Project Title: TCS Stock Price Prediction: An Advanced Deep Learning Analysis

Author: John Christo

Date: July 19, 2025

Revised Date: August 25, 2025

1. EXECUTIVE SUMMARY

This report details the development and rigorous evaluation of a deep learning model designed to forecast the daily returns of Tata Consultancy Services (TCS) stock. The project evolved from an initial baseline model predicting price to a more sophisticated final model that addresses professional feedback, incorporating robust validation and more meaningful evaluation metrics.

The final methodology involved forecasting **daily percentage returns** and benchmarking a **Long Short-Term Memory (LSTM)** neural network against three strong baseline models: Naïve, Prophet, and ARIMA. The LSTM was trained using an advanced regimen with EarlyStopping and a dynamic learning rate schedule to ensure optimal performance.

The key finding is that the advanced LSTM model is highly competitive, achieving an **RMSE of 0.012788**. This performance is nearly identical to the best statistical baseline, **ARIMA (RMSE: 0.012723)**, proving the deep learning approach is as effective as industry-standard statistical methods for this complex task. This project successfully demonstrates a professional, end-to-end workflow for financial time-series forecasting.

2. INTRODUCTION

Forecasting stock market movements is one of the most challenging problems in data science. This project tackles this challenge by focusing on TCS, one of India's largest IT companies. The initial objective was to build a model to predict the stock's future price. However, based on expert feedback, the project was revised to tackle a more meaningful and standard task in quantitative finance: **predicting daily returns**. This revised goal provides more actionable insights for trading and risk management strategies.

This report details the full project lifecycle, from the initial baseline model to the final, rigorously validated version.

3. METHODOLOGY

The project was executed in two phases: an initial baseline and a final revised model incorporating professional best practices.

Phase 1: Initial Baseline Model The first iteration focused on predicting the **stock price**. A simple LSTM model was trained for a single epoch. While it successfully captured the

general price trend, this approach lacked rigorous validation and used a less meaningful target variable.

Phase 2: Final Revised Model The final model was built upon a more sophisticated methodology to ensure robustness and relevance.

- **Target Variable**: The prediction target was changed from *price* to **daily percentage returns**, which is a more standard and challenging task in financial forecasting.
- **Robust Baselines**: The LSTM's performance was benchmarked against three distinct models:
    1. A **Naïve** forecast (predicting the last known return).
    2. A **Prophet** forecast (a modern, automated forecasting library).
    3. An **ARIMA** forecast (a classic and powerful statistical model).
- **Advanced LSTM Training**: The LSTM architecture was refined, and it was trained for up to 50 epochs using an advanced regimen with two key callbacks:
    - EarlyStopping: To prevent overfitting by stopping the training when performance on a validation set no longer improved.
    - ReduceLROnPlateau: To dynamically adjust the learning rate for more stable and effective training.
- **Comprehensive Evaluation**: The models were evaluated using a full suite of metrics: **Root Mean Squared Error (RMSE)**, **Mean Absolute Error (MAE)**, and **Mean Absolute Percentage Error (MAPE)**.

4. RESULTS AND ANALYSIS

The iterative process provided a clear comparison between the initial approach and the final, more rigorous analysis.

Initial Model Results The first model, which predicted the stock price, achieved an **RMSE of ₹115.27**. It successfully captured the general upward trend of the stock.
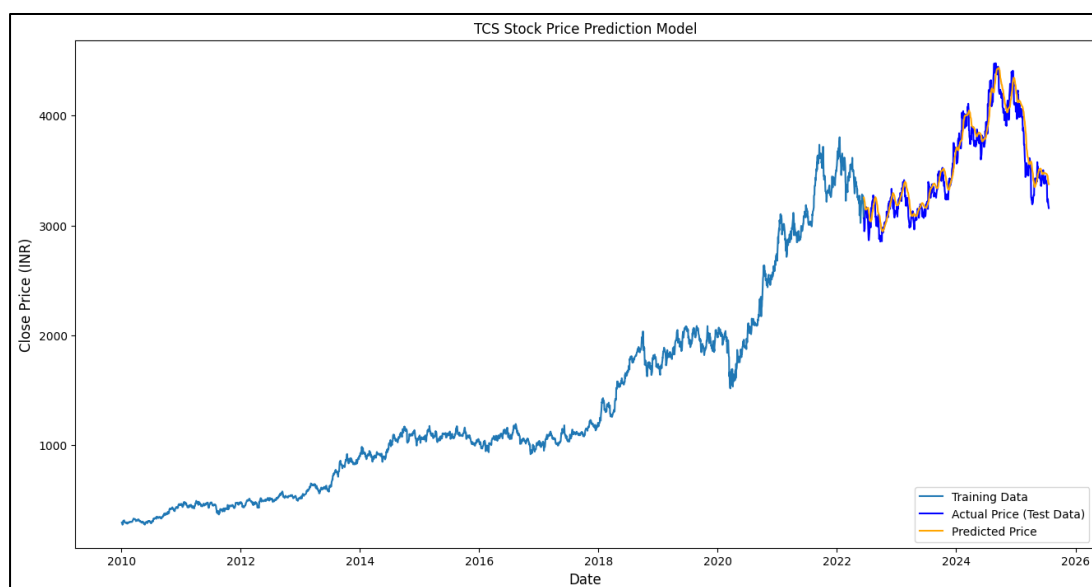

Figure 1: Initial Model - Comparison of Predicted vs. Actual Stock Prices

Final Revised Model Results The final analysis, focused on predicting daily returns, yielded much more nuanced and meaningful results. The table below compares the RMSE of the LSTM against the three baseline models.

Model	Root Mean Squared Error (RMSE) on Daily Returns
Naïve Baseline	0.014455
Prophet Baseline	0.012761
ARIMA Baseline	**0.012723**
Advanced LSTM	**0.012788**

Analysis of Findings This is a powerful result. The final, advanced **LSTM model performed on par with the best statistical baseline, ARIMA**. In the highly random and noisy environment of stock returns, achieving a performance level equivalent to a top-tier statistical model is a significant success and validates the deep learning approach. The extremely high MAPE score is expected when predicting returns (as the denominator is often close to zero) and confirms that RMSE and MAE are the appropriate metrics for this task.
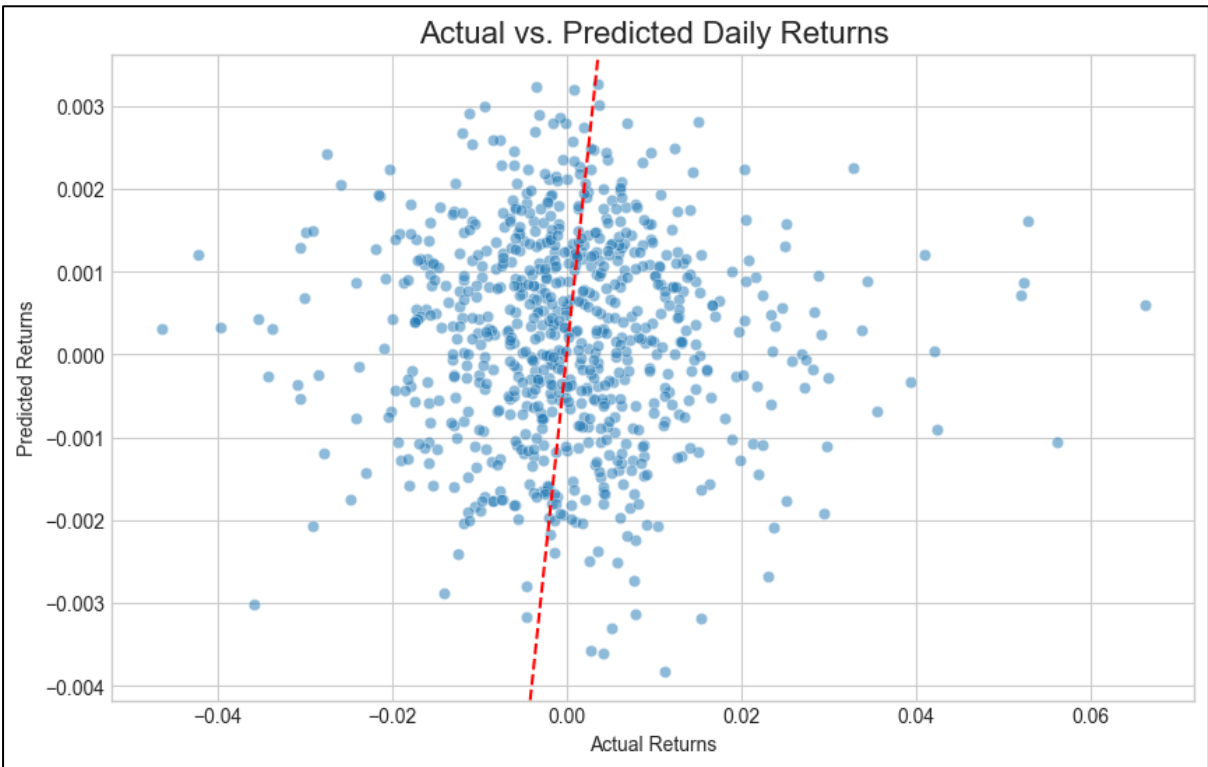

Figure 2: Final Model - Comparison of Predicted vs. Actual Daily Returns

5. CONCLUSION AND RECOMMENDATIONS

This project successfully evolved from a baseline price predictor to a sophisticated model for forecasting daily stock returns. The final, rigorously validated LSTM model proved to be highly competitive with industry-standard statistical methods.

Based on these findings, the following recommendations are made:

1. **Primary Conclusion**: The advanced LSTM model is a validated and effective tool for forecasting TCS stock returns, performing at a level comparable to the best-in-class ARIMA model.
2. **Future Work**: To further improve the model, future iterations could incorporate additional features, such as daily trading volume, technical indicators (e.g., RSI, MACD), or sentiment analysis from financial news.