**INTERNSHIP PROJECT REPORT**

**Project : Stock Price Predictor  
Intern Name: Anmol Kumar/ E-mail:kumaranmol1909@gmail.com  
Internship Duration: 19.05.2025 - 26.06.2025  
Organization: Tata Motors  
Department: Tata Technologies  
Mentor: Sanket Mohanty**

**Acknowledgement**

I would like to express my sincere gratitude to my mentor, Mr. Sanket Mohanty, and the entire team at Tata Technologies for providing me with this valuable internship opportunity. Their guidance, feedback, and support played a crucial role in my learning and successful completion of this project.

**Table of Contents**

1. Introduction
2. Problem Statement
3. Tools and Technologies Used
4. Methodology
5. Implementation
6. Results
7. Challenges Faced
8. Learnings
9. Conclusion
10. Visual Results

**1. Introduction**

This report outlines the work completed during my internship at Tata Motors, under the Tata Technologies division. The objective was to work on a machine learning project that predicts stock prices based on historical data using various models and techniques. As a solo contributor, I was responsible for end-to-end development and analysis.

**2. Problem Statement**

The objective of the project was to develop a system that could predict future stock prices by analyzing historical trends and financial indicators. The challenge involved was to implement a robust ML pipeline capable of extracting useful patterns from stock data.

**3. Tools and Technologies Used**

* **Programming Language:** Python
* **Libraries/Frameworks:** Pandas, NumPy, Scikit-learn, Yahoo Finance API, Streamlit
* **Models Implemented:** Linear Regression, Random Forest Regressor, Time Series Forecasting
* **IDE/Notebook:** Jupyter Notebook
* **Frontend:** Streamlit (for user interaction and data visualization)

**4. Methodology**

1. Collected historical stock data using Yahoo Finance API.
2. Cleaned and processed the data to handle missing values and anomalies.
3. Conducted exploratory data analysis to understand trends.
4. Engineered features for better model performance.
5. Split the data into training and testing sets.
6. Trained multiple ML models (Linear Regression, Random Forest).
7. Applied time series techniques for future price prediction.
8. Evaluated model performance using RMSE and R² scores.
9. Developed an interactive user interface using Streamlit for real-time predictions and data visualizations.

**5. Implementation**

The implementation involved:

* Fetching stock data using the yfinance library.
* Performing data preprocessing like handling nulls, formatting date columns, and feature scaling.
* Using technical indicators like moving averages and RSI as additional features.
* Applying regression and time series algorithms and plotting actual vs predicted prices.
* Creating a Streamlit dashboard with input forms for stock symbols, date range selection, and dynamic output charts.
* Comparing model performance for selection.

**6. Results**

The Random Forest Regressor outperformed the Linear Regression model with a better R² score and lower RMSE. The prediction graphs closely followed actual trends, demonstrating strong correlation and accuracy in the model. Time series forecasting provided an additional layer for long-term outlook. The Streamlit app enabled users to test predictions interactively, making the tool user-friendly and intuitive.

**7. Challenges Faced**

* Faced issues during data splitting, especially maintaining time series order.
* Encountered errors during technical analysis feature generation.
* Initial model performance was low due to unclean data.
* Faced minor bugs while integrating ML outputs with the Streamlit frontend.

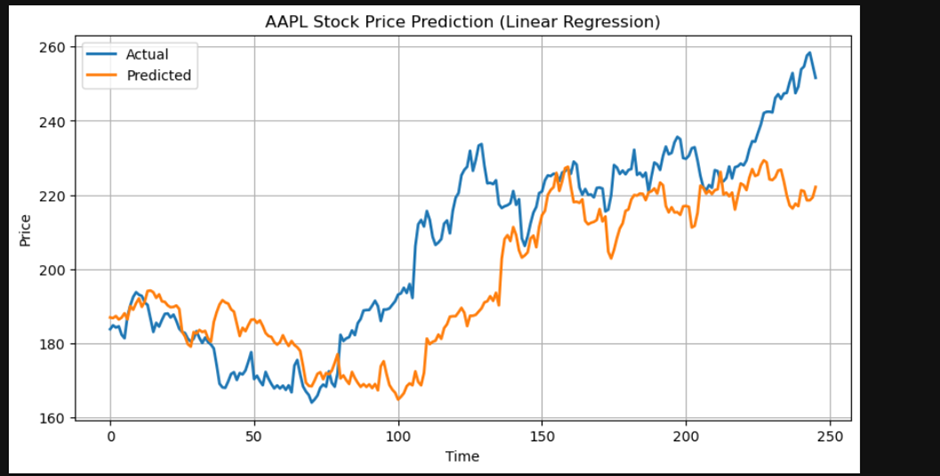
**8. Learnings**

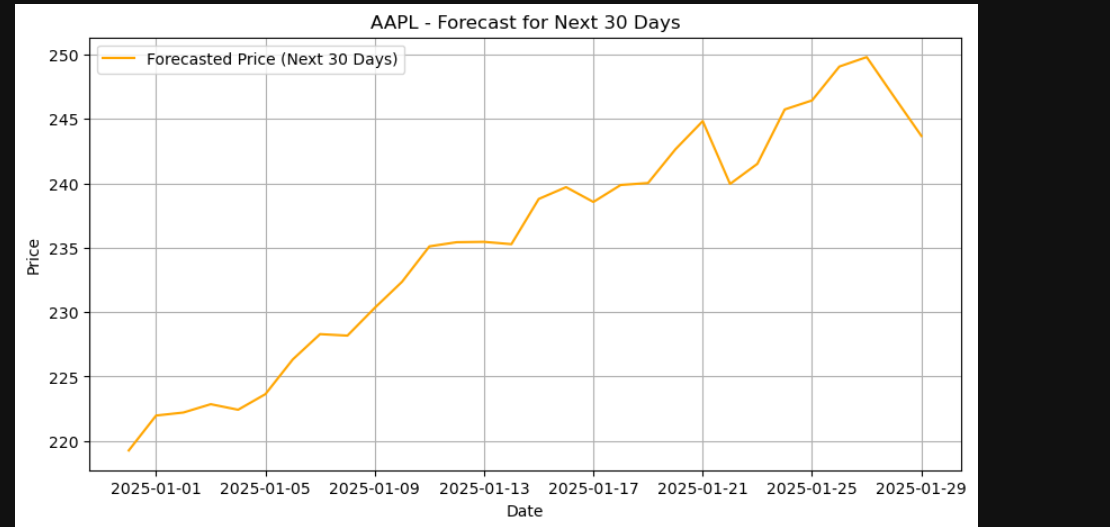
* Gained hands-on experience with machine learning workflows.
* Improved understanding of regression models and how to tune them.
* Learned to manage and clean large datasets effectively.
* Understood the basics of frontend deployment using Streamlit.
* Gained exposure to time series forecasting techniques.

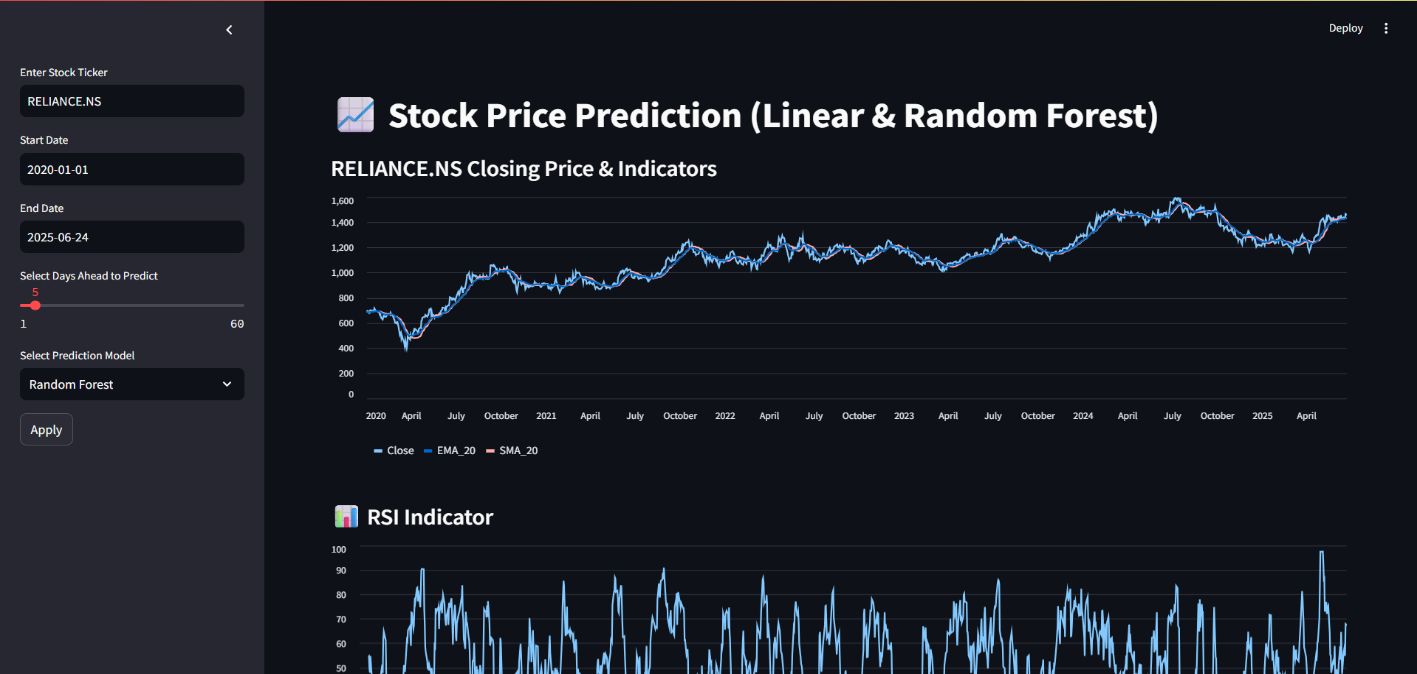
**9. Conclusion**

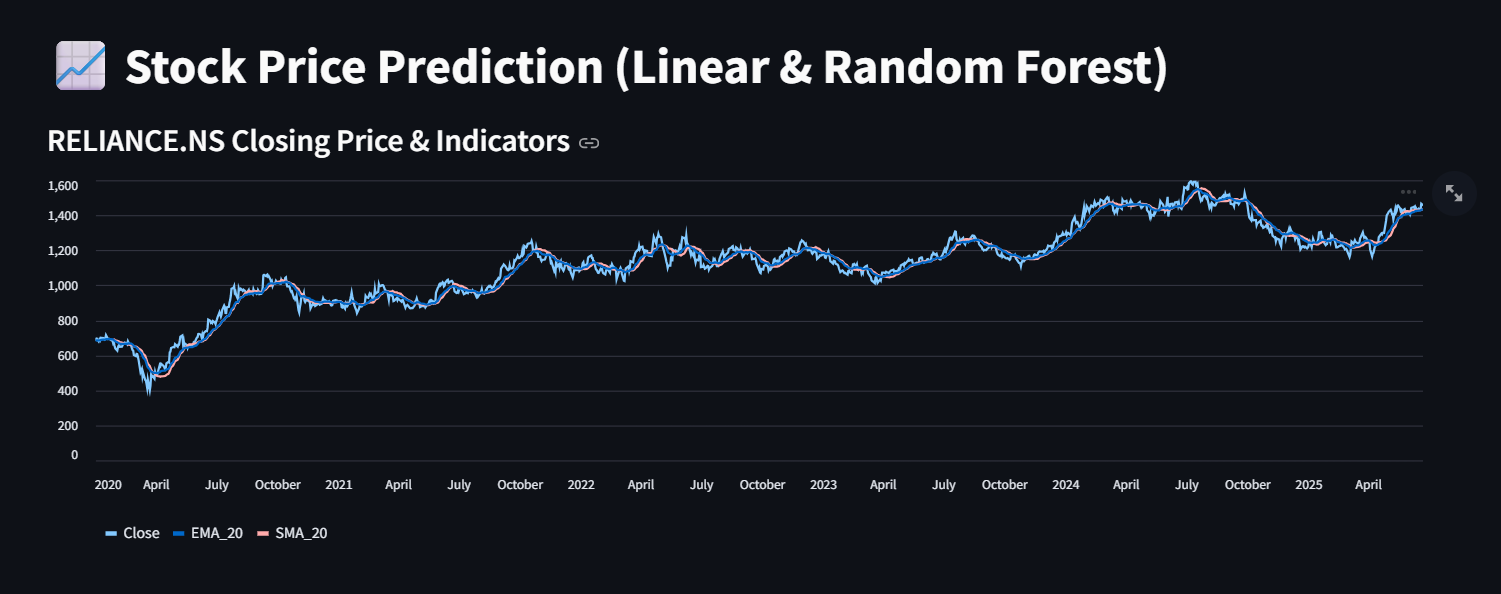
This internship has been a valuable learning journey, allowing me to apply machine learning techniques in a real-world scenario. The Stock Price Predictor project gave me exposure to end-to-end project execution from data collection to model evaluation and frontend integration using Streamlit. I am thankful to Tata Motors and Tata Technologies for this enriching experience.

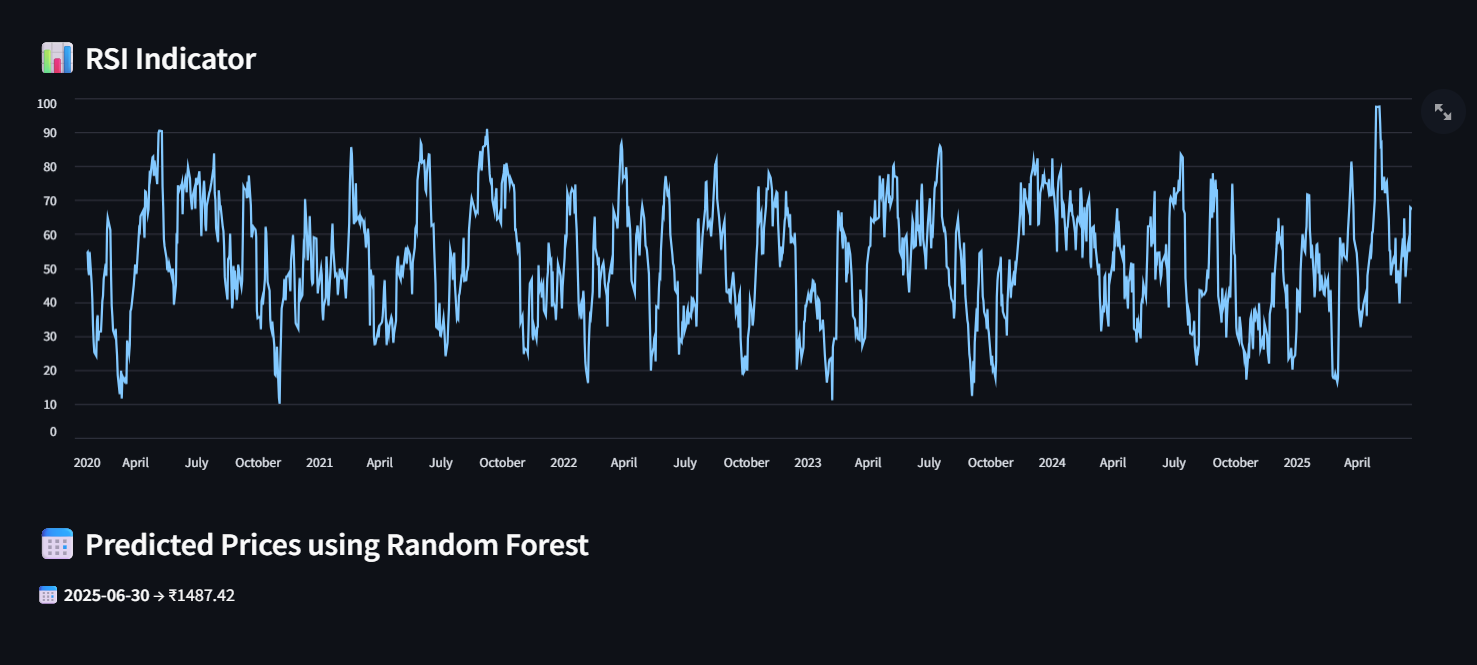
**10. Visual Results**

**Figure 1: Linear Regression – Actual vs Predicted Prices**  
****

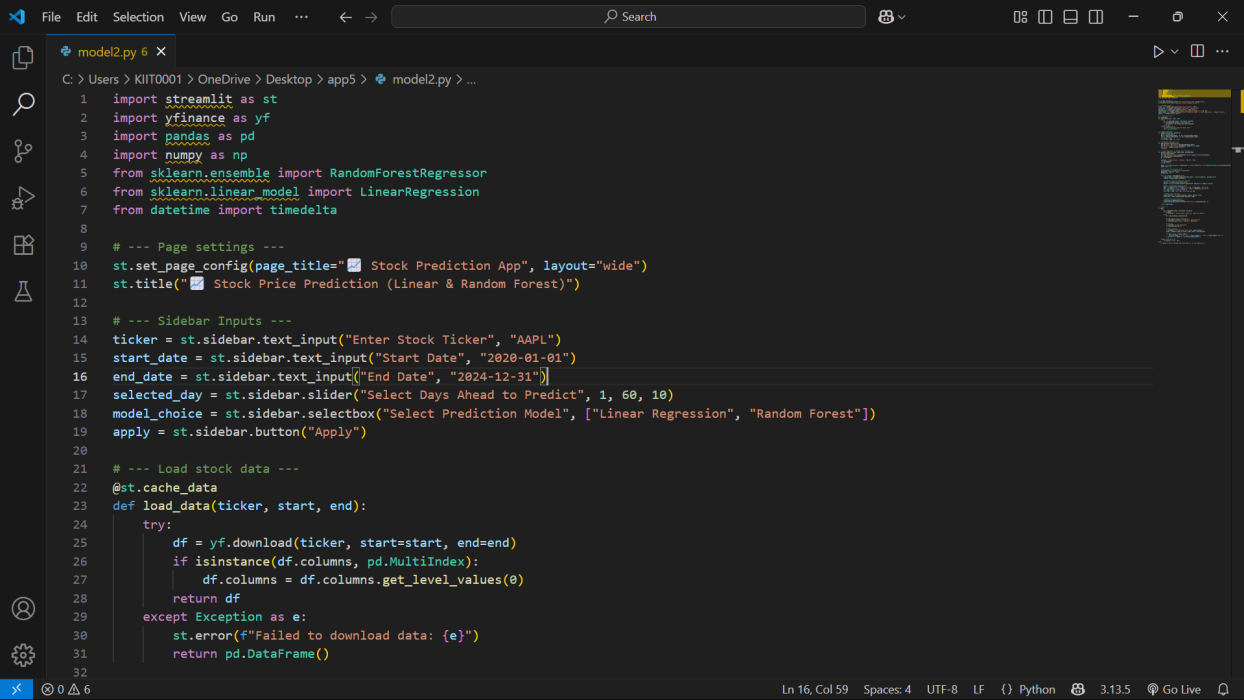
**Figure 2: Forecasted Prices for Next 30 Days**  


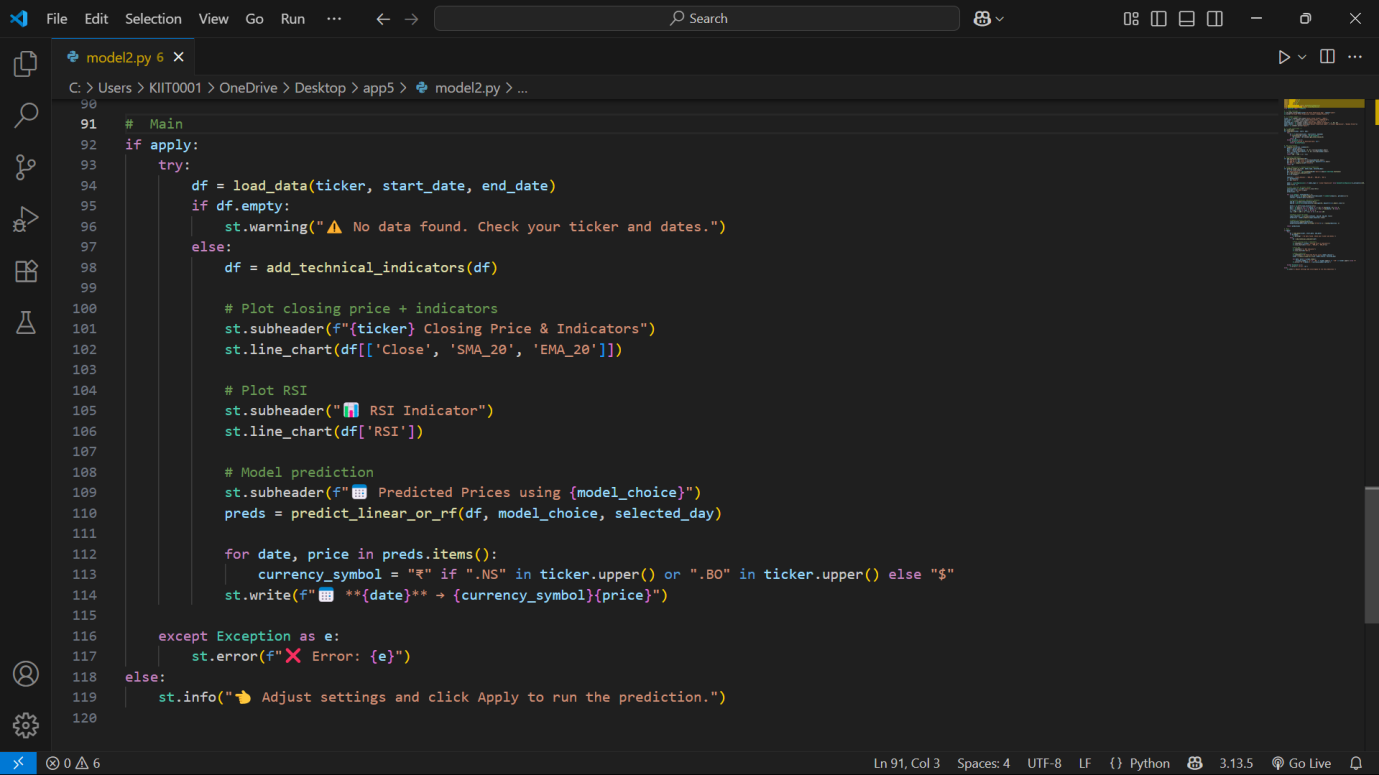
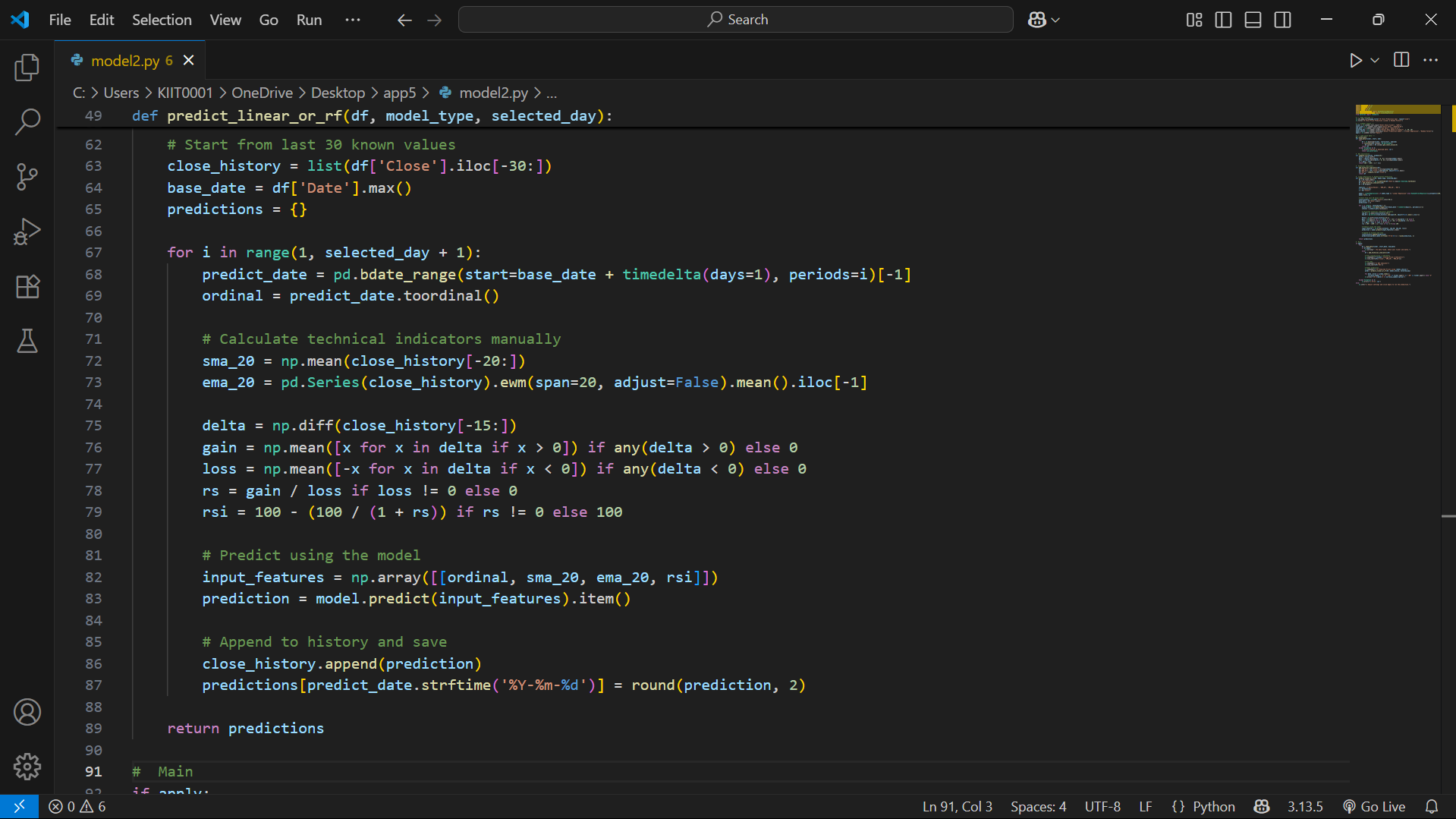
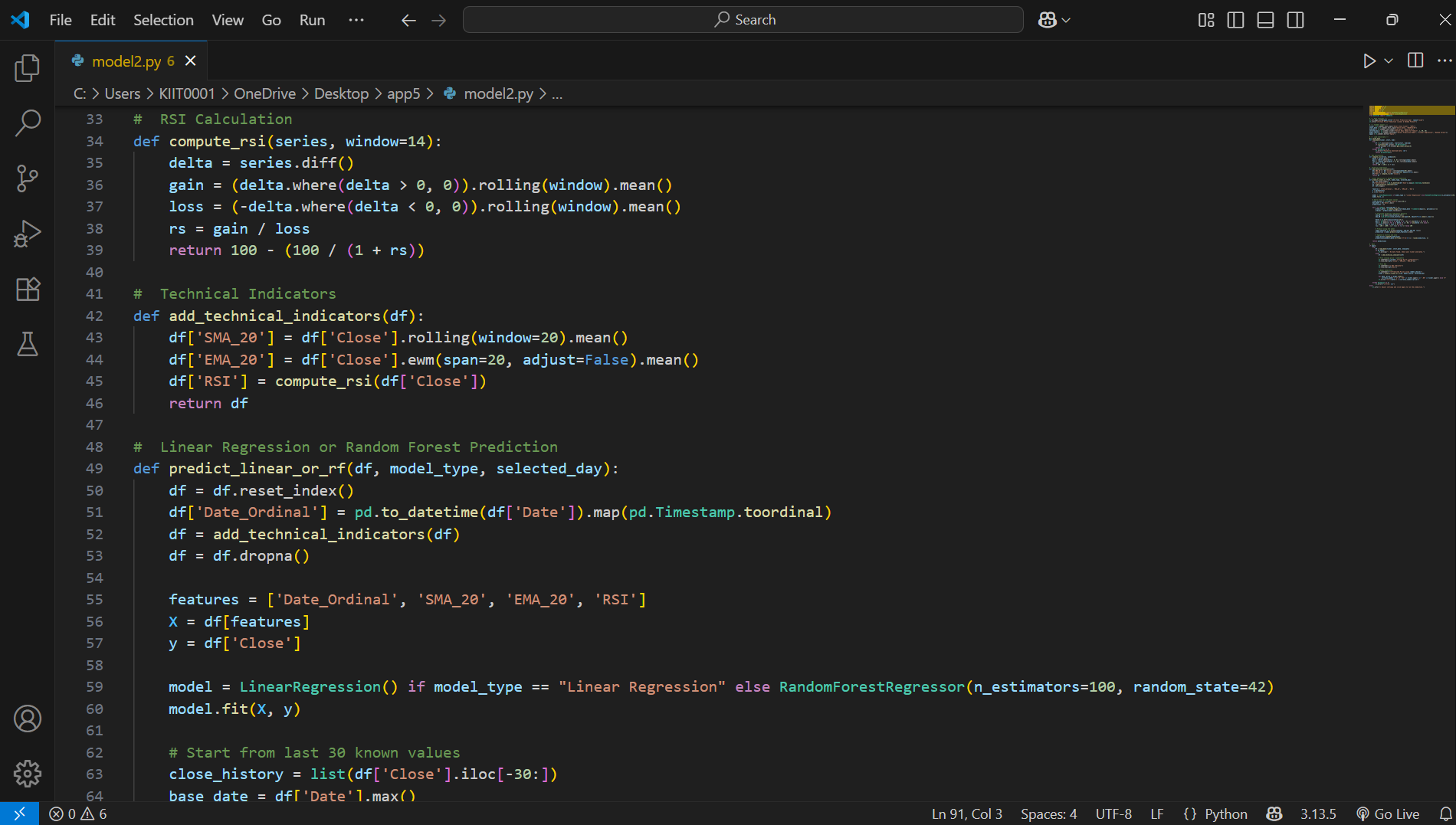
**Figure 3: Streamlit Dashboard - Prediction Interface and Chart View**  


**Figure 4: Technical Indicator View - SMA, EMA, RSI**  


**Figure 5: RSI Indicator with Predicted Output**  


**11. Code Appendix**

The core logic of this project, including data preprocessing, model training, prediction, and visualization, was implemented in Python using the Streamlit framework. Below is a simplified version of the key code components:



Below is the code appendix for the training and testing of the model in Jupyter Notebook