Stock Price Forecasting and Recommendation System using Machine Learning Techniques and Sentiment Analysis

Sushanth Kurdekar[1], Suhaas K P[2]

[1]Dept, of ISE, The National Institute of Engineering, Mysuru

[2Asst,] Professor, Dept., of ISE, The National Institute of Engineering, Mysuru

---***---

Abstract - *The stock market has always been bringing in people with its high returns however with its risks and also obstacles. A stock market illustrates savings and investments that are useful to raise the performance of the national economy. The future stock price has some anticipating relationships with the openly offered details of present and historical market indices. Regression is an analytical model that is understood to be reliable for time series projecting particularly for the short-term forecast. In this paper, we recommend a model for prognostication of the equities market fads based on the technical interpretation using historical financial market information and present view of trending key phrases. This model will help us to automate the process of prediction of future share price indices, thereby, aiding the monetary professionals to select much better timing for holding and marketing of supplies. The outcomes are received in terms of web-application as frontend and application will certainly be done by python maker finding out libraries. The obtained outcomes confess that the version has a strong ability for short-term forecasts of financial market fads.*

Key Words: Data, Stock Market, Stock Price Prediction, Machine learning, Regression, Python

1. INTRODUCTION

Forecasting the supply exchange information is a monetary subject area that involves a presumption that fundamental details openly available in history have some anticipating interconnection to the future financial market returns. Stock market projecting has discovered the market patterns, intending financial investment strategy, identifying the ideal time to buy the share. A financial business industry is a nondirected, nonparametric structure that is difficult to design with any kind of sensible exactness.

Since the problem in the diverse world of the supply market, Stock fad forecasting is taken into consideration as one of the most challenging jobs to achieve in cash relevant determining. Several investors in the financial market are finding a strategy that can guarantee very easy financial gain by forecasting the supply fads and also decreasing the threat of investing. This inspires the scientists in the domain name area to establish new forecasting versions [1].

Regression evaluation techniques make use of verifiable details as the facility for evaluating future outcomes. The twitter point of view, data can comprise the values accumulated at the end of each forecast. The objective is to discover if there is any type of connection between the data accumulated until now as well as in what means the data changes. To decrease the threat of financial investment, the exchange of equities between the seller and also buyers is also helped with the financial exchanges. A financial market is a federation or a location where the supply traders or financiers can take care of supplies. Some of the instances for stock market companies include NASDAQ, BSE, NSE, and so on [1][2].

2. SYSTEM ANALYSIS

A stock market shows savings and also stakes that are useful to improve the fruitfulness of nationwide economics. The subsequent financial returns have some predictive partnerships with the openly offered information about historic and existing stock exchange indices. The capitalists decide the best time to sell/buy/hold a share in the securities market based upon the previous partnership [2]. Every financier has an interest in anticipating the future financial prices, whether the financier may be a long-term financier or a day-trader. This has a significant difficulty in style and also establishes a reliable and reliable predictive design that helps the financiers to make appropriate choices [5].

Capitalists in the securities market have tried to discover a solution to estimate the supply patterns to choose the much better timing to sell or get or hold a share. Projecting the financial trends have been done both on qualitative as well as quantitative analysis. There are lots of statistical designs available for forecasting supply fads and also choosing a proper design for a projecting application relies on the style of the information [6].

2.1 Existing System

The authors of [3] offered a technique that utilized data mining approaches and semantic networks for forecasting securities market returns. Among the significant monetary topics that have fascinated the researcher's interest for many years is anticipating the supply returns. In the existing system, Investors can discover the data as well as anticipate the results just for foreign exchanges such as NASDAQ, NYSE, and also the very same prediction versions; they cannot offer NSE as well as BSE information as input to anticipate the data for Indian exchanges.

2.2 Proposed System

In this work, we suggest a forecast model for the Indian financial market such as NSE, BSE. This model will computerize the process of change of financial price indices based on technical interpretation and belief analysis of twitter information to choose the far better timing for acquiring as well as selling supplies. Information mining strategies are used to create the forecast design as well as Python programs language is made use of for visualization of results.

3. IMPLEMENTATION

Data mining can be taken as an expertise exploration process. Information extraction techniques are created to attend to the issues by offering a reputable version with data mining attributes. To create a model that investigates the supply patterns by using the historic supply data [2].

System style is a version that defines the habits of a system in the conceptual design. The significant systems are broken down right into subordinate systems to offer a comparable collection of services. The start format strategy of viewing these sub-systems and also developing a framework for sub-systems control as well as teamwork is called architecture style [3].
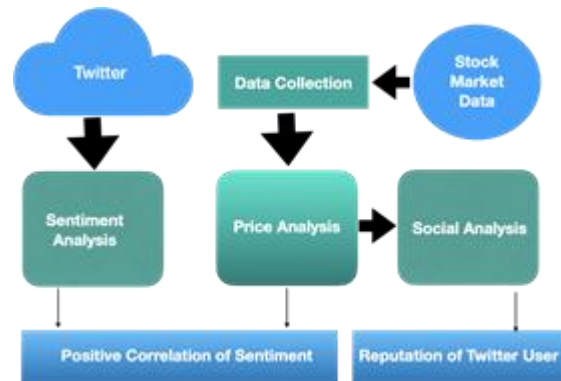

Fig-2: System architecture

Fig-3: Data Flow Diagram

3.1 Understanding the Objectives

The number one step in developing a project is to recognize the reason which involves a know-how of the cause and also fundamentals of a machine. This expertise is used as a trouble summary and a preparatory system to achieve the expectations. The reason for our activity is neither to develop a device that makes billions nor to throw away billions too. The cause is to create a machine that finds the direction of adjustment of supply price indices based totally at the correlations in between supply prices as nicely as useful resource the investors inside the deliver market in taking a choice whether to buy/sell/hold a deliver by supplying the consequences in phrases of visualizations [2].

3.2 Data Collection

As soon as the understanding of the intention is over, the subsequent movement is to accumulate the data. Data collection entails the understanding of initial monitoring of the information to determine the beneficial parts from theories of the concealed info. Here we use API to gather the data from Yahoo Finance and Twitter.

```python
import yfinance as yf
import pandas as pd

data = yf.download('GOOGL', start="2010-01-10",
                   end="2020-4-9")
[*********************100%***********************]  1
of 1 completed

df = pd.DataFrame(data=data)

df = df[['Open', 'High', 'Low', 'Close']]

df.head()
```

	Open	High	Low	Close
Date				
2010-01-11	302.532532	302.532532	297.317322	300.855865
2010-01-12	299.124115	299.379395	294.294281	295.535522
2010-01-13	288.533539	294.484497	287.237244	293.838837
2010-01-14	292.242249	297.397400	291.696686	295.220215
2010-01-15	296.966980	297.077087	289.309296	290.290283

Fig-4: Collecting Data

3.3 Data Pre-Processing

The statistics pre-processing segment involves all the sports to put together the final dataset from the primary raw facts. The records prep paintings jobs may be performed several times as there is no certain order. These duties consist of the selection of a record, table, quality, and cleaning of statistics for modeling devices [3].

1. Calculating Simple Moving Average.

$$\bar{p}_{\text{SM}} = \frac{p_M + p_{M-1} + \cdots + p_{M-(n-1)}}{n}$$

$$= \frac{1}{n} \sum_{i=0}^{n-1} p_{M-i}.$$

2. Calculating the mean using mean method

```
fb['MA10'] = fb['Close'].rolling(10).mean()
fb['MA50'] = fb['Close'].rolling(50).mean()
```


Fig-5: Moving Average

3.4 Data Processing: Training

In technological evaluation, financiers use the automobile regressive as well as moving average versions to anticipate the stock patterns. In this version, we utilize multiple algorithms such as Linear Regression, K-Nearest neighbor to contrast the ideal design for our forecast [3][4].

1. Regression

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j}{\sum_{i=1}^{n}(x_i^2) - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2}$$

$$= \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{\text{Cov}[x,y]}{\text{Var}[x]} = r_{xy}\frac{s_y}{s_x},$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x},$$

2. K-Nearest Neighbor for Regression

$$D_H = \sum_{i=1}^{k} |x_i - y_i|$$

$$x = y \Rightarrow D = 0$$

$$x \neq y \Rightarrow D = 1$$

3.5 Forecasting Results

The strategy of making predictions of the future via relying upon the historical and existing information is referred to as projects. Various prediction methods are made use of by way of the financial analysts to evaluate the destiny inventory trends value. Forecast additionally provides a vast criterion for businesses that have a long-term belief of activities [3].

Fig-6: Forecasting Future Price

3.6 Visualization

Data visualization is a graphical depiction of the numerical information. In our method, after anticipating the securities marketplace fads, we visualize the outcomes for short-term financial investment support in terms of line charts, bar charts, candlesticks-charts, as well as pie charts. Right here x-axis reveals the time in terms of year/months/days, as well as y-axis, reveals the supply rate worth [4].

Fig-7: Visualization

3.7 View & Analyze Results

As quickly as after outlining the results in-phrases of visualizations we will learn the connections to get the transient forecasts. In the following area, we offer many of the screenshots wherein the financier can evaluate and also anticipate the future inventory trends of an employer at a sure time. The investors inside the inventory market can use this as help to sell/buy/keep a share [3].

Fig-8: Returns

Fig-9: Moving Average Comparison

Fig-10: Regression Model Prediction

Fig-11: K-Nearest Neighbor for Regression

Fig-12: Bayesian Model Prediction

Fig-13: Confidence Score

4. CONCLUSION

In this paper, we have installed a prediction model for looking forward to the inventory market patterns based at the technical analysis using ancient stock marketplace statistics in addition to Twitter sentiments which informs the present pattern of the supply. The experimental results proved the prospective model to predict the inventory rate indices temporarily. This could direct the investors inside the inventory market to make profitable economic investment decisions whether or not to buy/sell/maintain a share [4].

REFERENCES

[1] Mahantesh C. Angadi, Amogh P. Kulkarni, "Time Series Data Analysis for Stock Market Prediction using Data Mining Techniques with R", International Journal of Advanced Research in Computer Science(IJARCS), Volume 6, No. 6, July-August 2015, ISSN No.0976-5697.

[2] K. Hiba Sadia, Aditya Sharma, Adarrsh Paul, SarmisthaPadhi, Saurav Sanyal, "Stock Market Prediction Using Machine Learning Algorithms", International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, Volume-8 Issue-4, April 2019.

[3] Vivek Rajput, Sarika Bobde, "STOCK MARKET FORECASTING TECHNIQUES: LITERATURE SURVEY", International Journal of Computer Science and Mobile Computing(IJCSMC), 6, June 2016, pg.500 – 506, ISSN 2320–088X.

[4] Li Zhe; "Research on China's stock exchange markets: problems and improvements", International Conference on Education and Management Technology, 2010. pp 465-469.

[5] https://www.analyticsvidhya.com

[6] https://www.wikipedia.org