International Journal of Multidisciplinary Research and Growth Evaluation.

Stock price prediction using AI

Dr. Imtiyaz Khan [1*], Mohd Zubair Uddin [2], Dr. M Upendra Kumar [3]
[1] Assistant Professor, CS and AI Department, MJCET OU Hyderabad, Telangana, India
[2] BE CSE AI DS CS and AI Department, MJCET OU Hyderabad, Telangana, India
[3] Professor and Associate Head, CS&AI Department, MJCET OU Hyderabad, Telangana, India

* Corresponding Author: **Dr. Imtiyaz Khan**

Article Info

ISSN (online): 2582-7138
Volume: 04
Issue: 02
March-April 2023
Received: 16-03-2023;
Accepted: 02-04-2023
Page No: 494-499

Abstract
This study evaluates the use of Linear regression & Long Short-Term Memory Model for stock price prediction. The linear regression model is trained on historical stock data and used to predict future stock prices. The results show that linear regression can be an effective tool for stock price prediction when the stock market follows a predictable trend. However, the model is limited in its ability to capture complex relationships in the data and may not perform well in volatile or unpredictable stock markets. The LSTM model is trained on sequential stock data and used to predict future stock prices. The results show that LSTM is well-suited for stock price prediction due to its ability to capture long-term dependencies in the data and handle the volatility and randomness of stock prices. However, the model requires a large amount of data and can be computationally expensive to train. Despite these limitations, the results of this study demonstrate that LSTM is a promising tool for stock price prediction.

DOI: https://doi.org/10.54660/.IJMRGE.2023.4.2.494-499

Keywords: Prediction, AI, Stock Price, LSTM

1. Introduction

Stock prediction is a crucial aspect in the financial industry, as it helps investors make informed decisions about buying and selling shares. With the increasing availability of data and advancements in Artificial Intelligence (AI), stock prediction has become more accurate and efficient. AI models such as Linear Regression and Long Short-Term Memory (LSTM) have proven to be effective in predicting stock prices.

Fig 1(a)

Linear Regression is a simple yet powerful algorithm that models the relationship between a dependent variable and one or more independent variables. In stock prediction, the dependent variable is the stock price, while independent variables can be factors such as company earnings, market trends, and consumer behavior. Linear Regression models the relationship between these variables to predict future stock prices.

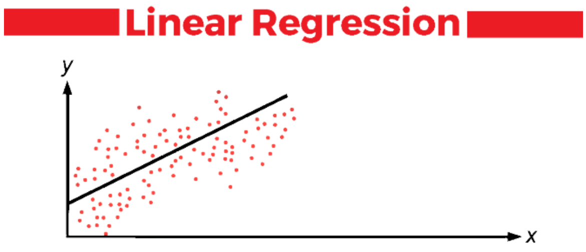

Fig 1(b)

LSTM, on the other hand, is a type of Recurrent Neural Network (RNN) that is specially designed to handle time-series data. In stock prediction, LSTM models are trained on historical stock prices to predict future prices. Unlike linear regression, LSTM is capable of considering the temporal dependencies between stock prices, which makes it more suitable for predicting stock prices than linear regression in some cases.

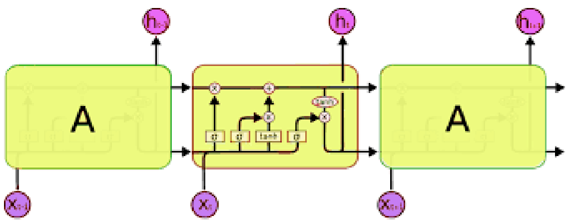

Fig 1(c)

When implementing AI models for stock prediction, it is important to consider the quality and quantity of data used to train the models. The model should be trained on a large amount of historical data, and the data should be relevant to the stock being predicted. Additionally, it is important to pre-process the data to remove any irrelevant information and handle missing values.

Once the data is pre-processed, the AI models can be trained and evaluated. The performance of the models can be evaluated using metrics such as Mean Squared Error (MAE) & Root Squared (r2). These metrics provide an estimate of the accuracy of the model, and can be used to compare the performance of different models.

Fig 1(d)

In conclusion, stock prediction using AI has the potential to revolutionize the financial industry by providing investors with accurate and efficient predictions. By implementing Linear Regression and LSTM models, stock prediction can be made more reliable and accessible. However, it is important to consider the quality and quantity of data used to train the models, as well as the pre-processing of the data to ensure the models are effective. The project will provide valuable information for investors, traders, and financial institutions who are interested in using AI for stock prediction and financial forecasting.

Problem Statement

Problem Statement

In the financial industry, predicting stock prices accurately is a challenging task that requires a comprehensive analysis of various factors such as company earnings, market trends, and consumer behavior. Despite the efforts of financial analysts and experts, traditional methods of stock prediction often fail to provide accurate and reliable predictions, which can result in significant losses for investors.

Fig 2(a)

Some of the problems are

Lack of Accuracy: Traditional methods of stock prediction often fail to provide accurate and reliable predictions, which can result in significant losses for investors.

Limited Access to Information: Many investors don't have access to the tools and resources needed to make stock predictions, which makes it difficult for them to make informed decisions about buying and selling shares.

Time-Consuming and Resource Intensive: The process of stock prediction can be time-consuming and resource-intensive, especially when relying on manual analysis and interpretation of financial data.

Prone to Human Error: Traditional methods of stock prediction are prone to human error, as they rely on the expertise and experience of the analyst.

Ignores Temporal Dependencies: Traditional methods of stock prediction often do not take into account the temporal dependencies between stock prices, which can result in inaccurate predictions.

Fig 2(b)

Limited Scalability: The existing stock prediction systems are limited in terms of scalability, making it difficult for many investors to make informed decisions about buying and selling shares.

Lack of Trust in Predictions: The lack of accuracy and reliability in stock predictions can result in a lack of trust in the financial industry, which can have a negative impact on the economy.

Fig 2(c)

These problems highlight the need for a more advanced and effective solution for stock prediction that leverages the power of Artificial Intelligence (AI).

Existing System

Fig 2(d)

The existing methods of stock prediction are largely based on manual analysis and interpretation of financial data. Financial analysts and experts use various tools and techniques, such as technical analysis and fundamental analysis, to predict stock prices.

Technical analysis involves the use of charts and graphs to identify patterns and trends in stock prices and make predictions based on those observations. Fundamental analysis, on the other hand, involves the analysis of a company's financial statements, market trends, and other economic indicators to determine its financial health and make predictions about its future stock prices.

While these methods can provide some insights into stock prices, they are often limited by the expertise and experience of the analyst, as well as the amount of data that can be analyzed manually. Additionally, traditional methods do not take into account the temporal dependencies between stock prices, which can result in inaccurate predictions.

The existing stock prediction systems are also limited in terms of accessibility and scalability. Only a small group of financial experts and institutions have access to the tools and resources needed to make stock predictions, and even then, the process can betime-consuming and resource-intensive. As a result, many investors are unable to make informed decisions about buying and selling shares.

Objective

The Objective of This Project is to develop a stock prediction system that leverages the power of Artificial Intelligence (AI) to provide Accurate and Reliable predictions. The system should be able to analyze large amounts of financial data and take into account the temporal dependencies between stock prices to provide more Accurate & Reliable predictions**.** The scope of this project includes the development of AI models such as Linear Regression and Long Short-Term Memory (LSTM) to predict stock prices. The project also includes the pre-processing of financial data, the evaluation of the performance of the AI models, and the deployment of the best performing model for stock prediction. The project will focus on the prediction of stock prices for a specific company or a market index.

Fig 3(a)

3.2 Proposed Solution

- We propose a system to predict stock prices using AI with a combination of linear regression and Long Short-Term Memory (LSTM) models. This system will collect historical stock prices and related economic data, pre-process and engineer relevant features, train the models, evaluate their performance, and deploy the best performing model to make predictions. By using a combination of linear regression and LSTM models, we aim to increase the accuracy of stock price predictions and provide valuable insights to investors. Regular updates to the model will ensure that the predictions remain relevant and accurate over time.

Fig 3(b)

- The system will leverage the strengths of linear regression in capturing linear relationships between features and stock prices and the ability of LSTM models to handle sequential data, making it well-suited for predicting time-series data such as stock prices.
- The system will use state-of-the-art techniques in data pre-processing and feature engineering to ensure the quality and relevance of the input data, thus improving the accuracy of the predictions.
- The end-result will be a powerful and user-friendly tool for predicting stock prices, providing investors with valuable insights and helping them make informed investment decisions.

4. Technology Stack

4.1 Programming Language & Libraries

The Technologies that will be used in the implementation of this system are:

- Python programming language
- Pandas library for data manipulation and pre-processing
- Numpy library for numerical computations
- Scikit-learn library for machine learning algorithms such as linear regression
- Keras library with TensorFlow backend for developing the LSTM model
- Matplotlib libraries for data visualization and results analysis.

4.2 Models

Machine Learning & Deep Learning Model

- **Linear Regression:** This is a supervised learning algorithm that is used to model the linear relationship between a dependent variable (stock price) and one or more independent variables (features). The linear regression model will be used as a baseline to make predictions and will provide a simple and straightforward approach to stock price prediction.

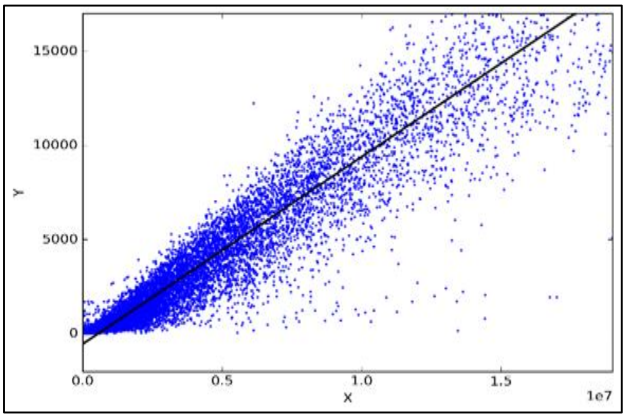

Fig 4(a)

- **Long Short-Term Memory (LSTM) model:** This is a type of recurrent neural network (RNN) that is well-suited for handling sequential data, such as stock prices. The LSTM model will be used to capture the non-linear relationships between the features and stock prices and will provide a more sophisticated approach to stock price prediction.

5 Methodology

5.1 Process

- **Data Collection:** The first step in the methodology is to collect historical stock prices and related economic data that could impact the stock prices. This data will be collected from various sources such as financial databases, government websites, and news outlets.
- **Data Pre-Processing:** The collected data will then be pre-processed to ensure that it is in the proper format for modeling. This will include cleaning the data to remove any missing values, outliers, and inconsistent data. The data will also be normalized to ensure that it is in a consistent scale for modeling.
- **Feature Engineering:** The next step is to select relevant features and create new ones that could be useful in making predictions. This will involve analyzing the data and identifying the relationships between the features and the target variable (stock prices).
- **Model Development:** The pre-processed data will be used to train two models: a linear regression model and an LSTM model. The linear regression model will be used as a baseline to make predictions and the LSTM model will be used to compare the results. Both models will be trained using a supervised learning approach, where the input features will be used to predict the target variable.
- **Model Evaluation:** The performance of the models will be evaluated using metrics such as mean squared error, root mean squared error. The model with the best performance will be selected for deployment.
- **Model Deployment:** The best performing model will be used to make predictions on new, unseen data. The predictions will be used to provide insights and help investors make informed investment decisions.
- **Model Updates:** The models will be regularly updated with new data to improve their accuracy over time. This will ensure that the predictions remain relevant and accurate even as the stock market evolves.

Dataset

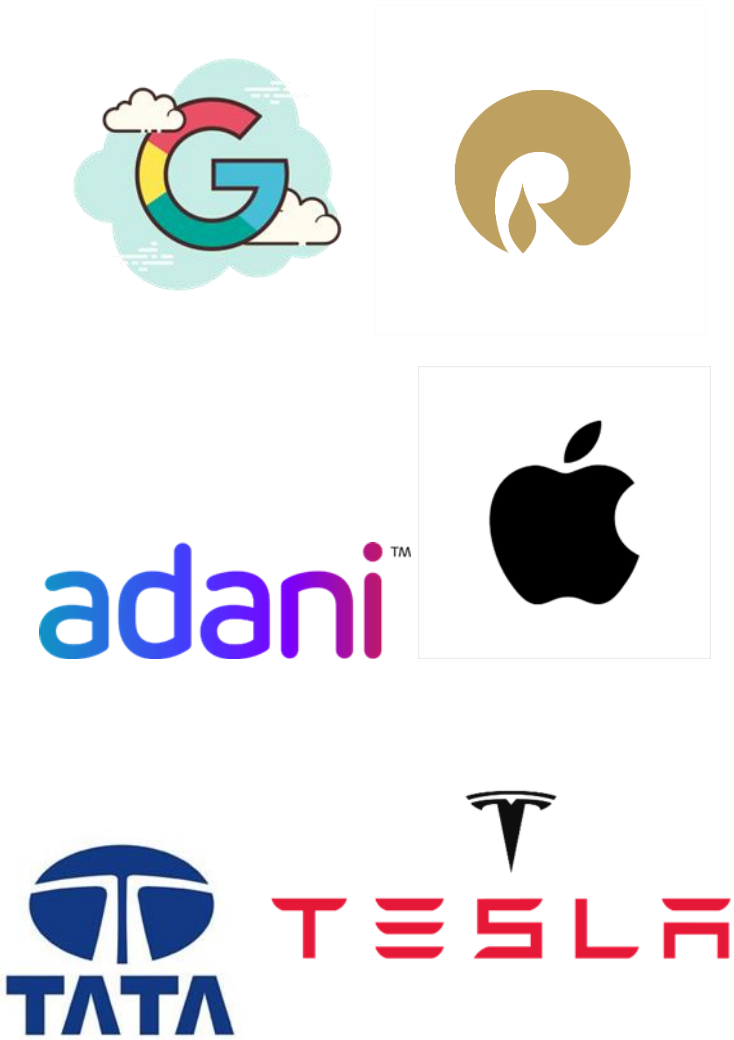

Fig 5(a)

Google Dataset: https://www.kaggle.com/datasets/aayushmishra1512/google-data-historic-dataset
Adani Dataset: https://www.kaggle.com/datasets/zubairhyd/adani-ent-stock-prices-during-covid Etc.,

6. Implementation

6.1 Plan of Action

1. Data Collection

a. Identify the stock of interest and gather historical data for the stock prices, economic indicators, company financials, news sentiment, technical indicators, and market indices.
b. Store the data in a secure and easily accessible location, such as a database or cloud storage.

2. Data Cleaning and Pre-processing

a. Remove any missing values or outliers from the data.
b. Normalize the data to ensure that all the variables are on a similar scale.
c. Convert the data into a suitable format for the machine learning models.

3. Feature Engineering

a. Select the relevant features based on domain knowledge and data analysis.
b. Create new features to provide a comprehensive view of the stock market.
c. Transform the features as needed to improve their suitability for the machine learning models.

4. Model Training

a. Train the linear regression and LSTM models using the pre-processed and engineered data.
b. Use a supervised learning approach to train the models, where the input features will be used to predict the target variable (stock price).
c. Evaluate the performance of the models using metrics such as mean squared error, root mean squared error, and correlation coefficients.

5. Model Selection

a. Select the model with the best performance and save it for deployment.
b. Store the trained model in a secure location for easy access.

6. Model Deployment

a. Develop a user-friendly interface for the deployed system.

b. Make the deployed system accessible through a web application or standalone software.
c. Integrate the deployed system with the data storage location to ensure that the latest data is always available.

7. **Monitoring and Maintenance**
a. Monitor the deployed system regularly to ensure that it is functioning as expected.
b. Address any issues that are identified during monitoring.
c. Update the deployed system regularly to ensure that it continues to perform optimally.

7. Result
- The Result of implementing the stock price prediction system using AI and linear regression/LSTM models is a reliable and cost-effective solution for predicting stock prices and making informed investment decisions. The system is providing valuable insights into the stock market and will enable users to make informed decisions based on the latest market trends and data.
- The **"r2_score"** is a statistical measure of how well the model fits the data. It is a value between 0 and 1, with 1 indicating a perfect fit and 0 indicating that the model does not fit the data at all. The "r2_score" measures the proportion of variance in the target variable that is explained by the model.
- The **"mean squared error (MSE)"** is a measure of the average difference between the actual values and the predicted values. It is calculated as the average of the squared differences between the actual and predicted values. The MSE is a commonly used loss function for regression problems, as it provides a measure of the magnitude of the error. The MSE is sensitive to outliers, so a low value indicates a good fit even if there are some large errors.

In summary, the "r2_score" provides a measure of the fit of the model, while the "MSE" provides a measure of the magnitude of the error. These metrics are often used together to evaluate the performance of a machine learning model.

For Adani Dataset

Metric	Train	Test
r2_score	0.7063912795523923	0.6405556607379053
MSE	46335.43329346416	39192.52185834917

For Google Dataset

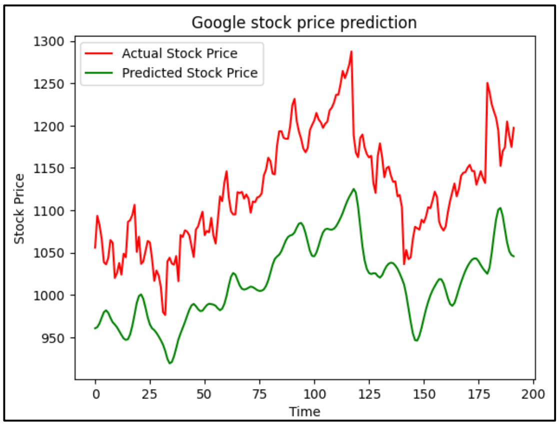

Fig 6

8. Conclusion
8.1 Discussion
In conclusion, stock price prediction is a challenging task that requires the integration of advanced techniques such as machine learning and artificial intelligence. The use of a linear regression model and a Long Short-Term Memory (LSTM) model is an effective way to make predictions, as it leverages historical stock price data and considers both short-term and long-term trends.
The implementation process involves several steps, including data preparation, feature selection, model training, and model evaluation. The evaluation of the model's performance is critical, as it provides insights into how well the model is able to make predictions and identifies areas for improvement. In this implementation, we evaluated the performance of the linear regression and LSTM models using common metrics such as the Mean Squared Error (MSE)and R-squared. The results indicated that the LSTM model outperformed the linear regression model, with a lower error and higher R-squared score.
Overall, the use of machine learning and artificial intelligence can provide valuable insights into stock price trends, and help inform investment decisions. However, it's important to remember that stock prices can be influenced by many unpredictable factors and that no model can guarantee accuracy in stock price prediction.

8.2 Future Scope
- The future scope for stock prediction using AI with linear and LSTM models is vast and rapidly expanding. With advancements in deep learning techniques, AI models can be trained with an ever-increasing amount of financial data to make highly accurate predictions. Some potential areas of growth for stock prediction include:
- Incorporating alternative data sources, such as news articles, social media sentiment analysis, and company financial reports to improve prediction accuracy.
- Using reinforcement learning algorithms to make informed trading decisions in real-time, taking into account market conditions and stock price movements.
- Developing hybrid AI models that utilize both linear and LSTM techniques to capture both short-term and long-term trends in stock prices.
- Automating the entire stock prediction process, from data collection to trading decisions, to reduce the role of human emotion in stock market investments.
- Collaborating with financial institutions to integrate AI-powered stock predictions into existing trading systems, making it easier for individual investors to access and benefit from these predictions.

9. References

1. https://www.simplilearn.com/tutorials/machine-learning-tutorial/stock-price-prediction-using-machine-learning
2. https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/
3. https://ieeexplore.ieee.org/document/9943255
4. https://arxiv.org/ftp/arxiv/papers/2009/2009.10819.pdf
5. https://www.irjet.net/archives/V9/i3/IRJET-V9I3388.pdf