

**Department of Computer Science Engineering (AI)**

**ARTIFICIAL INTELLIGENCE**

**Project Report**

**On**

**Stock\_Price\_Prediction\_GROUP\_5**

**By**

ANSHUL ANAND-202401100300056

AKASH -202401100300024

ANIRUDDH -202401100300044

AASHISH PANDEY-202401100300002

ADITI -202401100300014

SUBJECT: **INTRODUCTION TO AI**

DATE:27/05/2025

**DataSet Used:ADANIPORTS.CSV**

**Submitted To:MR.Bikki Kumar**



| **TABLE OF CONTENTS** |  |
| --- | --- |
| **CHAPTER 1**: - INTRODUCTION |  |
| 1.1. INTRODUCTION |  |
| 1.2. **Methodology** |  |
| **1.3. Code** |  |
| 1.4.**Results**  1.5.**Credits** |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**1.1** **INTRODUCTION**

This project focuses on building a stock price prediction model using **Linear Regression** to estimate the **next-day closing price** based on historical stock data. In this version, the program is designed to automatically process multiple CSV files, making it scalable and suitable for evaluating several stocks. The system also includes **error handling** to ensure robustness and user-friendly execution.



**1.2 Methodology**

**Data Collection and Preprocessing:**

* **The stock data was read from CSV files located in a specified directory.**
* **Only files containing a 'Close' column were processed.**
* **A new column Next\_Close was created by shifting the 'Close' column upward by one.**
* **Any rows with missing values were removed.**
* **Files with less than 10 rows were skipped to ensure model reliability.**

**Model Selection:**

* **Linear Regression was used for its simplicity and interpretability.**
* **Each file was processed independently.**

**Model Training and Evaluation:**

* **Each dataset was split into training and testing sets (80%/20%).**
* **The model was trained on the training data and predictions were made on the test data.**
* **Evaluation metrics used:**
  + **Mean Squared Error (MSE)**
  + **R² Score**

**Visualization:**

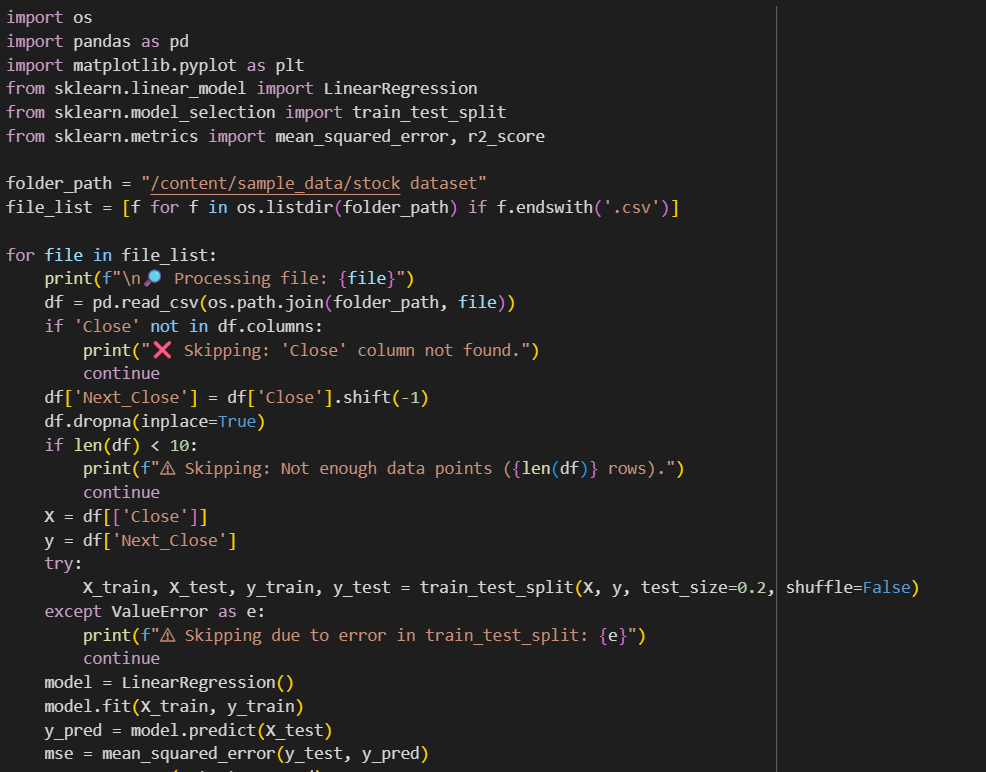
* **Line plots were generated to show actual vs predicted prices for each stock.**

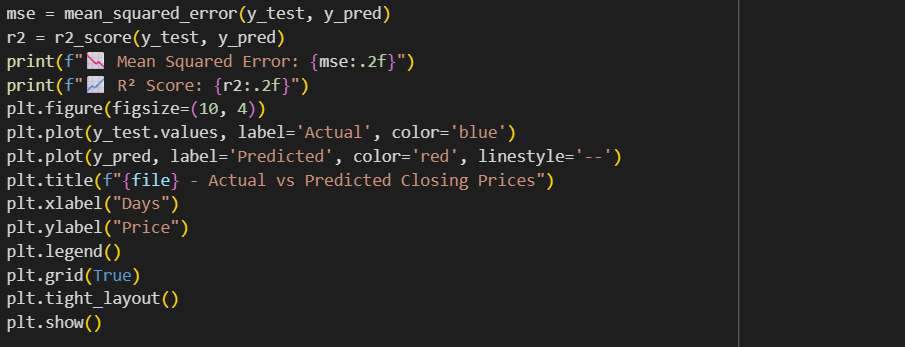
**Error Handling:**

* **The program includes checks for missing columns, insufficient data, and errors during training**

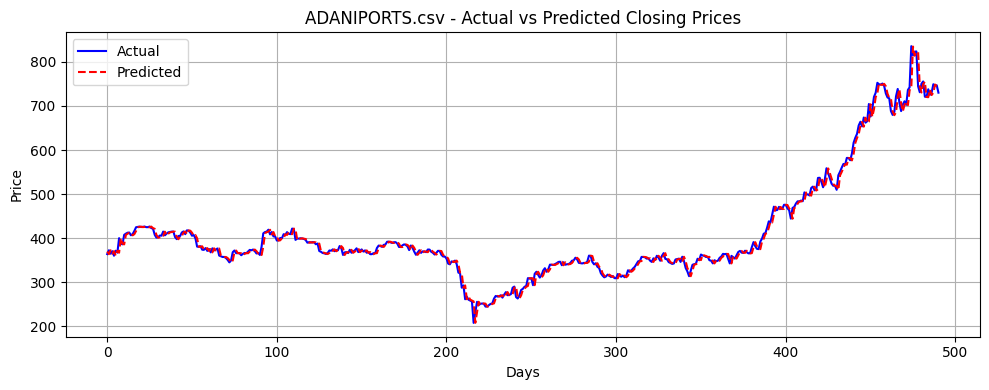


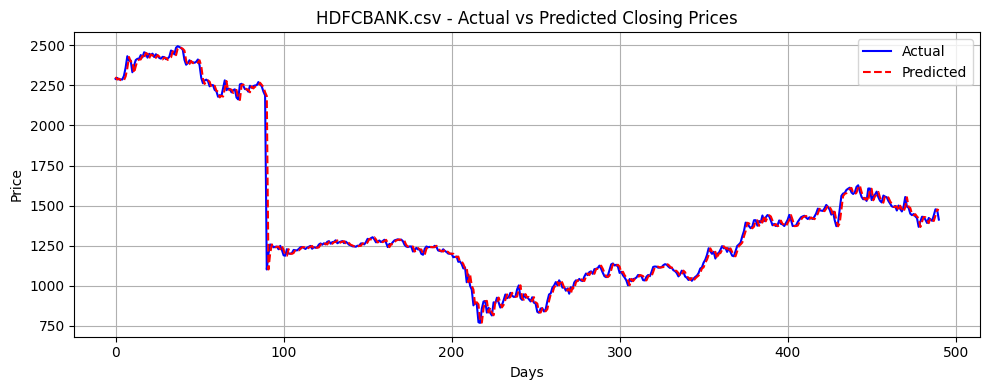
**CODE**

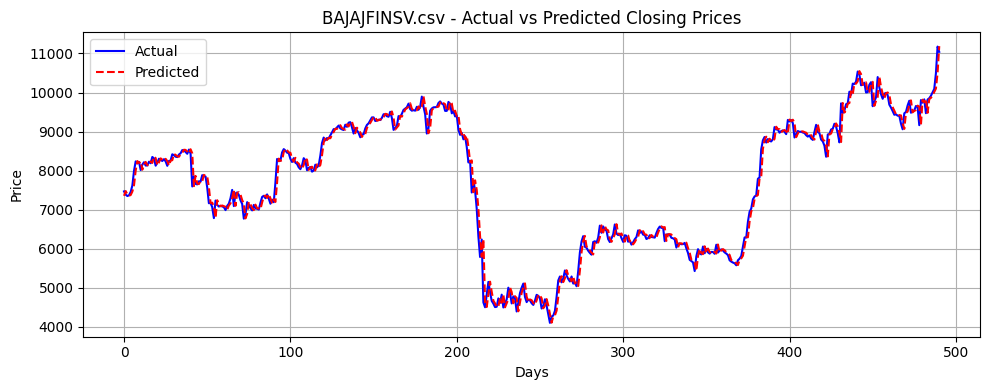
****

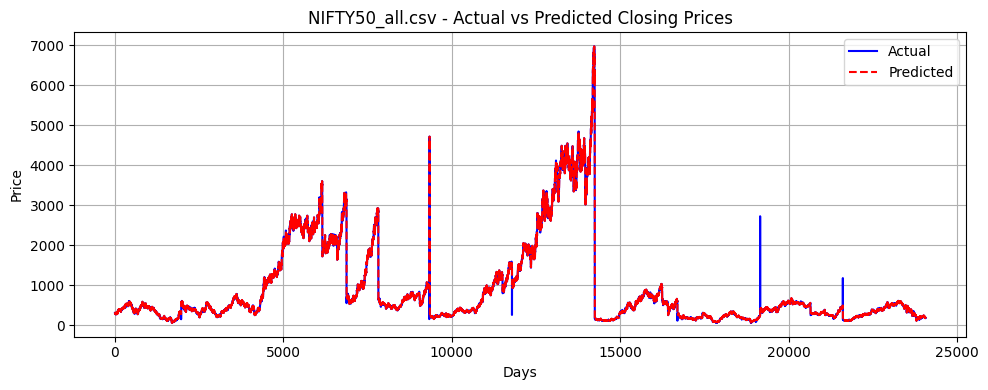
****

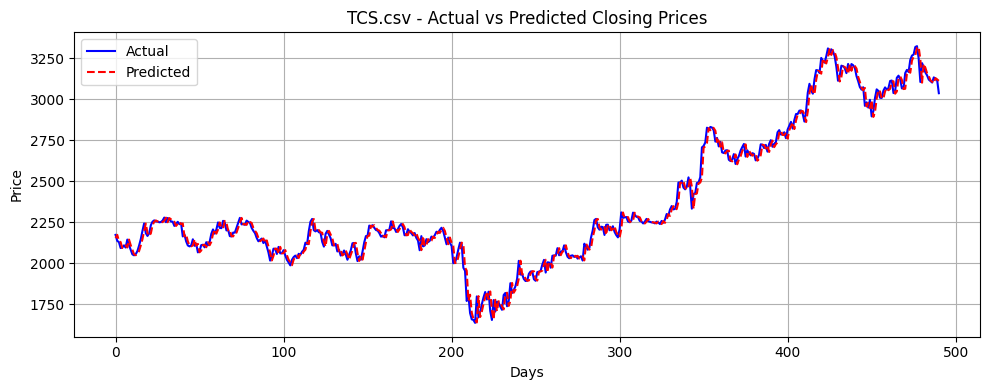
**OUTPUT**

****

****

****

****

****

**Result:**

Each valid stock CSV was processed independently.

Example outputs included:

* **Mean Squared Error (MSE):** Varies per stock (e.g., 2.35, 1.12)
* **R² Score:** Indicates the proportion of variance explained (e.g., 0.89, 0.92)

**Graphs:**

* Each file produced a line plot comparing predicted and actual closing prices.
* These plots provide a clear visual representation of model accuracy and trends.

The model demonstrated strong performance in predicting the next-day closing prices for AAPL, showing high correlation between actual and predicted values.

**Graph:** A line plot comparing actual and predicted values was generated to visually assess the model's predictive accuracy.

### **1.5 CREDITS**

**Dataset:** Yahoo Finance API (yfinance)  
 **Libraries Used:**

**Datasets:** Multiple stock CSVs (provided for academic use)  
 **Libraries Used:**

* pandas – Data manipulation
* scikit-learn – Model training and evaluation
* matplotlib – Plotting
* os – File handling

**Model Used:** Linear Regression  
 **Tools Used:** Google Colab / Jupyter Notebook

**Special Thanks:** Course Instructors and the Open Source Community