Nitte(DU) established under Section 3 of UGC Act 1965 | Accredited with 'A+' Grade by NAAC

A PROJECT SYNOPSIS ON

" StockPulse: AI Price Predictor "

as a component of

Project in Artificial Intelligence and Machine Learning
AM1602-1

SUBMITTED BY:

Anush C Rao	**NNM23AM009**
Akshay S Aithal	**NNM23AM004**

UNDER THE GUIDANCE OF

Ms. Disha D N

ASSISTANT PROFESSOR GD-II

DEPARTMENT OF AI & ML, NMAMIT

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING
2024-2025

CONTENTS

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to Ms. Disha D N, Assistant Professor Gd-II, NMAMIT Nitte, for their invaluable guidance, suggestions, and support throughout the course of this project. Their expertise and encouragement were instrumental in the successful completion of this work.

We would also like to extend our thanks to all the faculty members of the Artificial Intelligence and Machine Learning Dept. at NMAMIT Nitte for their insightful comments, which helped shape our understanding of the core concepts necessary for the completion of this project. Their courses and lectures provided a solid academic framework for this work, and their mentorship played a crucial role in improving our research and technical skills.

A special thanks goes to the lab assistants, whose technical assistance and feedback helped us address the challenges we faced during data preprocessing, model development, and evaluation.

We are equally grateful to the creators and contributors of the dataset used in this project, as well as the developers of the open-source libraries and tools (such as Python, Scikit-learn, Pandas, and Matplotlib), which made the implementation of this project possible.

We would also like to acknowledge our fellow team-mates for their encouragement and collaboration. Their valuable discussions and willingness to help provided additional motivation to continue our work.

ABSTRACT

Stock price prediction is a challenging task due to the volatile and nonlinear nature of financial markets. This project aims to develop a machine learning-based solution for predicting future stock prices using historical data. Specifically, the project leverages Long Short-Term Memory (LSTM), a type of recurrent neural network (RNN) that is wellsuited for time-series forecasting.

The project utilizes Yahoo Finance (yfinance) as the data source, from which historical stock prices are retrieved for companies like Apple (AAPL) and Google (GOOGL). The closing price data is preprocessed using MinMaxScaler, scaling the data to a range between 0 and 1 to normalize it for model training. A sequence of past 60 days of stock prices is used as input to predict the price for the next day, creating the necessary time-series input for the LSTM model.

The LSTM model is built using TensorFlow/Keras and consists of two LSTM layers followed by fully connected Dense layers to output the predicted stock price. The model is trained on 80% of the historical data and tested on the remaining 20%. The
performance of the model is visualized by plotting the predicted prices against the real prices, allowing us to assess the accuracy of the predictions.

The results show that the LSTM model can capture general trends in stock prices, although precise predictions remain a challenge due to the inherent unpredictability of financial markets. The project demonstrates the potential of machine learning, specifically LSTM networks, in timeseries forecasting and highlights the importance of proper data preprocessing and model selection in stock price prediction tasks.

Stock predictions

INTRODUCTION

Predicting stock prices is a challenging yet intriguing task in finance, primarily due to the volatile and dynamic nature of financial markets. With advancements in machine learning, especially deep learning techniques like **Long Short-Term Memory (LSTM)** networks, we can analyse historical data to uncover patterns that aid in forecasting future stock prices.

In this project, we utilize LSTM networks to predict stock prices based on historical data sourced from **Yahoo Finance** using the **yfinance** library. The data undergoes preprocessing with **MinMaxScaler** to normalize the values for better model performance. By analysing sequences of past 60 days' stock prices, we aim to predict the price for the next day.

This project showcases the potential of machine learning in stock price forecasting, highlighting how LSTM networks can be employed to analyse time-series data. While external factors influencing financial markets are not accounted for in this model, our approach provides a foundational understanding of using machine learning for stock price prediction.

OBJECTIVES

Data Collection:

- To gather historical stock price data for selected companies using the **yfinance** library, focusing on the closing prices for analysis.

Data Preprocessing:

- To preprocess the collected data by normalizing it with **MinMaxScaler** to ensure effective training of the LSTM model.
- To create sequences of stock prices that the model can use to learn patterns over time, specifically utilizing the past 60 days of data to predict the next day's price.

Model Development:

- To develop an LSTM neural network model capable of capturing temporal dependencies in the stock price data.
- To configure and compile the model with appropriate parameters, including the number of layers, units, and optimization techniques.

Model Training:

- To train the LSTM model on the prepared dataset, splitting it into training and testing sets to evaluate the model's performance effectively.

Prediction and Evaluation:

- To make predictions on unseen test data and evaluate the model's accuracy by comparing predicted stock prices against actual prices.

- To visualize the predictions alongside historical prices, providing insights into the model’s performance and effectiveness.

Analysis of Results:

- To analyze the model’s predictions and understand the limitations of using LSTM for stock price forecasting, considering external market factors that may influence stock prices.

Future Recommendations:

- To identify potential improvements for the model, such as incorporating additional features or exploring other machine learning techniques to enhance predictive accuracy.

METHODOLOGY

1. **Data Acquisition**:

- Historical stock price data is collected using the **yfinance** library. This library allows easy access to financial data from Yahoo Finance, specifically targeting the **closing prices** of selected stocks over a specified date range.

2. **Data Preprocessing**:

- **Normalization**: The collected stock prices are scaled to a range between 0 and 1 using **MinMaxScaler**. This normalization step is crucial as it helps the LSTM model learn effectively by ensuring that all input values are on a similar scale.

- **Sequence Creation**: To prepare the data for the LSTM model, sequences of stock prices are created. Specifically, for each target price (the price to be predicted), the preceding **60 days'** prices are used as input. This creates a dataset where each input (X) corresponds to the next day's price (y).

3. **Dataset Splitting**:

- The pre-processed data is divided into training and testing sets. Typically, **80%** of the data is used for training the model, while the remaining **20%** is reserved for testing and evaluating the model's performance.
-

4.**Model Development**:

- An LSTM neural network is constructed using **TensorFlow/Keras**. The model consists of:
  - Two LSTM layers to capture temporal patterns in the stock price data.
  - Dense layers that further process the output from the LSTM layers.
- The model is compiled using the **Adam optimizer** and **mean squared error** as the loss function to measure prediction accuracy.

5. **Model Training**:

- The LSTM model is trained using the training dataset for a specified number of epochs (iterations). The model learns to minimize the prediction error through backpropagation and gradient descent techniques.

6. **Prediction and Evaluation**:

- After training, predictions are made on the test dataset. The predicted stock prices are then rescaled back to their original scale using the inverse transformation of the **MinMaxScaler**.
- The model's performance is evaluated by comparing the predicted prices against the actual prices. Visualizations are created to show both predicted and actual prices, helping assess the accuracy of the predictions.

7.**Analysis of Results**:

- The results are analyzed to determine how well the LSTM model captures trends in stock prices. Limitations of the model are discussed, including the impact of external factors not accounted for in the predictions.

8. **Future Improvements**:

- Recommendations for future work include exploring additional features (such as technical indicators), experimenting with more complex architectures, or implementing other machine learning algorithms to enhance predictive accuracy.

CONCLUSION

This project demonstrates the application of **Long Short-Term Memory (LSTM)** neural networks to predict stock prices using historical data. By leveraging the sequential learning capabilities of LSTM, we were able to model the time-series behaviour of stock prices and make predictions for future prices. The use of **yfinance** allowed seamless access to real-world stock market data, while **MinMaxScaler** ensured proper data normalization for training the model.

Through this project, we observed that LSTM networks can capture general trends and patterns in stock prices over time, providing reasonably accurate predictions. However, the complex and highly volatile nature of stock markets makes exact prediction challenging. External factors, such as market sentiment, economic events, and company-specific news, are not captured in this model, which limits its predictive performance.

Despite its limitations, this project serves as a foundational step towards understanding the role of machine learning in financial forecasting. Future work can focus on incorporating additional features like technical indicators, market sentiment analysis, or the use of more sophisticated deep learning architectures to improve accuracy. Additionally, the use of more advanced optimization techniques, larger datasets, and tuning hyperparameters can further enhance the model's predictive capability.

In conclusion, while LSTM-based stock prediction provides valuable insights, it should be used in combination with other methods and domain expertise to make informed investment decisions.

REFERENCES:

- Stock price prediction dataset
  Yahoo Finance API Documentation (yfinance)
  Official documentation for the (yfinance) library used to download stock data in Python.