

Final Project Summary

Project Title:

Stock Price Prediction and Comparison between TCS & Infosys using Machine Learning (2015–2030)

Objective:

To build a machine learning model that predicts and compares future stock prices of two major Indian IT companies — **TCS (Tata Consultancy Services)** and **Infosys** — using **Linear Regression** and **Random Forest Regression**, and determine which stock is potentially safer for long-term investment.

Dataset:

- **Source:** Yahoo Finance
 - **Companies:** TCS and Infosys
 - **Time Period:** 2015 to 2025 (used for training)
 - **Attributes:** Date, Open, High, Low, Close, Volume
-

Tools & Technologies:

- **Languages:** Python
 - **Libraries:** Pandas, NumPy, scikit-learn, Matplotlib
 - **Models Used:**
 - Linear Regression
 - Random Forest Regressor
 - **Evaluation Metrics:** MAE (Mean Absolute Error), RMSE (Root Mean Squared Error)
-

Steps Performed:

1. Data Collection:

Downloaded TCS and Infosys stock data (2015–2025) from Yahoo Finance.

2. Data Preprocessing:

Cleaned data, extracted features, and prepared it for modeling.

3. Exploratory Data Analysis (EDA):

Visualized stock trends and compared price patterns.

4. Model Training & Evaluation:

- Trained Linear Regression and Random Forest models on both datasets.
- Evaluated models using MAE and RMSE.

5. Prediction Till 2030:

- Predicted future stock prices daily till **2030**.
- Stored and plotted the predictions.

6. Final Price Predictions for 2030:

TCS (2030) - Linear Regression : ₹4330.67

TCS (2030) - Random Forest : ₹4291.17

Infosys (2030) - Linear Reg. : ₹4313.30

Infosys (2030) - Random Forest : ₹1888.27

Insights:

- TCS stock is predicted to maintain a **more stable and steadily increasing trend** using both models.
- **Infosys** stock shows a strong prediction with Linear Regression but a noticeable drop using Random Forest.
- **TCS appears to be the safer long-term investment**, especially based on Random Forest's conservative estimation.

Conclusion: The models predict that both TCS and Infosys may grow by 2030, but **TCS is more consistent and reliable** based on historical data and model performance. Machine Learning models like Linear Regression and Random Forest provide valuable insights for long-term investment decisions.