**Stock Price Prediction**

TEAM MEMBERS

|  |  |  |
| --- | --- | --- |
| Sr.no. | Name | Registration no. |
| 1 | Himanshu kumar singh | RA2011047010137 |
| 2 | Rohan Chauhan | RA2011047010109 |
| 3 | Rishi Aslaliya | RA2011047010100 |

**Abstract**

Researchers have been studying different methods to effectively predict the stock market price. Useful prediction systems allow traders to get better insights about data such as: future trends. Also, investors have a major benefit since the analysis give future conditions of the market. One such method is to use machine learning algorithms for forecasting. This project’s objective is to improve the quality of output of stock market predicted by using stock value. A number of researchers have come up with various ways to solve this problem, mainly there are traditional methods so far, such as artificial neural network is a way to get hidden patterns and classify the data which is used in predicting stock market. This project proposes a different method for prognosting stock market prices. It does not fit the data to a specific model; rather we are identifying the latent dynamics existing in the data using machine learning architectures. In this work we use Machine learning architectures Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN) and Hybrid approach of LSTM + CNN for the price forecasting of NSE listed companies and differentiating their performance. On a long term basis, sling window approach has been applied and the performance was assessed by using root mean square error.

**INTRODUCTION**

**Stock Price Prediction**

Due to the high profit of the stock market, it is one of the most popular investments. People investigated for methods and tools that would increase their gains while minimizing the risk, as the level of trading and investing grew. Two stock exchanges namely- the National Stock Exchange (NSE) and the Bombay Stock Exchange (BSE), which are the most of the trading in Indian Stock Market takes place. Sensex and Nifty are the two prominent Indian Market Indexes. Since the prices in the stock market are dynamic, the stock market prediction is complicated. From gradually the very past years some forecasting models are developed for this kind of purpose and they had been applied to money market prediction. Generally, this classification

is done by:

1. Time series analysis

2. Fundamental analysis

3. Technical analysis

**Time Series Analysis**

The definition of forecasting can be like this the valuation of some upcoming result or results

by analysing the past data. It extents different areas like industry and business, economics and

finance, environmental science. Forecasting problems can be classified as follows:

* Long term forecasting (estimation beyond 2 years)
* Medium-term forecasting (estimation for 1 to 2 years)
* Short term forecasting (estimation for weeks or months, days, minutes, few seconds)

The analysis [1] of time consist of several forecasting problems. The designation of a time series is a linear classification of observations for a selected variable. The variable of the stock price in our case. Which can weather multivariate or univariate? Only particular stock is included in the univariate data while more than one company for various instances of time is added in multivariate. For investigating trends, patterns and cycle or periods the analysis of time series advantages in the present data. In spending money wisely an early data of the bullish or bearish in the case of the stock market. Also, for categorizing the best-performing companies the analysis of patterns plays its role for a specific period. This makes forecasting as well as time series analysis an important research area

Fundamental analysis

Fundamental Analysts are concerned with the business that reasons the stock itself. They assess a company's historical performance as well as the reliability of its accounts. Different performance shares are created that aid the fundamental forecaster with calculating the validity of a stock, such as the P/E ratio. Warren Buffett is probably the foremost renowned of all Fundamental Analysts.

What fundamental analysis within the stock market is making an attempt to reach, is organizing the true value of a stock, that then will be matched with the worth it is being listed on stock markets and so finding out whether or not the stock on the market is undervalued or not. Find out the correct value will be completed by numerous strategies with primarily a similar principle. The principle is that an organization is price all of its future profits. Those future profits has to be discounted to their current value. This principle goes on the theory that a business is all about profits and nothing else. Differing to technical analysis, the fundamental analysis is assumed as further as a long approach. Fundamental analysis is created on conviction that hominoid society desires capital to make progress and if the company works well, than it should be rewarded with an additional capital and outcome in a surge in stock price. Fundamental analysis is usually used by the fund managers as it is the maximum sensible, objective and prepared from openly existing data like financial statement analysis. One more meaning of fundamental analysis is on the far side bottom-up business analysis, it discusses the top-down analysis since initial analysing the world economy, followed by country analysis and also sector analysis, and last the company level analysis.

**Technical analysis**

Chartists or the technical analysts are not involved with any other of the fundamentals of the company. The long run price of a stock based generally exclusively on the trends of the past value (a form of time series analysis) that is set by them. The head and shoulders or cup and saucer are various numerous patterns that are employed. Also the techniques, patterns are used just like the oscillators, exponential moving average (EMA), support and momentum and volume indicators. Candlestick patterns, believed to have been initial developed by Japanese rice merchants, are nowadays widely used by technical analysts. For the short-term approaches, the technical analysis is used compare to long-run ones. So, in commodities and forex markets it is more predominant wherever traders target short-term price movements. There are basic rules are used in this analysis, first all significant about a company is already priced into the

**Applications**

* Business
* Companies
* Insurance company
* Government Agency
* This application is helpful for stock investors, sellers, buyers, brokers.

**Objectives**

A stock market prediction is described as an action of attempting to classify the future value of the company stock or other financial investment traded on the stock exchange. The forthcoming price of a stock of the successful estimation is called the Yield significant profit. This helps you to invest wisely for making good profits.

**Motivation**

The future price of a stock is the main motivation behind the stock price prediction. In various cases like business and industry, environmental science, finance and economics motivation can be useful. The future value of the company’s stock can be determining.

**Prerequisites**

1. Statistics
2. [Linear Algebra](https://www.edureka.co/blog/prerequisites-for-machine-learning/#Linear%20Algebra)
3. [Calculus](https://www.edureka.co/blog/prerequisites-for-machine-learning/#Calculus)
4. [Probability](https://www.edureka.co/blog/prerequisites-for-machine-learning/#Probability)
5. [Programming Languages](https://www.edureka.co/blog/prerequisites-for-machine-learning/#Programming%20Languages)

**Statistics**

Statistics contain tools that can be used to get some outcome from the data. There is descriptive statistics which is used to transform raw data in some important information. Also, inferential statistics can be used to get important information from a sample of data instead of using complete dataset.

To learn more about Statistics you can go through the following blogs:

1. All You Need To Know About Statistics And Probability
2. A Complete Guide To Maths And Statistics For Data Science

**Linear Algebra**

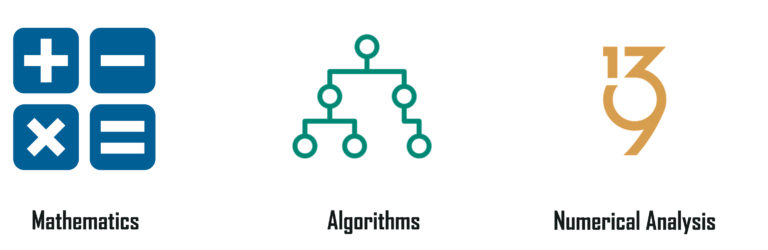
Linear algebra deals with vectors, matrices, and linear transformations. It is very important in machine learning as it can be used to transform and perform operations on the dataset.

**Calculus**

Calculus is an important field in mathematics and it plays an integral role in many machine learning algorithms. Data set having multiple features are used to build machine learning models as features are multiple multivariable calculus plays an important role to build a machine learning model. Integrations and Differentiations are a must.

**Probability**

Probability helps predict the likelihood of the occurrences, It helps us to reason the situation may or may not happen again. For machine learning, probability is a foundation.



**Programming language**

It is essential to know programming languages like R and Python in order to implement the whole Machine Learning process. Python and R both provide in-built libraries that make it very easy to implement Machine Learning algorithms.



Apart from having basic programming knowledge, it is also important that you know how to extract, process and analyze data. This is one of the most important skills that is needed for Machine Learning.

To learn more about the programming languages for Machine Learning, you can go through the following blogs:

1. A Comprehensive Guide To R For Data Science
2. Python for Data Science – How to Implement Python Libraries
3. The Best Python Libraries For Data Science And Machine Learning

Prerequisite links

|  |  |  |
| --- | --- | --- |
| Sr. no. | link | YouTube / github |
| 1 | <https://youtu.be/lncoLfue_Y4> | Youtube |
| 2 | <https://youtu.be/OXwZtlcTiuk> | Youtube |
| 3 | <https://github.com/scorpionhiccup/StockPricePrediction> | github |
| 4 | <https://github.com/tencia/stocks_rnn> | github |
| 5 | <https://github.com/IsaacChanghau/StockPrediction> | github |
| 6 | <https://github.com/wzchen/stock_market_prediction> | github |
| 7 | <https://www.youtube.com/watch?v=H6du_pfuznE> | Youtube |
| 8 | <https://www.youtube.com/watch?v=QIUxPv5PJOY> | Youtube |

**Survey**

Existing project in field

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sr. no. | Project name | Publishing year | Journal name | Author name |
| 1 | Stock Price Forecasting Using Data From Yahoo Finance and Analysing Seasonal and Nonseasonal Trend | 2018 | 2018 IEEE | Jai Jagwani, Hardik Sachdeva, Manav Gupta, Alka Singhal |
| 2 | Stock Market Prediction Using Machine Learning | 2018 | 2018 IEEE | Ishita Parmar, Ridam Arora, Lokesh Chouhan, Navanshu Agarwal, Shikhin Gupta, Sheirsh Saxena, Himanshu Dhiman |
| 3 | Multi-Category Events Driven Stock Price Trends Prediction | 2018 | 2018 IEEE | Youxun Lei, Kaiyue Zhou, Yuchen Liu |
| 4 | Short-term stock market price trend prediction using a comprehensive deep learning system | 2020 | SPRINGER | Jingyi Shen & M. Omair Shafiq |
| 5 | Stock Price Prediction using Machine Learning and Deep Learning | 2021 | IEEE | Pratheeth S &Vishnu Prasad R |
| 6 | Stock Price Prediction Based on LSTM Deep Learning Model | 2021 | IEEE | Kavinnilaa J ,Hemalatha E,Minu Susan Jacob, Dhanalakshmi R |
| 7 | Share Price Prediction using Machine Learning Technique | 2018 | IEEE | Jeevan B, Naresh E, Vijaya kumar B P, Prashanth Kambli |
| 8 | Forecasting stock price in two ways based on LSTM neural network | 2019 | IEEE | Jingyi Du, Qingli Liu, Kang Chen, Jiacheng Wang |
| 9 | Stock Price Prediction Based on Information Entropy and Artificial Neural Network | 2019 | IEEE | Zang Yeze,  Wang Yiying |
| 10 | Stock Market Prediction Using Machine Learning Techniques | 2016 | IEEE | Mehak Usmani, Syed Hasan Adil, Kamran Raza, Syed Saad Azhar Ali |

**Summary of the survey**

Here, I have reviewed various approaches for Stock price prediction. All approaches have

their own advantages and disadvantages. CNN & LSTM is a most popular algorithm to

prediction the stock price but there are some challenges in this method like use to need a lot

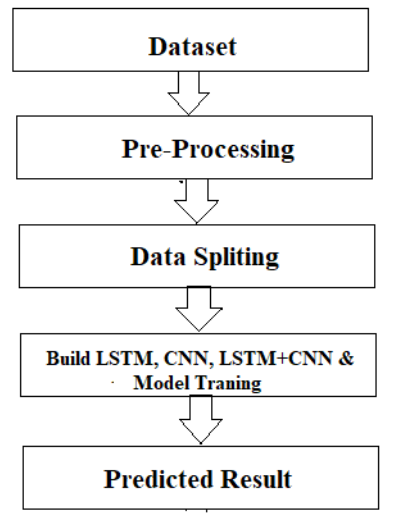
of training data, High computational cost, without GPU data quite slow to train, depend on

any previous information for prediction. A hybrid approach can be used to overcome these

issues. While machine learning is able to provide highly accurate prediction result using

standards tools and also outperforms all standard prediction methods

**Proposed work**



The system presented here composes of five modules:-

1. Input as Dataset

2. Pre processing

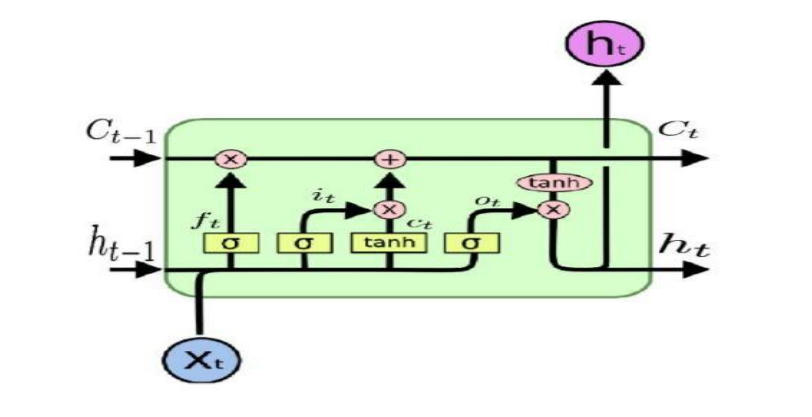
3. Data splitting

4. Build & Model train Lstm, CNN and Hybrid approach of LSTM+CNN

5. Output as Predicted Result

Attribute such as: price of open, high, low, close, adjusted close price taken from huge dataset are fed as input to the models for training to pre-process the data techniques like normalization & one hot encoding in applied on dataset. After this data is divided in two sets namely training & testing which are ratio of 80:20 respectively. Then, this set are used to train a model using 3 different approaches: LSTM, CNN and Hybrid approach of LSTM+CNNS. Finally, all these modules are evaluated using Root mean square error.

**Working of LSTM model**



Long Short Term Memory is a kind of recurrent neural network. In RNN output from the last

step is fed as input within the present step. It tackled the matter of long-term dependencies of

RNN within which the RNN will not predict the word hold on within the long term memory

however can offer additional accurate forecasts from the recent info. Because the gap length

will increases RNN does not offer an economical performance. LSTM will by default retain

the knowledge for a long period of time. It is used for processing, predicting and classifying

on the basis of time-series data.

* Structure of LSTM:
* LSTM has a chain organization that contains four neural networks and different memory
* blocks called cells.
* LSTM has a new structure called a memory cell. The memory cell makes the decisions about
* what information to store, and when to allow reading, writing and forgetting.
* A memory cell contains three main gates:

Input gate- a new value flows into the memory cell.

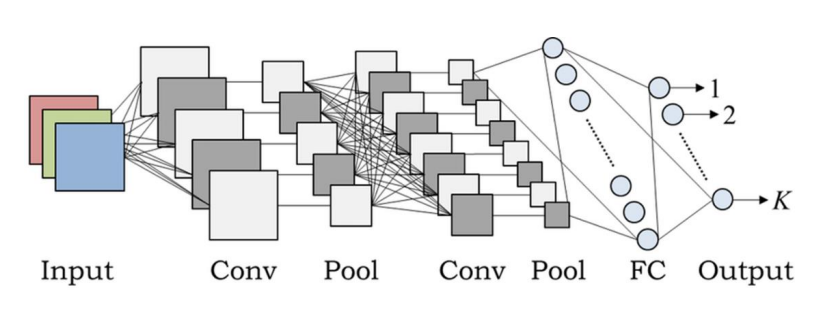
Forget gate- a value remains in the memory cell.

Output gate- value in the memory cell is used to compute the output

Applications of LSTM includes:

* Language Modelling
* Machine Translation
* Image Captioning
* Handwriting generation
* Question Answering Chatbot

**Working of CNN model**



Layer of CNN model:

* Convolution
* MAX Pooling
* Dropout
* Flatten
* Dense
* Activation
* Convolution:In the Convolution extract the featured from the input image. It given the output in matrix form.
* MAX Pooling: In the MAX polling it takes the largest element from a rectified feature map.
* Dropout: Dropout is randomly selected neurons are ignored during training.
* Flatten: Flatten feed output into a fully connected layer. It gives data in list form.
* Dense: A Linear operation in which every input is connected to every output by weight. It followed by a nonlinear activation function.
* Activation: It used sigmoid function and predict the probability 0 and 1.
* Applications of CNN includes:
* Decoding Facial Recognition
* Analyzing Documents

**Hybrid Approach of LSTM + CNN**

In the hybrid approach, the Convolutional Neural Networks (CNNs) offer benefits in choosing sensible options and Long Short-Term Memory (LSTM) networks have proven sensible skills to find out to learn sequential data. Each approaches are reported to produce improved result. CNNs to possess to convolute filters over every input layer so as to get the simple options and CNNs have shown enhancements in computer vision, natural language processing and different tasks . CNN may be a powerful tool to pick out features in order to improve the prediction accuracy . The capabilities of LSTMs in learning data series by considering the previous outputsThe multiple convolutional filters slide over the matrix to produce a new feature map and also the filters have numerous completely different sizes to generate different features. The Max pooling layer is to calculate the most value as a corresponding feature to a particular filter. The output vectors of the Max-pooling layer become inputs to the LSTM networks to measure the long-run dependencies of feature sequences. One in all the benefits of the LSTMs is that the ability to capture the sequential data by considering the previous data. This layer takes the output vectors from the dropout layer as inputs. This layer include a set number of units or cells and also the input of every cell is that the output from the dropout layer. The final output of this layer has the same number of units within the network the outputs from LSTMs are merged and combined in one matrix then passed to a fully connected layer. The array is converted into a single output in the range between 0 and 1 using the fully connected layer, in order to be finally classified using sigmoid function

|  |  |  |
| --- | --- | --- |
| Sr. no. | Name | contribution |
| 1 | Himanshu Singh | Model building, problem statement identification, algorithm building |
| 2 | Rohan Chauhan | Model testing and training |
| 3 | Rishi Aslaliya | Dataset collection |