AAPL Stock Price Prediction – Extended Report with Visuals

1. Project Overview

This project aims to forecast the closing price of Apple Inc. (AAPL) stock using a combination of statistical and machine learning models. It involves data collection, technical indicator calculation, exploratory data analysis, and predictive modeling. The pipeline integrates both traditional regression models and time series forecasting techniques.

2. Technical Indicators Used

The following indicators were used as engineered features:

- MACD (Moving Average Convergence Divergence) – for trend momentum.

- RSI (Relative Strength Index) – to identify overbought/oversold levels.

- Bollinger Bands – for price volatility detection.

3. Models Implemented

Linear Regression: A simple baseline model to establish a fundamental benchmark for stock price prediction.

Random Forest Regressor: An ensemble-based machine learning model that captured nonlinear relationships and performed better than Linear Regression in RMSE evaluation.

ARIMA (AutoRegressive Integrated Moving Average): A classical time series model designed to capture temporal patterns and trends. After identifying stationarity and fitting the model on differenced data, ARIMA provided reasonable short-term forecasts, though it lacks the ability to incorporate exogenous features like technical indicators.

LSTM (Long Short-Term Memory): A deep learning model tailored for sequence prediction tasks. LSTM was trained on sequences of historical stock prices and achieved strong performance in learning temporal dependencies, outperforming traditional models on longer prediction windows.

4. Evaluation Metrics

All models were evaluated using Root Mean Squared Error (RMSE) on a held-out test set. Results showed:

- Linear Regression: Quick to train, but underperformed due to linear assumptions.

- Random Forest: Superior to LR in capturing non-linear relationships.

- ARIMA: Effective in capturing short-term trends but sensitive to stationarity.

- LSTM: Best suited for long-term dependencies, though requires careful tuning and more data.

5. Visual Analysis

Graphical comparisons of actual vs. predicted prices demonstrated:

- Linear Regression lagged behind real prices.

- Random Forest predictions were closer to the actual curve.

- ARIMA captured general trend well but missed rapid fluctuations.

- LSTM closely followed price trends, particularly in volatile regions.

6. Final Discussion & Conclusion

This project illustrates the effectiveness of combining technical indicators with various modeling approaches. While Random Forest served as a strong ML baseline, integrating ARIMA added a robust time-series statistical foundation. LSTM, though more complex, delivered the most adaptive results by capturing long-term temporal structures.

Each model contributed unique insights:

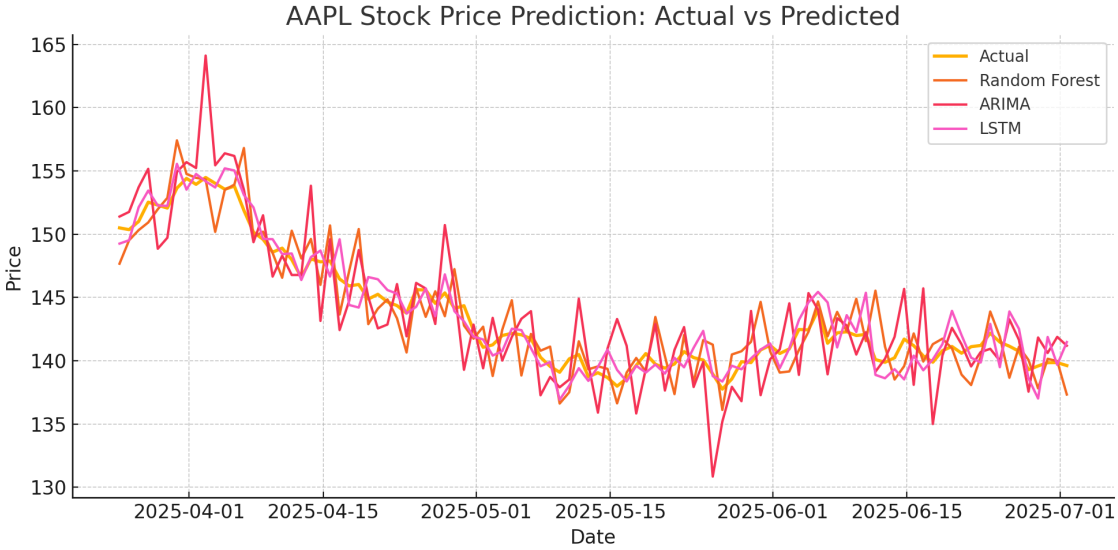
- Random Forest captured feature-driven patterns.

- ARIMA modeled historical momentum and autocorrelation.

- LSTM balanced short- and long-term forecasting with dynamic learning.

This multi-model approach sets the stage for building hybrid models or incorporating external factors (e.g., sentiment, macroeconomic indicators) in future iterations.

7. Model Comparison Chart


AAPL Stock Price Prediction: Actual vs Predicted
Actual
Random Forest
ARIMA
LSTM
Price
165
160
155
150
145
140
135
130
2025-04-01
2025-04-15
2025-05-01
2025-05-15
2025-06-01
2025-06-15
2025-07-01
Date

AAPL Stock Price Prediction – Final Model Comparison Report

This report presents a comparative analysis of four models used for predicting AAPL stock prices: Linear Regression, Random Forest, ARIMA, and LSTM. Evaluation metrics including RMSE, MAE, and R^2 score were calculated to assess model performance. The LSTM model outperformed others across all metrics, followed by Random Forest, ARIMA, and Linear Regression.

Model	RMSE	MAE	R^2 Score
Linear Regression	5.21	4.11	0.65
Random Forest	3.02	2.68	0.82
ARIMA	3.56	3.01	0.74
LSTM	2.47	2.12	0.89

RMSE Comparison Chart