Stockfolio

1833049 Sarath Ram M

1833055 Tilak Vijayaraghavan S

# Abstract:

Stockfolio is a multi-module application which aims to provide a variety of services for people who wish to gain knowledge to invest in stock market. This application aims to provide various details like stock trends, back testing, sentiment analysis, portfolio optimisation, stock movement visualisation etc.

# Introduction:

Stockfolio is a modular application which aims to provide its users with various stock information on demand. Stockfolio uses open-source API to track stock information and process them to create meaningful information for the end-user. By considering various angles to interpret stocks Stockfolio aims to create new insights for user to understand the movement of stocks in the market.

The tracker API tracks the movement of stocks using various techniques to generate possible buy / sell signal alerts. All the above paired with a sentiment analyser with data from twitter, enables the predictor to have a better understanding about the stock with regard to the stock’s impact on social media.

Built with open-source software, stockfolio is stable and scalable due to modularity. All this packed together in a deployable application can be used on any system provided the necessary libraries are included.

# Review of literature:

Usage of Technical analysis on the stock data have proven to yield substantial results to consider them for usage. Prevalent methods like candlestick charts, moving averages, convergence and divergence etc. prove useful[1]. A recent study by a group of researchers using a portfolio of stocks traded in the stock exchange of the BRICS countries done with the help of fundamental analysis and technical analysis proved to generate returns higher than the expected values[10]

Evaluation of a portfolio using a Sharpe ratio on a frontier guarantees the most mathematically optimised solution [2]. By evaluating a portfolio mathematically based on economic models and using a Ceteris Paribus way of evaluation, performance of an individual stock can also be recognised. A variance optimised portfolio often referred to as a VOP could either be a single variance or multi-variance optimisation technique to achieve the best possible portfolio according to modern portfolio theory[11].

Heikin-Ashi candlestick proves to be an easier way to understand market trend rather than a normal candlestick [3]. By capturing the trend in the data and smoothening the noise, Heikin-Ashi candlestick chart gives a good representation of the trend present in the data.

Bollinger Bands [4] use a simple yet powerful technique to provide a means to find significant point during a stock’s movement.[5]

Using Monte-Carlo simulation to deduce results prove better than other conventional retail models [6]. By using a Monte-Carlo simulation which uses randomness, simulation of risk and uncertainty becomes possible and estimation of stock rates becomes a possibility.

Sentiment Analysis on stocks prove useful in providing insights to track their movement and confirms the effect of social media on markets [7]. A recent analysis on stock news suggests the use of correlation as a factor to be considered while performing sentiment analysis. Correlation here refers to the connection that exists between the stock being monitored and a correspondingly correlated stock. This correlation can be traced using Deep learning model by processing various relationships that exists between stocks[9]. It is also proven that stock investors’ psychology influences the market, and hence by understanding their mood swings, predicting whether the market will rise or fall becomes possible [8].

# Problem Statement:

The aim of the project is to create a decision tool (stockfolio) which when deployed will be able to assist an individual to know about the stock market and formulate an investment strategy based on a set of technical analysis and market trackers.

# Objectives:

Stockfolio aims to provide its users with the following tools to gain insights, the tools provided include

1. Knowledge about stocks, equities, and other technical terms
2. A tracker with various techniques
3. A portfolio analyser and optimiser

# Dataset:

The dataset used for this project is the stock details obtained on the End of the Day (EOD) data. This also goes by other names like OHLCV data or Bhavcopy data

# Tools required:

The project is implemented using python and a set of its libraries. The libraries used include,

1. pandas
2. NumPy
3. pandas-ta
4. matplotlib
5. yfinance
6. PyPortfolioOpt
7. seaborn
8. nsepy
9. nltk
10. Wordcloud
11. Streamlit

other than the above-mentioned libraries, their dependencies are also required for the execution of the program.

# Tools Used:

Optimisation can be defined in many was one among them being, optimisation is the process of maximising or minimising a certain value while being subject to a set of mathematical constrains. The same can be re-phrased for optimising a portfolio, by considering the expected return and the percentage of risk willing to be taken, an area of possible solutions can be created and can be mathematically optimised to give the best possible solution. The boundary of the solution space is called the frontier and the solution often lies on the frontier as the frontier is the line that separates the set of possible and not possible solutions.

An efficient frontier is a solution line which spans at the edge of the solution space. This line when subject to further constrains will yield the desired result of optimisation. Based on the additional constrains placed on the frontier, the portfolio represented by the frontier takes different names like,

1. Minimum variance portfolio
2. Return-efficient portfolio
3. Parameter-efficient portfolio

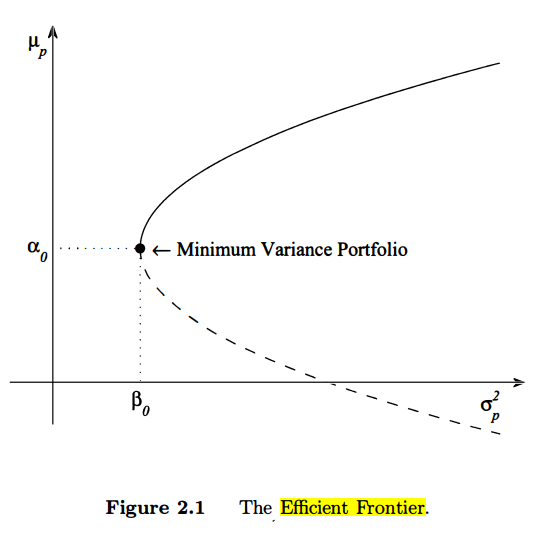


Figure 3.1 An illustration of an efficient minimum variant portfolio

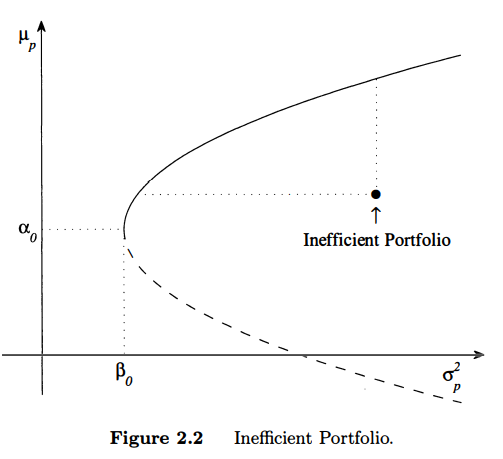


Figure 3.2 An illustration of an inefficient portfolio

The arbitrary constants and are the expected return and variance of the minimum variance portfolio.

By moving along the frontier, the point of maximum efficiency can be found, at the point where the frontier meets the imposed constrains which is usually determined by the Sharpe Ratio which can be defined as the ratio of every unit gain against the volatility the stock faces (risk).

Back-testing is a technique by which the performance of a given portfolio with the help of historical data. The testing technique is a heuristic technique built on the belief that “What once worked will work again.” Back-testing provides users with an estimate of their return based on historical data.

Visual representation is often preferred over raw numbers, candle stick charts prove to solve the problem by visualising the market OHLC data (Open, High, Low, Close of a stock’s price). There are two types of candle stick charts popularly used,

1. Normal Candle stick chart
2. Heikin-Ashi Candle stick chart

The normal candle stick charts provide the OHLC information about a stock and hence move along with the stock. This can cause the chart to capture all the stocks movement and noise. On the other hand, Heikin-Ashi charts which use a different formula tend to highlight stock movement trends along with their price movements.

Both the visuals have their pros and cons. The normal candle stick chart highlights the actual movement along with the noise while Heikin-Ashi candle sticks smoothen the chart but does not give the actual values of the stocks in question (because they smoothen the data, a bit of data loss/transformation occurs).

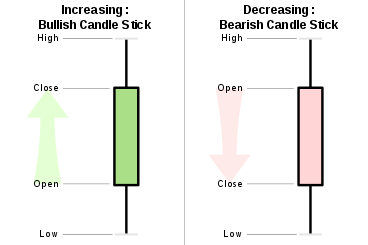
Figure 3.3 Classical representation of a candlestick

Figure 3.4 A candlestick chart for the Apple stock

Figure 3.5 A Heikin-Ashi candle stick for Apple Stocks

By comparing Figure 3.4 and 3.5, a difference can easily be seen between normal candlestick and a Heikin-Ashi candlestick. Figure 3.4 contains the actual stock prices and hence has more noise but the Heikin-Ashi candlestick has less noise but does not reflect the actual stock prices as the data used for the chart underwent smoothening.

“Make hay while the Sun shines” is a classical proverb which say about the importance of action done at the right time, similarly buying and selling apt is important. The knowledge of when to buy or sell can be achieved by performing technical analysis using different techniques. Few of them being,

1. Relative Strength Index
2. Moving averages
   1. Simple moving averages
   2. Moving average convergence / divergence
3. Bollinger bands

Relative Strength Index (RSI) is an index which describes the amount of oscillation a stock faces. Relative strength of a stock is calculated as

where Relative Strength (RS) is calculated by

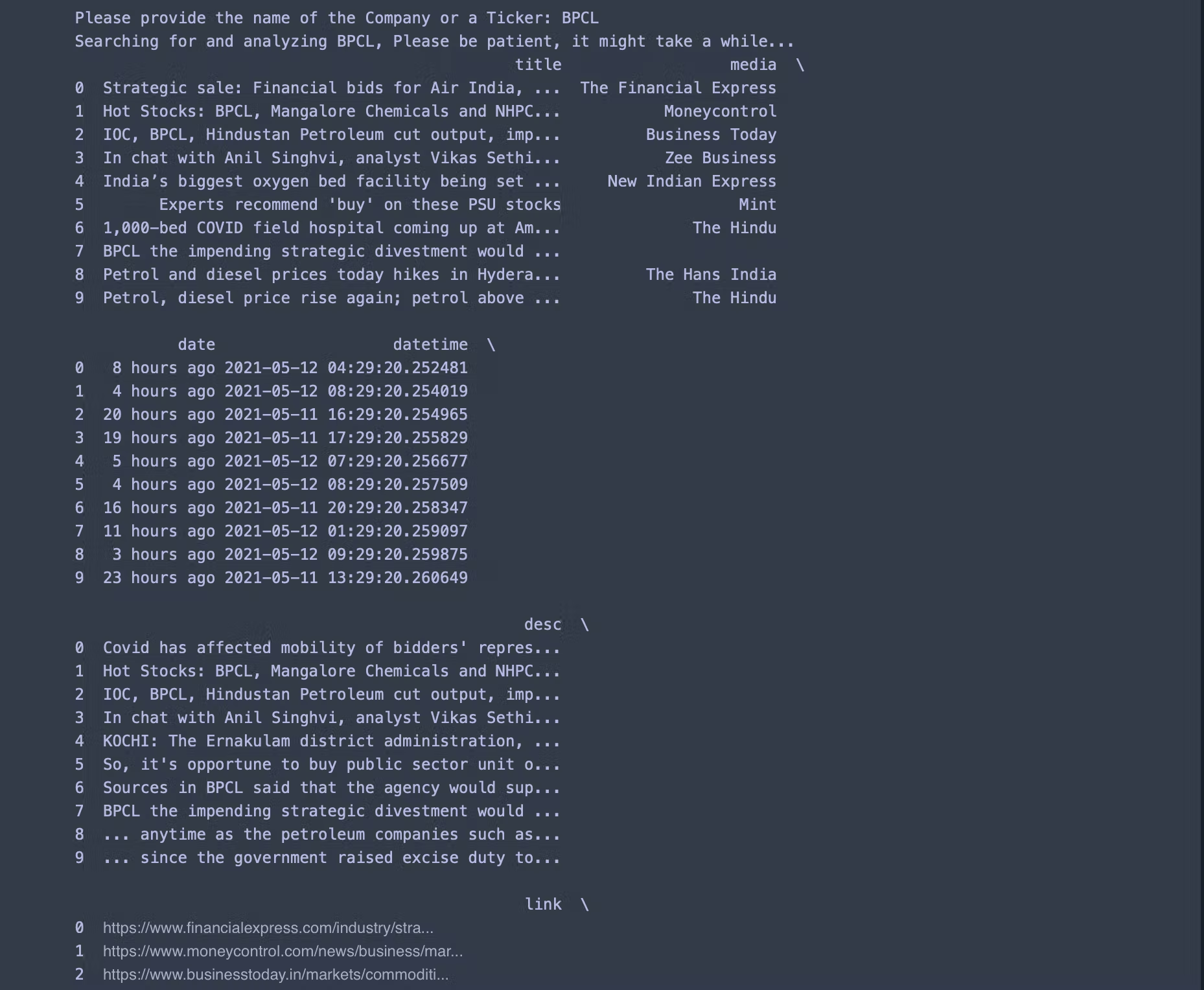
Where U and D are metrics calculated for change in closing price between two days. The function SMMA (), is a smoother or modified moving average taken for ‘n’ periods (days). If the market if rising, D is zero while U is the difference between the closing price of the stocks between two consecutive days, this logic is converse for a falling market where U is zero while D is the difference in closing price of consecutive days. RSI gives the current moment of the stock; this indicator is like a percent indicator. RSI values between 70 – 30 is considered as normal for a stock’s momentum while outside of the said bounds could me that the stock is over-bought or over-sold.

Moving averages as the name suggests are averages of the stock prices taken at an interval of ‘n’ days where ‘n’ is the number of consecutive days considered for taking the average. Buy-sell signal can be generated using the help of moving averages, consider two set of moving averages one long-term and one short-term moving average. Whenever the short term moving average is greater than the long term moving average and vice-versa, we can generate a signal to buy or sell respectively. This is a simple trading trick achieved by using moving averages. A more sophisticated trick is by using moving average convergence and divergence technique. This technique uses the idea of stock divergence to generate signals. Divergence is a term which is used to indicate that the “indicator” used and the stocks are moving in the opposite direction and a trend reversal is imminent.

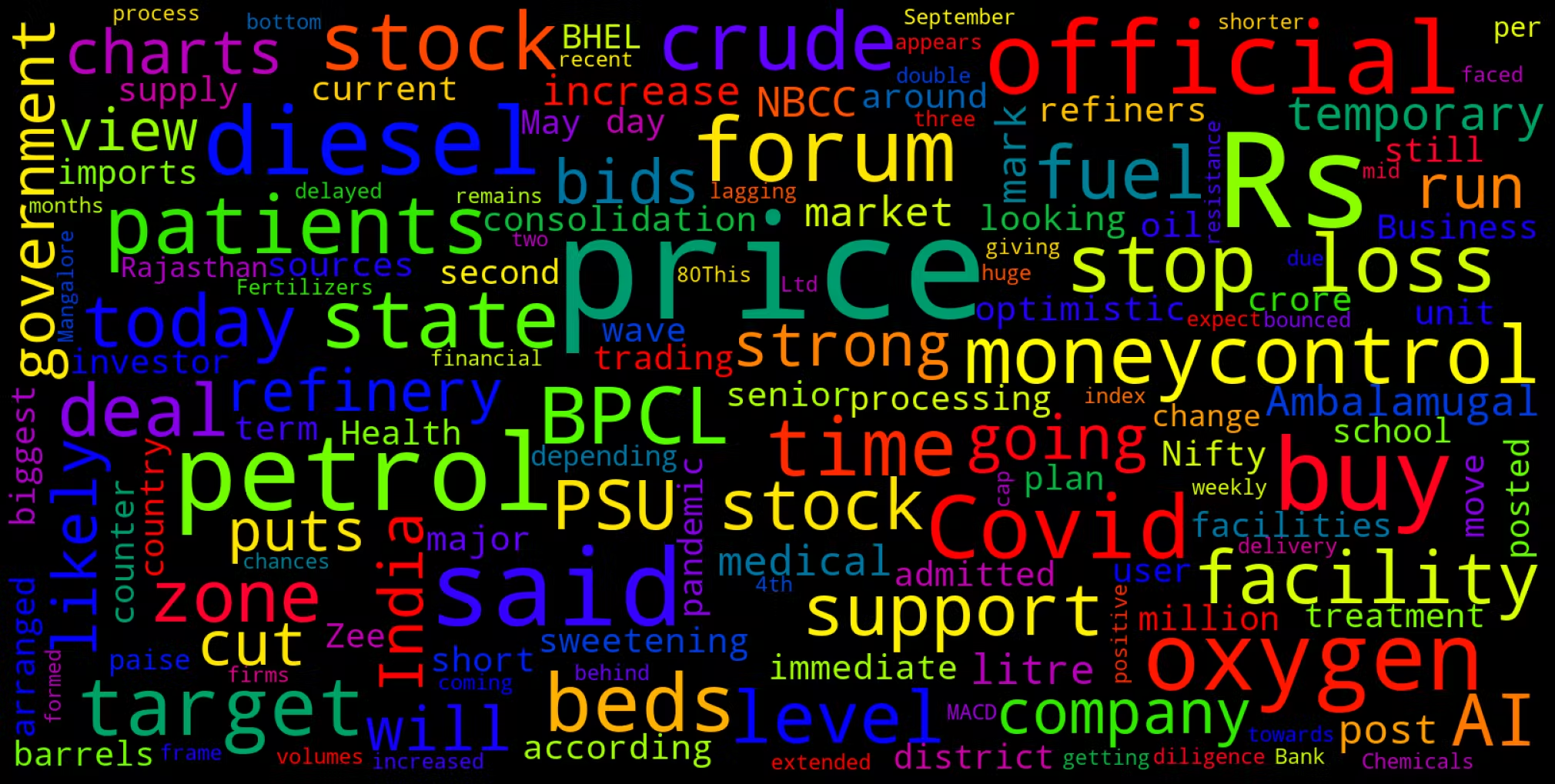
Bollinger bands are a simple and powerful technique to generate signals. By considering the stock price variation and setting a limit like ‘x’ standard deviations, we can create a boundary around the stock which when crossed signals a buy / sell.

Figure 3.6 A Bollinger Band chart

Internet is the current means through which information is traversing the world and social media has become the most influential network across the globe. By performing sentiment analysis on social media and news articles, we can capture an idea of the stock movement for a short term. The nltk library used VADER (Valence Aware Dictionary and Sentiment), a lexicon / glossary as well as a rule-based sentiment analysis tool that is specially tuned to sentiments expressed on social media. The API like snscrape and GoogleNews are used to gather information from the internet. Once all the raw data has been gathered, we proceed to summarize the aggregate information to make meaningful analysis and an interpretation.

Figure 3.7 A chunk of raw data collected from the internet using Google News API.

This info is useless by itself but, by generating a word cloud, it could make some useful interpretation. A word cloud is a representation which is used to represent words in the metadata on website or showcases the occurrence of words in a given phase or paragraph.

Figure 3.8 a word cloud generated from the chunk of raw data collected

This word cloud highlights certain phrases like ‘price, petrol, target, stock’ etc. Among other words. But, still the analysis can be made better by aggregating the emotion of the words / phrases by using nltk’s VADER[12] using which we can gather the percentage of positive, negative and neutral emotions expressed in the internet.

# References:

1. Technical Analysis of Stock Trends, By Robert D. Edwards, John Magee, W. H. C. Bassetti
2. Portfolio Optimization by Michael J. Best
3. Using The Heikin-Ashi Technique by Muhammad Iqbal
4. John Bollinger
5. Using Bollinger Bands by John Bollinger
6. Back testing Value-at-Risk: A Duration-Based Approach by Peter Christofferson, Denis Pelletier
7. Sentiment analysis on social media for stock movement prediction by Thien HaiNguyena, Kiyoaki Shiraia, JulienVelcinb
8. Security Analysis and Portfolio Management, Chandra Prasanna
9. Combining Enterprise Knowledge Graph and News Sentiment Analysis for Stock Price Prediction, Jue Liu, Zhuocheng Lu, Wei DU, Renmin University of China
10. Examination of the profitability of technical analysis based on moving average strategies in BRICS, Matheus José Silva de Souza, Danilo Guimarães Franco Ramos, Marina Garcia Pena, Vinicius Amorim Sobreiro & Herbert Kimura
11. A comprehensive review of deterministic models and applications for mean-variance portfolio optimization, Can B. Kalayci , Okkes Ertenlice , Mehmet Anil Akbay
12. Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.