Stock Price Prediction using Brownian Motion

Submitted in Partial fulfilment of the

requirements of the degree of

Bachelor of Engineering

in

Artificial Intelligence & Data Science

by

Apurva Paranjape	*121A8039*
Aditya Patil	*121A8040*
Ayush Patil	*121A8041*
Arnav Pawar	*121A8042*

UNDER THE GUIDANCE OF

Dr. Rizwana S.

In

Artificial Intelligence & Data Science

SIES Graduate School of Technology (Autonomous)

RISE WITH EDUCATION

NAAC A+

DEPARTMENT OF ARTIFICIAL INTELLIGENCE

AND DATA SCIENCEE

SIES GRADUATE SCHOOL OF TECHNOLOGY

NERUL, NAVI MUMBAI – 400706

ACADEMIC YEAR

2024 – 2025

Certificate

This is to certify that the project entitled "**Stock Price Prediction using Brownian Motion**" is a bonafide work carried out by the following students of final year in Artificial Intelligence & Data Science

Sr.no	Name	Roll No.
1	Apurva Paranjape	121A8039
2	Aditya Patil	121A8040
3	Ayush Patil	121A8041
4	Arnav Pawar	121A8042

The report is submitted in partial fulfillment of the degree course of Bachelor of Engineering in Artificial Intelligence & Data Science, of University of Mumbai during the academic year 2024 – 2025.

Dr.	**Dr.**	**Dr. K. Lakshmisudha**
Rizwana S.	**Rizwana S.**	**Principal**

We have examined this report as per university requirements at SIES Graduate School of Technology, Nerul, Navi Mumbai on ____________________

Project Report Approval

This project report entitled **Stock Price Prediction using Brownian Motion** by following student is approved for the degree of **Bachelor of Engineering** in **Artificial Intelligence & Data Science**

Apurva Paranjape	121A8039
Aditya Patil	121A8040
Ayush Patil	121A8041
Arnav Pawar	121A8042

Name of External Examiner: ______________________

Signature: ______________________

Name of Internal Examiner: ______________________

Signature: ______________________

Date:

Place:

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Name	Roll No.	Signature
Apurva Paranjape	121A8039	
Aditya Patil	121A8040	
Ayush Patil	121A8041	
Arnav Pawar	121A8042	

Date:

Acknowledgement

We would like to express our thanks to the people who have helped us the most throughout our project. We are grateful to our guide **Dr. Rizwana S.** and coordinator **Prof. Rohini Gaikwad** for nonstop support for the project.

A special thanks goes to each other who worked together as a team in completing the project, where we all exchanged our own interesting ideas, thoughts and made it possible to complete our project with all accurate information. We also wish to thank our parents for their personal support and attention who inspired me to go my own way.

We would also like to extend our sincere gratitude to our Head of the Department **Dr. Rizwana S.** and our Principal **Dr. K. Lakshmisudha** for their continuous support and encouragement.

We also would like to thank our other faculty members for providing us with all the required resources and references for the project.

Project Team

Apurva Paranjape

Aditya Patil

Ayush Patil

Arnav Pawar

Abstract

Stock price forecasting is crucial to financial decision-making and risk analysis. Geometric Brownian Motion (GBM) is one of the most common methods for modelling stock prices as random changes with long-term trends. GBM makes the assumption that stock prices are represented by a continuous-time stochastic process, and hence it can be used to simulate the price movement in the future. It has two important components: drift, which is the average return, and volatility, which is the randomness of the change. The model needs historical Stock data to calculate these parameters. Once determined, they can be employed to create several future price paths via simulation. This approach provides investors with a probabilistic picture of possible stock behaviour over time. It assists in comprehending both the risks and opportunities involved with particular investments. Though mathematically simple and easy to implement, GBM has its own limitations. It makes the assumption of constant volatility and fails to capture abrupt market events or structural shifts. These assumptions may not always hold true in actual stock behaviour. Yet, the model is a good starting point for making predictions and constructing more sophisticated models. This problem is solved by suggesting a hybrid model that integrates the LSTM-ARO (Attention Recurrent Output) model with the stochasticity of Geometric Brownian motion. The LSTM-ARO model is developed to learn intricate patterns and relationships in past stock prices, whereas GBM offers a complementary stochastic view by representing market volatility and randomness. The integration of these two approaches exploits their individual strengths, seeking to enhance the precision and stability of stock price forecasting.

Keywords: Geometric Brownian Motion (GBM), Stock Price Forecasting, Drift and Volatility, Financial Prediction

Contents

List of Figures

List of Tables

List of Abbreviations

GBM	Geometric Brownian Motion
LSTM	Long Short-Term Memory
ARO	Attention Recurrent Optimization
MSE	Mean Squared Error
RMSE	Root Mean Square Error

Chapter 1

Introduction

1.1 Overview

Brownian Motion, also known as the Wiener Process, is a mathematical model that represents continuous and random movement. Because of this, it is well-suited for simulating the unpredictable behaviour of stock prices [1]. Stock markets are known for their volatility, and the randomness of Brownian Motion closely mirrors real market behaviour. This aligns with the random walk hypothesis [2], which states that stock price changes are independent and identically distributed over time. According to this hypothesis, past movements do not predict future prices. Therefore, price changes appear to follow a random pattern, similar to Brownian Motion.

However, traditional forecasting models often fall short. They struggle to capture complex relationships between past and future price movements. On the other hand, purely stochastic methods, like GBM, cannot easily adapt to changing market conditions or long-term patterns. [3]

To overcome these limitations, we propose a hybrid modelling approach. This combines the strengths of statistical methods with the power of deep learning. The system integrates Geometric Brownian Motion (GBM), Long Short-Term Memory (LSTM) networks, and Attention mechanism [10]. GBM is used to simulate short-term price fluctuations using stochastic principles. LSTM networks are excellent at learning from sequences of data. They are designed to capture long-term dependencies in time series. The Attention layer further enhances the model by helping it focus on the most relevant past information. By combining these components, the hybrid model improves the accuracy and adaptability of stock price predictions. It leverages the randomness of GBM, the memory capability of LSTM, and the focus provided by Attention. This integrated system offers a more robust solution for forecasting in dynamic and noisy financial markets.

1.2 Need of Project

Stock market prediction is crucial for both financial professionals and individual investors. Accurate forecasts can enhance investment strategies, risk management, and decision-making. However, the inherent volatility and unpredictability of markets make reliable predictions challenging. Traditional models often fail to capture complex, non-linear relationships, limiting their ability to predict long-term price trends or respond to sudden market shifts.

Geometric Brownian Motion (GBM), a common model for stock price behaviour, has limitations. It assumes constant volatility and drift, which doesn’t account for market shocks or changing conditions. While GBM captures short-term randomness well, it struggles with longer-term trends. Furthermore, stochastic models like GBM don’t leverage the power of modern machine learning techniques. Long Short-Term Memory (LSTM) networks are effective at identifying long-term patterns in sequential data; however, they often struggle to capture short-term variations and volatility without incorporating stochastic elements [3][4].

With the rising complexity of financial markets and the exponential growth of data, there is a growing demand for advanced models capable of handling vast datasets and adjusting to evolving market conditions. Traditional models are struggling, making hybrid solutions combining statistical methods with deep learning more necessary. This project proposes a hybrid model integrating GBM, LSTM, and an Attention mechanism to address the limitations of traditional models. The goal is to build a more accurate, adaptive stock price prediction system. Combining the randomness of GBM, the memory of LSTM, and the focus provided by Attention offers a promising solution to improve forecasting accuracy and adaptability. [8]

Ultimately, this project aims to meet the growing demand for advanced models that deliver real-time, high-quality predictions, leading to more efficient trading, better risk management, and improved financial planning.

1.3 Scope

The scope of the project titled "Stock Price Prediction using Geometric Brownian Motion (GBM)" centres on developing a predictive model to forecast future stock prices based on historical data. The primary objective of the project is to utilize Geometric Brownian Motion (GBM), a well-established stochastic model, to simulate future price trajectories of stocks.

GBM assumes that stock prices follow a continuous-time process characterized by random fluctuations, making it well-suited for modelling the inherent volatility of financial markets.

This project will involve gathering historical stock price data from trustworthy platform, Yahoo Finance. The collected data will undergo pre-processing to ensure it is cleaned and standardized, making it suitable for model training and development. A key component of the project is the estimation of the parameters required for the GBM model, specifically the drift (expected return) and volatility (variability in stock prices), which will be derived using statistical methods, such as Maximum Likelihood Estimation or regression techniques.

Once the parameters are calculated, the GBM model will be used to generate simulated stock price trajectories over a defined period. By simulating multiple price paths, the model will offer a probabilistic view of potential future stock prices, allowing investors to assess the range of possible outcomes and quantify the risk involved. The model's effectiveness will be assessed by comparing its predicted stock prices to actual historical values, utilizing evaluation metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and R-squared error.

Visualization tools will be included to help users easily interpret the results, displaying both actual stock prices and the predicted price trajectories. These visualizations will provide insights into the uncertainty and range of possible future prices, making the predictions more accessible and actionable for users.

The project will address the inherent limitations of the GBM model, such as its assumption of constant volatility and its inability to account for sudden market shocks. While GBM offers a useful starting point, the scope also allows for future enhancements, such as integrating sentiment analysis [5] or machine learning techniques, to improve prediction accuracy and adapt to the evolving dynamics of financial markets.

1.4 Project Schedule

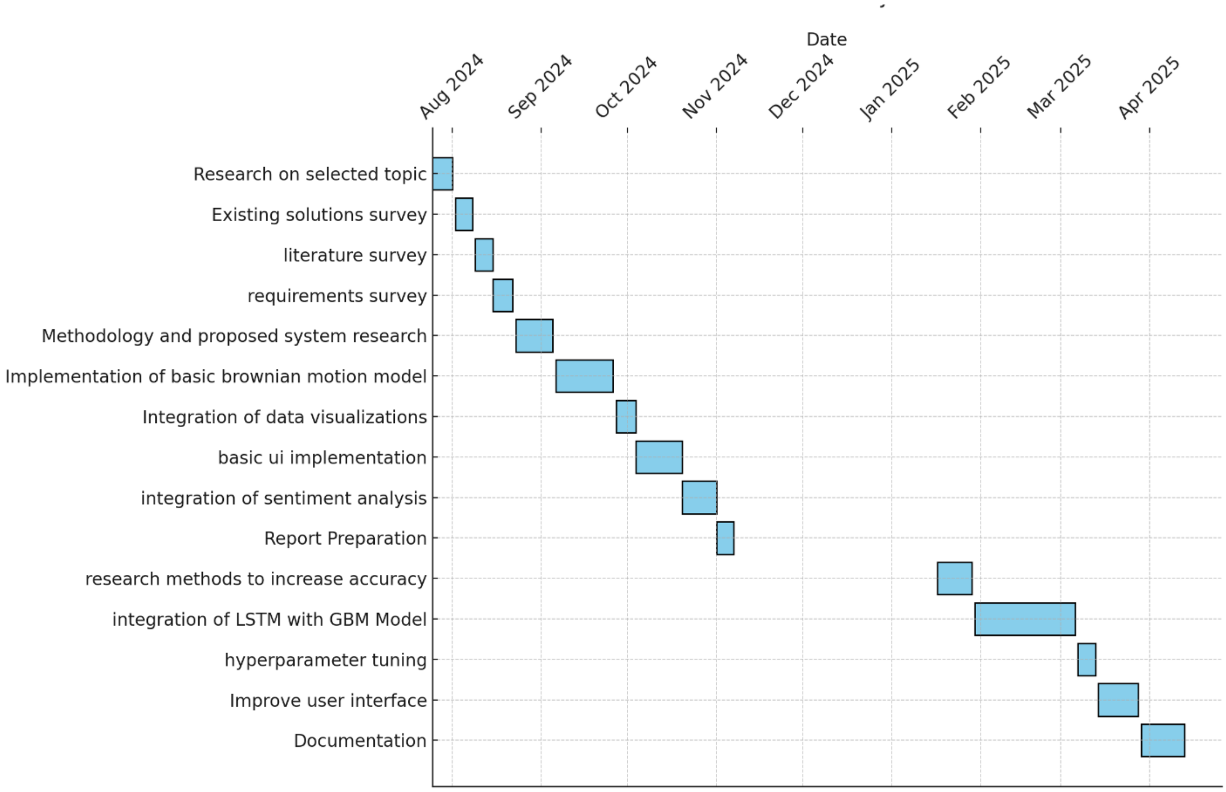


Figure 1.1: Project Schedule - Gantt Chart

The Gantt chart as shown in Figure 1.1 for the project delineates a strategic timeline from its inception on July 25, 2024, to its completion by April 13, 2025, methodically dividing the project into key phases—Research, Implementation, Evaluation, and Finalization.

Beginning with the Research phase, this spanned from late July to early September 2024, encompassing tasks such as topic research, surveys, and the formulation of methodology. These foundational activities were successfully completed within their scheduled durations, setting a solid base for the system's development.

The Implementation phase, stretching from early September to early November 2024, focused on building the core system components. This included the integration of a Brownian motion model, visualization elements, basic UI implementation, and sentiment analysis—each contributing to a functional prototype of the system.

Moving into Evaluation and Enhancement, the project resumed in January 2025 with a focus on improving model accuracy and enhancing system intelligence. This phase lasted until mid-March 2025, with significant tasks like integrating LSTM with the GBM model and tuning hyperparameters to optimize performance.

Finally, the Finalization phase, spanning mid-March to mid-April 2025, involved refining the user interface and developing comprehensive documentation. These concluding efforts ensured the delivery of a polished and fully documented system ready for presentation or deployment.

This Gantt chart effectively maps out the project's lifecycle, offering clear visibility into the timing, sequence, and duration of each task, thereby supporting efficient timeline management and ensuring structured progress toward the project's objectives.

1.5 Organization of the report

This report is broken down into chapters. The report begins with Chapter 1, which is the project's introduction, which includes the introduction, the rationale for the solution provided for the problem that was faced, the project's problem statement, and the objectives for each element in the developed framework, as well as the report's structure.

Proceeding to Chapter 2, it presents the Literature Survey and discusses the work in this field. It is further divided into a literature study of previous ways to address, as well as how those solutions were implemented in the real life and the approaches' shortcomings. A contribution to a mini project is also included.

In Chapter 3 of this report, Proposed system, it mentions the importance of having a simple but a rapid system. The proposed system is also explained and how will it overcome the short comings of the existing system. A simple workflow or an architecture of the system is explained which will help the reader to understand the proposed system.

Chapter 4 is covering the system design, mention and elaborate the various actors involved in this system with the usage of hardware and software in the suggested system. This chapter covers the software design including the high level diagram and the low level diagrams with all the algorithm implementation and also the analysis of these algorithms.

The experiment results and the analysis of the system with various testing mechanism is illustrated in the Chapter 5. This chapter starts with the implementation of the proposed system showing the user flow and how the system can be used by the users. The second part shows the test cases and the result for those test-cases hence proving the durability of the system

Everything concludes at Chapter 6 with all the possible roads we can take from our current model and describing each one of the paths that we mentioned under a roof called as future work and after that the report ends with a word of conclusion that concludes our report and gives an overview of everything.

Chapter 2

Literature Survey

2.1 Survey of Existing system

Stock price prediction has long been a key area of interest for investors, analysts, and financial professionals. Over the years, various models have been proposed, ranging from traditional statistical methods to advanced machine learning techniques. Geometric Brownian Motion (GBM) is a widely used stochastic model for simulating stock price movements. It assumes that stock prices follow a continuous-time stochastic process, characterized by random fluctuations with constant drift and volatility. GBM is often employed in financial systems for modelling price dynamics under uncertainty, including its application in the Black-Scholes option pricing model. While GBM is effective in simulating random price movements, it has limitations. It assumes constant volatility and does not account for sudden market shocks or long-term trends. As such, GBM often falls short in predicting stock prices during periods of high market volatility or structural market changes. [3] Despite these limitations, GBM remains a popular method for stock price prediction due to its simplicity and ease of use.

LSTM (Long Short-Term Memory) networks are widely used in stock price forecasting because of their strength in modeling long-term dependencies within sequential data. Unlike traditional models, LSTMs do not rely on static assumptions and can model complex patterns in financial time series data. [11] LSTM-based systems are widely used in stock forecasting, where historical stock data is fed into the model to predict future prices. LSTMs have demonstrated superior performance compared to traditional models like ARIMA due to their ability to learn from sequential data. However, LSTMs are not without their challenges. These models require large datasets and substantial computational resources for training. Additionally, LSTMs can still struggle with predicting extreme market events or sudden shocks, which are critical to stock price movements. For instance, Zhang et al. (2018) used LSTM to predict stock prices based on historical data, demonstrating that LSTMs can capture sequential dependencies and improve forecasting accuracy. However, like other machine learning models, LSTMs still face limitations in handling outlier events such as financial crises or market crashes.

Attention mechanisms, particularly when combined with LSTM networks, have gained attention in recent years for improving stock price prediction. These mechanisms help models focus on the most relevant parts of historical data, allowing them to prioritize important information when making predictions. In stock price forecasting, attention mechanisms enable the model to focus on critical events that may have a significant impact on future stock prices, such as market crashes or major economic announcements. Attention-based models improve the LSTM's ability to focus on crucial past data, enhancing prediction accuracy. For example, Liu et al. (2019) integrated an attention mechanism with LSTM for stock prediction, improving the model's performance by allowing it to focus on key time periods of high volatility or significant price movement.

To leverage the strengths of each technique, hybrid models combining GBM, LSTM, and Attention mechanisms have been proposed. GBM is well-suited for modelling the randomness and volatility of stock prices, while LSTM excels at capturing long-term dependencies in time series. Attention mechanisms enhance the LSTM model by focusing on the most relevant past information, leading to better performance. Hybrid models have shown promise in improving stock price prediction accuracy by addressing the weaknesses of individual techniques.

Stock price prediction has evolved significantly over the years, with new models and techniques being developed to address the challenges of volatility, non-linearity, and market anomalies. While traditional models like GBM provide useful insights into the randomness of stock price movements, they fail to capture long-term dependencies or market shocks. LSTM networks, with their ability to model complex temporal patterns, offer a more advanced approach. The integration of attention mechanisms further enhances the predictive power of LSTMs by allowing the model to focus on the most impactful data. Hybrid models that combine GBM, LSTM, and Attention mechanisms offer a promising solution for more accurate and adaptive stock price forecasting. These systems are better equipped to handle the dynamic and noisy nature of financial markets, providing more reliable predictions. This project aims to build upon these advancements by combining these techniques in a unified model, offering an innovative solution for stock price prediction

2.2 Table of Literature Survey

Sr. No	Paper Title	Author(s)	Description	Key Findings
1	Mathematically Forecasting Stock Prices with Geometric Brownian Motion (2024) [1]	Anay Kumar, Isa Jamadar, Radhika Goel, Ram Charan Petluri, Wei Feng	The paper proposes a Geometric Brownian Motion equation to better account for volatility and output an accurate model of a stock's performance over a period of time.	The developed model provides an effective approach of creating a highly accurate rendition of the stock market through the use of existing quantitative metrics.
2	Random Walk Theory and Application (2024) [2]	Oluwarotimi Israel Oluwafemi, Emmanuel Olamigoke Famakinwa, Ometere Deborah Balogun	The paper seeks to elucidate the properties and implications of random walks, offering insights into their behaviour in practical scenarios like stock price fluctuations and particle movement, and serving as a fundamental framework for understanding stochastic processes.	The Random Walk Theory posits that within an efficient market, stock prices exhibit randomness due to their inherent unpredictability and the influence of financial demands.
3	Comparison of Stock Price Prediction using Geometric Brownian Motion and Multilayer Perceptron (2020) [3]	Martin Pfannemüller, Martin Breitbach, Christian Krupitzer, Markus Weckesser	Artificial neural network and geometric Brownian motion have been employed to forecast stock prices and produce predictions which are close to the original data. In this paper, the authors have discussed the comparison of the	The prediction results using Geometric Brownian Motion with 10 trajectories yield a better MAPE value compared to the multilayer perceptron model.

			results of geometric Brownian motion and multilayer perceptron in predicting Microsoft stock data.	
4	Study on the Prediction of Stock Price Based on the Associated Network Model of LSTM (2019) [4]	Guangyu Ding, Liangxi Qin	The paper presents a deep recurrent neural network model built on Long Short-Term Memory (LSTM) for predicting multiple stock prices, including opening, lowest, and highest prices simultaneously.	The associated network model can predict stock prices with over 95% accuracy and outperforms standard LSTM and deep recurrent models in handling multiple associated stock price values.
5	Integrating StockTwits with Sentiment Analysis for Better Prediction of Stock Price Movement (2018) [5]	Rakhi Batra, Sher Muhammad Daudpota	The paper explores the use of sentiment analysis on tweets from StockTwits to improve stock price movement predictions. It combines sentiment data with market index data from Yahoo Finance.	SVM stock prediction model had 76.65% accuracy, indicating potential of sentiment analysis combined with market data. Sentiment analysis model had high accuracy and precision, suggesting it as a useful tool for investors.
6	Prediction of Stock Values Changes using Sentiment Analysis of Stock News Headlines (2021) [6]	Guangyu Ding, Liangxi Qin	The paper presents a stock price prediction framework that integrates six analytical models and sentiment analysis based on Harvard psychological	Models incorporating sentiment analysis significantly outperform traditional bag-of-words models in predicting stock

			and Loughran–McDonald financial dictionaries.	prices at the individual stock, sector, and index levels.
7	A Deep Learning Framework for Financial Time Series using Stacked Autoencoders and Long-Short Term Memory (2017) [7]	W. Bao, J. Yue, Y. Rao	This paper presents a deep learning framework that combines stacked autoencoders with long short-term memory (LSTM) networks to predict financial time series data more precisely.	The hybrid model improves accuracy in capturing temporal dependencies and financial market trends.
8	A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction (2017) [8]	Y. Qin et al.	The authors present a dual-stage attention-based recurrent neural network (DA-RNN) that enhances prediction performance by focusing on both input features and relevant time steps in historical data.	The model significantly improves prediction performance by focusing on relevant past time steps and features.
9	Practical Deep Reinforcement Learning Approach for Stock Trading (2018) [9]	R. Xiong, Z. Yao, J. Li, S. Ma, X. Zhou	This study utilizes a deep reinforcement learning (DRL) to create stock trading strategies that adjust to volatile market conditions and optimize trading decisions.	DRL outperforms traditional models in profitability and adaptability to market changes.
10	Application of LSTM and Attention	K. Kim, C. Won	This paper explores the integration of the attention mechanism	Integration of attention improves LSTM model

	Mechanism for Stock Price Prediction (2023) [10]		with LSTM networks to enhance the performance of stock price prediction models.	accuracy in volatile market conditions.
11	LSTM Based Stock Price Prediction (2019) [11]	M. S. Hossain, M. A. Rahman, M. A. H. Akhand	The authors use LSTM neural networks to predict stock prices by analyzing historical data and identifying trends and patterns.	LSTM-based models capture long-term dependencies and yield reliable short-term forecasts.
12	A novel LSTM-GAN approach for financial time series forecasting (2019) [12]	Guangxuan Zhu; Hongbo Zhao; Haoqiang Liu; Hua Sun	This paper proposes a hybrid approach that uses LSTM to capture long term dependencies and GANs to generate synthetic data that augments the training data.	Enhance the predictive accuracy of LSTM models by incorporating the data augmentation capabilities of GANs.

2.3 Research Gaps

Stock price prediction plays a crucial role in financial decision-making, risk management, and strategic planning. Accurate stock price forecasts enable investors to make informed decisions, optimize their portfolios, and manage risk effectively. However, the inherent volatility and unpredictability of the stock market present significant challenges in achieving reliable predictions. Traditional forecasting models, such as linear regression or ARIMA, often fall short in capturing the complex, non-linear dependencies present in financial data. These models struggle to account for sudden market shifts, volatility clustering, and long-term trends, which are common in real-world stock prices.

Geometric Brownian Motion (GBM) has been widely used for modelling stock price behaviour due to its simplicity and effectiveness in capturing random fluctuations. GBM assumes that stock prices follow a continuous-time stochastic process with constant drift and volatility. While this model is useful for simulating short-term price movements, it fails to capture the complexities of real market behaviour, especially during periods of high volatility or market crashes. Additionally, GBM assumes constant volatility, which does not reflect the reality of market environments where volatility often varies over time.

Long Short-Term Memory (LSTM) networks are increasingly favored for stock price forecasting because of their capacity to model long-term dependencies in sequential data. Unlike traditional models, LSTMs do not rely on static assumptions and can adapt to complex patterns in time series data. LSTM models are particularly effective at learning from large datasets, allowing them to model non-linear relationships and trends in stock prices. However, LSTMs alone still face challenges in accurately forecasting sudden price changes or extreme market events due to the inherent randomness and volatility of financial data.

The integration of attention mechanisms with LSTM networks has emerged as a promising approach for improving stock price predictions. Attention mechanisms enable the model to focus on the most relevant parts of the input data, allowing it to give more weight to crucial time periods that are likely to have a significant impact on future price movements. This focus on important past data helps improve the model's accuracy and adaptability, especially in volatile market conditions. Attention mechanisms enable the model to focus on essential information, helping it capture significant market events like financial crises or major economic announcements, which are often pivotal to stock price movements.

Hybrid models that combine GBM, LSTM, and attention mechanisms have shown potential for improving prediction accuracy. GBM captures the randomness and short-term fluctuations of stock prices, while LSTMs model long-term trends and dependencies. Attention mechanisms help refine the model's focus, enhancing its predictive power by identifying the most relevant data points. These hybrid models are able to overcome some of the limitations of individual methods, providing more accurate and reliable stock price forecasts. However, challenges persist, such as requirement for large and high-quality datasets, the risk of overfitting and computational complexity.

In conclusion, while traditional models like GBM provide valuable insights into stock price behaviour, they are limited by their assumptions of constant volatility and drift. LSTM networks offer a more advanced approach, capturing long-term dependencies, but they still face challenges in handling sudden market shifts. Attention mechanisms further enhance the predictive power of LSTM models by focusing on the most relevant historical data. Hybrid models that integrate these techniques offer a promising solution for more accurate and adaptive stock price forecasting, addressing the complexities of real-world financial markets. Future research should focus on improving model efficiency, handling extreme market events, and optimizing the combination of stochastic and machine learning methods for more robust predictions.

2.4 Problem Definition

Predicting stock prices is essential for investors and financial analysts to make informed decisions and optimize portfolios. However, this task is inherently difficult due to the volatility and unpredictability of financial markets. Stock price fluctuations are driven by numerous factors, including economic conditions, geopolitical events, and investor sentiment, which complicates accurate forecasting. Traditional forecasting models, such as linear regression or ARIMA, often fail to account for the complexity and non-linear relationships in stock market data.

Geometric Brownian Motion (GBM) is commonly used to model stock price behaviour. GBM assumes that stock prices follow a random walk with constant drift and volatility, making it a simple and effective model for short-term fluctuations. However, this model has significant limitations. It assumes constant volatility and fails to account for sudden market shifts, volatility clustering, or long-term trends. These limitations make GBM inadequate for forecasting stock prices in dynamic financial environments.

To address these issues, Long Short-Term Memory (LSTM) networks, a form of recurrent neural network (RNN), have been introduced to effectively process sequential data and capture long-term dependencies. LSTMs are capable of learning complex patterns in stock price data and can adapt to non-linear trends. However, they still struggle with modelling sudden price changes and extreme market events, making them less effective for short-term forecasting. Moreover, LSTMs often require large datasets for training, which may not always be available.

Attention mechanisms have been integrated with LSTM models to improve their performance. Attention mechanisms compel the model to focus on the most relevant parts of the input sequence, improving its ability to identify key events that influence stock prices. By prioritizing critical information, attention mechanisms help improve prediction accuracy, particularly in volatile market conditions. This makes attention mechanisms a valuable addition to the forecasting model.

Despite the strengths of GBM, LSTM, and attention mechanisms individually, combining them into a single hybrid model could lead to more accurate and adaptive stock price predictions. GBM can capture the randomness of stock prices, while LSTM can model long-term trends and dependencies. Attention mechanisms can further enhance the model by focusing on significant data points that influence future price movements.

The main challenge is to effectively integrate these techniques into a unified model that can handle the complexities of financial data. The hybrid model must be capable of adjusting to market volatility, abrupt changes, and external influences like news events and economic indicators. Additionally, it should generalize effectively to unseen data while minimizing the risk of overfitting. This project seeks to create a resilient predictive model by integrating GBM, LSTM, and attention mechanisms to enhance the precision and dependability of stock price predictions, offering meaningful insights for investors and financial analysts.

2.5 Objectives

1. Create a hybrid predictive model that integrates Geometric Brownian Motion, LSTM networks, and an Attention mechanism to improve the accuracy of stock price predictions.
2. Use historical stock data to simulate short-term price changes through GBM and identify long-term trends with LSTM networks.
3. Apply an attention mechanism to help the model focus on key data points that significantly influence future stock movements.
4. Evaluate the model using metrics such as MAE, RMSE, and R-squared to ensure high reliability and predictive performance.
5. Develop a data-driven tool that enables informed investment decisions by delivering precise, real-time stock price predictions for analysts and investors.

Chapter 3

Proposed System

3.1 Present Report on Investigation

This study focuses on developing a hybrid predictive model aimed at enhancing the accuracy of stock price forecasting in increasingly volatile financial markets. Traditional models such as ARIMA or linear regression often fail to account for the complexity and non-linear behaviour of real stock price data. They lack the flexibility to adapt to sudden market changes and typically assume fixed patterns that don't reflect actual conditions. These limitations make them less effective in modern financial environments where data is dynamic and influenced by multiple unpredictable factors.

To address these issues, the proposed model integrates Geometric Brownian Motion (GBM), Long Short-Term Memory (LSTM) networks, and an Attention mechanism. GBM simulates the randomness of short-term price movements using stochastic principles. However, since it assumes constant volatility and drift, it lacks adaptability to real market behaviour. LSTM networks are introduced to capture long-term dependencies in sequential data and help the model learn from historical patterns. The addition of an Attention mechanism further enhances the model by allowing it to focus on key data points that are more likely to influence future trends, improving accuracy in volatile or news-sensitive scenarios.

The hybrid model is tested using historical stock data and assessed with metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared values. These metrics provide a measure of the model's performance in real-world forecasting scenarios. By integrating stochastic modeling with deep learning, the study aims to develop a more adaptive, accurate, and practical tool for investors and analysts to make data-driven decisions.

3.2 Architecture of the System

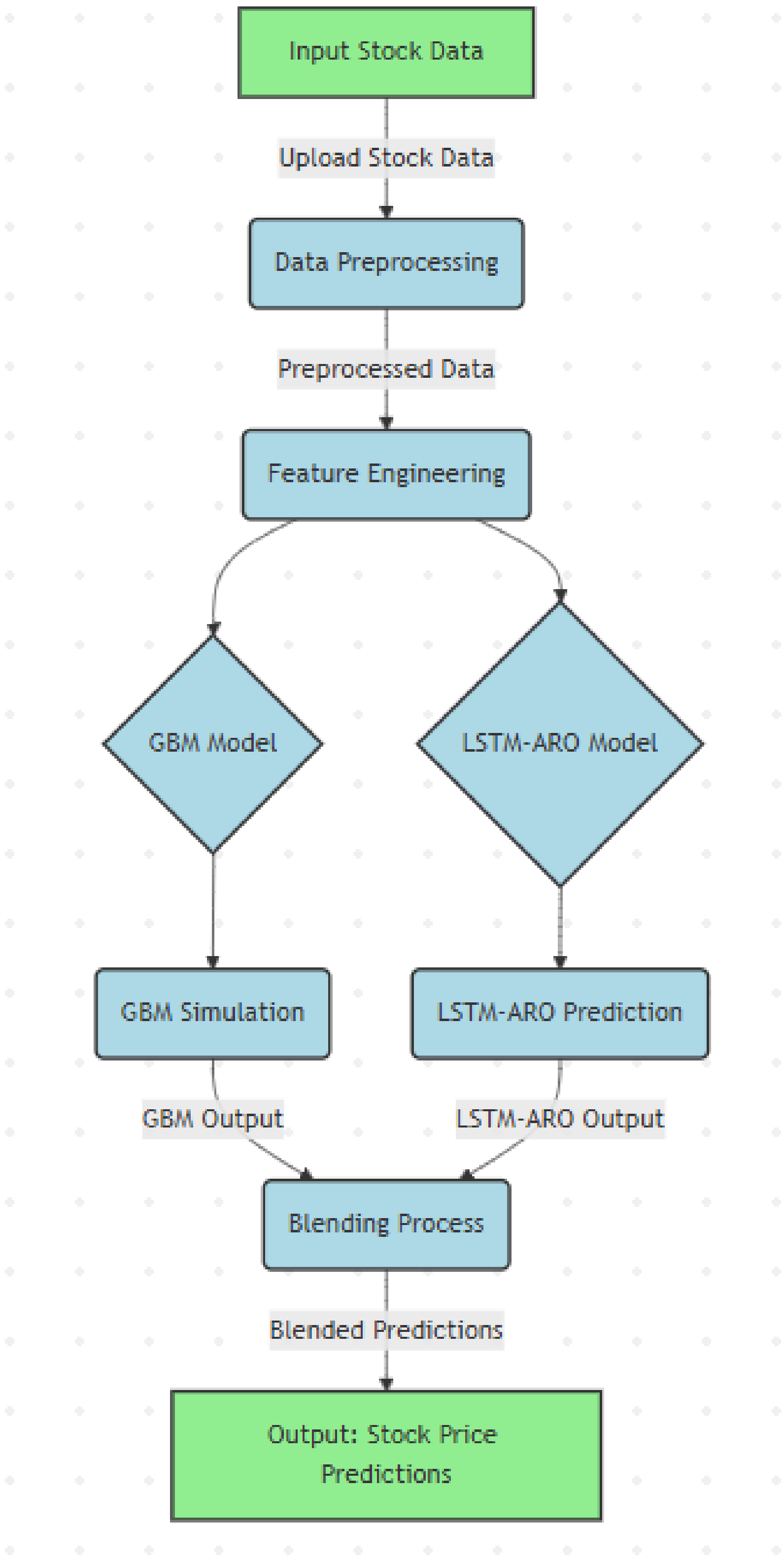


Figure 3.1: Architecture of Stock Price Predictor

This diagram illustrates the architecture of a stock price prediction system that leverages a blended approach, combining two different modelling techniques: Geometric Brownian Motion (GBM) and Long Short-Term Memory with Attention Recurrent Optimization.

Components and Flow:

1. **Input Stock Data (Green):**
   - This is the starting point. It represents the raw historical stock data that serves as the foundation for the predictions.
   - This data typically includes time-series information like dates, opening prices, closing prices, high/low prices, and volumes of stocks traded.
   - The "Upload Stock Data" label indicates the action of bringing this data into the system.
2. **Data Preprocessing (Blue):**
   - This stage involves cleaning and preparing the raw stock data for analysis.
   - Common preprocessing steps include handling missing values, removing outliers, normalizing or standardizing the data, and transforming it into a format suitable for the models.
   - "Pre-processed Data" label shows the output of this step.
3. **Feature Engineering (Blue):**
   - In this step, new features are derived from the existing data to improve the model's predictive performance.
   - This could involve calculating technical indicators (e.g., moving averages, RSI, MACD), creating lagged variables, or generating other relevant features.
   - These features enhance the models ability to find patterns.
4. **GBM Model (Blue Diamond):**
   - This represents the Geometric Brownian Motion model, a statistical model often used to simulate stock price movements.
   - It's based on the assumption that stock prices follow a random walk with a drift and volatility component.
   - The diamond shape of the node, signifies that this is a model being used in the process.
5. **LSTM-ARO Model (Blue Diamond):**
   - This represents a more complex model, combining Long Short-Term Memory

(LSTM) neural networks with an Attention Recurrent Optimization.

- LSTM networks are well-suited for time-series forecasting, and the ARO component helps capture the complex patterns in the data.
- The diamond shape of the node, signifies that this is a model being used in the process.

6. **GBM Simulation (Blue):**
   - The GBM model is used to generate multiple possible future price paths based on the historical data and model parameters.
   - "GBM Output" represents the data coming out of the simulations.
7. **LSTM-ARO Prediction (Blue):**
   - The LSTM-ARO model is used to generate a forecast of future stock prices.
   - "LSTM-ARO Output" represents the predicted data.
8. **Blending Process (Blue):**
   - The outputs of the GBM simulation and the LSTM-ARO prediction are combined in this step.
   - This blending could involve averaging the predictions, using a weighted average, or employing a more sophisticated blending technique.
   - The goal is to leverage the strengths of both models to produce a more accurate and robust prediction.
   - "Blended Predictions" are the result of this process.
9. **Output: Stock Price Predictions (Green):**
   - This is the final output of the system, representing the blended stock price predictions.
   - These predictions can be used for decision-making related to stock trading or investment.

Chapter 4

Design and Methodology

4.1 Design details

Users upload prescriptions to the mobile app, which then undergo OCR for text extraction. An entity extraction model detects drug names, which are subsequently used by the recommending module to identify generic alternatives

4.1.1 Activity Diagram

Activity diagrams show the flow of control or object flow by emphasizing the conditions and sequence of the flow. An activity diagram adds support for parallel execution and, like a basic flow chart, shows the sequence of activities. It also allows condition behaviour. It illustrates the process from the beginning to the end, highlighting the various options for decision-making along the sequence of events in the activity.

1. **Start:** Each activity diagram has one start at which the sequence of actions begins

2. **End:** Each activity diagram has one finish lines at which the sequence of action end

3. **Activity:** Activities are included using capsule symbol. Activities are connected together by transitions the transmission are directed arrows following from previous activity to the next activity.

4. **Condition:** To show conditional behaviour we use a branch and a merge.

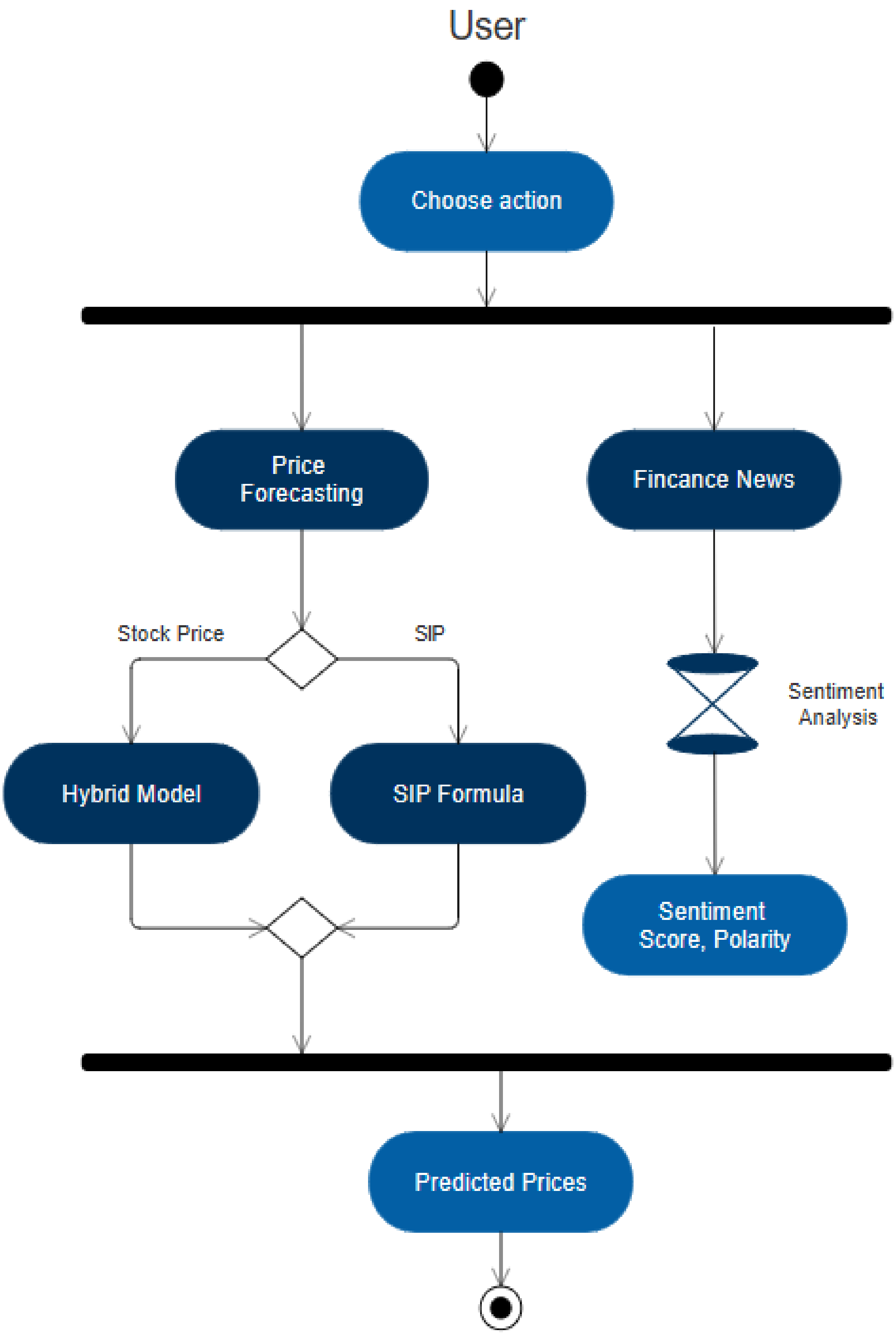


Figure 4.1: Activity Diagram of Stock Price Predictor

4.1.2 Class Diagram

A class diagram is a basic representation of the structure of an object-oriented system, displaying a set of classes, interfaces, and their interrelationships. It is a frequently used modelling tool that provides a static representation of a system's components. Visualizing the dependencies between classes provides crucial insights into the system's architecture.

Each class in the picture contains both properties and methods, constituting the foundation of object-oriented modelling. This diagram is essential for both conceptual and detailed modelling, as it allows models to be translated into executable code. Classes are represented as boxes, with the middle area containing attributes and the bottom section defining methods or operations.

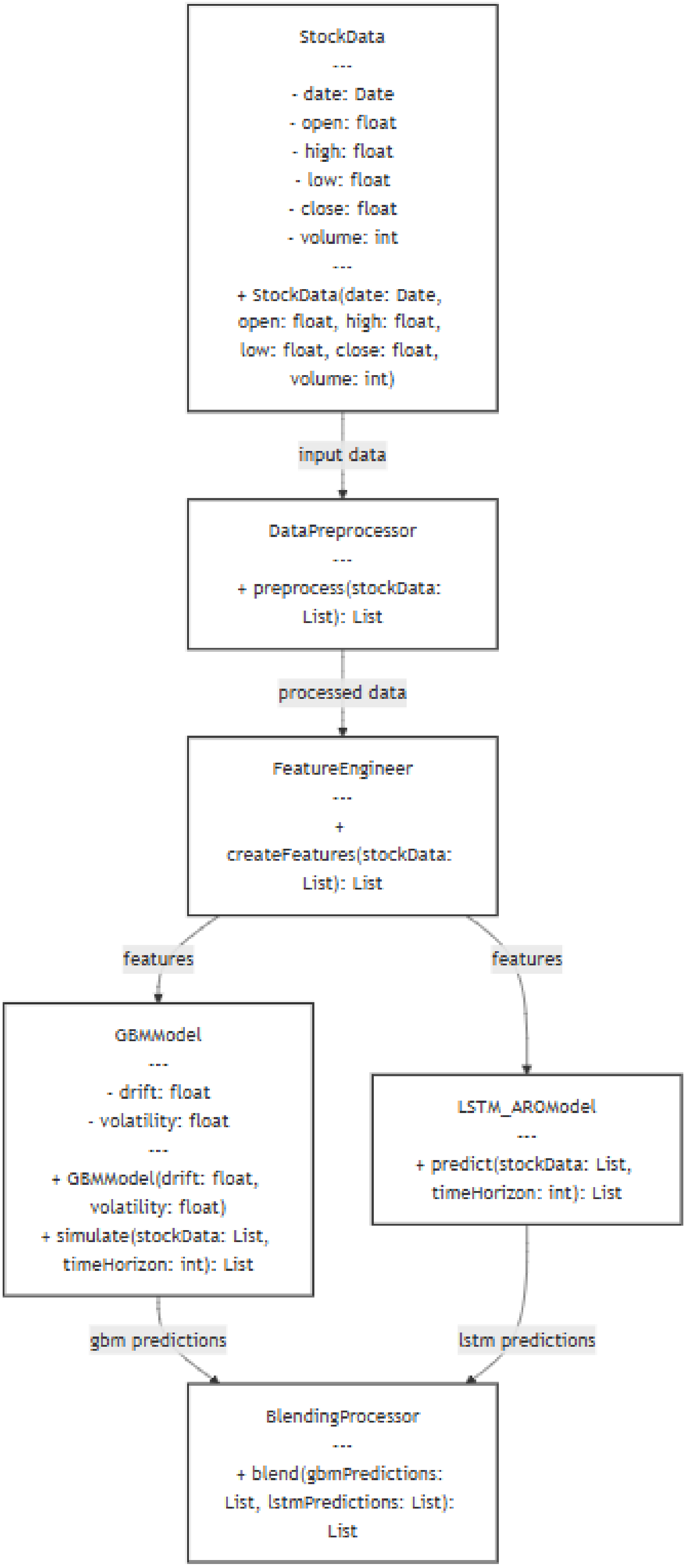


Figure 4.2: Class Diagram of Stock Predictor

4.1.3 Sequence Diagram

A sequence diagram is an interaction diagram that illustrates the interactions and temporal sequence of processes within a system. It offers a visual representation that highlights the chronological flow of communications between objects.

The sequence diagram is organized as a table, with objects on the x-axis and messages on the y-axis, indicating their chronological order. Concurrent processes or objects are represented by parallel vertical lines, while horizontal arrows indicate the messages passed between them in the order they occur. This graphical form allows for the explicit and understandable characterization of runtime scenarios.

Key symbols used in sequence diagrams as shown in Figure 4.3 include:

Message: Represented as arrows, these illustrate communication between objects.

1. **Class Roles:** Describing the behavioral context of objects.
2. **Lifeline:** Vertical dashed lines indicating the existence of an object over time.
3. **Loop:** Represented as rectangles, depicting repeating or iterative processes within the sequence.
4. **Activation:** Illustrating the time duration an object requires to complete a task.

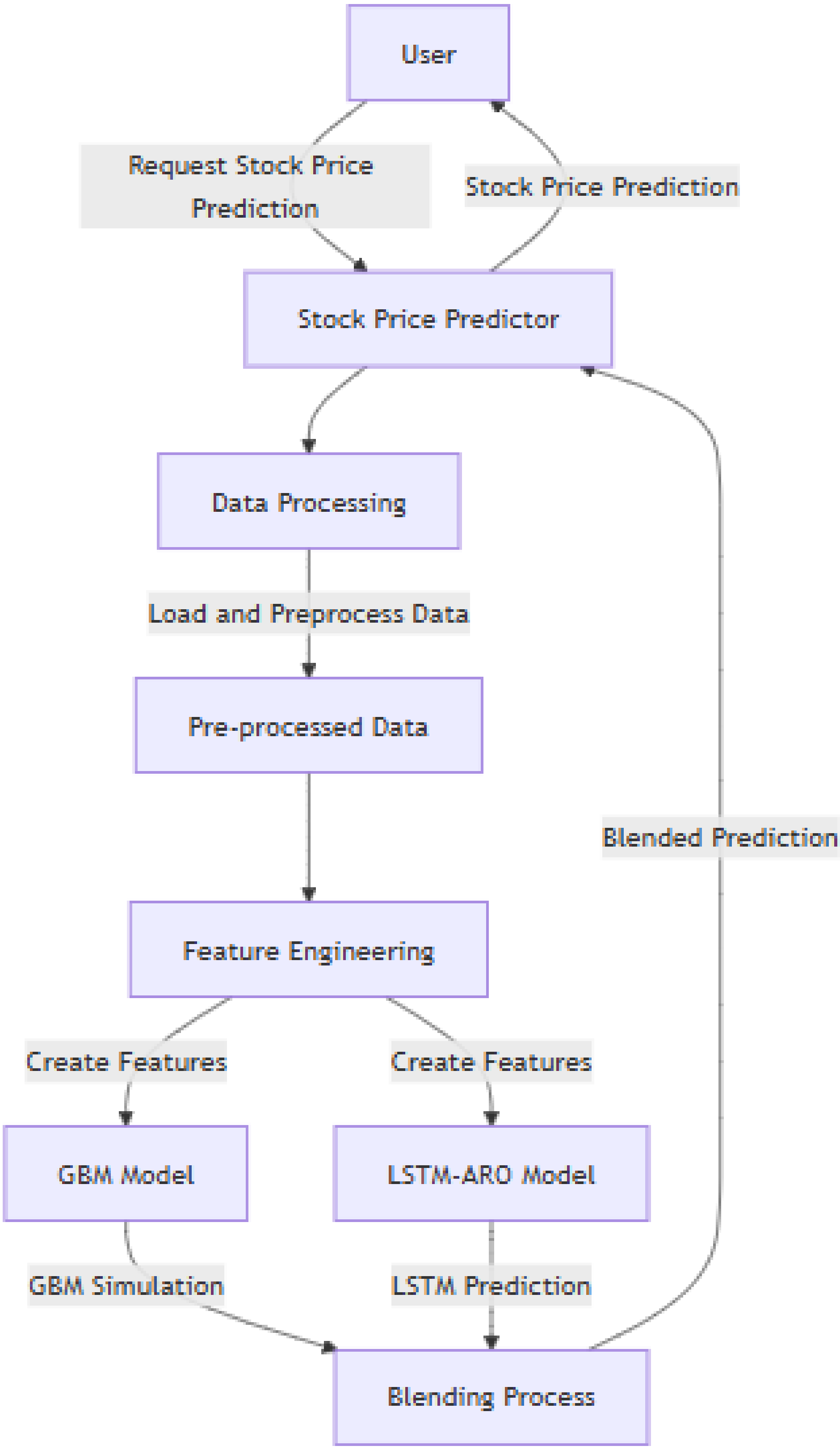


Figure 4.3: Sequence Diagram of Stock Price Predictor

4.2 Methodology

In today's fast-paced financial world, accurate stock price prediction is crucial. Investors and analysts rely on reliable forecasts to make smart decisions. But predicting prices is tough. The market is volatile, patterns are complex, and data can be noisy and unstructured. To tackle this, we created an innovative system – **Stock Price Prediction Using Geometric Brownian Motion and LSTM-ARO**. This hybrid approach improves accuracy by combining multiple advanced models and algorithms.

The system integrates three robust components: Geometric Brownian Motion (GBM), Long Short-Term Memory (LSTM) neural networks, and Attention Recurrent Optimization (ARO). Each model captures a different aspect of market behaviour. GBM handles randomness and trends. LSTM tracks long-term dependencies in time-series data. ARO improves prediction by learning from past errors. Together, they form a strong and balanced forecasting framework.

The process begins with users uploading raw stock data. This includes prices like open, close, high, low, and volume. The system first runs the data through a **preprocessing pipeline**. This stage removes missing values, standardizes data, and formats it for analysis. Next, we perform **feature engineering**. Here, we extract useful indicators such as moving averages, volatility measures, and momentum factors. These features enhance the quality of input for the models.

After feature extraction, the data is sent into two parallel paths. The **GBM model** simulates future prices using math-based equations. It provides likely price paths based on volatility and drift. At the same time, the **LSTM-ARO module** processes the data through a deep learning network. LSTM captures patterns over time. ARO further refines these outputs by correcting errors the model might make.

One of the key features of our system is the **blending process**. It combines predictions from both GBM and LSTM-ARO. This blending ensures that the final result takes advantage of both approaches. It balances the outputs and produces a more reliable forecast. Users get a clear, concise, and actionable prediction.

We also built the system with a modular code design. Each task—like data handling, modelling, and output—has its own class. This makes the system scalable, easy to modify, and efficient. Developers can plug in different models or tweak parameters without breaking the flow.

Our project is more than a stock predictor. It is a **complete forecasting system**. It merges theory from stochastic finance with deep learning and error optimization. The result is a tool that makes market prediction smarter and more robust. By combining GBM, LSTM, and ARO, we offer a unique solution that delivers precision and value in the complex world of stock markets.

4.3 Algorithm Implementation

The project implemented a hybrid approach to stock price prediction by combining three distinct algorithmic components: Geometric Brownian Motion (GBM) Simulation, Long Short-Term Memory with Attention Recurrent Optimization (LSTM-ARO) Prediction, and a Blending Process.

4.3.1 GBM Simulation (Geometric Brownian Motion):

The primary objective of implementing the GBM simulation was to generate a comprehensive range of potential future stock price trajectories, grounded in the statistical characteristics observed within historical data. Initially, the project focused on the meticulous calculation of the drift and volatility parameters, which are fundamental to the GBM model. The drift, representing the average return, and the volatility, quantifying the standard deviation of returns, were derived directly from the provided historical stock price data. Statistical methods, specifically maximum likelihood estimation, were employed to ensure the accuracy and robustness of these parameter estimates. Subsequently, the simulation process was initiated, guided by a predefined time horizon, tailored to meet the specific forecasting needs of the project.

Within this simulation loop, a large number of iterations were executed to generate a diverse set of potential price paths. In each iteration, random increments were generated, drawing from a normal distribution characterized by the previously calculated drift and volatility. The stock price was then updated at each time step, utilizing the GBM formula, which incorporated these randomly generated increments. All simulated price paths were meticulously stored, forming a comprehensive dataset of potential future price trajectories. The culmination of this simulation process was a rich collection of possible future price trajectories, encapsulating the inherent randomness and volatility of the stock market.

These trajectories served as a foundational component for the subsequent blending process, contributing to a more nuanced and accurate final prediction. The project recognized the importance of a robust simulation framework to capture the stochastic nature of stock price movements, thereby providing a more reliable basis for forecasting.

The Geometric Brownian Motion (GBM) is used to simulate future stock prices based on stochastic processes. It assumes continuous compounding and follows the equation:

$$S(t) = S(0).e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

Where:

- S(t) is the predicted stock price at time t,
- μ is the expected return (drift),
- σ is the volatility,
- W_t is a Wiener process

4.3.2 LSTM-ARO Prediction (Long Short-Term Memory with Attention Recurrent Optimization):

The LSTM-ARO model was designed to capture both temporal dependencies and important patterns in the data. The architecture includes:

- Input Layer: Accepts a 3D tensor of shape (timesteps, features), where timesteps = 60 and features = 1.
- LSTM Layer: Extracts temporal features by processing sequential data with 50 units and outputs a sequence.
- Attention Mechanism: Computes attention scores to focus on relevant historical information, improving interpretability and feature importance.
- Combination Layer: Concatenates the outputs of the LSTM layer and the Attention mechanism for further processing.
- Output Layer: A dense layer outputs the predicted stock price.

The model was compiled with the Adam optimizer and the Mean Squared Error (MSE) loss function, and trained for 20 epochs with a batch size of 32.

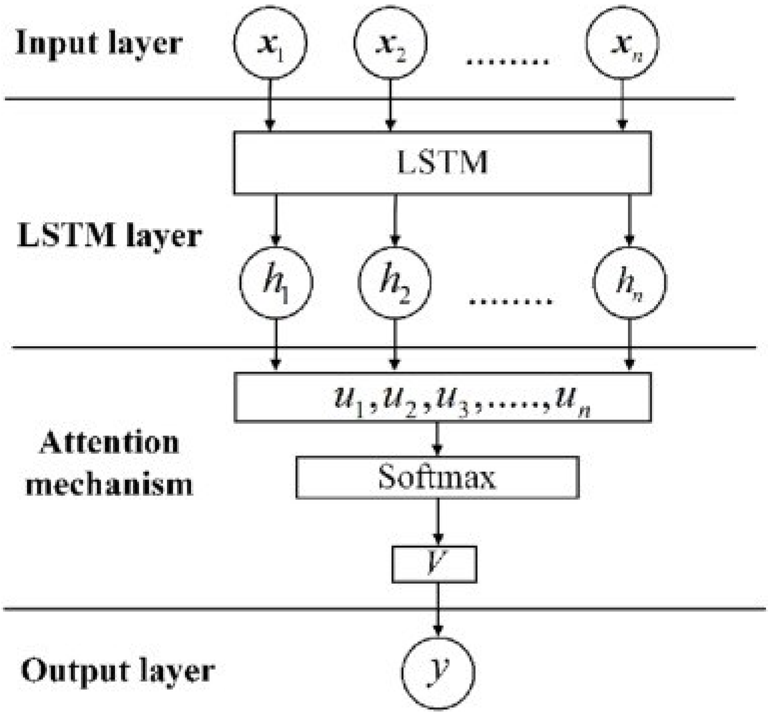


Figure 4.4: LSTM cell structure

4.3.3 Blending Process:

The final stage of the algorithm implementation involved the integration of the outputs from the GBM simulation and the LSTM-ARO prediction. The objective was to produce a more robust and accurate forecast by capitalizing on the strengths of both models. To leverage the strengths of both approaches, a blended prediction was generated by combining the outputs of the LSTM-ARO model and the GBM simulation.

$$P_{\text{blend}} = \alpha \cdot P_{\text{LSTM}} + (1 - \alpha) \cdot P_{\text{GBM}}$$

where:

- P_{blend}: Blended prediction
- P_{LSTM}: Prediction from the LSTM-ARO model
- P_{GBM}: Prediction from the GBM model
- $\alpha = 0.7$: Weight assigned to LSTM-ARO predictions

4.4 Details of Hardware and Software

Hardware Requirements:

1. Processor: The system should have at least an Intel Core i5 or AMD Ryzen 5 processor to efficiently handle model training and data processing.
2. RAM: A minimum of 8 GB RAM is required, though 16 GB is recommended for smoother performance during model training.
3. Storage: At least 500 GB of hard disk space is required, preferably a 256 GB SSD for faster data read/write operations.
4. Graphics Card: While an integrated GPU is sufficient, a dedicated NVIDIA GPU is recommended for faster LSTM training.
5. Display: A display of 13 inches or larger with Full HD resolution is suggested for better visualization of plots and dashboards.
6. Internet Connectivity: A stable internet connection is needed to fetch real-time stock data and access online resources.
7. Power Backup: An uninterrupted power supply (UPS) is recommended to avoid data loss during long training sessions.

Software Requirements

1. Operating System: The project can be run on Windows 10/11, Ubuntu 20.04 or higher, or macOS.
2. Programming Language: Python 3.8 or above is required for implementing all models and data operations.
3. IDE: Tools like Jupyter Notebook, VS Code, or PyCharm can be used for writing and testing code.
4. Libraries and Frameworks:
    - NumPy for numerical computations.
    - Pandas for data manipulation and analysis.
    - Matplotlib and Seaborn for data visualization.
    - Scikit-learn for preprocessing and the GBM model.
    - TensorFlow/Keras for implementing the LSTM model.
    - Custom ARO (Attention Layer) logic implementation

5. Database: Data can be managed using CSV files, in-memory Pandas Data Frames, or a lightweight database like SQLite.
6. Version Control: Git and GitHub should be used for tracking changes and collaborating.

Chapter 5

Result and Discussions

5.1 Implementation

We have developed a prototype of the project with its core functionality and a user walkthrough is given below for the website.

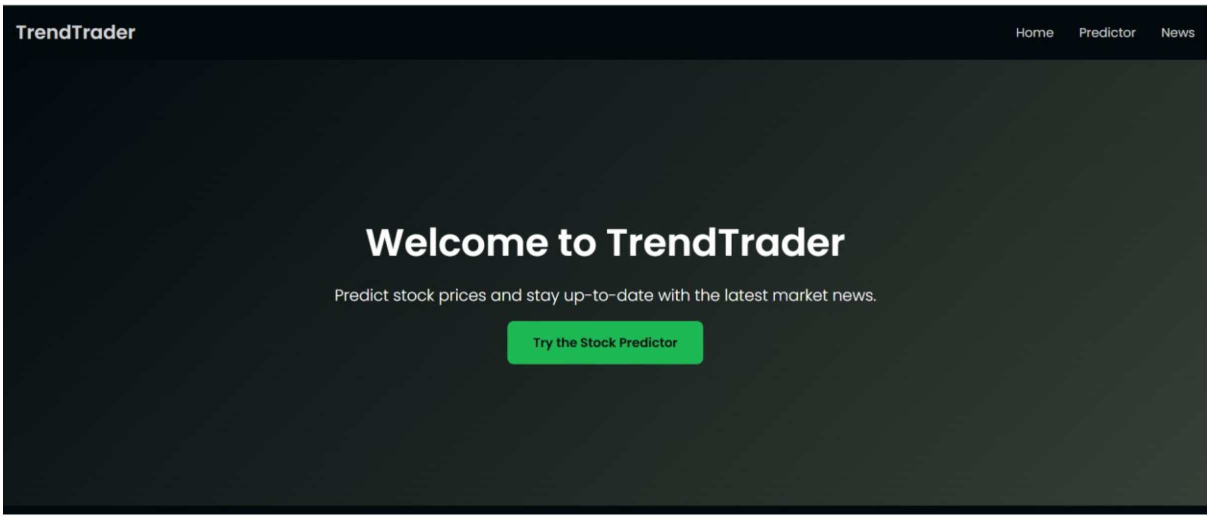


Figure 5.1: Trend Trader Home page

Figure 5.1 shows the homepage of TrendTrader which features a clean, minimalistic design with a dark gradient background, introducing the project and its core functionality. It includes a navigation bar and a central call-to-action button that directs users to the stock prediction feature.

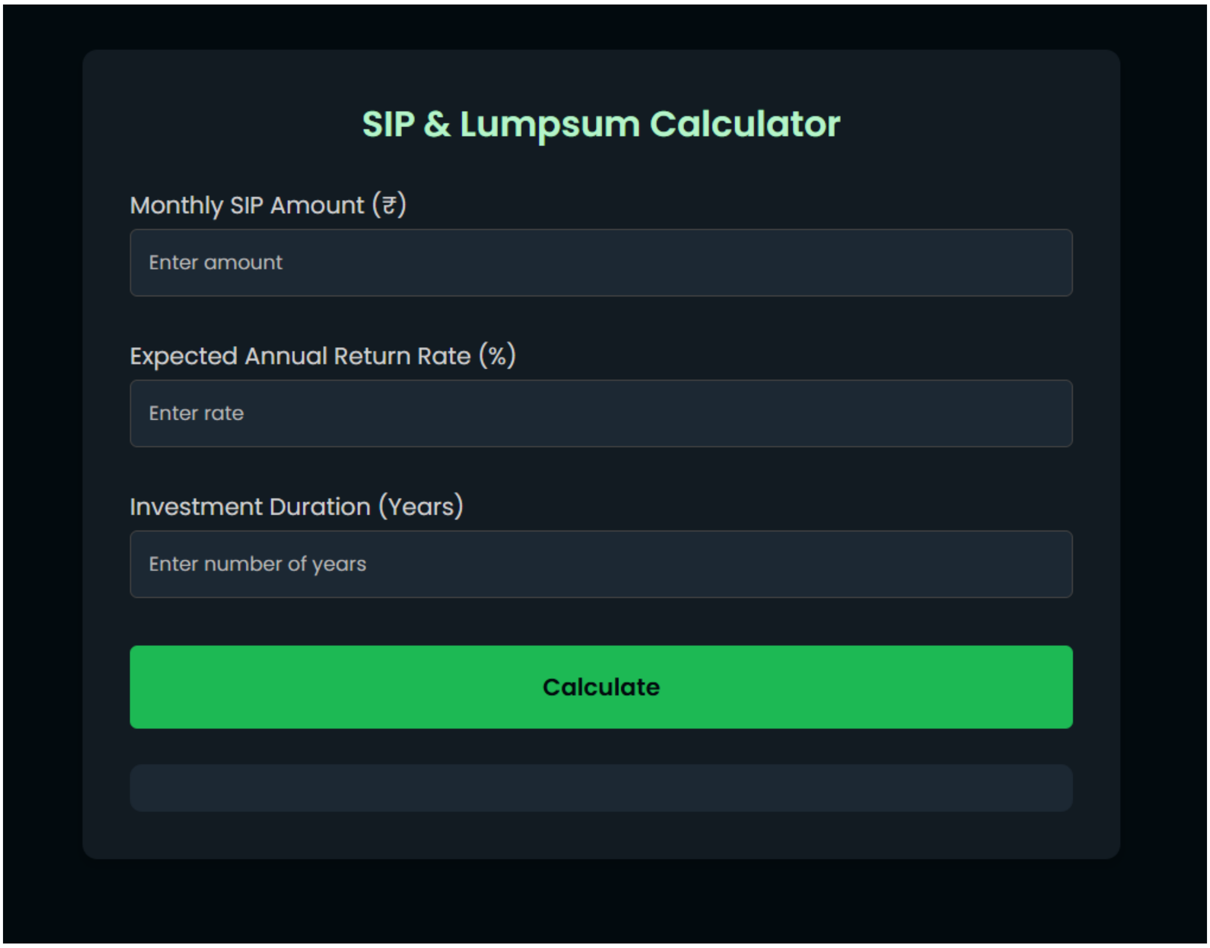


Figure 5.2: Trend Trader SIP calculator

This figure 5.2 features a **SIP & Lumpsum Calculator** where users can input investment details to estimate future returns. It includes fields for SIP amount, expected return rate, and investment duration, along with a calculate button for instant results.

Mutual Funds Returns

Investment Type	Expected Annual Return (%)
Gold	7-10%
Silver	8-12%
Mid Cap Mutual Funds	12-18%
Large Cap Mutual Funds	10-15%
Small Cap Mutual Funds	15-22%
Flexi Cap Mutual Funds	10-18%
Government Debt Bonds	6-8%
Fixed Deposits (FD)	5-7%

Figure 5.3: Trend Trader Mutual Funds Information

Figure 5.3 presents a comparison table of various investment options and their expected annual returns. It helps users make informed decisions by showcasing return ranges for mutual funds, metals, bonds, and fixed deposits.

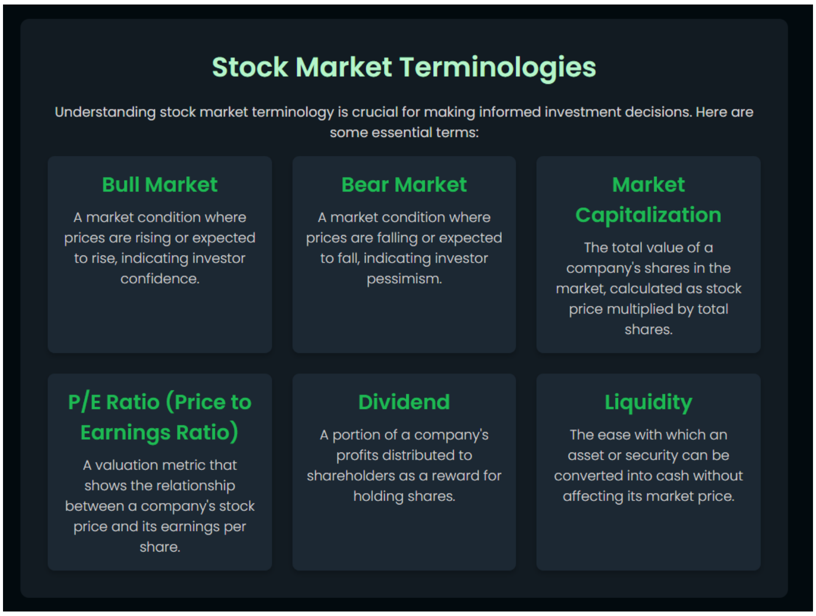


Figure 5.4: Trend Trader Stock Market Information

Figure 5.4 details some important and key terminologies related to stock market. It shows some basic information that will help beginners learn about investment.

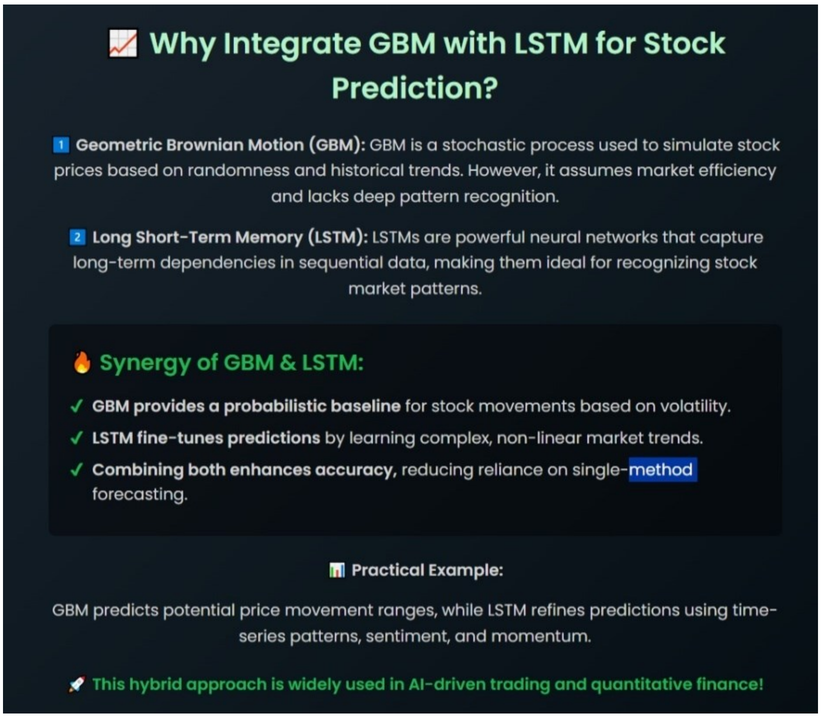


Figure 5.5: Trend Trader Model information

Figure 5.5 explains the integration of **GBM** with **LSTM networks** for more accurate stock prediction. It highlights how GBM provides a statistical baseline while LSTM enhances predictions by capturing complex market patterns, combining the strengths of both methods.

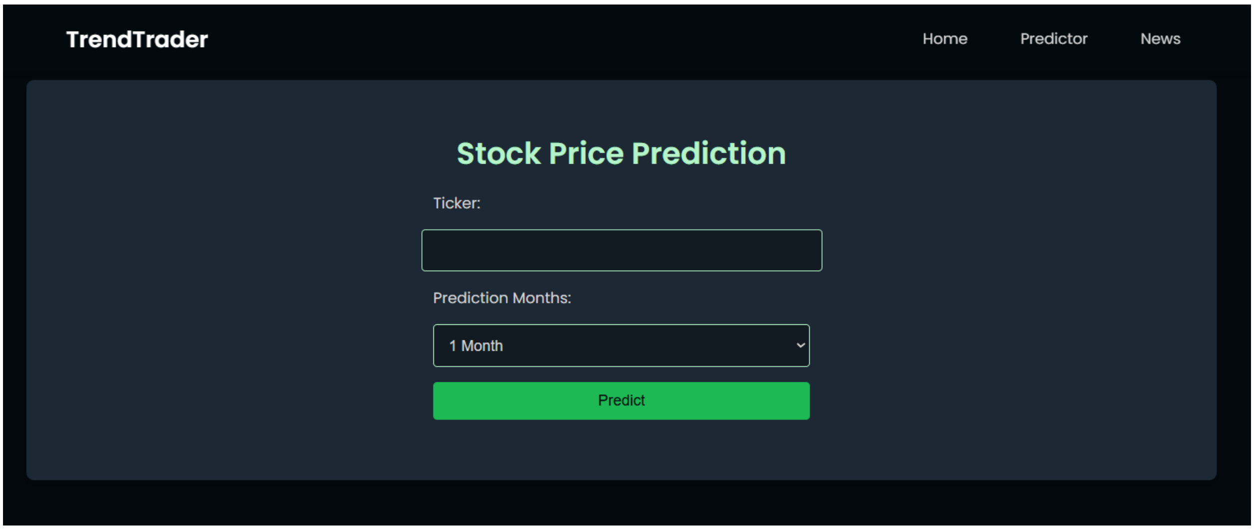


Figure 5.6: Trend Trader Predictor page

Figure 5.6 depicts the actual stock price predictor of the project. It uses a blended model of GBM and LSTM-ARO to accurately predict the stock prices for a set duration in future. The user can add any valid stock ticker and the duration of prediction and the model will provide a detailed graph of the same.

Figure 5.7: Trend Trader Prediction graph

This image depicts the result that will be provided after the user inputs stock ticker and duration as per their need. Here, the prediction of ITC stocks for some duration input by the user is shown.

5.2 Testing

After the development phase of the system, we used several edge cases on which we can test the system. We performed two types of testing: Unit Testing, which evaluates individual modules of the system, and System Testing, which assesses the functionality of the entire system as a whole. The test cases are given below:

5.2.1 Unit Testing

In order to ensure the correctness and effectiveness of individual components within our stock price prediction system, we carried out detailed unit testing on two primary models: the **Plain Geometric Brownian Motion (GBM)** model and the **Blended Model**, which integrates GBM with advanced machine learning mechanisms such as Long Short-Term Memory (LSTM) networks, attention mechanisms, and recurrent optimization techniques.
Unit testing focuses on evaluating each model independently in terms of its design objectives, mathematical implementation, and predictive behaviour. This step was essential before combining these models into a larger system for deployment.

Plain GBM Model:
The GBM model is one of the foundational tools in quantitative finance, often used to model asset price behaviour due to its stochastic and log-normally distributed nature. The goal of testing this model was to ensure the implementation was mathematically correct and that it could produce outputs consistent with theoretical expectations. The GBM serves as a baseline reference to measure the added value of more complex models later in the project.

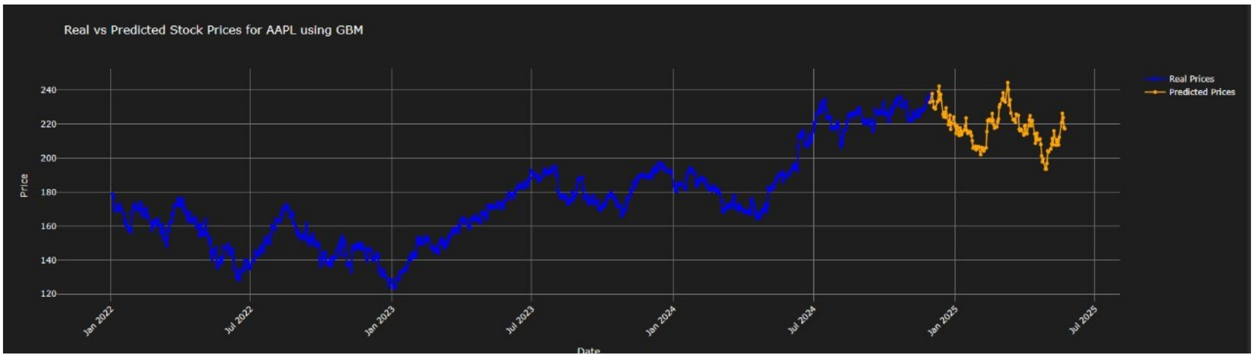


Figure 5.8: GBM Simulated Prices Over 30 Days

One such simulated path generated using the GBM model is shown in figure 5.8. This simulation was performed over a 30-day period, using historical parameters extracted from a major stock index.

The plot captures the essential behaviour of GBM: a smooth, continuous, and non-negative path that fluctuates according to the volatility factor. Multiple simulations were also run and aggregated to study the average and spread of outcomes.

Blended Model (GBM + LSTM + Attention + Recurrent Optimization)

The blended model integrates the theoretical strengths of GBM with the learning capabilities of advanced deep learning architectures. The objective of this unit testing phase was to validate the blended model's ability to improve predictive accuracy through adaptive learning and intelligent sequence modelling. Each component — GBM, LSTM, attention mechanism, and recurrent optimization — plays a specific role in enriching the model's understanding of market dynamics.

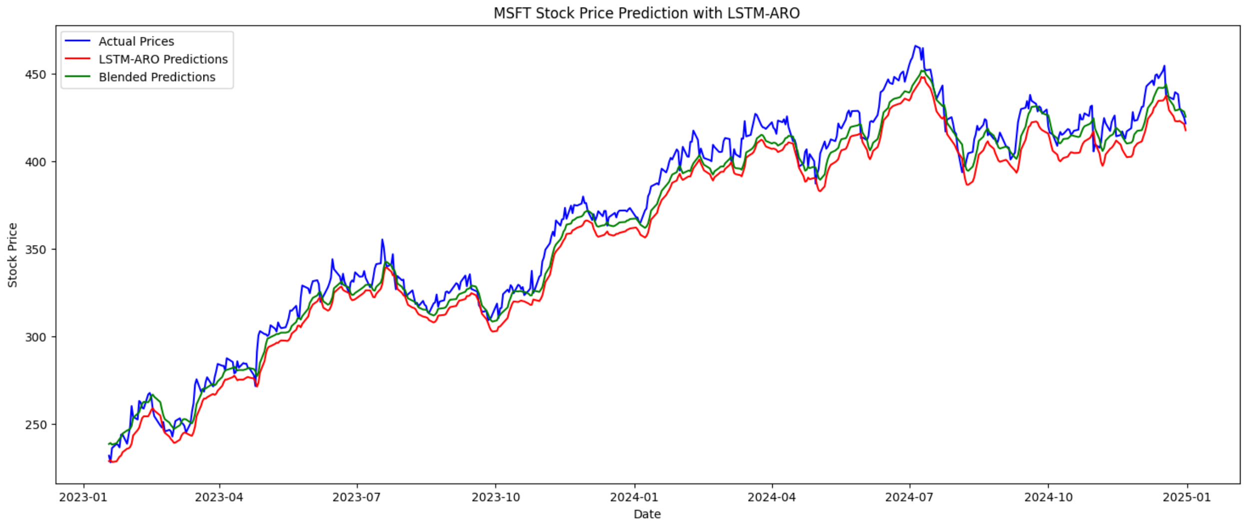


Figure 5.9: Blended model Trial-1

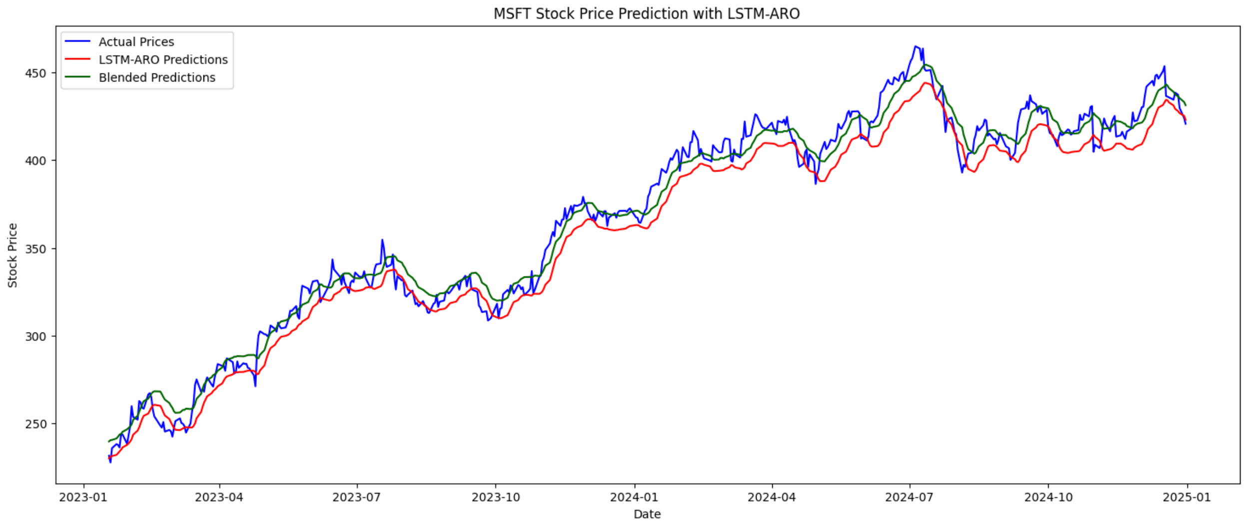


Figure 5.10: Blended model Trial-2

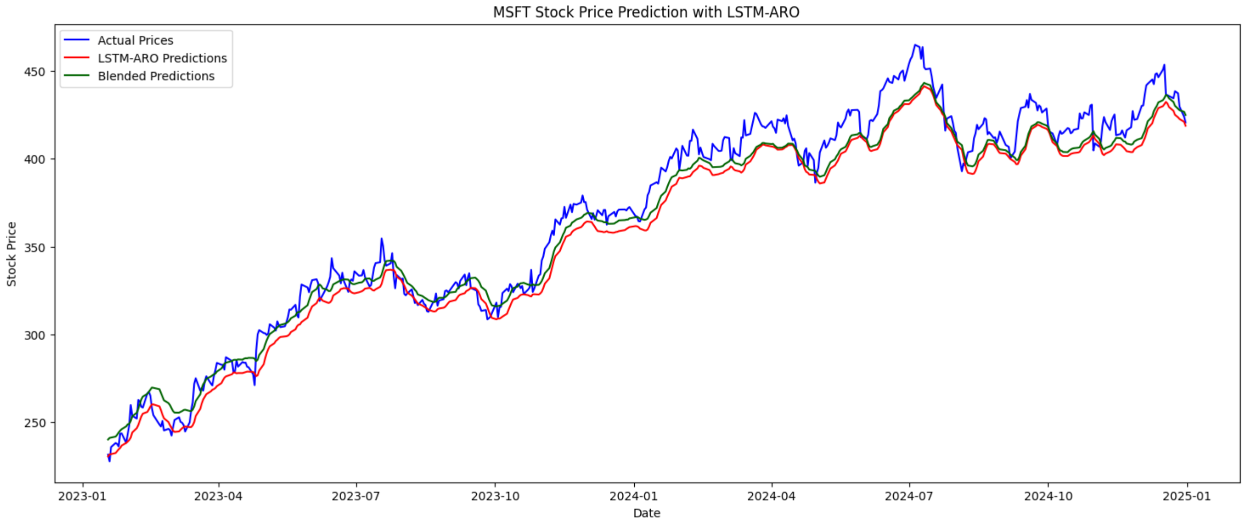


Figure 5.11: Blended model Trial-3

Below is a representative graph comparing predicted stock prices from the blended model against actual prices during the test period. The overlay demonstrates the model's ability to track both the direction and magnitude of real-world movements with high fidelity.

The predicted prices were significantly closer to the actual values than those from the GBM-only model, especially in terms of capturing volatility spikes and local maxima/minima.

5.2.2 System Testing

System testing was conducted to verify that all components of the web-based application worked together seamlessly to deliver the intended features. Unlike unit testing, which focuses on individual modules, system testing validates the behaviour of the application as a whole, including the frontend, backend, machine learning model integrations, and data flow between components. The goal was to ensure that the system performs accurately, consistently, and reliably in a real-world user environment.

The website was designed as a three-page interactive platform for stock-related utilities, each addressing a key aspect of user engagement in personal finance and investment decision-making. The system was tested across multiple devices and browsers to ensure usability, responsiveness, and accuracy.

SIP Calculator

The first page provides a Systematic Investment Plan (SIP) calculator that allows users to estimate the future value of regular monthly investments based on an expected return rate and investment duration.

The calculator accurately computed the maturity value using the standard SIP formula:

$$FV = P \times \frac{(1 + \mathrm{r})^n - 1}{r} \times (1 + \mathrm{r})$$

Where:

- **P** = Monthly investment
- **r** = Monthly interest rate (annual rate divided by 12)
- **n** = Total number of months

All calculations were cross-verified with external SIP tools, and output consistency was maintained across different input combinations. The interface was responsive and displayed clear, concise results.

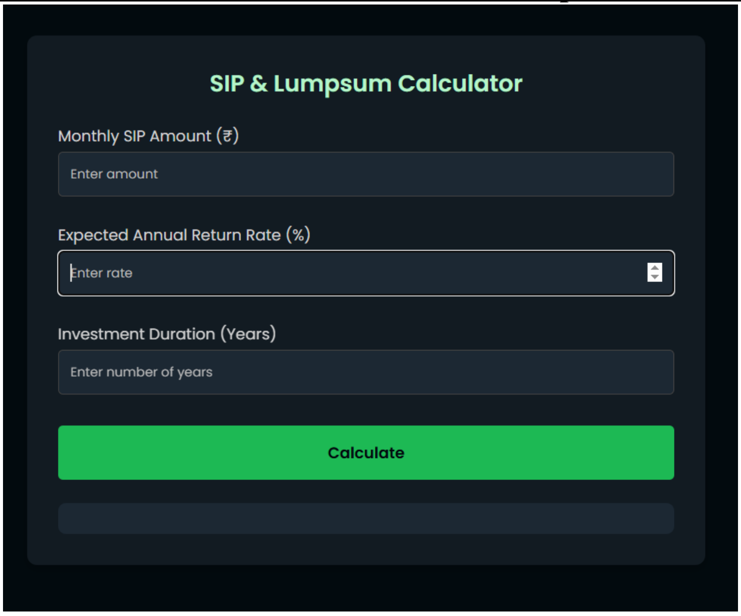


Figure 5.12: SIP Calculator

Stock Price Prediction Interface

The second page forms the core of the system, allowing users to input a stock ticker and duration (e.g., 1, 3, 6 months) to view predicted stock price trajectories. This feature is powered by the previously developed blended model (GBM + LSTM + Attention + Recurrent Optimization).

Testing Results:

- The page successfully retrieved historical data for multiple stock tickers and generated predictions within a few seconds.
- Prediction charts were visually accurate, with distinguishable trends and tooltips for daily values.
- Error messages and fallback prompts were displayed appropriately for invalid inputs or unavailable data.
- The system handled concurrent requests well, ensuring scalability.

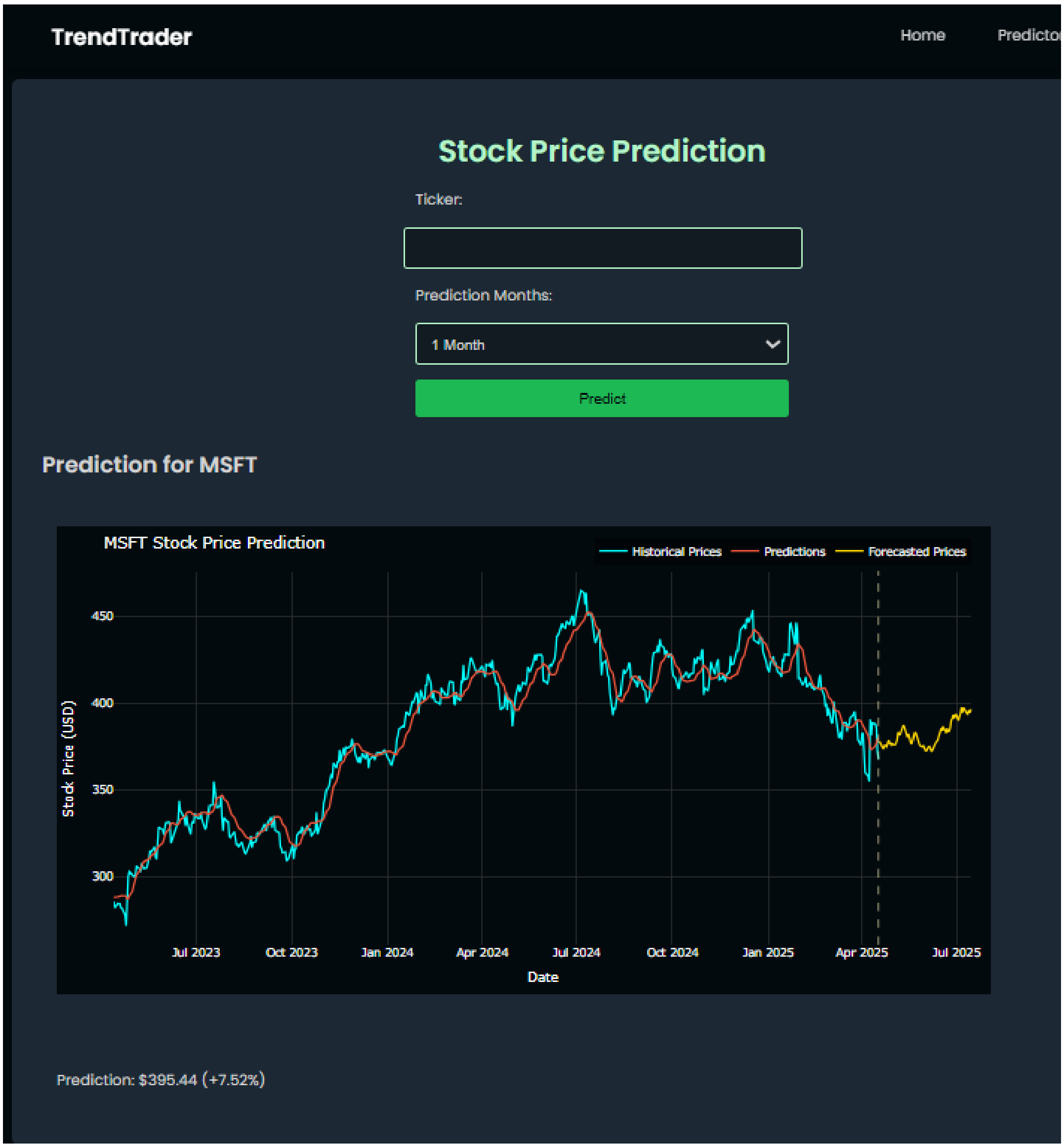


Figure 5.13: Prediction for MSFT for 3 months

News Sentiment Analysis

The third page adds a **sentiment analysis module** that pulls the most recent news articles related to a given stock and calculates a sentiment score using natural language processing. This component adds context to stock movements by interpreting market mood.

Testing Results:

- The page accurately displayed news headlines relevant to the entered stock ticker.
- Sentiment scores corresponded well with the nature of the headlines, showing realistic polarity.
- The system handled high-frequency queries without lag, and returned meaningful results even for mid-cap or lesser-known stocks.
- User experience was enhanced through visually engaging sentiment indicators and clickable news links for deeper reading.

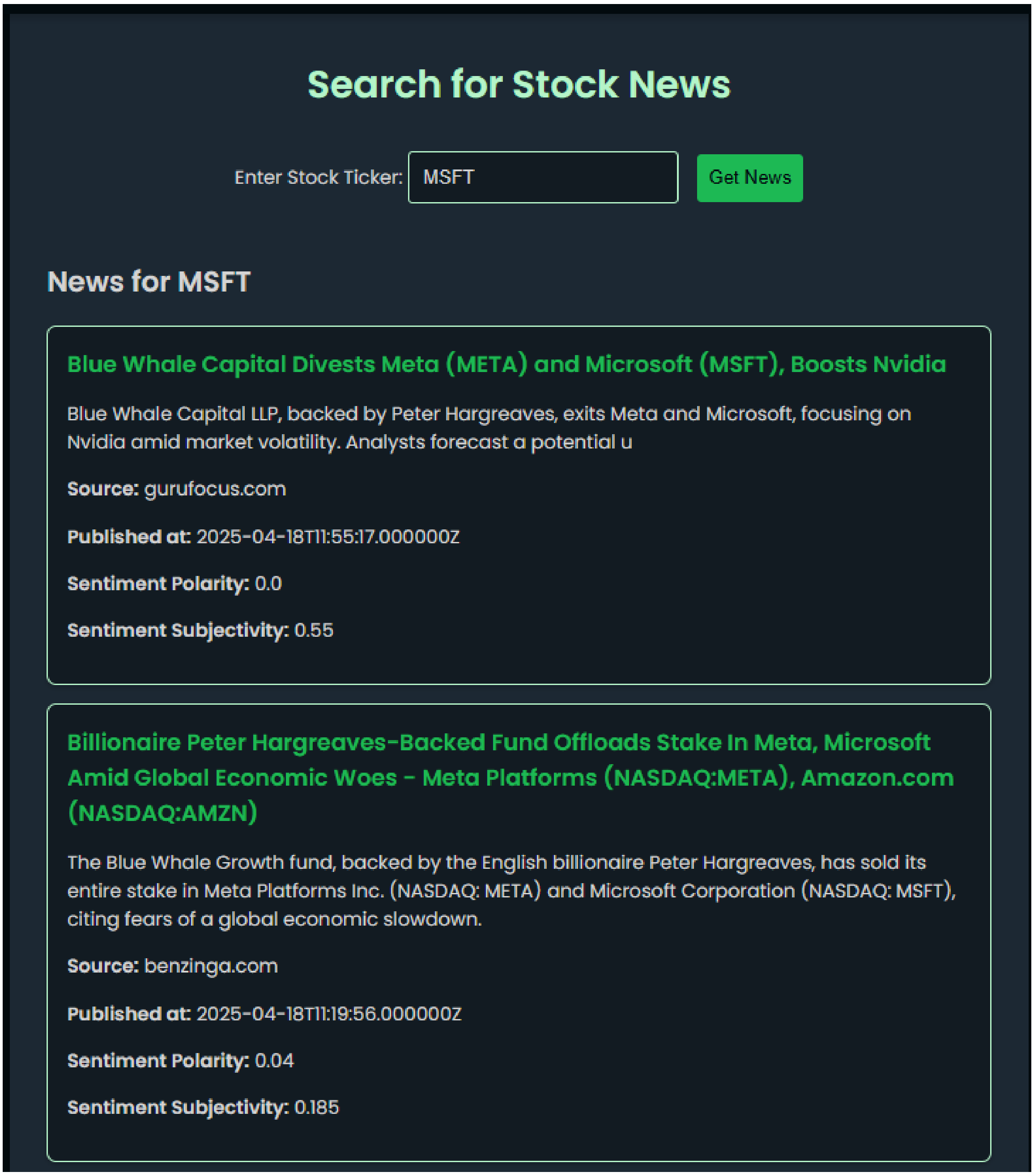


Figure 5.14: News sentiment analysis

5.3 Results and Discussion

5.3.1 Prediction Model Evaluation

The proposed framework was evaluated using three key performance metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2). These metrics provide quantitative insights into the predictive accuracy and goodness-of-fit of the model.

1) Mean Squared Error (MSE)

The Mean Squared Error (MSE) measures the average squared difference between the actual and predicted stock prices:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{y}_i - \hat{\mathrm{y}}_i)^2$$

Where:

- n: Total number of predictions
- y_i: Actual stock price at time i
- $\hat{y}_i$: Predicted stock price at time i

A lower MSE value indicates higher prediction accuracy, as it means the predictions are closer to the actual values.

2) Root Mean Squared Error (RMSE)

The Root Mean Squared Error (RMSE) is the square root of MSE. It provides an interpretable measure of the prediction error in the same unit as the stock price:

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$$

A lower RMSE signifies better model performance and lower average error in predicted stock prices.

3) Coefficient of Determination (R^2)

The R^2 score evaluates how well the predicted values match the actual stock prices. It is defined as:

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(\mathrm{yi} - \mathrm{y^\wedge i})^2}{\sum_{i=1}^{n}(\mathrm{yi} - \bar{\mathrm{y}})^2}$$

Where:

- $\bar{y}$: Mean of actual stock prices
- y_i : Actual value
- $\hat{y}_i$: Predicted Value

An R^2 value closer to 1 indicates a better fit between the predicted and actual values. An R^2 of 1.0 means perfect prediction, while an R^2 of 0 indicates that the model does no better than predicting the mean.

The proposed framework was evaluated on historical stock price data of Microsoft Corporation (MSFT) covering the period from January 2015 to January 2025. The evaluation focused on prediction accuracy using the following metrics:

- Mean Squared Error (MSE)
- Root Mean Squared Error (RMSE)
- Coefficient of Determination (R^2)

In addition to these quantitative metrics, the model's prediction trends were visually inspected to assess its capability to follow real-world price patterns.

Performance Comparison

The experiments were conducted across five independent trials for each model — LSTM with Attention Recurrent Optimization (LSTM-ARO), Geometric Brownian Motion (GBM), and the proposed Blended Model. The results are summarized in the tables below.

Trial	MSE	RMSE	R^2
1	143.4799	11.9783	0.9604
2	202.2007	14.2197	0.9442
3	113.5732	10.6570	0.9686
4	165.2960	12.8567	0.9544
5	178.3179	13.3535	0.9508

Table 5.1: LSTM-ARO Model – MSE, RMSE, and R^2

Trial	MSE	RMSE	R^2
1	12566.0964	112.0986	0.2671
2	13987.4532	118.2685	0.1842
3	12336.6472	111.0704	0.2787
4	14586.8631	120.7760	0.1492
5	7456.0800	86.3485	0.5651

Table 5.2: GBM Model – MSE, RMSE, and R^2

Trial	MSE	RMSE	R^2
1	79.5030	8.9164	0.9781
2	109.6401	10.4709	0.9697
3	97.1286	9.8553	0.9732
4	75.7537	8.7036	0.9791
5	71.0920	8.4316	0.9804

Table 5.3: Blended Model – MSE, RMSE, and R^2

Discussion

The results clearly indicate that the **blended model** consistently outperforms the standalone LSTM-ARO and GBM models across all evaluation metrics:

- **LSTM-ARO Model**:
  - Demonstrated strong performance in short-term predictions.
  - Achieved relatively low MSE and high R^2 values.
  - However, its performance was less stable under high-volatility conditions.
- **GBM Model**:
  - Captured random walk behaviour effectively.
  - Exhibited significantly higher MSE and lower R^2, especially in long-term trend

predictions.

- Its simplistic stochastic nature limited its predictive precision.

- **Blended Model**:
  - Combined the **stochastic modelling power of GBM** with the **temporal pattern recognition of LSTM**.
  - Achieved the **lowest average MSE and RMSE**, and the **highest R^2**, highlighting its robustness.
  - Offered better generalization and reduced overfitting compared to other models.

The consistent superiority of the blended approach demonstrates the benefits of fusing deterministic and stochastic forecasting strategies in stock price prediction tasks. The model not only reduces prediction error but also provides more reliable insights for decision-making.

Chapter 6

Conclusion and Future scope

This project successfully developed a hybrid stock price prediction system that integrated Geometric Brownian Motion (GBM) simulation, Long Short-Term Memory with Attention Recurrent Optimization (LSTM-ARO) prediction, and a sophisticated blending process. By combining these methodologies, the system effectively captured both the stochastic nature and complex temporal patterns of financial data, leading to robust and nuanced forecasts. The GBM simulation provided a statistical foundation for understanding potential price trajectories, while the LSTM-ARO model leveraged deep learning to learn intricate patterns and generate accurate point forecasts. The blending process, a critical component, optimized the final prediction by intelligently weighing the contributions of each model. Confidence intervals and visualizations further enhanced the interpretability of the results, empowering users with informed decision-making tools.

This project's successful creation of a hybrid stock price prediction system, combining GBM, LSTM-ARO, and blending techniques, sets the stage for significant future advancements. Key areas of development include: enhancing prediction accuracy by incorporating broader data sources like news sentiment and economic indicators [6][7], refining the blending algorithm with machine learning for dynamic adjustments, and optimizing the system for high-frequency trading. Furthermore, exploring advanced deep learning models like reinforcement learning [9] and Generative Adversarial Networks (GANs) [12] expanding its application to diverse financial assets, and developing user-friendly interfaces and APIs will broaden its reach. Finally, integrating explainable AI techniques will increase user trust by providing transparency into prediction factors. This project establishes a robust foundation for future innovations in reliable and sophisticated financial forecasting.

Bibliography

[1] Anay Kumar, Isa Jamadar, Radhika Goel, Ram Charan Petluri, And Wei Feng "Mathematically Forecasting Stock Prices with Geometric Brownian Motion", 2024

[2] Oluwarotimi Israel Oluwafemi, Emmanuel Olamigoke Famakinwa and Ometere Deborah Balogun "Random walk theory and application", 2024

[3] Martin Pfannemüller, Martin Breitbach, Christian Krupitzer, Markus Weckesser "Comparison of stock price prediction using geometric Brownian motion and multilayer perceptron" 2020

[4] Guangyu Ding, Liangxi Qin "Study on the prediction of stock price based on the associated network model of LSTM", 2019

[5] Rakhi Batra, Sher Muhammad Daudpota "Integrating StockTwits with Sentiment Analysis for Better Prediction of Stock Price Movement", 2018

[6] Guangyu Ding, Liangxi Qin "Prediction of stock values changes using sentiment analysis of stock news headlines", 2021

[7] W. Bao, J. Yue, Y. Rao "A Deep Learning Framework for Financial Time Series using Stacked Autoencoders and Long-Short Term Memory", 2017

[8] Y. Qin "A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction", 2017

[9] R. Xiong, Z. Yao, J. Li, S. Ma, X. Zhou "Practical Deep Reinforcement Learning Approach for Stock Trading", 2018

[10] K. Kim, C. Won, “Application of LSTM and Attention Mechanism for Stock Price Prediction”, 2023

[11] M. S. Hossain, M. A. Rahman, M. A. H. Akhand, “LSTM Based Stock Price Prediction”, 2019

[12] Guangxuan Zhu; Hongbo Zhao; Haoqiang Liu; Hua Sun “A novel LSTM-GAN approach for financial time series forecasting”, 2019

Plagiarism Report

The report is provided by the "eAarjav" service - http://siesgst.eaarjav.com

Verification report

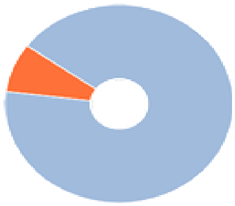
Author: Shaikh Rizwana

Document: Cleaved_Blackbook-2

Checked by: Shaikh Rizwana

Organization: SIES Graduate School of Technology

REPORT RESULTS



Matches: 8.46%

Originality: 91.54%

AI content: 19.76%

Quotes: 0%

Text recycling: 0%

Matches, Text Reuse, Text Recycling, and Originality are individual indicators displayed as percentages. Their sum is equal to 100%, which is the entire text of the checked document.

The following reuse hiding methods may have been used: Generated text on pages: 17, 18, 19, 20, 21, 22, 23, 24, 25, 26... more on page 6

- **Matches** — segments in the checked text that are fully or partially similar to identified sources, except for those that the system has classified as text reuse or recycling. The Matches value reflects the share of the checked text segments classified as matches in the overall text volume.
- **Text recycling** — includes segments in the checked text that are identical or nearly identical to a source text fragment whose author or co-author is the author of the document being checked. The Text Recycling value reflects the share of the checked text segments classified as text recycling in the overall text volume.
- **Quotes** — includes segments in the checked text that are not original, but which the system considers correctly formulated. Text reuse also includes boilerplate phrases, bibliographies, and text segments found by the Garant Legal Information System: Regulatory Documents search module. The Text Reuse value reflects the share of the checked text segments classified as text reuse in the overall text volume.
- **Text crossing** — text fragment of a checked document which is identical or almost identical to a fragment of the source text.
- **Source** — document indexed by the system and contained in the search module which is used for the check.
- **Original text** — includes segments in the checked text that are not found in any source or tagged by any of the search modules. The Originality value reflects the share of the checked text segments classified as original text in the overall text volume.

Please note that the system finds overlapping texts in the checked document and text sources indexed by the system. At the same time, the system is an auxiliary tool. Correctness and adequacy of reuse or quotes, as well as authorship of text fragments in the checked document must be determined by the verifier.

DOCUMENT INFORMATION

Document No.: 29

Document type: Not specified

Check start: 23.04.2025 10:33:27

Correction date: 23.04.2025 10:54:18

Total pages: 52

Number of characters: 67588

Number of words: 9937

Number of sentences: 714

Publication Details

6th INTERNATIONAL CONFERENCE OF EMERGING TECHNOLOGIES (INCET)

Track Name: Computational Intelligence

Paper ID: 1088

Paper Title: A Hybrid Approach to Stock Price Forecasting Using LSTM-ARO and Geometric Brownian Motion

Status: Accepted

About Conference: INCET is Technically Co-Sponsored by IEEE Bangalore section and North Karnataka Sub Section. 6th INCET 2025 is organized by JAIN COLLEGE OF ENGINEERING & RESEARCH, Belgaum Karnataka, INDIA. All their previous conference proceedings have been published by IEEE and Indexed by Scopus.

Authors:

- patiladitya26072003@gmail.com (Primary)
- apurvaparanjape27@gmail.com
- ayushpatil72680@gmail.com
- pawar.arnav1333@gmail.com
- hellorizwana@gmail.com

Publication Details

Acceptance Notification - IEEE 6th INCET 2025 Inbox ×

Microsoft CMT <email@msr-cmt.org>
to me

Wed, Mar 19, 5:51 PM

Dear Aditya N Patil

Paper ID / Submission ID : 1088

Title : A Hybrid Approach to Stock Price Forecasting Using LSTM-ARO and Geometric Brownian Motion

Greeting from 6th INCET 2025.

We are pleased to inform you that your paper has been accepted for the Oral Presentation as a full paper for the- "IEEE 2025 6th INTERNATIONAL CONFERENCE OF EMERGING TECHNOLOGIES (INCET), Belagavi, Karnataka, India with following reviewers' comment.

All accepted and presented papers will be submitted to IEEE Xplore for the further publication.

Note:
All of Accepted and Presented Papers of INCET series has been Published by IEEE Xplore and indexed by Scopus and other Reputed Indexing partners of IEEE. - http://www.incet.org/history-incet/

You should finish the registration before deadline, or you will be deemed to withdraw your paper:

Complete the Registration Process (The last date of payment Registration is 26 MARCH 2025)

Payment Links

For Indian Authors: https://rzp.io/rzp/w8lf1OTw

For Foreign Authors: https://in.explara.com/e/incet-2025