Stock Price Prediction Using ML

A Project Work Synopsis

Submitted in the partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE WITH SPECIALIZATION IN CLOUD COMPUTING

Submitted by:

Ashish Kumar 21BCS11125

Ayush Sharma 21BCS4010

Yuvika 21BCS3690

Under the Supervision of:

Ms. Shaveta Jain

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413, PUNJAB

September, 2022

Abstract

This synopsis delves into the integration of historical stock price data, technical indicates, and relevant market-related information as input features to various ML algorithm. The stock market is a highly dynamic and complex system that can be influenced by various factors such as economic trends, political events, global events, and company-specific news. Predicting the stock market's behavior is crucial for investors and traders to make informed decisions about buying, selling, or holding stocks. Machine Learning techniques can help in predicting the stock market by analyzing historical data and identifying patterns in the market behavior. Stock price prediction is a crucial area of research and application in financial markets. The volatile and complex nature of stock prices poses a challenging task for investors and traders to make informed decisions. This abstract highlights a comprehensive approach to stock price prediction utilizing machine learning (ML) techniques.

Keywords: stock price prediction using ML, python, database, user-friendly platform, machine learning, python libraries, stock prediction, machine learning

Table of Contents

1. INTRODUCTION

1.1 Problem Definition

The problem addressed in this research is that stock price prediction is a critical challenge financial markets due to the complex and dynamic nature of stock price movements. Accurate prediction of stock prices has the potential to provide valuable insights to investors, traders, and financial analysts, aiding them in making informed decisions to optimize their investment strategies.

Key Challenges:

Nonlinear and Dynamic Nature: Stock prices are influenced by a multitude of factors, including economic indicators, company performance, market sentiment, and external events. The complex and nonlinear relationships between these factors make accurate predictions challenging.

Noisy Data: Financial data can be noisy and prone to outliers, missing values, and data inconsistencies. Handling these issues effectively is crucial to ensure the reliability of the predictive model.

1.2 Problem Overview

The primary objective of this project is to develop a machine learning model that can predict the stock market's behavior accurately. Specifically, we aim to:

1. Develop a machine learning model that can accurately predict the stock prices for the next day, week, or month.

2. Identify the critical factors that affect stock prices and incorporate them into our model.

3. Analyze the model's accuracy and compare it with traditional methods of stock market analysis.

4. Develop a user-friendly interface that allows investors and traders to interact with the model and receive predictions.

1.3 Hardware Specification

CPU Core i3-2100 or higher, Minimum 2 G.B. RAM, Internet Connection

1.4 Software Specification

Windows 10 or above 64-bit O.S., python (using latest version 3.12)

2. LITERATURE SURVEY

2.1 Existing System

The existing stock price prediction exhibit nonlinear and dynamic nature, noisy data, complex and dynamic nature of stock price movement. These limitations impede dynamic pricing, comprehensive information access, and user satisfaction. This study aims to enhance the system using python and their libraries, rectifying these issues and improving user experiences.

2.2 Proposed System

The envisioned system seeks to revolutionize stock price prediction by seamlessly integrating Python and their libraries. Through dynamic pricing algorithms, efficient data structuring, and an intuitive user interface, customers can predicte stock prices, access comprehensive stock details, and effortlessly manage their data. This innovative approach aims to significantly enhance user satisfaction, streamline operational processes, and elevate the overall experience. By combining technological advancements with mediocre design, the proposed system aims to set a new standard in the realm of stock price prediction.

2.3 Literature Review Summary

Year and Citation	Article/ Author	Tools/ Software	Technique	Source	Evaluation Parameter
2017	Rodolfo Toríbio Farias Nazário, Herbert Kimura, Jéssica Lima e Silva, Vinicius Amorim Sobreiro	Not specified	Literature Review	A Literature Review Of Technical Analysis On Stock Markets DOI:10.1016/j.qref.2017.01.014	Technical Analysis On Stock Market
2020	Anusha J Adhikar,Apeksha K Jadhav, Charitha G, Karishma KH	Not specified	Literature Review	LITERATURE SURVEY ON STOCK PRICE PREDICTION USING MACHINE LEARNING	Stock Price Prediction Using Machine Learning
2019	Xinyi Li, Yinchuan Li, Hongyang Yang, Liuqing Yang, Xiao-Yang Liu .	Not Specified	Financial News	DP-LSTM: Differential Privacy-inspired LSTM for Stock	DP-LSTM: Differential Privacy-inspired

				Prediction Using Financial News	LSTM for Stock Prediction
2020	Sidra Mehtab, Jaydip Sen, Abhishek Dutta	Not specified	Regression,Deep Learning	Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models .	Payment Security, Transaction Integrity
2020	Obthong, M., Tantisantiwong, N., Jeamwatthanachai, W., & Wills, G.	Not Specified	Survey	A survey on machine learning for stock price prediction: Algorithms and techniques.	Algorithm and techniques.

3. PROBLEM FORMULATION

The problem formulation addressed by this synopsis is centered around the deficiencies prevalent within contemporary stock price prediction systems. These systems, although widely used, exhibit inherent limitations that hinder the efficiency and user experience associated with the stock price prediction process. The stock market is a complex system that is influenced by a variety of factors, including economic conditions, investor sentiment, and company news. Stock prices are often volatile and can change rapidly. It is difficult to predict the future with certainty, even with a large amount of historical data. Despite these challenges, there are a number of machine learning algorithms that have been used to predict stock prices with some success. These algorithms typically learn from historical data to identify patterns that can be used to predict future prices. The identified shortcomings encompass static pricing models, fragmented data management practices, and inadequate user interfaces. Fundamental analysis involves analyzing the financial statements of a company to determine its intrinsic value. Technical analysis involves studying historical price charts to identify trends and patterns. Crowd sourcing involves gathering opinions from a large

number of people to predict the future price of a stock. No single approach to stock price prediction is guaranteed to be successful. The best approach will vary depending on the specific problem and the availability of data.

4. OBJECTIVES

The objective of stock price prediction is to develop a predictive model that can forecast the future price movements of a given stock or financial instrument based on historical price data, technical indicators, and potentially relevant market information. This synopsis are outlined to address the shortcomings within existing stock price prediction and propose a comprehensive solution using python and their libraries. The specific objectives are as follows:

Inputs:

Historical Price Data: A time series dataset containing the historical prices of the target stock. This data includes the opening, closing,

high, and low prices for each time step (e.g., daily, hourly).

Technical Indicators: Derived from the historical price data, these indicators include metrics like moving averages, relative strength index (RSI), moving average convergence divergence (MACD), and others. They capture different aspects of price trends, momentum, and volatility.

Market News and Sentiment: Optionally, textual data from news articles, social media, or other sources can provide market sentiment, event impacts, and other qualitative factors that may influence stock prices.

Output:

The output of the prediction model is an estimation of the future price movement, which can be one of the following:

Binary Classification: Whether the stock price will go up or down within a specified time frame.

Regression: The predicted numerical value of the stock price for a future time point.

5. METHODOLOGY

The methodology section of this synopsis outlines the systematic approach used to design, develop, and evaluate the proposed stock price prediction system. It encompasses the tools, techniques, and procedures employed to achieve the research objectives. The following details the methodology:

1. Data Preprocessing:

Organize and clean historical price data, handling missing values and outliers.

Calculate relevant technical indicators from the historical price data.

If applicable, preprocess and analyze market sentiment data using natural language processing techniques.

2. Feature Selection and Engineering:

Choose appropriate features from the historical price data and technical indicators based on their relevance and potential predictive power.

Perform feature engineering to create new features that capture additional patterns or relationships.

3. Model Selection:

Select suitable machine learning or statistical models for the prediction task, considering the problem type (classification/regression) and the nature of the data.

Commonly used models include linear regression, support vector machines, decision trees, random forests, neural networks, and LSTM networks.

4. Training and Validation:

Split the historical data into training and validation sets, ensuring that the training data covers a significant portion of the historical timeline.

Train the selected models on the training data, adjusting hyper parameters as needed.

5. Evaluation:

Evaluate the trained models on the validation set using appropriate evaluation metrics such as accuracy, precision, recall, F1-score for classification tasks, or mean squared error, mean absolute error for regression tasks.

6. Model Tuning and Optimization:

Fine-tune hyperparameters and model configurations to improve prediction performance.

Consider techniques like cross-validation to assess model generalization.

7.Testing and Deployment:

Once satisfied with the model's performance on the validation set, apply the model to test data that represents unseen future data.

Deploy the trained model to make real-time predictions or provide insights to users.

6. EXPERIMENTAL SETUP

The experimental setup of the research paper involves a systematic configuration of software

tools, programming languages, and methodologies to evaluate the proposed stock price prediction system.. The objective is to assess its functionality, accuracy, and user experience. The following describes the experimental setup in detail:

1. **Programming Languages and Tools:**

    - Python: The core application is developed using python programming language to create the user interface and interact with the database.

2. **Data Collection and Preprocessing**:

    Libraries like pandas are used to collect, clean, and preprocess historical stock price data. These libraries enable efficient handling of large datasets and facilitate tasks like data cleaning, normalization, and feature extraction.

3. **Technical Indicators Calculation:**

    numpy and pandas are often used to calculate technical indicators from historical price data. For instance, moving averages, RSI, MACD, and Bollinger Bands can be computed using these libraries.

4. **Visualization & Feature Engineering:**

matplotlib and seaborn are widely used for data visualization. These libraries help analysts and researchers visualize historical price trends, technical indicators, and model predictions.

Python allows for advanced feature engineering using libraries like scikit-learn. This includes creating lag features, generating rolling statistics, and engineering domain-specific features.

5. **Machine Learning and Deep Learning Models:**

Libraries like scikit-learn, TensorFlow, and Keras provide a rich set of tools for building various predictive models, ranging from traditional machine learning algorithms (linear regression, SVM, etc.) to advanced deep learning models (LSTM, CNN, etc.).

6. **Model Training and Evaluation:**

scikit-learn offers functionalities for splitting datasets into training and validation sets, training models, and evaluating their performance using metrics like accuracy, precision, recall, and F1-score.

7. **Hyperparameter Tuning:**

   Libraries like scikit-learn and optuna help optimize model hyper parameters to improve predictive performance.

8. **Time Series Analysis:**

   Libraries like statsmodels provide tools for time series analysis, including ARIMA and SARIMA models, as well as methods for understanding and modeling time-dependent patterns.

9. **Sentiment Analysis:**

   Python, along with libraries like nltk and spaCy, can be used to perform sentiment analysis on textual data like news articles and social media content, which can be incorporated as features in prediction models.

10 **Web Scraping and Data Retrieval:**

   Libraries such as beautifulsoup and requests can be used to scrape financial news, market data, and other relevant information from websites.

11 **Interactive Dashboards and Visualization:**

Python frameworks like Plotly and Dash enable the creation of interactive dashboards and web applications for visualizing real-time stock data and model predictions.

7. CONCLUSION

In conclusion, this project aims to develop a machine learning model that can predict the stock market's behavior accurately. The model's accuracy can be improved by identifying the critical factors that affect stock prices and incorporating them into the model. We expect that the developed model will be useful for investors and traders to make informed decisions about buying, selling, or holding stocks.

Stock price prediction is a complex and challenging endeavor that has garnered substantial attention from researchers, analysts, and investors alike. The dynamic nature of financial markets, influenced by a multitude of unpredictable factors, underscores the inherent uncertainty in predicting stock prices accurately. However, through the integration of sophisticated techniques and approaches, valuable insights can be gleaned to assist decision-making in investment strategies.

Successful stock price prediction necessitates a comprehensive approach that integrates historical price data, technical indicators, sentiment analysis, and potentially alternative data sources. The combination of quantitative and qualitative factors enriches the prediction process. Machine learning algorithms, ranging from traditional regression models to cutting-edge deep learning networks, offer powerful tools for capturing intricate patterns and trends in historical data. The adaptability of these models to various time series patterns enhances their effectiveness.

The careful selection and engineering of features play a pivotal role in model performance. Effective feature engineering enables the models to encapsulate the underlying dynamics of stock price movements and their interactions with market variables. It is essential to recognize the limitations of predictive models and the uncertainty inherent in financial markets. Integrating risk management strategies, such as diversification and considering worst-case scenarios, is crucial for prudent decision-making.

Time series analysis techniques, such as autoregressive models and exponential smoothing, capture temporal dependencies and seasonality inherent in stock price data.

These methods can complement machine learning models and improve predictive accuracy. The financial landscape is dynamic, and models need to evolve to accommodate changing market dynamics, emerging trends, and unexpected events. Regular retraining and adaptation are essential to maintaining model relevance. Ensuring transparency and ethical use of predictive models is paramount. Clear communication of the limitations of predictions and responsible use of models in decision-making are critical aspects of their deployment.

In summary, while stock price prediction remains inherently uncertain due to the interplay of complex market forces, predictive models provide valuable tools for understanding historical trends and making informed decisions. The integration of advanced technologies, domain expertise, and a cautious approach to risk management contributes to more effective utilization of predictive insights in financial markets. As financial technology and data science continue to advance, the field of stock price prediction will undoubtedly witness further innovation, ultimately enhancing our understanding of market behavior.

REFERENCES

[1] Masoud, Najeb MH. "The impact of stock market performance upon economic growth." *International Journal of Economics and Financial Issues* 3.4 (2013): 788-798.

[2] Schöneburg, Eberhard. "Stock price prediction using neural networks: A project report." *Neurocomputing* 2.1 (1990): 17-27.'

[3] Obthong, M., Tantisantiwong, N., Jeamwatthanachai, W., & Wills, G. (2020). A survey on machine learning for stock price prediction: Algorithms and techniques.

[4] Yu, P. and Yan, X., 2020. Stock price prediction based on deep neural networks. *Neural Computing and Applications*, *32*, pp.1609-1628.

[5] Adebiyi, Ayodele A., et al. "Stock price prediction using neural network with hybridized market indicators." *Journal of Emerging Trends in Computing and Information Sciences* 3.1 (2012).

[6] Lu, Wenjie, Jiazheng Li, Jingyang Wang, and Lele Qin. "A CNN-BiLSTM-AM method for stock price prediction." *Neural Computing and Applications* 33 (2021): 4741-4753.

[7] Ariyo, A. A., Adewumi, A. O., & Ayo, C. K. (2014, March). Stock price prediction using the ARIMA model. In *2014 UKSim-AMSS 16th international conference on computer modelling and simulation* (pp. 106-112). IEEE.

[8] Adebiyi, Ayodele A., Charles K. Ayo, Marion Adebiyi, and Sunday O. Otokiti. "Stock price prediction using neural network with hybridized market indicators." *Journal of Emerging Trends in Computing and Information Sciences* 3, no. 1 (2012).