Stock Price Prediction

SRINIDHI S

Problem Statement

Stock market prediction is a long-standing challenge due to volatile and non-linear price behaviour. Investors require reliable forecasts to make informed buy/sell decisions. Traditional methods often fail to capture market dynamics and hidden patterns. This project uses historical stock data to predict future prices using machine learning.

Tech Stack

- **Programming Language:** Python
- **Libraries:**
    - ✓ Pandas, NumPy – for loading, cleaning, and structuring data
    - ✓ Matplotlib – for visualizing trends and stock movements
    - ✓ Scikit-learn – for training regression models
- **Development Environment:** Jupyter Notebook

Solution Approach

1. **Data Handling**:
    - Loaded stock market data with Date, Open, High, Low, Close, Volume, and Price.
    - Converted date formats, handled missing values, and scaled numerical features.
2. **Feature Engineering**:
    - Created lag features like previous day's closing price, moving averages, etc.
    - Removed irrelevant or redundant columns to prevent data leakage.
3. **ML Algorithms Used**:
    - **Linear Regression**: For baseline modeling of price vs. time-based features.
    - **Random Forest Regressor**: Captured non-linear dependencies and improved accuracy.

4. **Model Evaluation**:
    - Used Root Mean Square Error (RMSE) to evaluate model predictions.
    - Also calculated R^2 score to measure how well the model explains price variance.
5. **Visualization**:
    - Plotted predicted vs actual prices for visual validation.
    - Analyzed error distribution to ensure robust forecasting.