TCS Stock Data Analysis and prediction

**Table of Contents**

* Problem Statement
* Dataset Overview
* Data Analysis & Visualization
* Machine Learning Models
* Challenges Faced
* Tools, Software, and Libraries

**Problem Statement**

This project aims to analyze the historical stock data of Tata Consultancy Services (TCS) to identify key trends, visualize stock behaviour, and predict future stock prices using Machine Learning models. The primary objectives include :

* Understanding TCS stock price movements over time.
* Identifying key patterns such volatility and moving averages.
* Building predictive models to forecast future stock price.
* Providing insights that could support investment strategies.

**Dataset Overview**

The dataset consists of historical stock trading data for TCS, including :

* **Date –** Trading date
* **Open –** Opening stock price
* **High –** Highest stock price of the day
* **Low –** Lowest stock price of the day
* **Close –** Closing stock price
* **Volume –** Number of shares traded
* **Dividends –** Dividends paid
* **Stock Splits –** Number of stock splits

The dataset enables us to perform time-series analysis and build predictive models of stock price forecasting.

**Data Analysis & Visualization**

**Steps Involved :**

1. **Data Preprocessing :** Handling missing values, scaling features, and structuring data for analysis.
2. **Exploratory Data Analysis (EDA) :**

* Visualizing price trends, stock fluctuations, and trading volume.
* Identifying stock volatility using moving averages and rolling statistics.
* Analyzing stock return distributions.

**Machine Learning Models**

Three different machine learning models were used for stock price prediction :

1. **Linear Regression**

* A simple statistical model that finds relationships between stock features and the closing prices.
* Helps in understanding how each feature impacts stock prices.

1. **LSTM (Long Short-Term Memory) Neural Network**

* A deep learning model capable of learning sequential dependencies in time-series data.
* Help capture long-term trends in stock prices for better forcasting.

**Challenges Faced**

* **Data Cleaning Issues :** Missing values and inconsistencies required imputation and smoothing.
* **LSTM Computational Requirements :** Training deep learning models required GPU acceleration for efficiency.
* **Stock Market Unpredictability :** External factors such as news events and economic trends impact prices beyond historical patterns.

**Tools, Software, and Libraries**

* **Python –** Data analysis and machine learning
* **Pandas, NumPy –** Data manipulation
* **Matplotlib, Seaborn –** Data visualization
* **Scikit-learn –** Machine learning algorithms
* **TensorFlow/Keras –** Deep learning models (LSTM)
* **VS Code –** Interactive data exploration