**TCS Stock Data Analysis and Prediction Using Machine Learning**

Contents

[Abstract 2](#_Toc208258079)

[Introduction 2](#_Toc208258080)

[Dataset Overview 2](#_Toc208258081)

[Preprocessing 2](#_Toc208258082)

[Exploratory Data Analysis 2](#_Toc208258083)

[Feature Engineering 3](#_Toc208258084)

[Model Building 3](#_Toc208258085)

[Results 4](#_Toc208258086)

[Conclusion 4](#_Toc208258087)

[Future Enhancements 4](#_Toc208258088)

**Abstract:**

This project aims to analyze historical stock data of Tata Consultancy Services (TCS) to identify trends and forecast future prices using machine learning techniques. It includes data preprocessing, exploratory data analysis (EDA), feature engineering, and predictive modeling using Linear Regression and LSTM.

**Introduction:**

TCS is a global IT services company with significant market capitalization. Understanding its stock behavior can help investors and analysts make informed decisions. This project leverages historical data to build predictive models.

**Dataset Overview:**

The dataset includes:

* Date
* Open, High, Low, Close prices
* Volume
* Dividends
* Stock Splits

It spans over 4,000 entries and is clean with no missing values.

**Pre-processing:**

* Converted date formats
* Handled nulls using forward fill
* Converted price columns to numeric
* Extracted time-based features (Year, Month, Day, Day of Week)
* Created lag features like previous day’s close

**Exploratory Data Analysis:**

* Time series plots of closing prices
* Volume and dividend trends
* Moving averages (50-day, 200-day)
* Correlation analysis
* Scatter plots for Close vs Volume, Dividends, and Stock Splits

**Feature Engineering:**

* Added temporal features
* Created lag variables
* Calculated moving averages
* Generated buy/sell signals using crossover strategy

**Model Building**

**Linear Regression:**

* Features: Open, High, Low, Volume, Prev\_Close, Day\_of\_Week, Month
* Evaluation: MSE and R² score

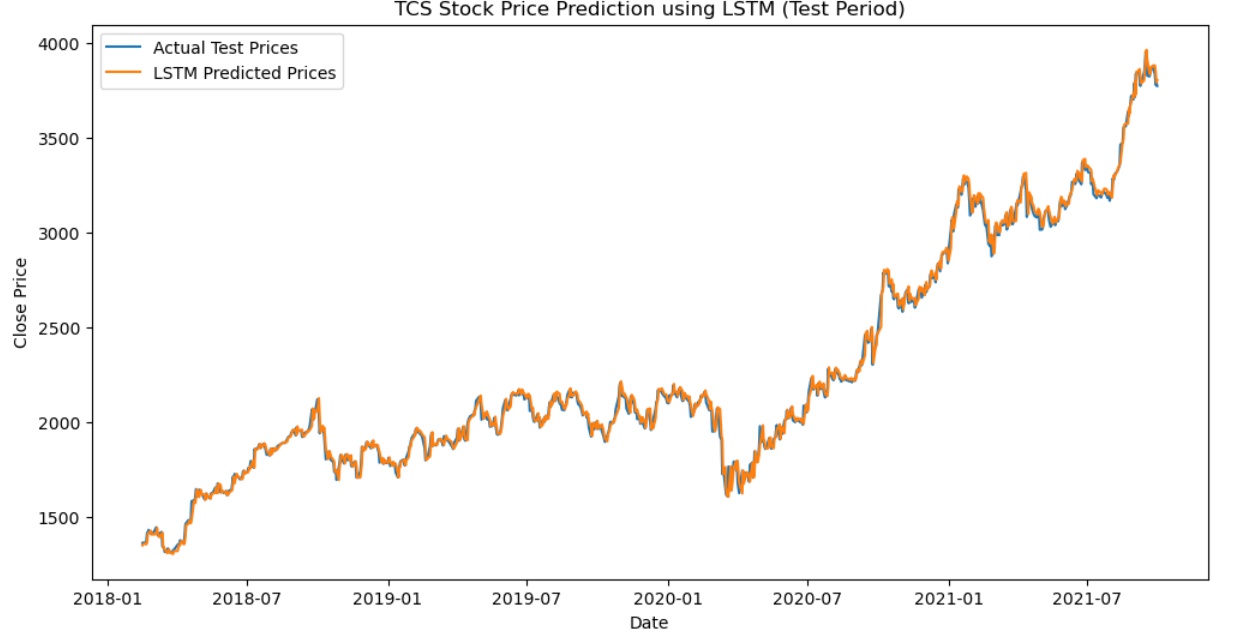
**LSTM:**

* Scaled data using MinMaxScaler
* Reshaped for LSTM input
* Trained with 30 epochs and batch size of 15
* Forecasted future prices

**Evaluation:**

* Linear Regression: Provided baseline predictions
* LSTM: Achieved better accuracy with MAE ≈ 69.5
* Visual comparison of actual vs predicted prices

**Results:**

****

* LSTM outperformed Linear Regression in capturing temporal patterns
* Moving averages and crossover strategies provided actionable insights

**Conclusion:**

The project successfully demonstrated stock price prediction using machine learning. LSTM proved effective for time-series forecasting.

**Future Enhancements:**

* Incorporate models like Random Forest, XGBoost
* Apply hyperparameter tuning
* Explore ARIMA and Prophet for time-series modeling
* Integrate external factors like news sentiment