STOCK PRICE PREDICTION

TEAM MEMBERS

MANIKANDAN.B

KARTHICK.M

MATHAN.S

SHIBIN.G

PRASANNA.KM



INTRODUCTION

In this notebook I will show you how to write a python program that predicts the price of stocks using a machine learning technique . This program is really simple and I doubt any major profit will be made from this program, but it’s slightly better than guessing! Remember the stock price can be affected by many different things.

A It's important to note that stock price prediction is a challenging and highly volatile field. While some models may produce accurate short-term predictions, long-term predictions tend to be more uncertain due to the complex unpredictable nature of financial markets. Additionally, factors like market sentiment, external events, and unexpected news can have a significant impact on stock prices, making it difficult to make precise forecasts. Investors should always exercise caution and diversify their portfolios to manage risk effectively.

Stock price prediction is a complex problem that involves forecasting the future price of a particular stock or a set of stocks in financial markets. This problem is of great interest to investors, traders, financial analysts, and researchers, as it can provide valuable insights for making informed investment decisions. The problem of stock price prediction can be outlined as follows:

PROBLEM STATEMENT

1. \*\*Non-linearity and Volatility:\*\* Stock prices are influenced by a multitude of factors, including market sentiment, economic conditions, company performance, and external events. These factors often exhibit non-linear and volatile behavior, making it challenging to predict stock prices accurately.

2. \*\*Time Series Data:\*\* Stock price data is typically presented as a time series, with daily, hourly, or even minute-by-minute data points. This time-dependent nature of the data requires specialized techniques to capture trends, seasonality, and other temporal patterns.

3. \*\*High-Dimensional Data:\*\* Stock price prediction involves considering a wide range of features and variables, such as historical stock prices, trading volumes, news sentiment, financial ratios, and macroeconomic indicators. Managing and selecting relevant features is crucial for building accurate models.

4. \*\*Data Noise and Anomalies:\*\* Financial markets can be noisy, and data anomalies can occur due to various reasons, including data errors, market manipulations, or sudden events. Handling noisy data and identifying anomalies is essential for reliable predictions.

5. \*\*Limited Predictive Power:\*\* It's important to acknowledge that stock prices are influenced by both rational and irrational factors, and market movements can be unpredictable. This inherent uncertainty limits the predictive power of any model.

6. \*\*Information Asymmetry:\*\* Market participants often have varying levels of information about a stock, and insider trading can significantly impact prices. Predictive models may not have access to all relevant information, making it challenging to outperform the market consistently.

7. \*\*Market Sentiment and News Impact:\*\* News articles, social media, and public sentiment can play a significant role in stock price movements. Sentiment analysis and incorporating news data into models can be critical for accurate predictions.

8. \*\*Long-Term vs. Short-Term Prediction:\*\* Stock price prediction can be categorized into short-term (intraday or daily) and long-term (weekly, monthly) forecasting. The choice of prediction horizon impacts the modeling approach and techniques used.

9. \*\*Model Selection and Evaluation:\*\* Various machine learning and statistical models can be used for stock price prediction, including time series models, regression models, neural networks, and more. Selecting an appropriate model and evaluating its performance is crucial.

10. \*\*Overfitting and Generalization:\*\* Due to the noisy nature of financial data, models can be prone to overfitting, where they perform well on training data but poorly on unseen data. Ensuring that models generalize to new data is a key challenge.

11. \*\*Risk Management:\*\* Predictions are only one aspect of investing. Managing risk, such as setting stop-loss orders and diversifying portfolios, is equally important to mitigate potential losses.

12. \*\*Ethical and Regulatory Considerations:\*\* Financial markets are subject to regulations and ethical considerations, and certain prediction strategies may not be allowed or may raise ethical concerns.

Solving the problem of stock price prediction involves a combination of data preprocessing, feature engineering, model development, and continuous monitoring and adaptation to changing market conditions. It's an ongoing and evolving field that leverages advances in data science, machine learning, and domain expertise to improve prediction accuracy.

DESIGN THINKING PROCESS

Applying the design thinking process to stock price prediction involves a user-centered, iterative approach to problem-solving that focuses on understanding the needs of investors, traders, and other stakeholders. Here's a simplified version of the design thinking process adapted for stock price prediction:

1. \*\*Empathize: Understand User Needs\*\*

- Identify the target users (e.g., investors, traders, financial analysts).

- Conduct interviews, surveys, and observations to understand their goals, challenges, and data preferences.

- Empathize with the users to gain a deep understanding of their needs and motivations in stock trading.

2. \*\*Define: Problem Statement and Goals\*\*

- Distill the insights gathered from user research into a clear problem statement related to stock price prediction.

- Define specific goals, such as improving accuracy, reducing risk, or providing actionable insights.

3. \*\*Ideate: Generate Solution Ideas\*\*

- Brainstorm a wide range of potential solutions to the defined problem.

- Explore various data sources, modeling techniques, and technologies that could be used in stock price prediction.

4. \*\*Prototype: Develop Conceptual Models\*\*

- Create conceptual models or mockups of potential stock price prediction solutions.

- Design sample interfaces, visualizations, or decision-support tools to illustrate how the solution might work.

5. \*\*Test: Collect Feedback\*\*

- Present prototypes to users and stakeholders for feedback and validation.

- Iterate on the prototypes based on feedback, making necessary adjustments to address user preferences and needs.

6. \*\*Implement: Build the Solution\*\*

- Develop the stock price prediction model and associated tools based on the validated prototype.

- Implement data collection and preprocessing pipelines, machine learning algorithms, and user interfaces as required.

7. \*\*Test Again: Evaluate Model Performance\*\*

- Assess the model's performance using historical data and, if possible, with a simulation environment.

- Identify any discrepancies or issues and make refinements to the model as needed.

8. \*\*Deploy: Make the Solution Accessible\*\*

- Deploy the stock price prediction solution in a user-friendly and accessible manner, which could include web or mobile applications, APIs, or integration into existing trading platforms.

9. \*\*Iterate and Improve\*\*

- Continuously monitor the model's performance in real-time trading scenarios.

- Gather user feedback and adapt the model and tools as market conditions change or as new data becomes available.

10. \*\*Educate and Train Users\*\*

- Provide training and educational resources to help users understand how to interpret and use the stock price predictions effectively.

- Ensure that users are aware of the model's limitations and risks.

11. \*\*Ethical and Regulatory Compliance\*\*

- Ensure that the stock price prediction system complies with financial regulations and ethical guidelines.

- Address any ethical concerns and communicate transparency in the prediction process.

12. \*\*Feedback Loop\*\*

- Maintain an ongoing feedback loop with users to gather insights and adapt to changing user needs and market conditions.

Throughout this process, it's crucial to involve domain experts, data scientists, and software developers to collaborate on developing and refining the stock price prediction system. Additionally, the design thinking process encourages flexibility and user-centric design, allowing for continuous improvement and adaptation to the dynamic nature of financial markets.

PHASE OF THE DEVELOPMENT

Phase 1: Problem Definition and Design Thinking

In this part we understand the problem statement and created a document on what we have understood and how will we proceed ahead with solving the problem. We think on a design and present in form of the document.

Phase 2: Innovation:

In this section we put our design into innovation to solve the problem. Created a document around it.

Phase 3: Development Part 1:

In this section we begin building our project by loading and preprocessing the dataset.

Phase 4: Development Part 2

In this section continue building the project by performing different activities like feature engineering, model training, evaluation etc as per the instructions in the project.

Phase 5: Project Documentation & Submission

In this section you will document the complete project and prepare it for submission.

DATA SET

Data source link: (https://www.kaggle.com/datasets/prasoonkottarathil/microsoft-lifetime-stocks-dataset)

Several datasets are commonly used for stock price prediction and analysis. These datasets typically include historical stock price and trading data, along with various additional features. Here are some of the common datasets used in stock price prediction:

1. \*\*Yahoo Finance Data:

Yahoo Finance provides historical stock price and trading volume data for a wide range of publicly traded companies. This data is freely available and can be accessed through various APIs or downloaded in CSV format.

2. Alpha Vantage:

Alpha Vantage offers APIs and downloadable datasets that include historical stock prices, real-time data, technical indicators, and fundamental data for various global stock exchanges.

3. Quandl:

Quandl is a data provider that offers financial and economic data, including historical stock price data for thousands of companies. They offer both free and premium datasets.

4. IEX Cloud:

IEX Cloud provides access to real-time and historical stock price and trading data, along with other financial market data, through their API.

5. Kaggle:

Kaggle is a popular platform for data science competitions and datasets. It hosts various stock price prediction datasets, including competitions like the "Two Sigma Financial Modeling Challenge."

6. AlphaQuants Dataset:

AlphaQuants is a commercial provider of financial datasets, including historical stock price data, financial statements, and news sentiment data.

7. Google Finance:

Google Finance provides historical stock price and financial data. While it has undergone changes and restrictions over time, it remains a potential source of historical stock price data.

8. Intrinio: Intrinio offers financial market data, including historical stock price data, company fundamentals, and market news.

9. Financial News Sources:

Datasets from financial news sources like Reuters, Bloomberg, or news APIs can be used to incorporate sentiment analysis and news-related features into stock price prediction models.

10. Custom Data:

In some cases, financial institutions and researchers may generate custom datasets by combining multiple sources of data, including proprietary trading data, economic indicators, and news sentiment data.

It's important to note that the quality and accuracy of the data are critical for stock price prediction models. When using these datasets, it's crucial to clean, preprocess, and validate the data carefully. Additionally, historical stock price data should be adjusted for stock splits, dividends, and other corporate actions to ensure accurate modeling and backtesting.

DATA PREPROCESSING:

Data preprocessing is a crucial step in preparing data for stock price prediction. It involves cleaning, transforming, and organizing the data to make it suitable for modeling. Here are common data preprocessing steps for stock price prediction:

1. \*\*Data Collection:\*\* Collect historical stock price data from reliable sources, such as Yahoo Finance, Alpha Vantage, or other financial data providers.

2. \*\*Data Cleaning:\*\*

- Handle missing data: Check for missing values in the dataset and decide on an appropriate strategy for dealing with them. Common approaches include interpolation, forward-fill, backward-fill, or dropping incomplete rows.

- Handle outliers: Identify and address outliers, which can significantly impact the model's performance. You can use techniques like winsorization or replace outliers with more reasonable values.

- Adjust for corporate actions: Account for stock splits, dividends, and other corporate actions that can affect stock prices. Adjust historical prices accordingly to ensure consistency.

3. \*\*Feature Engineering:\*\*

- Create relevant features: Generate additional features that may help in prediction, such as moving averages, technical indicators (e.g., Relative Strength Index, Moving Average Convergence Divergence), and other financial metrics.

- Lag features: Include lagged versions of stock prices or other relevant features to capture temporal patterns. For example, you can use past stock prices to predict future prices.

- Calculate returns: Calculate returns or price changes, which are often used in time series analysis and modeling.

4. \*\*Time Alignment:\*\*

- Ensure that all data, including features and target variables, are aligned in time. This is essential for time series forecasting and avoids data leakage.

- Align datasets: If you're using additional data sources like news sentiment or economic indicators, align their timestamps with the stock price data.

5. \*\*Scaling and Normalization:\*\*

- Normalize features: Scale numerical features to have zero mean and unit variance (z-score normalization) or scale them to a specific range, such as [0, 1] (min-max normalization).

- Standardize target variable: Standardize or normalize the target variable to improve model convergence and interpretability.

6. \*\*Splitting Data:\*\*

- Split data into training, validation, and test sets. The training set is used to train the model, the validation set is used for hyperparameter tuning, and the test set is used for final model evaluation.

- Ensure that the split is done chronologically to simulate real-world scenarios.

7. \*\*Handling Categorical Variables:\*\*

- Encode categorical variables: Convert categorical variables into numerical representations using techniques like one-hot encoding or label encoding, if applicable.

8. \*\*Data Transformation:\*\*

- Log transformations: Apply log transformations to variables with skewed distributions to make them more symmetric.

- Differencing: In time series analysis, differencing can be used to make the data stationary (constant mean and variance) if necessary.

9. \*\*Handling Imbalanced Data:\*\* If stock price data is imbalanced (e.g., due to long periods of stable prices), consider techniques like oversampling, undersampling, or using different evaluation metrics to account for the imbalance.

10. \*\*Data Visualization:\*\* Visualize the preprocessed data to gain insights into the relationships between features and the target variable. Visualization can help you identify patterns and anomalies.

11. \*\*Data Serialization:\*\* Save the preprocessed data in a suitable format, such as CSV or a database, for easy access in modeling and future analysis.

12. \*\*Documentation:\*\* Maintain documentation that records all preprocessing steps, as this is essential for reproducibility and troubleshooting.

Data preprocessing is often an iterative process, and you may need to revisit and refine these steps as you explore and model the data. Careful data preprocessing can significantly impact the quality and performance of your stock price prediction models.

EXTRACTION TECHNIQUES:

To predict stock prices accurately, you can employ various data extraction techniques to gather information and features that are relevant to the prediction task. These techniques involve collecting data from multiple sources and processing it for analysis. Here are some key data extraction techniques for stock price prediction:

1. \*\*Historical Stock Price Data:\*\*

- Retrieve historical stock price data, including opening, closing, high, and low prices, as well as trading volumes for the stock or stocks of interest.

- This data can be obtained from financial data providers, stock exchanges, and APIs like Yahoo Finance or Alpha Vantage.

2. \*\*Fundamental Data:\*\*

- Gather fundamental financial data for the company, such as earnings reports, revenue, earnings per share, and other financial indicators.

- Information like balance sheets, income statements, and cash flow statements can provide insights into a company's financial health.

3. \*\*Market and Economic Indicators:\*\*

- Collect macroeconomic indicators, such as GDP growth, inflation rates, interest rates, and unemployment rates, which can impact the overall stock market.

- Incorporate market-specific data, including market indices (e.g., S&P 500), sector indices, and volatility indices (e.g., VIX).

4. \*\*News and Sentiment Data:\*\*

- Utilize news articles and sentiment analysis to capture market sentiment and public perception of the stocks. Positive or negative news can impact stock prices.

- APIs and news aggregators can be used to gather relevant news and social media sentiment data.

5. \*\*Technical Indicators:\*\*

- Calculate technical indicators like moving averages, Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands from historical price and volume data.

- These indicators can help identify trends and potential buy/sell signals.

6. \*\*Options and Derivatives Data:\*\*

- Extract options and derivatives data for the stock, including open interest, implied volatility, and options trading volume.

- These data points can provide insights into market expectations and hedging activities.

7. \*\*Company News and Reports:\*\*

- Access company-specific news, earnings reports, and financial statements from sources like company websites, investor relations sites, and news sources.Extract key information from quarterly and annual reports.

8. \*\*Social Media Data:\*\*

- Monitor social media platforms and financial forums for discussions and sentiment related to the stock.

- Tools and APIs are available for scraping and analyzing social media data.

9. \*\*Alternative Data Sources:\*\*

- Explore alternative data sources, such as satellite imagery, credit card transaction data, supply chain data, and more, which can offer unique insights into a company's performance and prospects.

10. \*\*Blockchain and Cryptocurrency Data:\*\*

- For companies involved in blockchain or cryptocurrency, collect data related to blockchain transactions, token prices, and trading volumes.

11. \*\*Regulatory Filings:\*\*

- Access regulatory filings such as 10-Ks and 10-Qs filed with the U.S. Securities and Exchange Commission (SEC) to gather detailed financial and operational information about publicly traded companies.

12. \*\*Market Order Book Data:\*\*

- For high-frequency trading or intraday prediction, access order book data, including bids and asks at different price levels, to analyze market depth and liquidity.

13. \*\*Weather Data:\*\*

- In certain cases, weather data can be relevant, especially for industries sensitive to weather conditions, such as agriculture or energy.

14. \*\*Global Events Data:\*\*

- Collect data related to geopolitical events, natural disasters, and other global occurrences that can influence stock markets.

It's important to consider the availability, quality, and timeliness of data when selecting sources for stock price prediction. Additionally, integrating multiple data sources and performing feature engineering are critical for building comprehensive predictive models. Data extraction techniques should be flexible to adapt to changing market conditions and to capture the most relevant information for accurate predictions.

MACHINE LEARNING ALOGORITHM:

INTRODUCTION

The modern advances in artificial intelligence have led to the creation of new mathematical tools like deep learning and reinforcement learning. Businesses use data science and analytics to obtain solutions for various business issues. Stock trading is one of the most important activities in the finance world. Stocks are an equity investment that denotes a part of ownership in an organization or a company; it entitles you to be a part of that company's earnings and assets. Stock market prediction can be defined as trying to determine the future value of a stock or other financial instrument that is traded on a financial exchange. The successful prediction of a stock's future price can lead to hefty profits. Financial information is a significant component of all electronic data. An average stock exchange creates nearly trillions of Gigabytes (GB) of trade and order book data in a month. In recent years, the rising popularity of machine learning in various industries has enlightened many traders to apply machine learning techniques to the field, and some of them have produced quite promising results. Machine Learning (ML) provides a unique perspective to us on understanding the stock market and financial data.

ALGORITHMS USED

Multiple Linear Regression

Multiple linear regression is an extended version of the simple linear regression algorithm. The objective is the prediction of the value of a variable based on the value of two or more other variables. The independent variables are needed to predict the value of the dependent variable. The input independent variables could be of continuous or categorical type. The variable for the prediction is called the dependent variable. The regression coefficient means how the dependent variable changes due to a unit change in the independent variable.

Polynomial Regression

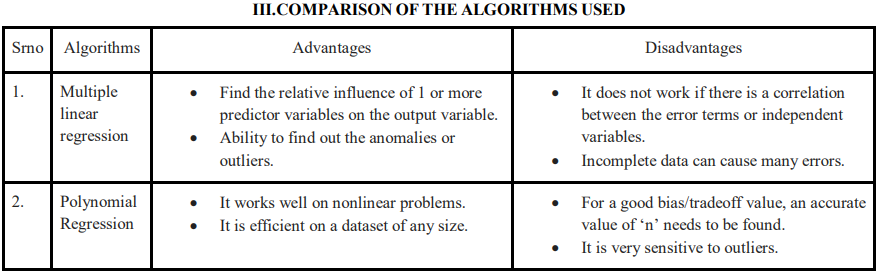
In statistics, a form of regression analysis in which the relationship between the dependent variable y and the independent variable x is modelled like an nth degree polynomial in x is called a Polynomial Regression. It has extra predictors which are obtained by increasing each of the original predictors to a power. The amount of higher-order terms rises with the increasing value of n. Users can fit a non-linear line to a data set using it. This is done via the use of higher-order polynomials such as square, cubic, quadratic, etc. to one or more predictor variables.

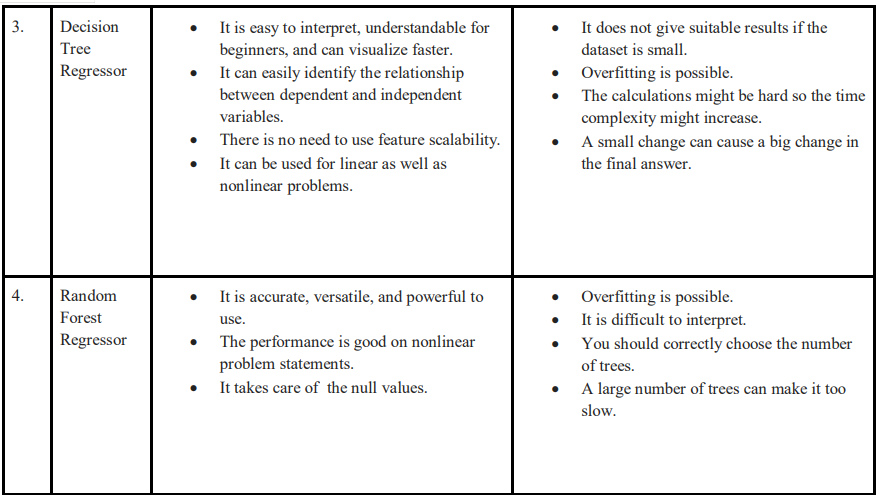
Decision Tree Regressor

A Decision Tree is a very prominent practical approach for supervised learning. It can be employed to solve both Regression and Classification tasks. Both continuous and categorical output variables can be predicted. This algorithm is very useful for resolving decision-related problems. Applications are evaluating growth opportunities for businesses, use of demographic data to find potential clients, etc. The features of an object are analysed to train a model as a tree and produce meaningful continuous output.

Random Forest Regressor

Random forest is a supervised learning algorithm. It is an ensemble learning method for regression and classification problems. An ensemble method uniquely combines the predictions from many ML algorithms together to provide more accurate predictions than any individual model. You can get higher accuracy through cross-validation. Random forests do not overfit so the user can run as many trees as possible





METHODOLOGY

The entire code is in the format of the Python Notebook. Python libraries like Pandas, Numpy load the dataset and perform the mathematical calculations on the dataset. Sklearn is used to implement the four different machine learning algorithms. Matplotlib and Seaborn are needed to visualize the data in an interactive way. The historical data of the last 5 years was downloaded from the Yahoo finance website. The stock in consideration is Tata Consultancy Services - TCS.

The dataset available has the following attributes:

1)Date

2) Open

3) High

4) Low

5) Close

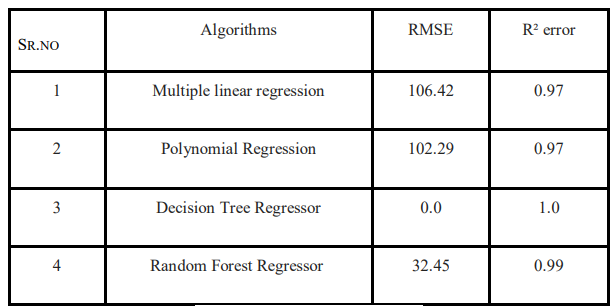
6) Volume

7) Adj. Close

a) Open Price:

First price of the stock at the beginning of a trading day b) Closing Price: The price of the stock at the end of a trading day c) Adj Close: Stock’s value after distributing dividends (True value of the stock) d) High Price: The highest price of the stock in the trading day e) Low Price: Lowest price of the stock in the trading day f) Volume: Number of stocks traded for security in all the markets during a given time

RESULTS



The performance of regression models must be denoted in the form of an error. The performance might change in case of a change of datasets and yield different results. You cannot find the accuracy score. Root mean squared error (RMSE) and the coefficient of determination (R² error) can be used in this case. 1) R² error: It is the proportion of variance in the dependent variable that can be told by the independent variable. It is used to measure the goodness of fit. A larger value means that it is a better model. 2) RMSE: It is the standard deviation of the prediction errors i.e residuals. It gives you a relatively high weight to the large errors. A smaller value means it is a better model.

FUTURE SCOPE

Even though good scores are achieved using ML algorithms, there can be an improvement. Adding more data helps the algorithm to learn better. Hyperparameter optimization is another method of tuning the hyperparameters to get the best performance on the data set provided. It can be implemented using the Scikit-learn machine learning library. The two famous algorithms which can be used are:

1) Grid Search: In this method, a search space as a bounded domain of hyperparameter values is defined then random points are sampled within the bounded domain.

2) Random Search: In this method, a search space as a bounded domain of hyperparameter values is defined, and then random points are sampled in the bounded domain.

Deep Learning algorithms can be implemented to predict accurate results. Deep learning is a branch of machine learning where neural networks algorithms are inspired by the human brain. Long-short term memory (LSTM) can be implemented to predict the stock price. LSTM can learn order dependence in sequence prediction problems. Artificial Neural Network (ANN) is also an extremely recognized method for predictive finance. ANNs are multi-layer fully connected neural nets. Convolutional Neural Networks (CNN) are made up of neurons with biases and learnable weights. CNNs, which are designed to map image data to an output variable, can help to improve predictions. The future prospects include building a Machine learning web app in Python where the user can simply input a stock dataset and get appropriate output with the highest accuracy. The machine app should take in the dataset correctly and choose the algorithm that gives the lowest error rate. The predictions should get printed on the screen. The user interface should be easy and user-friendly for beginners. The app can then be deployed on servers like Heruko to see the model in action

CONCLUSIONS:

Machine learning has applications related to the recommendation of financial products, customer sentiment analysis, etc. In order to predict the prices of stocks, we need the historical data of the stock. This paper analyses the Data Set with 7 attributes and makes a prediction using different regressors to find the future price. It can be seen that the highest accuracy is obtained using the Decision Tree Regressor model with the R² error being 1.0 and the RMSE being 0.0. We can summarize by saying that the Decision Tree Regressor gave the best results out of all the models used for the stock price prediction.

EVALUATION METRICS:

Evaluating the performance of stock price prediction models is crucial to assess their accuracy and effectiveness. Various evaluation metrics can be used to measure how well a model performs in forecasting stock prices. The choice of metrics depends on the specific goals of the prediction task, the prediction horizon (short-term or long-term), and the characteristics of the data. Common evaluation metrics for stock price prediction include:

1. \*\*Mean Absolute Error (MAE):\*\*

- MAE is the average of the absolute differences between the predicted and actual stock prices. It provides a measure of the magnitude of prediction errors.

- MAE is robust to outliers and gives equal weight to all errors.

2. \*\*Mean Squared Error (MSE):\*\*

- MSE is the average of the squared differences between predicted and actual prices. It gives more weight to large errors and penalizes them.

- RMSE (Root Mean Squared Error) is the square root of MSE and provides the error in the same units as the target variable.

3. \*\*Mean Absolute Percentage Error (MAPE):\*\*

- MAPE calculates the percentage difference between predicted and actual prices, making it easier to interpret the accuracy relative to the magnitude of stock prices.

- MAPE is particularly useful for assessing the percentage accuracy of predictions.

4. \*\*Directional Accuracy (DA) or Correct Sign Predictions:\*\*

- DA measures how often the model correctly predicts the direction of price movement (e.g., up or down) without considering the magnitude of price changes.

- It is especially relevant for traders interested in market direction.

5. \*\*Coefficient of Determination (R-squared, R²):\*\*

- R² quantifies the proportion of the variance in the dependent variable (stock prices) that is predictable from the independent variables (features) in the model.

- A higher R² value indicates that a larger portion of the variance in stock prices is explained by the model.

6. \*\*Sharpe Ratio:\*\*

- The Sharpe ratio evaluates the risk-adjusted return of a trading strategy. It considers not only the accuracy of predictions but also the risk taken in making trading decisions.

- It is commonly used to assess the performance of trading strategies that utilize stock price predictions.

7. \*\*Information Ratio:\*\*

- The Information Ratio, like the Sharpe ratio, measures the risk-adjusted return but compares it to a benchmark index or portfolio. It assesses whether the model's predictions outperform a reference.

8. \*\*F1 Score:\*\*

- F1 Score is a metric used when dealing with imbalanced datasets or when precision and recall are important. It combines precision and recall into a single value and is useful for binary classification tasks.

9. \*\*Profit and Loss (P&L) Metrics:\*\*

- P&L metrics measure the financial performance of a trading strategy based on stock price predictions. These include metrics like profit, loss, return on investment (ROI), and drawdown.

10. \*\*Calibration Metrics:\*\*

- Calibration metrics assess the confidence of the model's predictions. These include measures like Brier Score and calibration plots.

11. \*\*Kappa Statistic:\*\*

- The Kappa statistic measures the agreement between observed and predicted price movements, accounting for the agreement expected by chance.

12. \*\*Custom Metrics:\*\*

- Depending on the specific trading strategy or risk management approach, custom metrics may be defined to assess the model's performance in a way that aligns with the desired objectives.

It's important to choose evaluation metrics that align with the goals of your stock price prediction task and consider the practical implications of your predictions in real trading scenarios. Additionally, it's common to use a combination of metrics to gain a comprehensive understanding of the model's performance.

CONCLUSION:

The stock market plays a remarkable role in our daily lives. It is a significant factor in a country's GDP growth. In this tutorial, you learned the basics of the stock market and how to perform stock price prediction using machine learning.

Do you have any questions related to this tutorial on stock prediction using machine learning? In case you do, then please put them in the comments section. Our team of experts will help you answer your questions

This study has several limitations that can provide new directions for future studies. First, we only collected social media text data from one platform. Although we collected as much data as possible from large companies, the investors of other platforms may present different emotions and one website is less representative. We will try to collect more financial social media documents from different platforms in the future. Second, only one stock is selected for prediction in our study.