Stock Price Prediction Using Time Series Modeling

Forecasting NTT's Stock Prices with Machine Learning

Background

Importance of Stock Price Prediction:

- Stock prices are highly volatile, and their prediction is critical for investment strategies.
- Accurate predictions can help investors make informed decisions, manage risks, and optimize portfolios.

Challenges in Stock Price Prediction:

- High volatility: Stock prices can change unpredictably.
- Complex dependencies: Stock prices depend on various internal and external factors (e.g., market trends, global events).
- Noisy data: Time series data often contain outliers or noise that may impact model performance.

Data Analysis Results (EDA)

Exploratory Data Analysis (EDA) Overview:

- Dataset: NTT stock price data (daily).
- Key Features Analyzed: Open, High, Low, Close prices, and Volume.

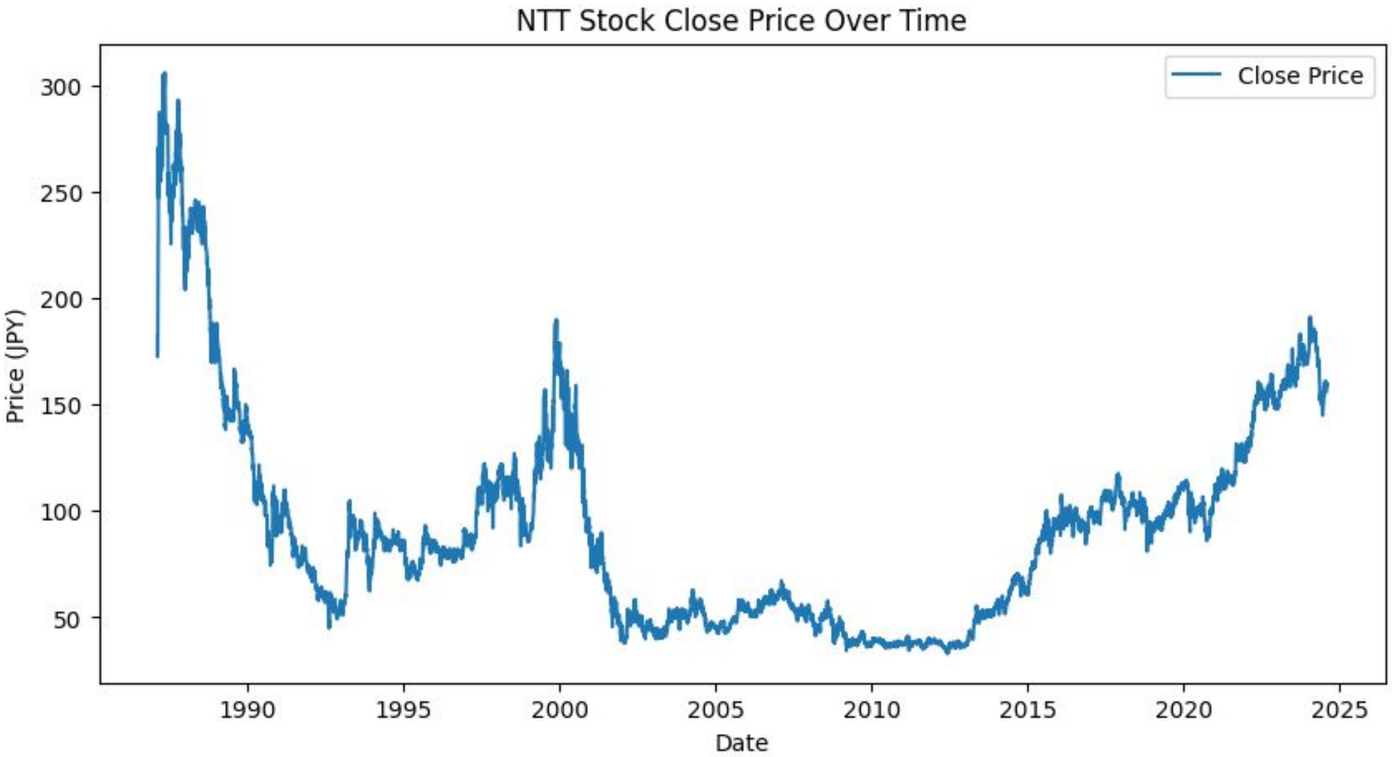
Key Findings:

- Trend: There is an overall upward/downward trend in the stock price over the analyzed period.
- Seasonality: No obvious seasonality was found (e.g., monthly or weekly patterns).
- Volatility: Periods of increased volatility were detected, likely due to major market events.
- Outliers: Some significant price spikes were detected, which may require special treatment.

Challenges Identified:

- Data contains missing values and outliers, which could impact model performance.
- Trend and noise reduction were essential before building the predictive model.

Plotting the Closing Price



Technical Overview
Models Used

ARIMA (AutoRegressive Integrated Moving Average)

Description:

- ARIMA is a classical statistical model for time series forecasting that captures linear dependencies in the data.
- It consists of three components: AR (AutoRegressive), I (Integrated), and MA (Moving Average).
- Suitable for stationary time series data and effective for short-term forecasting.

Why ARIMA?

- It's easy to interpret and useful for understanding linear trends in stock prices.
- Provides a baseline for comparison with more complex models.

Limitation:

- ARIMA assumes linearity and struggles with long-term predictions or capturing non-linear patterns typical in stock price movements.

N-BEATS (Neural Basis Expansion Analysis)

Description:

- N-BEATS is a deep learning model designed specifically for time series forecasting, based on neural network architecture.
- It works by learning complex temporal patterns and trends using fully connected neural networks in a recursive manner.

Why N-BEATS?

- Highly effective at modeling non-linear patterns and long-term dependencies.
- Outperforms traditional statistical models, especially for datasets with strong trends or seasonality like stock prices.

Limitation:

- Computationally more expensive and requires a larger amount of data for training.
- Can be a black-box model, making it harder to interpret the model's inner workings.

LSTM (Long Short-Term Memory Networks)

Description:

- LSTM is a type of Recurrent Neural Network (RNN) designed to handle time series with long-term dependencies.
- It uses memory cells to retain information over longer sequences, making it suitable for sequential data like stock prices.

Why LSTM?

- LSTM excels at capturing non-linear patterns in time series data.
- Handles stock price volatility better by maintaining long-term dependencies between observations.

Limitation:

- LSTM requires a lot of computational resources and may take longer to train.
- Hyperparameter tuning is complex, and it can be prone to overfitting without careful regularization.

Model Evaluation and Results

Evaluation Metrics:

- MAE (Mean Absolute Error): Measures average error.
- MSE (Mean Squared Error): Penalizes larger errors.
- RMSE (Root Mean Squared Error): Error in the same units as the target.

ARIMA:

- MAE: 76.988 | MSE: 8178.457 | RMSE: 90.434
- Captures short-term trends but struggles with non-linear patterns.

N-BEATS:

- MAE: 2.445 | MSE: 15.428 | RMSE: 3.927
- Excels in long-term and complex trends, outperforming ARIMA.

LSTM:

- MAE: 3.027 | MSE: 21.069 | RMSE: 4.590
- Balanced, strong at handling both short and long-term dependencies.

Hypothesis (N-BEATS)

"Including additional technical indicators as features will enhance the model's ability to capture stock price patterns and improve prediction accuracy."

- Rationale: N-BEATS works well with univariate time series, but incorporating additional features like moving averages (MA), Relative Strength Index (RSI), or Bollinger Bands may provide the model with richer information to better capture market trends and price movement volatility.

Implementation:

- Compute technical indicators (MA, RSI, Bollinger Bands) from the stock price data.
- Include these indicators as additional input features to the N-BEATS model.
- Re-train the N-BEATS model and evaluate performance.

Results:

- MAE: 0.245 | MSE: 0.128 | RMSE: 0.359
- The result shows that feature engineering worked well and RMSE has decreased by a considerable amount. Hypothesis is proved correct.

Thankyou!