Phase-1

CRACKING THE MARKET CODE WITH AI DRIVEN STOCK PRICE PREDICTION USING TIME SERIES ANALYSIS

Student Name: Nasreen Farzana. S

Register Number: 511723104007

Institution: Pallavan College of Engineering

Department: B.E. (CSE)

Date of Submission: 29/04/2025

1. Problem Statement

The stock market is a dynamic and volatile system that poses a significant challenge for accurate price forecasting. Investors, traders, and financial institutions constantly seek predictive insights to make informed decisions and mitigate risk. Traditional methods rely heavily on technical indicators or manual interpretation, which can lead to inconsistency and human error.

This project addresses the problem of stock price forecasting using AI-based models. By applying time series analysis through deep learning techniques like Long Short-Term Memory (LSTM), the project aims to enhance the reliability and accuracy of stock price predictions. This has potential real-world applications in algorithmic trading, portfolio optimization, and financial analytics.

2. Objectives of the Project

- To develop a time series forecasting model using LSTM that predicts future stock prices based on historical data.
- To preprocess and visualize stock market trends using effective data handling techniques.
- To evaluate model performance using error metrics such as MSE and RMSE.
- To compare the LSTM model with traditional forecasting methods (like ARIMA).
- To explore deployment options such as dashboards or web apps for practical use.

3. Scope of the Project

Planned Features:
- Data collection from Yahoo Finance using APIs (yfinance).
- Visualization of price trends (e.g., line charts of predicted vs actual).
- Deep learning model (LSTM) for sequential prediction.
- Model performance analysis with error metrics.
- Basic user interface or notebook demo for forecast visualization.

Limitations:
- Focus is on closing price only (univariate analysis).

- Does not include external factors like news or financial indicators.
- Limited to educational/research use, not live market trading.
- Predicts short-term trends; long-term forecasting may be less accurate.

4. Data Sources
- Source: Yahoo Finance via yfinance Python package.
- Nature: Public, free-to-access.
- Type: Time series (daily OHLC + volume data).
- Format: Static (historical stock data downloaded during the project timeframe).

5. High-Level Methodology
● Data Collection
- Collect historical stock prices (e.g., Apple, AAPL) using yfinance.

● Data Cleaning
- Handle missing values, remove null entries.
- Normalize data using MinMaxScaler for LSTM compatibility.

● Exploratory Data Analysis (EDA) -
Plot price trends over time.
- Visualize moving averages and volatility.
- Compare model prediction vs actual price graphs.

● Feature Engineering
- Create time-step sequences (e.g., 60-day windows) for model input.
- Optionally include technical indicators like moving average.

● Model Building
- Use LSTM (deep learning RNN variant) due to its effectiveness in sequential data modeling.
- Compare with a baseline ARIMA model if needed.

● Model Evaluation
- Use RMSE and MAE to assess performance.
- Visual comparison of predicted vs actual prices.

● Visualization & Interpretation
- Plot outputs using matplotlib or seaborn.
- Use graphs to show trends, prediction performance.

● Deployment
- Optional demo using Streamlit or notebook interface for user interaction.
- Users can input a stock ticker and receive a prediction plot.

6. Tools and Technologies

● Programming Language
- Python

● Notebook/IDE
- Google Colab (for GPU support)
- Jupyter Notebook (for local development)

● Libraries
- pandas, numpy – Data processing
- matplotlib, seaborn – Visualization
- yfinance – Data collection
- scikit-learn – Data preprocessing
- tensorflow, keras – Model development

● Optional Tools for Deployment
- Streamlit (for interactive dashboards)
- Flask or Gradio (for API or UI)

7. Team Members and Roles

NAME	ROLE & RESPONSIBILITY
Keerthana. D	Project Coordinator – Oversees end-to-end project execution, model validation, documentation.
Anitha . V	Data Handler – Collects and preprocesses stock data using Python.
Divya. K	EDA and Visualizations – Performs data analysis and generates graphs.
Ananthi. R	ML Model Developer – Builds and trains the LSTM model.
Nazreen Farzana . S	UI & Reporting – Handles visualization presentation and prepares project reports.