



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 12, December 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.524

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Analyzing the Stock Price with the Help of Machine Learning for N Number of Days

Mr. Shree Kantesh K H¹, Dr. Vikas S²

PG Studies, Department of CSE (MCA), Visvesvaraya Technological University, Mysore, India^{1,2}

ABSTRACT: Forecasting the stock price presents a significant challenge because of market volatility and the influence of numerous economic, political, and psychological factors. This study leverages machine learning models, specifically LSTM networks and Facebook's Prophet (fbprophet) model, to predict Indian stock prices across the Bombay Stock Exchange (BSE) and National Stock Exchange (NSE) for an extended period of over 20 years. Using historical data, including periods of major economic disruptions such as the 2008 recession and the 2020 COVID-19 pandemic, we evaluate the efficacy of these models in capturing trends and making predictions. Techniques of series analysis on time and neural methodology of networks are applied to enhance the potential of deep learning for short-term stock price forecasting. Our findings indicate that LSTM networks, despite longer training times, outperform traditional models in accuracy and predictive reliability. This study also marks the use of indicators of technical analysis and emotion analysis of news events as integral factors in model optimization. These insights contribute to developing a robust model that minimizes human emotional bias, providing investors with data-driven decision support for market trading.

KEYWORDS: LSTM, fbprophet, time series analysis, stock price prediction, trend forecasting, technical indicators

I. INTRODUCTION

Data analysis is the way of gathering bare data and converting it into useful insights that the users will use to make decisions. In order to find important information, make inferences, and support decision-making, it focuses on examining, cleaning, manipulating, and modeling data. In today's business world, The data insights contributes significantly to scientific decision-making, helps organizations use data-driven decision-making, and improves their operational efficiency.

II. LITERATURE SURVEY

The mentioned papers were gone through to get an overview of the methods that were before used to forecast the analysis of stocks market.

LSTM Fully Convolutional Networks for Time Series Classification

-FazleKarim, SomshubraMajumdar, HoushangDarabi and Shun Chen [1]

With the proposed models with the classification, we significantly improve the state-of-the-art time series using deep neural networks. Our baseline models are trainable end-to-end with nominal preprocessing, both with and without fine-tuning, and can reach significantly superior performance.

The performance of FCN models can be greatly enhanced by LSTM-FCNs with just a slight increase in the number of parameters. With an LSTM-FCN, which provides a strong baseline by itself, the decision-making process of the LSTM RNN can be visually examined. As a general technique, fine-tuning a model can further improve its performance.

Learning Long term Dependencies with Gradient Descent is difficult

-Yoshuabengio, Patrice Simard and Paolo Frasconi [10]

The background information provided by recurrent networks is often superior to that of static networks. They may not be sufficiently trained utilizing the factor off gradient descent of an error criteria for a task involving long-term

dependencies.

It has been found that the system would not be robust to input noise or efficiently function when the long-term context is required.

trainable using gradient descent. This work presents a theoretical result that is applicable to any error criterion, not just mean square error.

Enhancing the RSI Financial Indicator's N Calculation using Neural Networks

-Ricardo Colomo-Palacio Giner, Alejandro Rodríguez-González, and Fernando GuldrisIglesiasPosada-Gomez Ruben and Alor-Hernandez [8]

In recent times, trading analytics and decision driven support systems have gained more attention. Despite its volatility, it is either highly complex and volatile or nonlinear and dynamic, rather than completely random. Stock movement is influenced by both random (like emergent events or daily operating fluctuations) and term (like a gradual change in strength between the buying and selling sides) factors.

Stock Trend Prediction Using Simple Moving Average[10] Supportedby News Classification

-Stefan Lauren Dra. Harley S., M.Sc. [5]

The simple moving average is one of several approaches and strategies for evaluating time series. Time series analysis is a timely organized data processing technique for identifying statistics or significant features for other applications. By figuring out The movement and trend of stocks are displayed by the simple moving average, which displays the average value of stock prices over a given time period. The prices that are applied at the end of the day are known as closing prices.

This method smoothes the trend movement by reducing noise.

The main objective of financial news classification is to categorize each news item and assess its sentiment worth. In trade, a sentiment value greater than 0 signifies good news, whereas a sentiment number less than 0 implies unfavorable news. News with a sentiment value of 0 will not be included since its neutrality has no effect on the stock trend.

VISUALIZING AND UNDERSTANDING RECURRENT NETWORKS

-Andrej Karpathy, Justin Johnson, Li Fei-Fei [4]

Among the error kinds seen in recurrent neural networks are character-level language models, training dynamics, representations, and predictions. These test beds can be interpreted. In particular, the qualitative visualization experiments, cell activation statistics, and comparisons to finite-horizon n-gram models demonstrate that these networks are capable of learning and frequently interpreting long-range interactions on real-world data; however, further architectural innovation may be necessary to rectify the remaining errors.

III. MOTIVATION

Since money evokes emotion and individuals attempt to make decisions by appealing to their emotions, we in the twenty-first century rely on human efforts to predict which stocks will rise or become the target of bulls they often make poor decisions. For this reason, using machine learning—which is, the ability of a system to learn from data to make prediction tasks easier is becoming popular because data scientists and analysts must series the sequence of raw data in order to provide insightful information or predictions.

Because computers can get data on time every second and have trained algorithms that provide investors with the best information possible without allowing their emotions to interfere, machine learning techniques are employed in stock price prediction.

By improving productivity, decreasing manufacturing costs, and replacing humans in numerous industries, machines have made life simpler for people.

The financial markets provide a unique trading and investment platform that can be used by any device with an Internet connection. People now have several options to increase the growth of their money thanks to the emergence of stock markets. Furthermore, it created a variety of fund types, including index funds, mutual funds, and hedge funds, so that individuals and organizations may invest their money in accordance with their risk tolerance.

IV. PROBLEM DOMAIN

Every minute of every day, the stocks are in the news. Every time it hits a new high or a new low, you hear about it. If an algorithm is developed to forecast the short-term price of a particular company, the rate of investment, profit, and business prospects in the stock market will increase significantly over time as the machine learns from the data. Convolutional and artificial neural networks, which were utilized in previous stock prediction, often had an error loss of 15% to 20%.

The purpose of this article is to ascertain whether a recurrent neural network model can be created to forecast stock prices with a reduced error rate. Additionally, if the answer is yes, we will look at the reliability of the model.

4.1 PROBLEM FORMULATION

Making market predictions is the fascinating task. There are numerous ways to accomplish this task. Techniques range widely from quite informal to highly professional. These technologies are classified as:

Prediction methods, traditional time series, technical analysis methods, and machine learning methods
Techniques for Basic Analysis The prerequisites for this category are Deep Learning, this kind of data, and the instruments that these techniques employ to show the market. The method is similar in that it helps the market by forecasting its future behavior.

V. SOLUTION METHODOLOGIES OR PROBLEM SOLVING

5.1 CHOOSING THE DATASET

For this project, we choose to use Google stocks. Google stock is one of the main indices that are traded on the New York Stock Exchange. Every company in the index is a sizable publicly listed business that leads its industry. The index covers a wide range of industries and includes companies such as Microsoft, Visa, Boeing, and Walt Disney. It is essential to use a predefined collection of organizations rather than a specially chosen one in order to prevent approach errors or accusations of fishing expeditions.

One may contend that the set was specifically designed to improve our results if we had selected a specific set of companies. as creating a model of stock markets in general is the aim of the research. Due of its extensive recognition, Google was chosen. The components offer a respectable balance between accessible data and computational feasibility.

5.2 GATHERING THE DATASETS

A primary dataset will be used in the study. The dataset will include the daily percentage change in the stock price. Thankfully, daily stock price data is easily accessible. Users may obtain real-time CSV files with a full 14-day pricing history from the websites of both Google and Yahoo. These are helpful for looking at criteria-specific data for firms, but they are challenging to apply when getting a lot of data from a lot of stocks. For this reason, Quandl was used instead of search engines like Google and Yahoo to collect the data.

With a focus on economic data, Quandl is a free website that holds and maintains enormous volumes of quantitative and numerical data, including stock market data that is sponsored by Google and Yahoo. Quandl also offers a useful little Python package for database access that may be done programmatically. The library provides a simple technique for calculating the daily % change in prices.

Quandl also provides a useful little Python package for database access that may be done programmatically. The library provides the finest and most popular way to calculate the daily % change in prices.

5.3 TECHNICAL ANALYSIS METHODS

Almost any exchangeable or tradable asset that is regularly affected by buying and selling and related to supply and demand, including stocks, bonds, futures, and currency pairs, may have its price movement predicted using technical analysis. The study of supply and demand dynamics in connection to changes in the market price of securities is known as technical analysis. Although it is most frequently used in relation to price fluctuations, some analysts may also monitor other indicators, such as trading volume or open interest data.

In an effort to precisely predict future price changes, experts have created a variety of technical indicators over the years. While some indicators assess the strength and likelihood of the trend's continuance, others concentrate on identifying the present market trend, including regions of support and resistance.

Moving averages and trendlines, Momentum indicators like the moving average convergence divergence (MACD) indicator are examples of technical indicators that are often used.

VI. STEP BY STEP APPROACH AND DESIGN

The method we used to create the prediction model using ml algorithms is explained in the graphic below.

Fig 1: Process of predicting IPL team

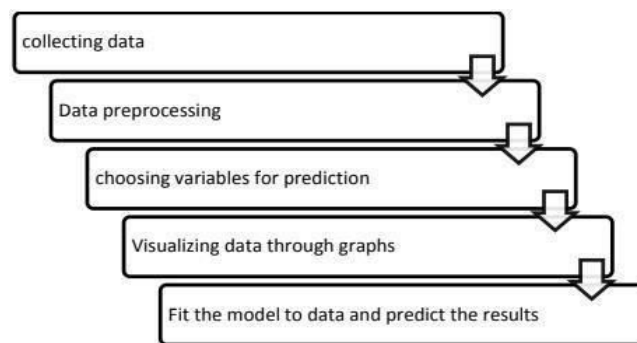


Fig 1: process of predicting IPL team

6.1 Data Collection

Information assemblage is the process of gathering and assessing data from several sources. to use the information we gather to create useful machine learning solutions. By gathering information, you may keep track of previous occurrences and do data analysis to identify patterns. Prediction models, which search for patterns and make decisions based on the information and potential changes, are created using machine learning algorithms.

A large set/model of the project's information was sourced from the Indian Stock Analytcis main website. A module named Lovely Soup was used to collect and properly store the data. In order to correspond with the numbers, Indian Stock Analyze season year, location and stadium name, match winner details, teams participating, margin of victory, umpire details, player of a match, etc., the dataset requires columns.

6.2 Data Preprocessing

6.2.1 Data cleaning

- Because the winner, city, venue, and other variables have null values, the dataset cannot be appropriately categorized.
- To replace the null values in various columns with imaginary values.

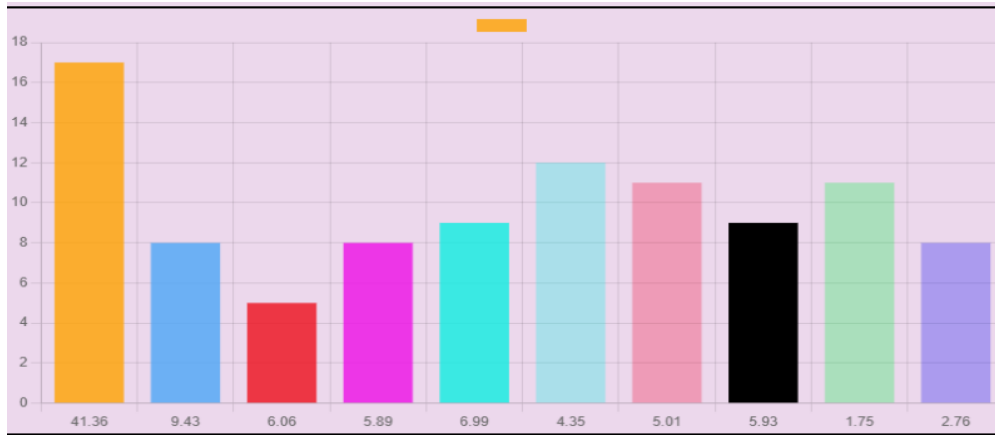
6.2.2 Choosing Attributes

- This step's primary objective is to remove any dataset columns that aren't necessary for determining which team won the game.
- This is evaluated based on the characteristic's relevance.
- The attributes in question have the following feature relevance.

6.3 Data Visualization

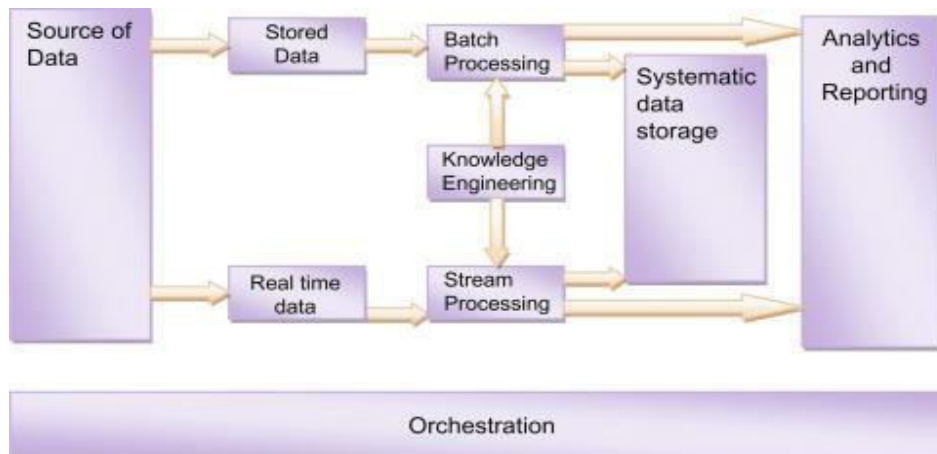
To increase understanding, the information is visualized using the collected data.

Fig 2 - Relation in the fig in terms of winning and loosing



6.4 Architecture of System

Fig 3: System Architecture



Based on user needs and a careful analysis of the existing system, a new system must be developed. The design of the system is currently under progress. Design is the most effective way to accurately translate a client's needs into the finished software product. In addition to providing information about the software data structures, interfaces, architecture, and other elements needed to construct a system, the design creates a representation or model. The logical system design derived from the systems analysis is used to build the physical system design.

VII. DATA MODEL

Data collection is the steps of gathering and measuring information from a difeferent range of sources. To develop practical machine learning solutions, we collect data.

By collecting information, you can record past events and use data analysis to spot recurring patterns. You build predictive models based on those patterns using machine learning algorithms that look for trends and predict future changes.

The primary source of data for this project is the official Indian Stock Analysis website. A Python module called Beautiful Soup was used to scrape the data from the website and preserve it in the proper format. The columns in the dataset

Fig 4: Data set Format

	Date	Open	High	Low	Close	Volume	Turnover	P/E	P/B	Div Yield
0	2000-01-03	1482.15	1592.90	1482.15	1592.20	25358322	8.841500e+09	25.91	4.63	0.95
1	2000-01-04	1594.40	1641.95	1594.40	1638.70	38787872	1.973690e+10	26.67	4.76	0.92
2	2000-01-05	1634.55	1635.50	1555.05	1595.80	62153431	3.084790e+10	25.97	4.64	0.95
3	2000-01-06	1595.80	1639.00	1595.80	1617.60	51272875	2.531180e+10	26.32	4.70	0.94
4	2000-01-07	1616.60	1628.25	1597.20	1613.30	54315945	1.914630e+10	26.25	4.69	0.94
5	2000-01-10	1615.65	1662.10	1614.95	1632.95	45013949	2.375350e+10	26.57	4.74	0.93
6	2000-01-11	1633.25	1639.90	1548.25	1572.50	49120254	2.596950e+10	25.59	4.57	0.96
7	2000-01-12	1572.30	1631.55	1571.70	1624.80	38364961	1.895000e+10	26.44	4.72	0.93
8	2000-01-13	1627.85	1671.15	1613.65	1621.40	44738447	2.237610e+10	26.38	4.71	0.93
9	2000-01-14	1622.15	1627.40	1591.40	1622.75	43292009	1.979980e+10	26.41	4.71	0.93

7.1 COMPARISON OF RESULTS:

Fig 5: Choose the file to analyze data

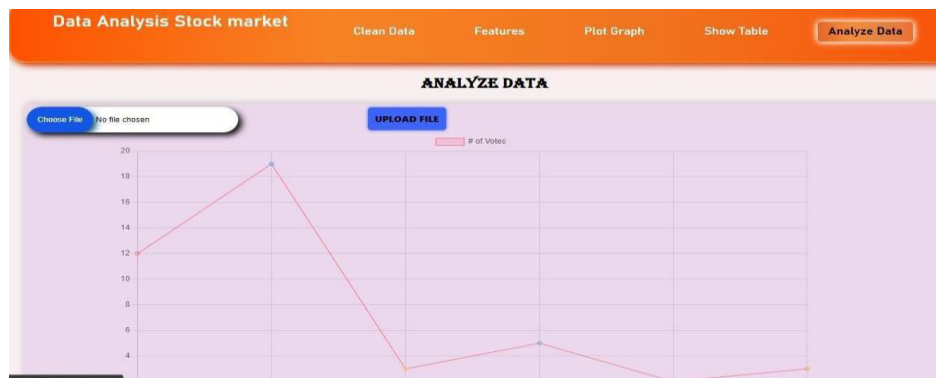
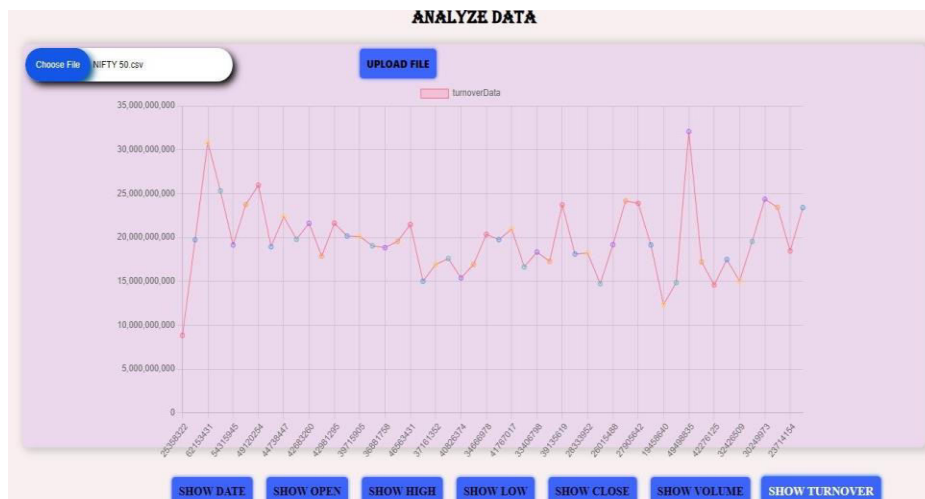
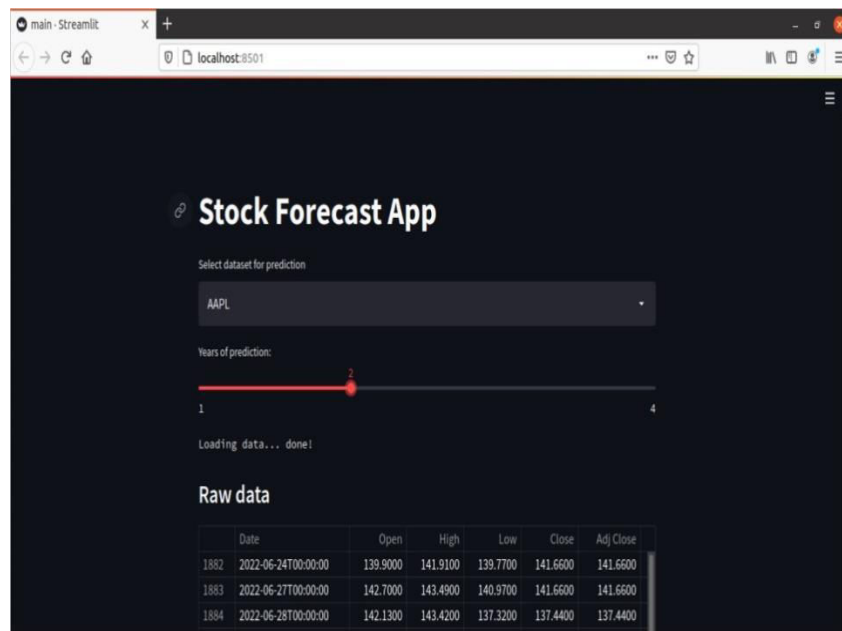


Fig 6: Shows the data in line graph



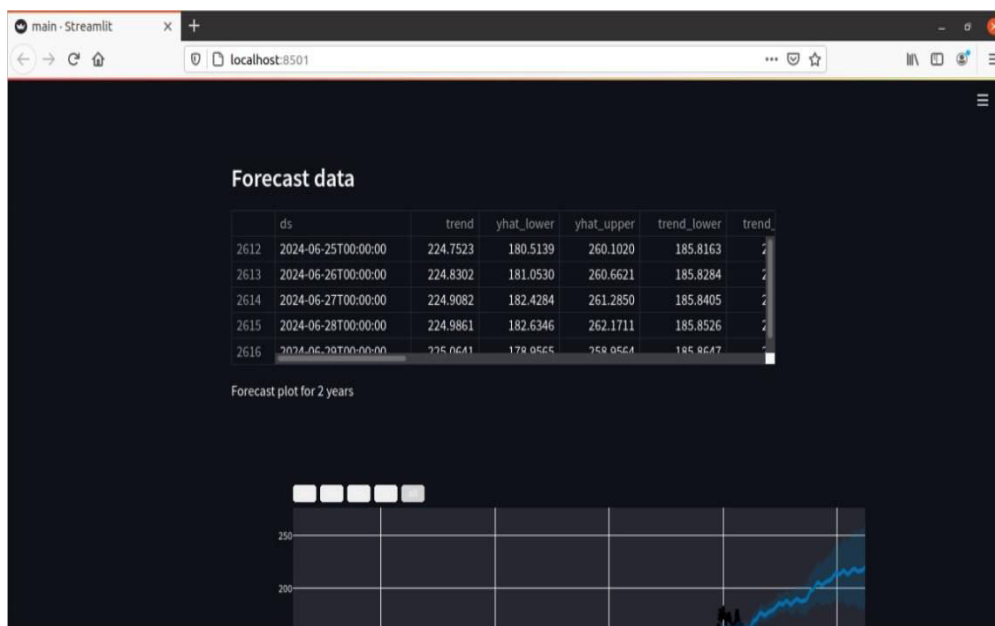
Choosing the name of the stock from the given option placed in dropdown:

Fig 7: Choosing The Stocks From option given in dropdown



To assist investors in manually assessing the value, selected stock data will be shown in the table:

Fig 08: Stocks Price From option given



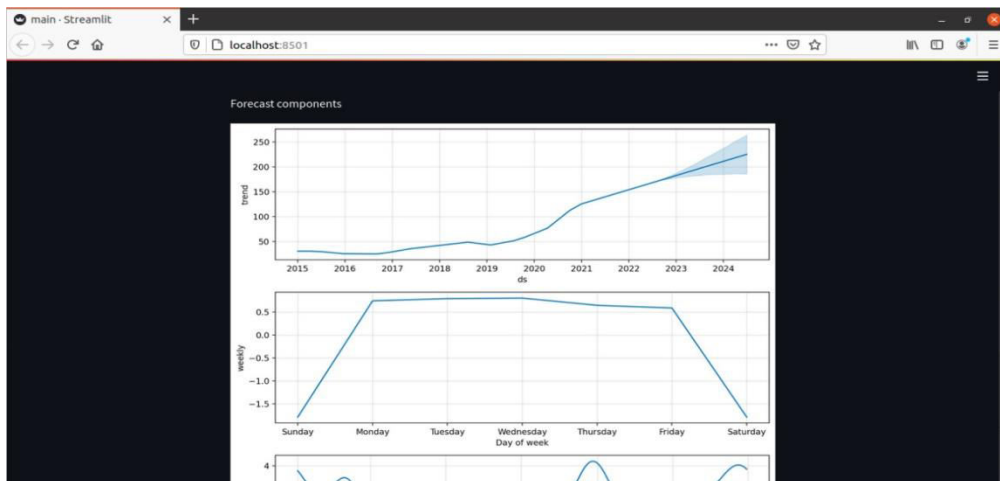
Red line and Blue line represent- stock closing price and opening price respectively.

Fig 9: Stock Price Moving Analysis



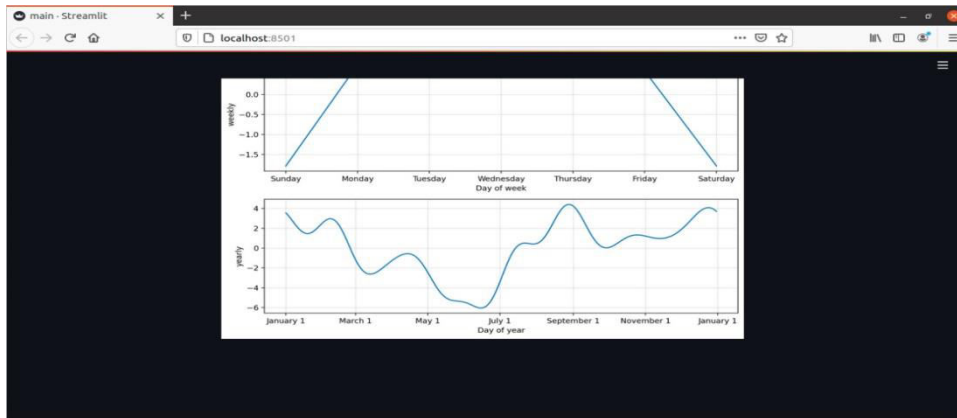
Forecast the trends of the stock for a single day or a monthly basis:

Fig10: Forecast trends of the stock for a single day or a monthly basis



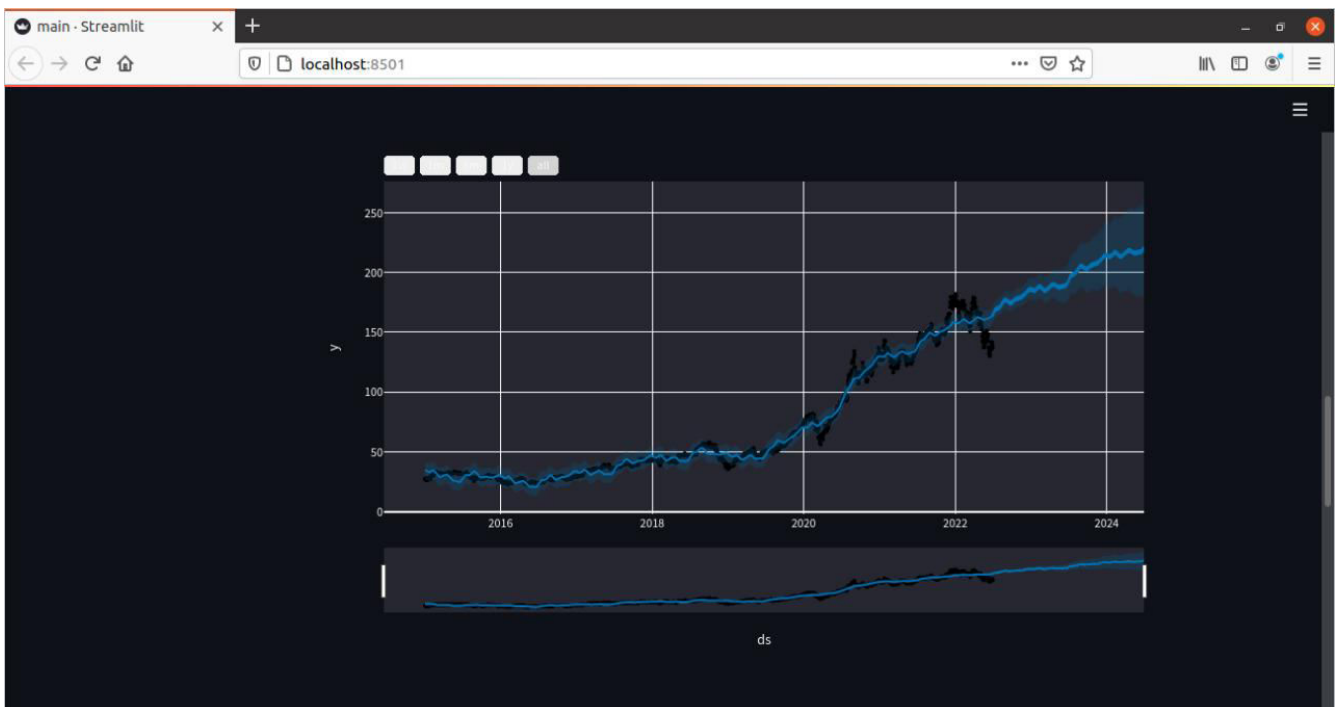
7.1.1 Forecasting moving of stock for a yearly basis:

Fig11: Forecasting moving of stock for a yearly basis



7.1.2 The black line shows the actual stock price movement, while the blue line shows the prediction:

Fig12: black line shows the actual stock price movement, while the blue line shows the prediction



Dropout is a useful tool to aid with this, but even with the best dropout selection, proper validation findings cannot be guaranteed. even though performance indicators like accuracy, sensitivity, and specificity produce favorable results.

Even though it takes a lot longer to train, the LSTM fared somewhat better than the RNN, albeit not enough to be statistically significant. The obvious performance advantages of parallelizing machine learning algorithms on a GPU are demonstrated by the 70.7% improvement in training the LSTM model. Additionally, the model should be improved by the ability to generate predictions utilizing streaming data.

7.2 JUSTIFICATION OF THE RESULTS

- There are null values in a few of the dataset's columns, such as winner, city, venue, etc.
- The presence of such null values precludes accurate classification.
- To replace the null values in different columns with dummy values.
- This is the main phase where we can eliminate dataset columns that aren't required for determining the victorious team.
- This is estimated using feature importance.
- The attributes in question have the following feature importance.
- The information is visualized using the collected data to enhance understanding.
- The LSTM model's performance improved by 70.7% during training; the model should function better if it can forecast utilizing streaming data.

VIII. CONCLUSION

The financial markets provide a unique trading and investment platform that can be used by any device with an Internet connection. The rise of stock markets has given people a number of ways to boost the growth of their money. To help people and organizations invest their money according to their risk tolerance, it also developed a number of fund kinds, such as hedge funds, mutual funds, and index funds.

To boost returns for everybody, the governments of the majority of nations invest a portion of their healthcare, employment, or retirement assets in stock markets. The way that consumers purchase and sell stocks has already been drastically altered by online trading platforms. The financial markets have expanded quickly to become a robust, globally integrated market. These developments create new possibilities, and although data science methods have many benefits, they also present a number of new difficulties.

This work offers a thorough literature assessment of the most recent algorithms and techniques that are often employed for stock market prediction, in addition to suggesting a taxonomy of computational approaches to stock market analysis and prediction. It also talks about some of the current problems in this area that need greater focus and offer chances for further advancement and study.

Unlike earlier systems, the current stock market is built using a range of technologies, such as machine learning, expert systems, and big data, which work together to improve decision-making. Simultaneously, the internet's ability to connect people worldwide has made the stock market more vulnerable to harmful attacks, shifting news, and consumer mood. This is where additional study might be crucial in laying the groundwork for future stock market analysis and strengthening.

Since markets function like weighing machines over time, with less noise and greater predictability, investigating different algorithms to see if they are strong enough to make long-term predictions is an interesting line of inquiry. For stock prediction, hybrid strategies that incorporate machine learning and statistical methods are likely to be more effective.

REFERENCES

- [1] Yaser S., Abu-Mostafa, and Amir F. Atiya. 1996. A summary of forecasting in finance. *Applied Intelligence* 6, 205–13.
- [2] Francisco Guijarro, Jorge García, Rubén Arévalo, Alfred Peris, and Rubén Arévalo. 2017. a dynamic trading rule based on the identification of filtered flag patterns for stock market price predictions. *Expert Systems with Applications* 81, pp. 177–92.
- [3] Babu, M. Suresh, N. Geethanjali, and B. Satyanarayana. 2012. a clustering-based method for stock market forecasting. *International Journal of Advanced Networking and Applications* 3: 1281
- [4] Dirk Van den Poel, Michel Ballings, Nathalie Hespels, and Ruben Gryp. 2015.
- [5] Assessing many classifiers for predicting the direction of stock prices. *Applications of Expert Systems* 42: 7046–56. [Reference]

- [6] Wei Bao, YuleiRao, and Jun Yue. 2017. A deep learning system that uses long-short term and stacked autoencoders for financial time series
- [7] Francisco Guijarro, Roberto Cervelló-Royo, and Karolina Michniuk. 2015. Technical analysis and pattern recognition are the foundations of the stock market trading rule that forecasts the DJIA index using intraday data. *Expert Systems with Applications* 42, 5963–75. [Source]
- [8] Jonathan Creighton and Farhana H. Zulkernine. 2017. presented at the 2017 IEEE International Conference on Big Data (Big Data), which was held in Boston, Massachusetts, USA, from December 11–14. The presentation was titled "Towards Developing a Hybrid Stock Index Prediction Model."
- [9] Eugene F. Fama. 1970. A Review of Theory and Empirical Research on Effective Capital Markets. 25: 383–417 in *The Journal of Finance*. [Reference]
- [10] Stefan Lauren; S. Dra. Harlili Stock trend prediction using simple moving average supported by news classification



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details