**Technical Notes Related to This Model Building**

Few important points need to be discussed for this project. Those are below:

**1) Data preprocessing steps**: Renamed the column name from %Deliverble' to 'Deliverble'.

The dataset is cleaned and preprocessed by handling missing values, removing outliers, and converting date columns to date time format.

**2. Technical Analysis: Bollinger Bands to assess price deviation from a moving average for ADANIPORTS. It helps to analyze the volatility and unpredictable nature of the stock in best way.**

**3. Feature Engineering**: Various technical indicators such as Moving Averages, Bollinger Bands, and Relative Strength Index (RSI) are calculated. Additionally, lagged features, rolling window statistics, and time-based features are engineered.

**4. Hyper parameter Tuning:** (n\_estimators=100, random\_state=42, n\_jobs=-1)

**5. Machine Learning**: A Random Forest Regressor model is trained on the preprocessed data to predict future stock prices.

**6. Algorithm choices**: Various machine learning techniques could have been be applied and explored to stock market data, especially for trading algorithms and learning time series models. But I had chosen Random Forest Regressor for it. **The reason of it is:**

**a) Capturing Complex Relationships:** Stock prices are influenced by many interacting factors (lagged prices, volumes, technical indicators). Random Forest can effectively model these complex, non-linear relationships that might not be captured by simpler models.

**b) Handling Diverse Features:** Feature engineering created a variety of features (lagged values, moving averages, technical indicators). Random Forest can handle this diverse set of features without requiring extensive preprocessing like scaling (though you did scale your data, which is generally good practice).

**c) Identifying Important Factors:** The feature importance output of the Random Forest helps you understand which technical indicators, lagged prices, or other features are most influential in predicting the stock's closing price. This is valuable for gaining insights into the stock's behavior.

**d) Robustness:** Stock data can sometimes contain noise or outliers, and Random Forests are relatively robust to these issues.

**e) Predictive Performance:** Random Forests are generally known for their good predictive accuracy on many real-world datasets, making them a solid choice for a forecasting task like this.

**7. Model Evaluation**: The model is evaluated using metrics such as R2 score and Root Mean Squared Error (RMSE). The R2 score for the model was 90% approx.

**Key Findings:**

🡪Promising Stocks: Based on the analysis, Mundra Port emerges as a promising stock for investment due to its high average closing price and volatility.

🡪Predictive Model: The Random Forest Regressor model is able to predict future stock prices with a reasonable degree of accuracy, as indicated by the R2 score and RMSE.

After applying technical indicator the no. of features was increased up to 109. Before splitting the data I removed the Close and Symbol columns.

**Technical Indicators Used:**

1. Moving Averages: Simple Moving Average (SMA) and Exponential Moving Average (EMA)

2. Bollinger Bands: Upper and Lower Bands

3. Lagged Features: Close, Volume, and Volatility

4. Rolling Window Statistics: Mean and Standard Deviation

**Why Feature Importance was used?**

**Tree-based Feature Importance**: Calculates feature importance based on the feature importance scores of decision trees in order to understand which features drive the predictions, allowing for more informed decision-making.