy.explain(token.dep_))
```

```
print(headlines[4])
processed_hline = nlp(headlines[4].text)
for token in processed_hline:
    print(token.text, "---", token.pos_, "---", token.dep_, "---",
          spacy.explain(token.pos_))
```

```
spacy.displacy.render(processed_hline, style='dep', jupyter=True, options={'distance'
: 120})
```

```
spacy.displacy.render(processed_hline, style='ent', jupyter=True, options={'distance'
: 120})
```

```
!pip install yfinance
```

```
from google.colab import files
```

```
files.upload()
```

```
stocks_df = pd.read_csv("./ind_nifty500list.csv")
```

```
stocks_df.head()
```

```
import yfinance as yf
stock_info = yf.Ticker("3MINDIA")
```

```
stock_info
```

```
stock_info.info
```

```
stock_info = yf.Ticker("MSFT")
stock_info.info
```

```
stock_info = yf.Ticker("3MINDIA.NS")
stock_info.info
```

```
for ent in processed_hline.ents:
    print(ent)
```

```
for ent in processed_hline.ents:
    print(ent.label_)
```

3.1.6 app.py

```
import pandas as pd
import requests
import spacy
```

```

import streamlit as st
import os

from spacy import displacy
from bs4 import BeautifulSoup
import yfinance as yf
# from spacy_streamlit import visualize_ner
import matplotlib

st.title('Buzzing Stocks :zap:')

## get data from RSS feed

def extract_text_from_rss(rss_link):
    """
    Parses the XML and extracts the headings from the
    links in a python list.
    """
    headings = []
    r1 =

requests.get('https://economictimes.indiatimes.com/markets/stocks/rssfeeds/2146842.cms')
    r2 = requests.get(rss_link)
    soup1 = BeautifulSoup(r1.content, features='lxml')
    soup2 = BeautifulSoup(r2.content, features='lxml')
    headings1 = soup1.findAll('title')
    headings2 = (soup2.findAll('title'))
    print(headings2)
    headings = headings1 + headings2
    return headings

```

```

token_dict = {
    'Org': [],
    'Symbol': [],
    'currentPrice': [],
    'dayHigh': [],
    'dayLow': [],
    'forwardPE': [],
    'dividendYield': []
}

nlp = spacy.load("en_core_web_sm")

def stock_info(headings):
    """
    Goes over each heading to find out the entities
    and link it with the nifty 500 companies data.
    Extracts the data
    """

    stocks_df = pd.read_csv("../data/ind_nifty500list.csv")

    for title in headings:
        doc = nlp(title.text)

        for token in doc.ents:
            try:
                if stocks_df['Company Name'].str.contains(token.text).sum():
                    symbol = stocks_df[stocks_df['Company Name'].\
                                        str.contains(token.text)][ 'Symbol' ].values[0]
                    org_name = stocks_df[stocks_df['Company Name'].\
                                        str.contains(token.text)][ 'Company Name' ].values[0]
                    token_dict['Org'].append(org_name)
                    print(symbol+".NS")
                    token_dict['Symbol'].append(symbol)
                    stock_info = yf.Ticker(symbol+".NS").info
                    token_dict['currentPrice'].append(stock_info['currentPrice'])
                    token_dict['dayHigh'].append(stock_info['dayHigh'])
                    token_dict['dayLow'].append(stock_info['dayLow'])
                    token_dict['forwardPE'].append(stock_info['forwardPE'])

```

```
token_dict['dividendYield'].append(stock_info['dividendYield'])

    else:
        pass

    except:
        pass

output_df = pd.DataFrame(token_dict)
return output_df

## add an input field to pass the RSS link

user_input = st.text_input("Add your RSS link here!",
"https://www.moneycontrol.com/rss/buzzingstocks.xml")

## get the financial headings

fin_headings = extract_text_from_rss(user_input)

## output the financial info through a dataframe

output_df = stock_info(fin_headings)
output_df.drop_duplicates(inplace=True)
st.dataframe(output_df)
```



```
## display the news in an expander section

with st.expander("Expand for Financial News!"):
    for h in fin_headings:
        st.markdown("* " + h.text)
```

3.1.7 app_utils.py

```
# @st.cache

def process_rss(headings):
    nlp = spacy.load("en_core_web_sm")
    for title in headings:
        doc = nlp(title.text)

        # spacy.displacy.render(doc, style='ent', jupyter=True, options={'distance': 120})
        visualize_ner(doc, labels=['ORG', 'PERSON', 'DATE', 'GPE', 'PERCENT'],
                      show_table=False)
        print("Done-----")
    return

def viz_ner():
    with open("./data/History_102.txt") as file:
        text = file.read()
    print(type(text))
    doc = nlp(text)
    visualize_ner(doc, labels=['ORG', 'PERSON', 'DATE', 'GPE', 'PERCENT'],
                  show_table=False)
    return

# txt_file = st.file_uploader("Upload a doc", type=['txt'])
```

```
# if txt_file is not None:
#     stringio = StringIO(txt_file.getvalue().decode('utf-8'))
#     string_data = stringio.read()
#     viz_ner()
# else:
#     st.warning("Please upload an image.")
#     st.stop()
```

4. RESULT

It is wonderful project with the understanding of NLP and NEL. Following is a preview of the project result helping in the completion of the BSNF application.

```
resp = requests.get("https://economictimes.indiatimes.com/markets/rssfeeds/1977021501.cms")
resp

<Response [200]>
```

Fig 4.1. Requesting RSS feed link giving total response

```
[<title>Markets-Economic Times</title>,
<title>Economic Times</title>,
<title>Ahead of Market: 12 things that will decide stock action on Tuesday</title>,
<title>Trade Setup: Nifty50's up move may continue; 18130-190 levels to act as resistance</title>,
<title>MSCI likely to include 7 Indian stocks to its standard index</title>,
<title>Wall Street opens higher as industrials rise on infrastructure bill</title>,
<title>Market Watch: What were the key reasons for today's rally?</title>,
<title>V-Mart Retail Q2 results: Co posts net loss at Rs 14 cr</title>,
<title>Sundaram Finance Q2 results: Net profit falls 8%, but co confident on recovery</title>,
<title>Inspira Enterprise gets Sebi's go-ahead to raise Rs 800 cr via IPO</title>,
<title>It will not be easy for Britannia to pass on cost inflation: Naveen Kulkarni</title>,
<title>Market Movers: What drove investors to cement producers' counters today?</title>,
<title>NSE completes delivery in gold mini options contract</title>,
<title>Sebi allows FPIs to write off debt securities</title>,
<title>Sobha Q2 results: Profit jumps nearly 3-fold to Rs 48.3 cr</title>,
<title>Sensex rises! These stocks gained over 10% on BSE</title>,
<title>Stock market update: Nifty Pharma index falls 0.7%</title>,
<title>Sensex gains 478 pts, Nifty ends above 18,050; IT, metal, PSU banks rally</title>,
<title>Stock market update: Nifty Realty index advances 1.38%</title>,
<title>Tech View: Nifty50 signals trend reversal; more upside likely</title>,
<title>Stock market update: Power stocks up as market rises</title>,
<title>Karur Vysya Bank Q2 results: Net profit jumps 44% to Rs 165 crore</title>,
<title>Sensex rises! These stocks rallied over 15% on BSE in today's trade</title>,
<title>Time for broad-based sector picks over, go for execution kings now: Trideep Bhattacharya</title>,
<title>Stock market update: Stocks that hit 52-week highs on NSE in today's trade</title>,
<title>Stock market update: Nifty Bank index falls 0.34%</title>,
<title>Stock market update: Nifty Auto index advances 0.66%</title>,
<title>Anmi asks NSE to frame policy to detect prohibitive acts carried out by authorised persons</title>,
<title>Gold marginally lower; silver gains Rs 216</title>,
<title>Share market update: Most active stocks of the day in terms of traded value</title>,
<title>SBI Card to raise Rs 2,000 cr by issuing bonds</title>,
<title>Stock market update: Nifty IT index advances 1.27%</title>,
<title>The New Masters: How auction houses are chasing crypto millions</title>,
```

Fig 4.2 Extracted headline from RSS feed of economic times

```
<title>MSCI likely to include 7 Indian stocks to its standard index</title>
MSCI --- PROPN
likely --- ADV
to --- PART
include --- VERB
7 --- NUM
Indian --- ADJ
stocks --- NOUN
to --- ADP
its --- DET
standard --- ADJ
index --- NOUN
```

Fig 4.3. The pipeline performs all the tasks from tokenization to NER. These are the tokens with pos_ attribute.

```
<title>MSCI likely to include 7 Indian stocks to its standard index</title>
MSCI --- PROPN --- ROOT --- None
likely --- ADV --- amod --- adjectival modifier
to --- PART --- aux --- auxiliary
include --- VERB --- xcomp --- open clausal complement
7 --- NUM --- nummod --- numeric modifier
Indian --- ADJ --- amod --- adjectival modifier
stocks --- NOUN --- dobj --- direct object
to --- ADP --- prep --- prepositional modifier
its --- DET --- poss --- possession modifier
standard --- ADJ --- amod --- adjectival modifier
index --- NOUN --- pobj --- object of preposition
```

Fig 4.4 these are the token with dep_ attribute and label.

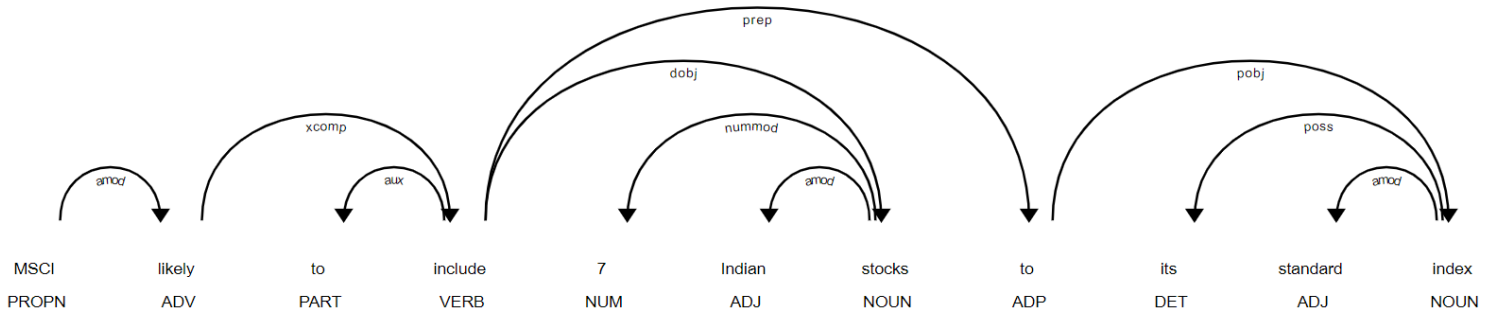


Fig 4.5 Graph visualizing the relationship dependencies among the tokens using displacy render()

MSCI likely to include **7** **CARDINAL** **Indian** **NORP** stocks to its standard index

Fig 4.6 Important entities of the sentence (entity extraction)

	Company Name	Industry	Symbol	Series	ISIN Code
0	3M India Ltd.	CONSUMER GOODS	3MINDIA	EQ	INE470A01017
1	ABB India Ltd.	INDUSTRIAL MANUFACTURING	ABB	EQ	INE117A01022
2	ABB Power Products and Systems India Ltd.	INDUSTRIAL MANUFACTURING	POWERINDIA	EQ	INE07Y701011
3	ACC Ltd.	CEMENT & CEMENT PRODUCTS	ACC	EQ	INE012A01025
4	AIA Engineering Ltd.	INDUSTRIAL MANUFACTURING	AIAENG	EQ	INE212H01026

Fig 4.7 Table extracted the list of companies to get their symbols for further process.

'phone': '425 882 8080',
'preMarketPrice': 337.19,
'previousClose': 336.06,
'priceHint': 2,
'priceToBook': 16.53888,
'priceToSalesTrailing12Months': 14.261033,
'profitMargins': 0.38515,
'quickRatio': 1.961,
'quoteType': 'EQUITY',
'recommendationKey': 'buy',
'recommendationMean': 1.6,
'regularMarketDayHigh': 337.615,
'regularMarketDayLow': 334.68,
'regularMarketOpen': 337.3,
'regularMarketPreviousClose': 336.06,
'regularMarketPrice': 334.78,
'regularMarketVolume': 7515628,
'returnOnAssets': 0.14589,
'returnOnEquity': 0.49303,
'revenueGrowth': 0.22,
'revenuePerShare': 23.395,
'revenueQuarterlyGrowth': None,
'sector': 'Technology',
'sharesOutstanding': 7507979776,
'sharesPercentSharesOut': 0.0064999997,
'sharesShort': 48556022,
'sharesShortPreviousMonthDate': 1631664000,
'sharesShortPriorMonth': 44204519,
'shortName': 'Microsoft Corporation',
'shortPercentOfFloat': 0.0064999997,
'shortRatio': 1.86,
'startDate': None,
'state': 'WA',
'strikePrice': None,

```
'pegRatio': None,
'phone': '91 80 2223 1414',
'preMarketPrice': None,
'previousClose': 26669,
'priceHint': 2,
'priceToBook': 16.1556,
'priceToSalesTrailing12Months': 10.278885,
'profitMargins': 0.078260005,
'quickRatio': None,
'quoteType': 'EQUITY',
'recommendationKey': 'none',
'recommendationMean': None,
'regularMarketDayHigh': 27800,
'regularMarketDayLow': 26700,
'regularMarketOpen': 26700,
'regularMarketPreviousClose': 26669,
'regularMarketPrice': 27396.65,
'regularMarketVolume': 7000,
'returnOnAssets': None,
'returnOnEquity': None,
'revenueGrowth': 1.149,
'revenuePerShare': 2665.486,
'revenueQuarterlyGrowth': None,
'sector': 'Industrials',
'sharesOutstanding': 11265100,
'sharesPercentSharesOut': None,
'sharesShort': None,
'sharesShortPreviousMonthDate': None,
'sharesShortPriorMonth': None,
'shortName': '3M INDIA LTD',
'shortPercentOfFloat': None,
'shortRatio': None,
'startDate': None,
'strikePrice': None,
```


Fig 4.8.1 using ticker function to extract information of MICROSOFT.

Fig 4.8.2 using ticker function to extract information of 3MINDIA.

7
Indian

CARDINAL
NORP

Fig 4.9 ENTITIES of the above sentence with label of the entities.

```
resp = requests.get("https://economictimes.indiatimes.com/markets/rssfeeds/1977021501.cms")  
resp
```

```
<Response [200]>
```

5. CONCLUSION

In conclusion, there are many challenges related to producing, gathering, analyzing, reporting and publishing data in condensed timelines required during the daily changes in stock. We certainly did not mention all of them, but we hope that researchers willing to contribute to our research related to stocks news will help to address those other issues as well. In this paper, with the help of the Goggle Colab/Jupyter Notebook, we have tried as much as possible to aware the public and invest their money in stocks. In future, this application can help people to know about the current situation of stock market. It is very easy to understand. This application is upgraded in such a way that it will let you known the current pricing of the stocks in the companies. This application provides detailed information about the pricing.


6. REFERENCES

- [1] RSS feed description- <https://en.wikipedia.org/wiki/RSS>
- [2] Introduction to Stocks - <https://www.nerdwallet.com/article/investing/what-is-a-stock>
- [3] <https://ijmr.net.in/current/CmrLMmi5Tvm9Elo.pdf>
- [4] <http://www.tjprc.org/publishpapers/2-35-1534487570-81IJAFMRAUG20181.pdf>
- [5] Economic Times xml- <https://economictimes.indiatimes.com/markets/rssfeeds/1977021501.cms>
- [6] Money control xml - <https://www.moneycontrol.com/rss/buzzingstocks.xml>
- [7] Nifty500 companies lists- https://www1.nseindia.com/products/content/equities/indices/nifty_500.htm
- [8] Yahoo finance library- <https://pypi.org/project/yfinance/>
- [9] https://en.wikipedia.org/wiki/Meta_Content_Framework

7. PUBLICATION/COPYRIGHT/PRODUCT

← 📄 🕒 🗑️ 📧 🕒 🗑️ 📄 🗑️ ⋮

IEEE IAS Global Conference on Emerging Technologies (GlobConET) : Submission (79) has been created. 🖨️ 📄

 **Microsoft CMT** 12:05 PM (35 minutes ago) ☆ ↶ ⋮
to me, Yashvardhan, ▾

Hello,

The following submission has been created.

Track Name: Track VIII: AI, AIoT, IIoT, Deep Learning, and Machine Learning.

Paper ID: 79

Paper Title: The buzzing stocks news using NLP

Abstract:
This paper aims to develop an online aggregator of news feed from the market finance such as Yahoo Finance etc. The most important factors influencing the stock market are stock issues. This study is an attempt to create a model that predicts how you can use NLP to improve the stock market news supply of online stock exchanges and news polarity that could affect changes in stock price trends. The novelty of our approach is to accurately predict the sensory model in all real-time news that is available in the public domain and getting insights from a raw and unstructured data is very important. Our approach uses emotion analysis in Really Simple Syndication (RSS) news feeds that impact on stock market prices. Therefore RSS news feed data is collected over a period of time. This proposed paper uses an emotion analysis algorithm to generate emotion in RSS news feeds. In our experimental study, the algorithm approach suggests that stock news sentiment can be used to predict a decline in stock prices, either up or down..

Created on: Tue, 07 Dec 2021 12:31:42 GMT

Last Modified: Tue, 07 Dec 2021 12:31:42 GMT

Authors:
- yivek.singh.scsebtch@kgotiasuniversity.edu.in (Primary)
- Yashvardhan.singh.scsebtch@kgotiasuniversity.edu

Secondary Subject Areas: Not Entered

Submission Files: BT4115 IEEE Research Paper.docx (818 Kb, Tue, 07 Dec 2021 12:31:42 GMT)

Submission Questions Response: Not Entered