PROJECT REPORT

ON

TIME SERIES ANALYSIS AND FORECASTING FOR STOCK MARKET

ABSTRACT :

This project focuses on predicting Apple Inc. (AAPL) stock prices using three time series forecasting models: LSTM, ARIMA, and SARIMA. The dataset contains historical closing prices, which were preprocessed and used for model training. LSTM, a deep learning model, captures long-term patterns in sequential data. ARIMA and SARIMA are statistical models suited for trend and seasonal data. The performance of all models is compared to evaluate accuracy and forecasting capability. This study helps understand which model works best for financial time series prediction.

INTRODUCTION :

Stock market prediction is a critical area of research in the field of financial analytics, offering valuable insights for investors, analysts, and institutions. The inherent volatility and complexity of financial time series data demand the use of advanced forecasting models capable of capturing both short-term fluctuations and long-term patterns.

This project focuses on predicting the closing prices of Apple Inc. (AAPL) stock using three distinct approaches: Long Short-Term Memory (LSTM), AutoRegressive Integrated Moving Average (ARIMA), and Seasonal ARIMA (SARIMA). LSTM, a deep learning model, is effective in modeling sequential dependencies, while ARIMA and SARIMA are statistical models known for their reliability in linear and seasonal time series forecasting.

By implementing and comparing these models, the project aims to evaluate their performance, highlight their strengths and limitations, and identify the most suitable technique for accurate stock

OBJECTIVE :

Perform data preprocessing and visualization of stock prices.

- Apply ARIMA and SARIMA models for statistical forecasting.

- Implement LSTM (Long Short-Term Memory) neural networks for deep learning-based prediction.

- Compare the performance of these models.

- Forecast future stock values based on the best-performing model.

METHODOLOGY :

This project is divided into 3 main phases:

PHASE 1: DATA COLLECTION AND PREPROCESSING

- Collected historical stock data (e.g., from Yahoo Finance or NSE).
- Handled missing values, checked stationarity using ADF test.
- Visualized trends, seasonality, and autocorrelation (ACF, PACF)

PHASE 2: MODELING AND TRAINING

1. ARIMA (Auto Regressive Integrated Moving Average)

- Selected order (p,d,q) using AIC/BIC and grid search.
- Trained ARIMA model and validated using Mean Squared Error.

2. SARIMA (Seasonal ARIMA)

- Captured both trend and seasonal components.
- Selected seasonal order (P,D,Q,s) along with (p,d,q).
- Trained SARIMA model with improved performance on seasonal data.

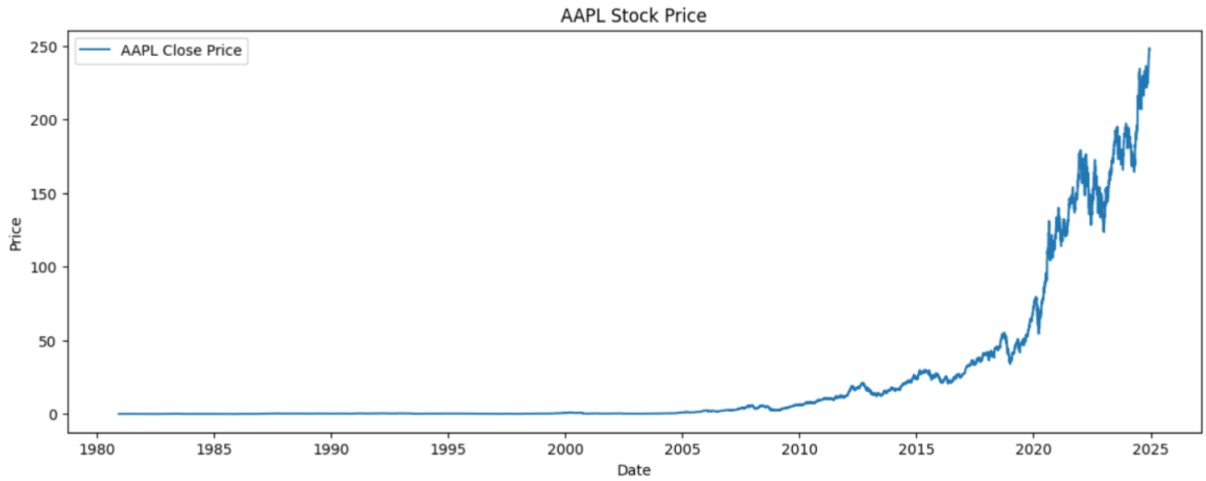
3. LSTM (Long Short-Term Memory)

- Scaled data using MinMaxScaler.
- Created sequences of data to feed into LSTM.
- Used stacked LSTM layers and Dense output layer.
- Trained using Adam optimizer, validated on test set.

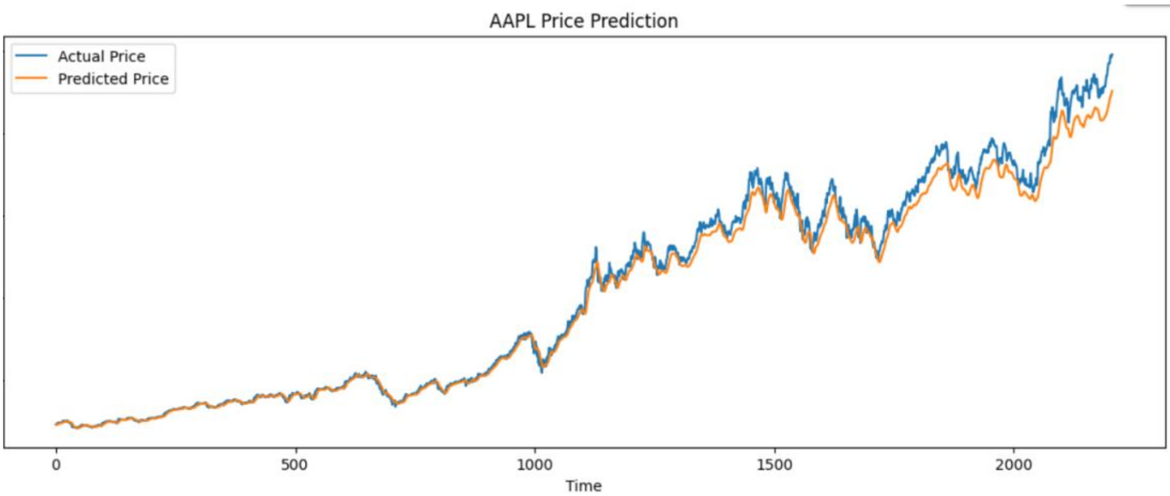
PHASE 3: EVALUATION AND FORECASTING

- Evaluated all models using RMSE and MAE.
- Compared predicted vs actual prices visually.
- Forecasted future prices (e.g., 30 days ahead) using the best model.

OUTPUT :




```
276/276 ━━━━━━━━━━━━━━━━━━━━ 16s 47ms/step - loss: 1.1817e-04
Epoch 2/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 21s 47ms/step - loss: 1.7896e-06
Epoch 3/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 20s 47ms/step - loss: 2.0885e-06
Epoch 4/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 20s 47ms/step - loss: 1.7360e-06
Epoch 5/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 21s 48ms/step - loss: 1.7027e-06
Epoch 6/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 21s 48ms/step - loss: 1.7932e-06
Epoch 7/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 21s 49ms/step - loss: 1.2841e-06
Epoch 8/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 20s 48ms/step - loss: 1.6066e-06
Epoch 9/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 13s 48ms/step - loss: 1.4822e-06
Epoch 10/10
276/276 ━━━━━━━━━━━━━━━━━━━━ 20s 48ms/step - loss: 1.1308e-06
69/69 ━━━━━━━━━━━━━━━━━━━━ 1s 17ms/step
```



RESULTS :

ARIMA performed well on stationary series but struggled with seasonality.

- SARIMA handled seasonal data effectively.
- LSTM achieved the best accuracy due to its ability to capture nonlinear patterns.
- Forecasting visualizations showed LSTM as most accurate for longer horizons.

CONCLUSION :

This project demonstrates that while traditional models like ARIMA and SARIMA are effective, LSTM offers superior performance for complex, nonlinear stock data. Time Series Forecasting is valuable in financial planning, investment strategies, and market risk assessment.

FUTURE SCOPE:

- Implement hybrid models combining statistical and neural approaches.
- Deploy a real-time forecasting dashboard using Streamlit or Flask.
- Extend the project to include multiple stocks and financial indicators.

REFERENCES:

- Box, Jenkins: Time Series Analysis (ARIMA)
- TensorFlow and Keras Documentation
- Time Series Forecasting with Python Jason Brownlee
- NSE/Yahoo Finance stock datasets