Machine Learning in Stock Market

Bhavin Bajaj

27/05/2023

1)Abstract

This report explores the utilization of Artificial Intelligence (AI) and Machine Learning (ML) techniques to predict stock prices, aiding traders and high-fund companies in making informed decisions. The report highlights the significance of AI and ML in analyzing stock market data, improving analysis accuracy, and formulating strategies for generating higher profits. By employing advanced algorithms and predictive models that will be tested on previous stocks price datasets, the study demonstrates the potential of ML in extracting valuable insights from historical data, identifying trends, and forecasting future stock prices. The results underscore the substantial benefits of integrating ML methodologies into stock market analysis, enabling traders and high-fund companies to optimize their investment strategies and enhance profitability, this can be done by creating an app where traders can use strategies for trading also they can test their strategies on our dataset.

2)Problem Statement

Traders and investment companies grapple with significant challenges in the stock market. Traditional strategies, often untested, fail to provide reliable predictions, exposing investors to substantial risk. With over 80% of active fund managers unable to consistently outperform market indices, there is a pressing need for a more effective approach. Machine Learning (ML) offers a solution by leveraging historical data to forecast stock prices accurately. ML algorithms identify intricate patterns, empowering traders to make data-driven decisions and develop strategies that optimize profitability. By integrating ML, traders and investment companies can overcome the unpredictability of the market, reduce risk, and achieve better returns on their investments.

3) Market / Customer/ Business Need Assessment:

The market is in need of innovative solutions that leverage ML to improve stock market analysis. With the increasing recognition of ML's potential and as technology continues to

advance rapidly, making it more accessible and affordable, the growing demand for data-driven insights, businesses offering ML-based models have substantial growth potential in catering to the needs of traders and investment companies.

Below are some of points that highlight the needs of customers and the market in general and how our product is a great market fit catering to the needs of the consumers:

1. **Enhanced Analysis , Decision Making and predictions:** Traders and investment companies require ML-based models to analyse vast amounts of stock market data quickly and accurately. These models can identify patterns, trends, predict prices and correlations that may not be possible through traditional analysis methods, enabling more informed decision-making and investments and higher profits.

2. **Growing Complexity of the Market:** The stock market is becoming increasingly complex, with numerous variables and factors influencing stock prices. ML-based models can effectively handle and analyze large datasets, providing traders and investment companies with valuable insights into market dynamics and helping them navigate the complexity more efficiently.

3. **Lack of knowledge :** Most of the people who trade ,lack knowledge about the stock market and have very basics and fundamental understanding of the market. making strategies and ml models will help them have better understanding and trade risk free

4. **Growing Complexity of the Market:** The stock market is becoming increasingly complex, with numerous variables and factors influencing stock prices. ML-based models can effectively handle and analyse large datasets, providing traders and investment companies with valuable insights into market dynamics and helping them navigate the complexity more efficiently.

4)Target Specification and characterization:

We aim to target a large group of audience from all overall the country that actively participate in the stock market and trading. We have a diverse pool of audience ranging from small traders to large investment companies each having different specification and needs. Our product aims to cater the need of all. I have listed down some of our most important potential users.

1)**Retail Traders**-

These are people like you and me and active individuals who trade on stock market on daily basis with small amount of funds. These traders need stability in their investments, lower risk

on their investment and maximum profits we can help them by testing their strategies and provide them with prediction and analysis of the market based on our algorithms.

2) High Networth Individuals (HNIs)

Investors' with over two crores of investible assets are generally considered as HNIs. Such individuals seek for high profits on their diversed portfolios and investment, we can help them by managing their funds and investing them based on our strategies and prediction generated through our tested ml models and algorithms thus generating high profit margins.

3) Institutions:

It includes domestic and foreign institutions and investment companies

Domestic-

1)India investment and asset management companies- These are Indian mutual funds that pool huge sum of money from individual investors and make investments. These investments are headed by fund managers.Companies can make investment using ml models for lower risk and higher profits.

2)Banks- Scheduled commercial banks also invest a small portion of the deposits they receive in the stock market.

Foreign-

1)Sovereign wealth funds: These are government owned investment funds. Surplus reserves are one of their popular sources of funding. The benefits of fund's investment are used for the citizens of that sovereign nation. Ex: Government of Singapore has large investments in Indian equity market.

5)External Search

https://upstox.com/learning-center/share-market/various-types-of-market-participants/

https://builtin.com/machine-learning/machine-learning-stock-prediction

https://www.simplilearn.com/tutorials/machine-learning-tutorial/stock-price-prediction-using-machine-learning

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3871022

https://www.forbes.com/sites/forbesfinancecouncil/2021/10/18/making-machine-learning-work-for-financial-market-prediction/?sh=7691866f2cce

6)Benchmarking:-various companies have started using Machine learning and artificial intelligence in the financial market to analyse the market news,identifying predicting future trends and stok prices.

Below are some of the companies that have used ML to break into the financial space and generate high profits.

1) Kavout-It is a company focused on applying AI for trading with its predictive equity ranking scores. Predictive equity ranking refers to the process of generating stock rankings based on a variety of input data, trading signals, and machine learning algorithms.

2) Precision Alpha- Precision Alpha uses six months of closing-price measurements and the mathematics of machine learning to calculate exact, closed-form expressions and numerically evaluate Market Probabilities, Market Energy, Market Power, Market Resistance, Market Noise, Market Temperature and Market Free Energy

3) MLQ.ai: In addition to traditional fundamentals, the platform features the following alternative data and machine-learning-based insights-1)Sentiment Analysis,2)ML estimates,3) Predictive Equity Rankings

7)Applicable patents

1) US20170315066A1 - Method and system for predicting stock prices using machine learning-This patent describes a method and system for predicting stock prices using machine learning. The method involves training a machine learning model on historical stock data and then using the model to predict future stock prices.

2) US20050091146A1- This patent describes a system and method for predicting stock prices using a combination of technical analysis and fundamental analysis.

3) US20180003513A1 - Method and system for predicting stock prices using deep learning- This patent describes a method and system for predicting stock prices using deep learning. Deep learning is a type of machine learning that uses artificial neural networks to learn from data.

4) US8285619B2 - Stock market prediction using natural language processing- his patent describes a method of using natural language processing (NLP) techniques to extract information from online news feeds and then using the information so extracted to predict changes in stock prices or volatilities.

8)Applicable Regulation

some of the most important applicable rules, regulations, and standards of organizations such as SEBI and NSE that will be needed to keep in mind while building an app to predict stock market prices, analyse market news, and predict trends using machine learning models and algorithms on previous stock data:

- SEBI (Stock Brokers and Sub-Brokers) Regulations, 1992 - These regulations regulate the activities of stock brokers and sub-brokers, who are individuals or firms that facilitate the trading of securities
- SEBI (Investment Advisers) Regulations, 2012 - These regulations regulate the activities of investment advisers, who are individuals or firms that provide investment advice to clients.
- NSE (Trading Regulations) - These regulations govern the trading of securities on NSE. They include provisions for the opening and closing of the market, the settlement of trades, and the prevention of market abuse.

In addition to the regulations of SEBI and NSE, we must also need to comply with other government policies and safety regulations such as-

- The Information Technology Act, 200-The Information Technology Act, 2000 is a comprehensive law that regulates the use of information technology in India. It includes provisions for protecting the privacy of users and for dealing with cybercrime.

9)Constraints

- Data collection from NSE and BSE and cleaning large amounts of datasets.
- Market Competition: The financial technology (fintech) industry is highly competitive.
- Security and Privacy: Handling sensitive financial data requires strict security measures to protect against cyber threats and unauthorized access.
- Budget: Consider the costs associated with data acquisition, hardware, software, licensing, infrastructure, personnel, and ongoing maintenance. Limited budget may restrict the extent of data that can be access or the computational resources you can employ.

10)Business opportunity-

With icreasing use of AI and ML to automate processes and services and with Higher risk while investing in stock markets due to lack of knowledge, untested strategies and unpredictivness of the market calls for a soln. There is a

- Increasing demand for tested strategies which will be based on ml models
- Demand for a deep market analysis based on the companies activities.
- Need of system to test strategies developed by traders on previous datasets of stocks

<u>We can generate Revenue through the following Streams-</u>

1)Charging commission-Charging commissions from individuals and companies to manage and invest their funds based on our tested strategies build on Ml models and algorithms.

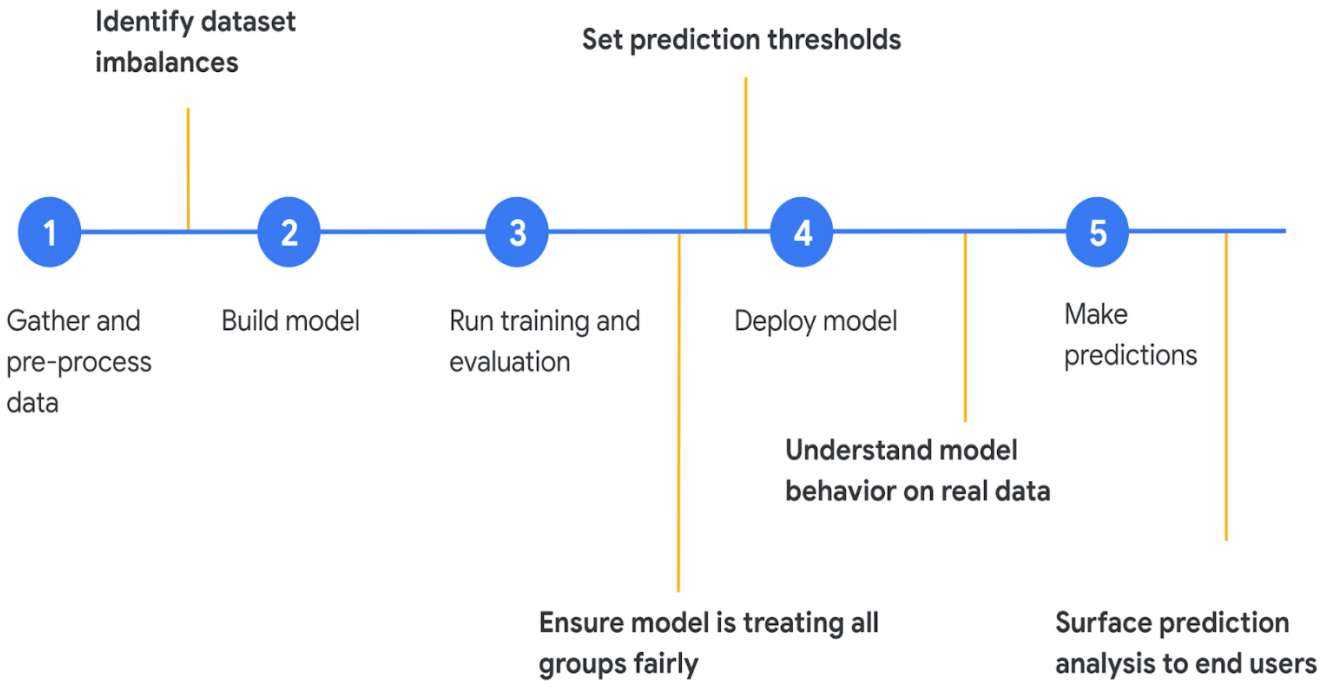
2)Charging Traders for testing their strategies on our large cleaned datasets system.

3)Subscription model-this plan will provide users with in depth knowledge of the market generated and trends in the stock prices using NLP models.

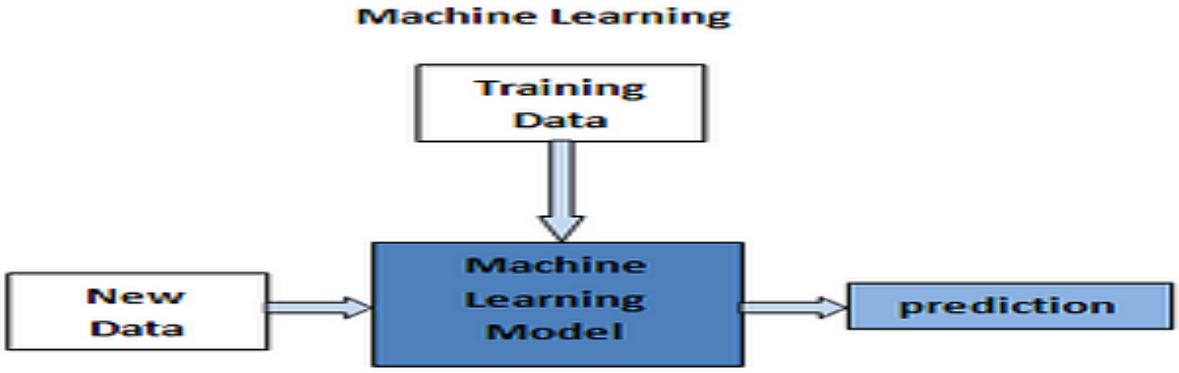
11)Concept Generation & Development

The basic concept is to train a ml model and make predictions on the stocks dataset extracted from various government organisation such as NSE. An ML model is a mathematical representation of a problem or system created using machine learning algorithms. It operates by taking training data and cleaning it, removing noise or outliers, to enhance its quality. The model is then developed by training it on the prepared data, where it learns patterns and relationships. During the training process, the model adjusts its internal parameters to minimize the difference between its predictions and the actual output labels. Once trained, the model can be used to make predictions or decisions based on new, unseen data by applying the learned patterns and associations.

1)Prediction model

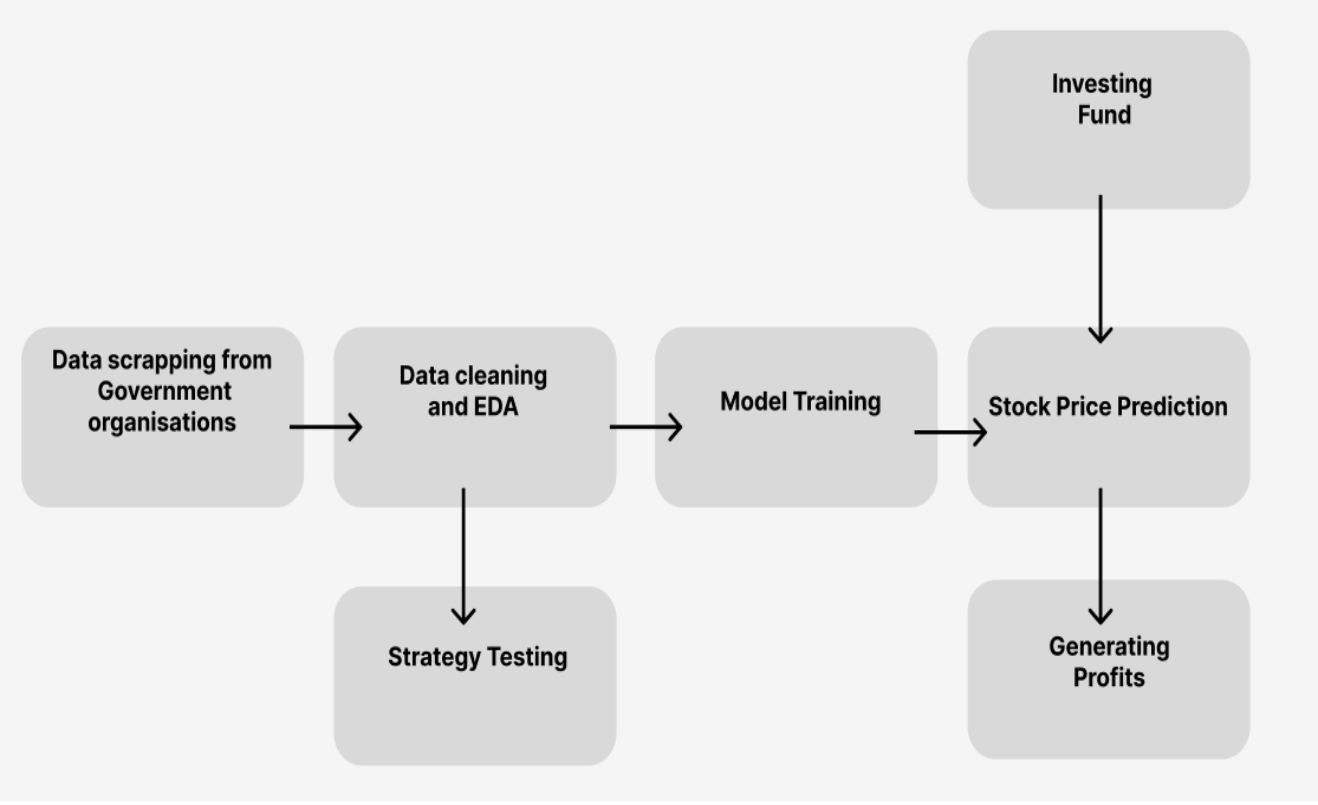


2)Strategy testing model



The ML model can be deployed in app with other features such as testing strategies and market analysis subscription model

12)Final Product Prototype



A product prototype is an early model of a product that will be built to see how it will work and looks, allowing us to test and improve it before making the final version. It helps bring ideas to life and ensures the product meets the desired requirements.

The above image is a basic model of how our stock price prediction will work and how traders can test their strategies.

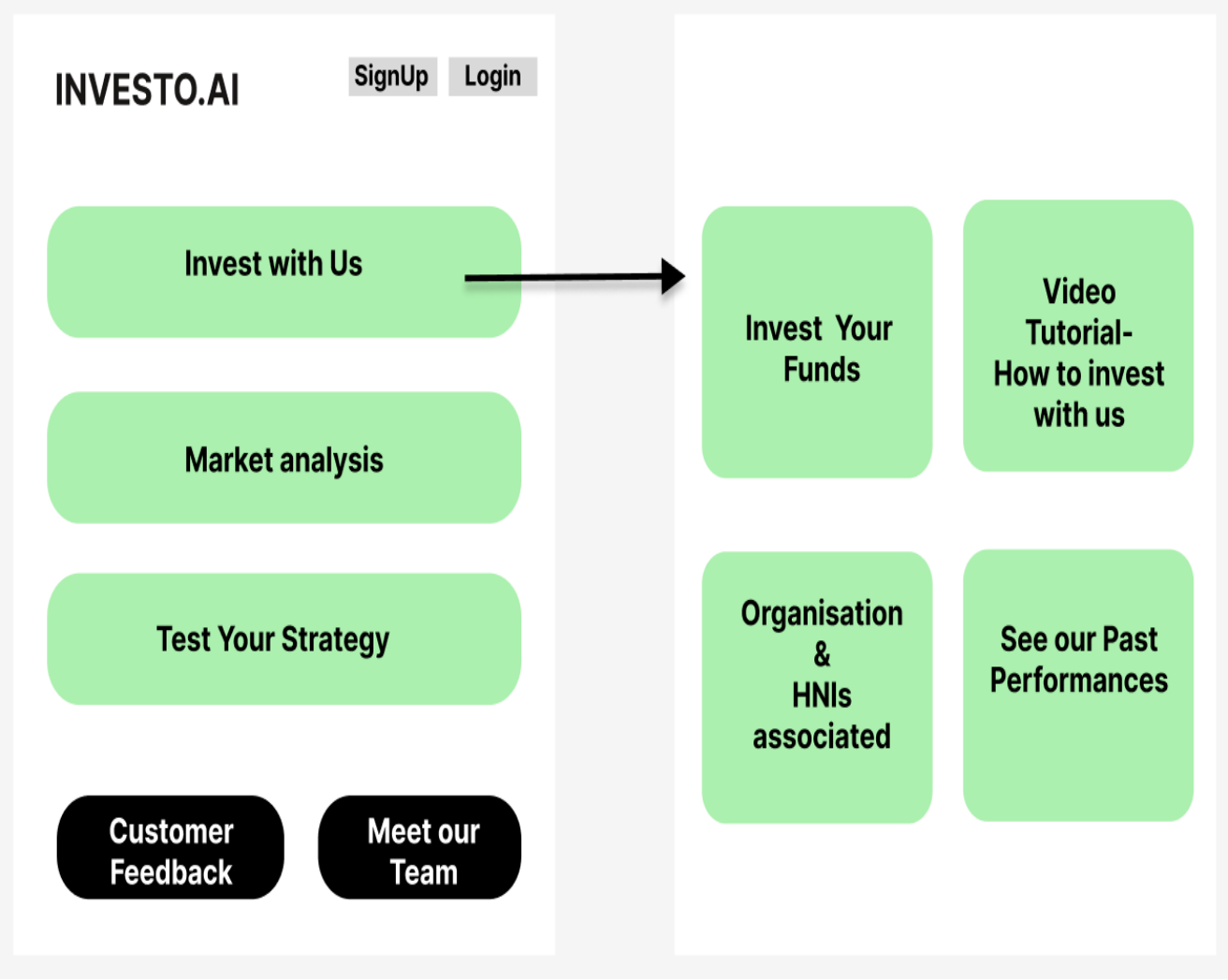
1)The data will be collected from the government organisations such as NSE & BSE.

2)The data will be cleaned properly , removing outliers and missing values etc for better analysis, on this datasets traders can test their strategies.

3)ML models and algorithms will be trained on the cleaned datasets.

4)Predictions will be made using various variables present in the dataset and then based on the prediction and strategies funds can be invested.

Below is a picture of the very basic **UX** of our **APPLICATION.**



The app will be based on three models as shown above:-

1)Fund Investment

2)In-Depth Market analysis

3)Testing your strategy

The In-Depth Market analysis will be prepared by a combination of NLP models ,trained on trusted news about companies and stocks and Expert financial advisory.

Other features in the app

- Customer Feedback-We will also have the section of Customer feedback where people can see the Feedback of our users
- Meet the team-Users can see the data scientist that will be working to devise the algorithms for our model
- Video tutorials-Tutorials that will explain how to invest with us and test out strategies.

13)Product Details

13.1)Data Sources

In India, there are several data sources from where we can collect datasets regarding stock prices. Here are some common data sources:

1) National Stock Exchange (NSE): The NSE provides historical and real-time stock price data for companies listed on the Indian stock market.

2) Bombay Stock Exchange (BSE): Similar to the NSE, the BSE is another major stock exchange in India that provides stock price data for listed companies

3) Financial Data Providers: Financial data providers such as Bloomberg, Refinitiv, and FactSet also cover stock prices for Indian companies. These providers often require a subscription or paid access to their data services.

4) Yahoo Finance: Yahoo Finance offers historical stock price data for Indian companies. Although the coverage may not be as extensive as dedicated data providers, it can still be a useful source for basic stock price information.

13.2)Team Required

1) Data Engineer
2) Data scientist/machine learning developer
3) Software developer
4) UI/UX designer

5) Financial Advisors
6) Fund Managers

14)Code Implementation

Data:-

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Symbol	Open	High	Low	Last Traded Price	Change	%Change	Traded Volume(lacs	Traded Value(crs)	52 Week High	52 Week Low	365 Days % Change	30 Days % Change
2	INDUSINDBK	1184.9	1222	1176.15	1217	60.75	5.25	44.47	538.82	1275.8	763.2	8.68	9.84
3	ADANIPORTS	804	826.3	790.1	825.3	40.9	5.21	104.46	849.55	987.85	651.95	10.9	3.01
4	BAJFINANCE	7335	7505	7306.35	7480	308.2	4.3	11.91	883.73	8050	5220	2.79	3.94
5	COALINDIA	218.3	224	218.1	223.7	8.75	4.07	108.62	240.54	240.5	139.15	17.77	3.41
6	TCS	3029.95	3098	3023	3097	112.05	3.75	21.46	659.42	4043	2926.1	17.92	1.16
7	UPL	678	691.7	674	689.8	24.7	3.71	23.72	162.04	848	607.5	1.87	7.64
8	HEROMOTOCO	2580	2627.9	2561.1	2614.95	89.4	3.54	3.91	101.94	2950	2146.85	8.43	7.99
9	BAJAJFINSV	1682	1705	1670.05	1702.9	56.3	3.42	18.06	304.96	1932.5	1072.72	90.44	90.2
10	JSWSTEEL	633.8	645	632.15	644.7	21.15	3.39	13.84	88.64	790	520.05	5.35	5.46
11	HINDALCO	396	400.9	390.3	393.8	12.85	3.37	90.17	355.67	636	308.95	22.13	8.27
12	TATASTEEL	100.3	101.5	100	101.4	3.05	3.1	401.87	405.04	142.66	82.7	92.34	5.14
13	HDFCBANK	1429.5	1458	1426.15	1456.15	42.95	3.04	57.69	835.6	1725	1271.6	8.17	2.6
14	HDFC	2328.7	2353.9	2315.25	2350.9	67.85	2.97	23.57	552.56	3021.1	2026	13.73	4.29
15	GRASIM	1667.8	1688.6	1651.4	1685	48.55	2.97	5.14	86.21	1929.8	1276.6	2.67	1.51
16	LT	1835.3	1880.95	1835.3	1873	53.7	2.95	16.38	305.67	2078.55	1456.35	9.2	4.84
17	WIPRO	399.95	406	399	405.6	11.1	2.81	62.96	254.21	739.85	384.6	36.73	0.02
18	ITC	328.7	334	328	333.1	8.7	2.68	88.03	291.78	349.55	207	40.79	1.29
19	AXISBANK	733.4	744.5	732	742	19.25	2.66	54.39	403.02	866.9	618.25	4.75	1.86
20	TATAMOTORS	406	408.95	402.9	407.9	10.25	2.58	127.61	518.75	536.7	335	19.27	11.13
21	HDFCLIFE	522.2	533.1	522.2	531.9	13.2	2.54	22.58	119.64	741.7	497.05	26.8	7.28
22	SBIN	533	535.8	527.8	532.95	13.2	2.54	87.46	465.62	578.5	425	15.07	1.1
23	INFY	1419.7	1432.25	1414	1428	34.25	2.46	66.31	946.06	1953.9	1355	14.94	2.28
24	EICHERMOT	3524	3579.9	3516.65	3548	84.65	2.44	9.61	341.58	3787.25	2159.55	29.21	4.34
25	TECHM	1022	1036	1017.9	1030	24.5	2.44	27.15	278.92	1838	943.7	26.38	3.29
26	KOTAKBANK	1808	1827.8	1802	1825.6	41.7	2.34	16.69	303.49	2253	1631	8.31	5.79
27	ICICIBANK	865	872.65	860.1	868	19.65	2.32	89.75	779.02	936.65	642.15	23.96	1.64
28	ADANIENT	3230	3291.2	3125	3229	71.7	2.27	56.42	1822.43	3885	1367.7	111.51	3.51

The data consists of various features, some of the most important among these are:-

1)opening & closeing price of stocks

2)Highest & lowest of the stocks on a day

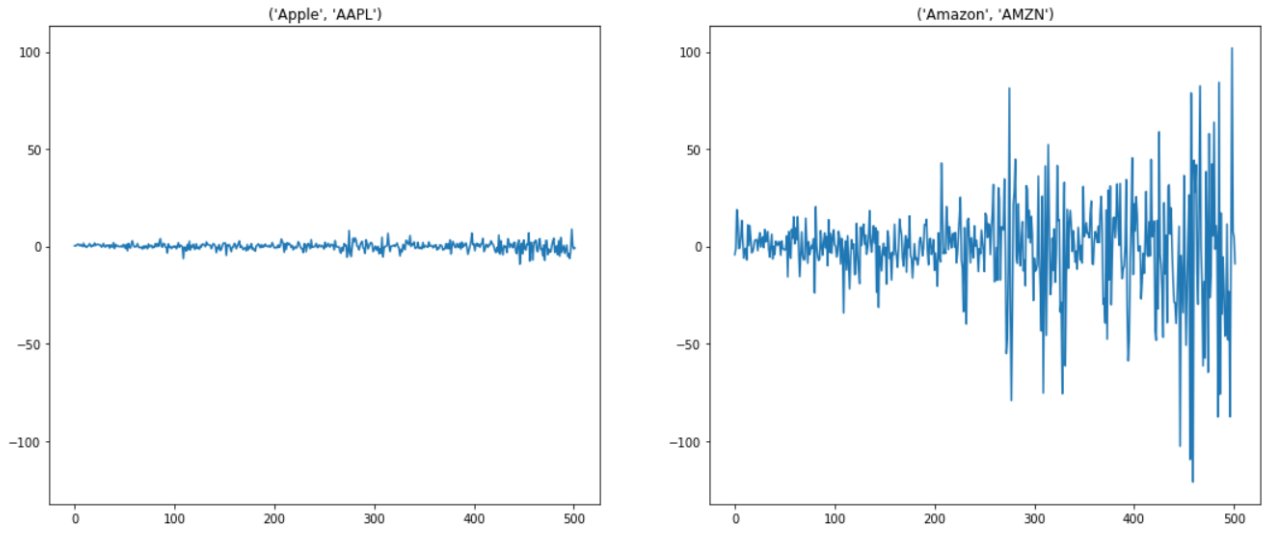
3)%change in price of the stocks

4)Volume of stocks traded

5)52 week low and high price

Code:-

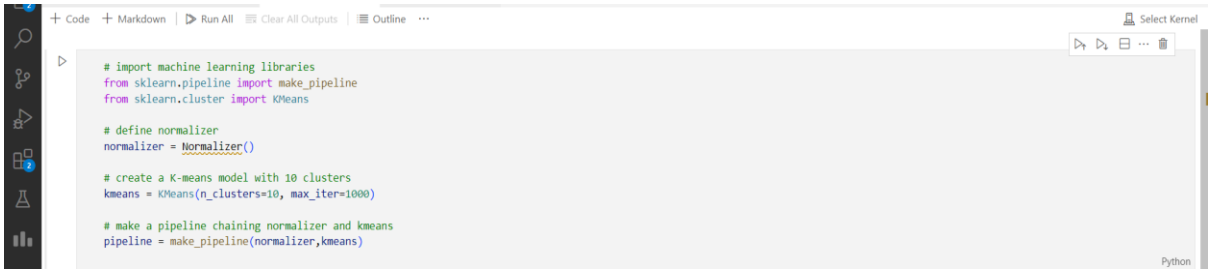
Data visualisation



Model Training:-

I have used a very basic algorithm here for the analysis:-

1)K-means clustering



Predictions-

	labels	companies
18	0	(Navistar, NAV)
2	1	(American Express, AXP)
4	1	(Bank of America, BAC)
0	2	(Apple, AAPL)
1	2	(Amazon, AMZN)
24	2	(Texas Instruments, TXN)
21	2	(Sony, SNE)
17	2	(Microsoft, MSFT)
14	2	(MasterCard, MA)
10	2	(Intel, INTC)
19	3	(Northrop Grumman, NOC)
13	3	(Lockheed Martin, LMT)
3	3	(Boeing, BA)

Other important Algorithms:-

1) **Support Vector Machines (SVM)**: SVM is a supervised learning algorithm that can be used for stock prediction. It attempts to find a hyperplane that separates the data into different classes (e.g., price increase or decrease) based on features such as historical prices, trading volumes, and technical indicators.
2) **Random Forest:** Random Forest is an ensemble learning method that combines multiple decision trees to make predictions. It can be used for stock analysis by considering features such as historical prices, trading volumes, market indicators, and company fundamentals.
3) **Long Short-Term Memory (LSTM):** LSTM is a type of recurrent neural network (RNN) that can capture sequential dependencies in data. It is commonly used for time series analysis and can be applied to stock price prediction by considering historical price data as input.
4) **Reinforcement Learning (RL):** RL is a branch of machine learning where an agent learns to make decisions based on rewards received from the environment. RL has been used for developing automated trading strategies in the stock market.
5) **Convolutional Neural Networks (CNN):** CNNs are mainly used for image analysis, but they can also be applied to stock prediction by treating historical price data as an image-like input. CNNs can learn spatial patterns in the data and extract relevant features.

   LSTMS and **CNN** are amongst the most used and popular algorithms in the industry for stock prediction.

15)Conclusion

Artificial intelligence (AI) and machine learning (ML) have revolutionized various industries, and the stock market is no exception. In this report, we explored the use of machine learning in stock price prediction.

With the increasing complexity of financial markets, accurate stock price prediction is crucial for investors, traders, and financial institutions. This report aims to address the problem statement of developing a reliable machine learning-based solution for predicting stock prices, catering to the market needs for informed investment decisions.

The report covers an analysis of target audiences, existing competitors, patent requirements, revenue models, and a product prototype for stock price prediction using machine learning algorithms. It provides valuable insights into the potential of AI/ML in the finance domain and highlights the importance of data quality and model robustness.

Moving forward, the future scope of this business lies in further refining the prediction models, incorporating more advanced algorithms, expanding the dataset coverage, and exploring real-time market analysis. With continued advancements in AI/ML technologies, the potential for accurate stock price prediction can greatly benefit investors and financial professionals in making informed decisions.