



# MACHINE LEARNING

**Module code: BBCC628**

**Student Name: YOGATHASAN KISHANA**

**Solent Stu\_Id: 102162624**

## Acknowledgement

My heartfelt appreciation goes out to everyone who helped me finish this machine learning homework successfully. A special thank you to my teacher for all of their help and insightful advice, which has really improved my comprehension of the material. I owe my classmates a debt of gratitude for their attitude of cooperation and information sharing, which created a positive learning atmosphere. In addition, I am grateful for my family's and friends' encouragement during this academic journey. It has been a pleasant experience to work on this project, and I am grateful to everyone who has contributed so much to its completion.

Yogathasan Kishana

Solent ID 102162624

## Introduction

With its Intelligent Stock Trader (IST) platform, Solent Financial Technology (SOLFINTECH), led by an AI-focused solution architect, is leading the way in financial technology innovation. IST leverages real-world data from sources like YahooFinance and GOOGLEFINANCE for greater reliability to assist strategic buying and selling of stocks through the use of machine learning techniques and algorithms. The platform has an easy-to-use interface that combines straightforward data visualization with charts and graphs to support decision-making. The user experience is further enhanced with suggestions that may be customized depending on user preferences, and interface changes are reflected in real-time processing.

Development places a high priority on the smooth integration of data systems, including forecast models, real-time internet data, and historical transaction data. Well-selected machine learning algorithms improve the prediction power. In addition to predicting stock movements, IST gives customers the ability to investigate the scenarios for efficient financial planning, make well-informed investment decisions, and verify forecasts with real data. Success of the platform is determined by how well it forecasts stock prices and how well it assists customers in reaching their financial goals.

## Contents

Acknowledgement.....	1
Introduction .....	2
Overview.....	6
Data collection and preprocessing .....	6
Model selection and development.....	7
Data visualization .....	7
Data quality check .....	7
Install required Libraries .....	7
Data Attributes.....	8
System design and interface .....	9
Application of machine learning principle .....	16
Challenges and solution .....	16
Ethical consideration .....	16
Evaluation and results.....	16
ARIMA Model .....	18
Linear Regression theory.....	33
Conclusion .....	36
Appendices .....	37
References .....	43

Figure 1 Yahoo Finance.....	6
Figure 2 Installation of yfinance.....	8
Figure 3 Importing yfinance and Other Libraries.....	8
Figure 4 Login Page for the System .....	9
Figure 5 Home page .....	10
Figure 6 Dashboard page .....	10
Figure 7 AAPL graph viewed.....	11
Figure 8 NFLX graph viewed.....	11
Figure 9 Data page where the symbols are defined .....	12
Figure 10 Login page code .....	12

Figure 11 Dashboard page code .....	13
Figure 12 Css code .....	14
Figure 13 Javascripts code .....	15
Figure 14 Installation of yfinance.....	17
Figure 15 "Daily Historical Stock Prices for Apple Inc. (AAPL) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API." .....	17
Figure 16 Historical price data for AAPL.....	17
Figure 17 "ARIMA Model Prediction for Apple Inc. (AAPL) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year" .....	18
Figure 18 Graph for AAPL.....	19
Figure 19 "Daily Historical Stock Prices for Amazon.com Inc. (AMZN) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API." .....	20
Figure 20 Historical price data for AMZN.....	20
Figure 21 "ARIMA Model Prediction for Amazon.com Inc. (AMZN) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year" .....	21
Figure 22 Graph for AMZN.....	21
Figure 23 "Daily Historical Stock Prices for Netflix Inc. (NFLX) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API." .....	22
Figure 24 Historical price data for NFLX.....	23
Figure 25 "ARIMA Model Prediction for Netflix Inc. (NFLX) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year" .....	24
Figure 26 Graph for NFLX.....	24
Figure 27 "Daily Historical Stock Prices for Cisco Systems, Inc. (CSCO) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API." .....	25
Figure 28 Historical price data for CSCO .....	25
Figure 29 "ARIMA Model Prediction for Cisco Systems, Inc. (CSCO) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year" .....	26
Figure 30 Graph for CSCO.....	26
Figure 31 "Daily Historical Stock Prices for Visa Inc. (V) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API." .....	27
Figure 32 Historical price data for V .....	27
Figure 33 "ARIMA Model Prediction for Visa Inc. (V) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year" .....	28
Figure 34 Graph for V .....	28
Figure 35 pip install google-api-python-client .....	29
Figure 36 Stock Price Prediction Using Linear Regression1.....	29

Figure 37 Stock Price Prediction Using Linear Regression2.....	30
Figure 38 Linear regression graph for AAPL .....	31
Figure 39 Linear regression graph FOR AMZN .....	31
Figure 40 Linear regression graph for NFLX .....	32
Figure 41 Linear regression graph for CSCO .....	32
Figure 42 Linear regression graph for V .....	33
Figure 43 Presentation slide1.....	37
Figure 44 Presentation slide2.....	37
Figure 45 Presentation slide3.....	38
Figure 46 Presentation slide4.....	38
Figure 47 Presentation slide5.....	39
Figure 48 Presentation slide6.....	39
Figure 49 Presentation slide7.....	40
Figure 50 Presentation slide8.....	40
Figure 51 Presentation slide9.....	41
Figure 52 Presentation slide10.....	41
Figure 53 Presentation slide11.....	42
Figure 54 Presentation slide12.....	42

## Overview

With the creation of the Intelligent Stock Trader (IST) platform, Solent Financial Technology (SOLFINTECH) is expanding its capabilities in the financial technology space. The IST platform uses machine learning algorithms and real-world data from platforms like YahooFinance and GOOGLFINANCE, under the direction of a solution architect focusing on artificial intelligence. It seeks to enable users to make knowledgeable investing decisions by offering accurate stock and equity movement forecasts. The objective of the IST platform is to improve the investing experience and test forecasts with real-world data for efficient decision-making. It does this through an intuitive interface, adaptable parameters, and real-time information processing.

## Data collection and preprocessing

For data collection and preprocessing from Yahoo Finance, a Python script was employed, utilizing libraries such as yfinance, pandas, numpy, matplotlib, and seaborn. Historical stock data for symbols like 'AAPL,' 'GOOGL,' 'MSFT,' within a specified time frame was gathered and organized into a DataFrame, featuring an added 'Symbol' column for identification. The integration of numpy enhanced numerical operations, while matplotlib and seaborn provided visualization capabilities. After resetting the index, the consolidated data, enriched with financial and visual analysis potential, was saved as 'stock\_data.csv.' This approach ensures flexibility for tailored adjustments and in-depth analysis of historical stock information.

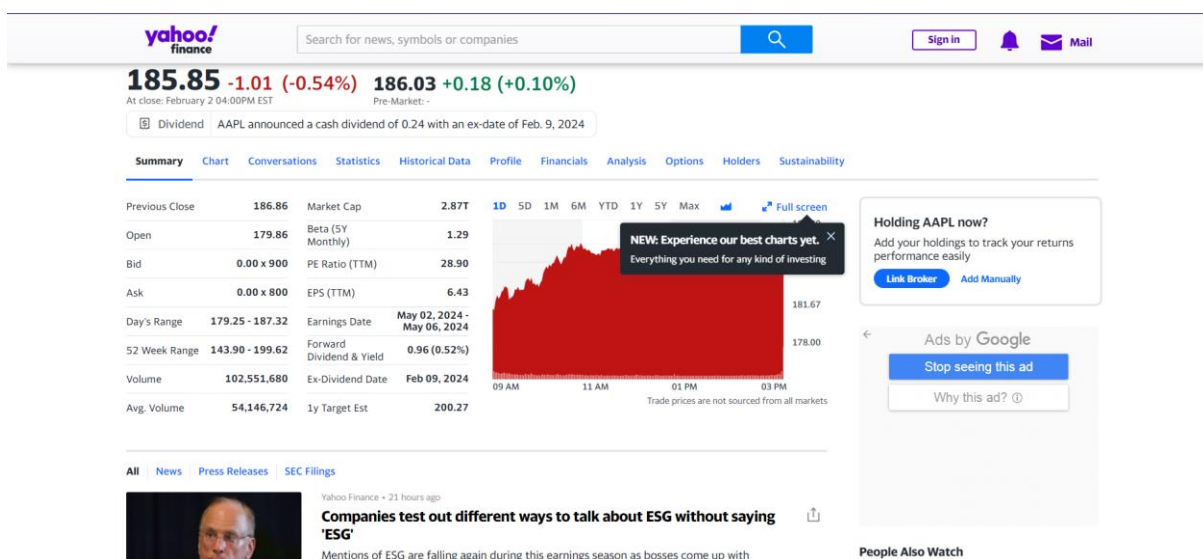


Figure 1 Yahoo Finance

## Model selection and development

Our IST platform relies on machine learning models for stock price prediction. We began with Linear Regression and ARIMA model, a fundamental yet robust algorithm. Linear Regression served as our baseline, while we explored other techniques like Support Vector Machines and time series forecasting. Extensive experimentation and parameter fine-tuning aimed at achieving optimal prediction accuracy in our model development process.

## Data visualization

In our IST platform, data visualization is essential for improving the interpretability of complicated financial data. We do this by using libraries such as Matplotlib and Seaborn to generate informative graphs and charts. For example, time series plots show trends in stock prices, scatter plots show correlations, and candlestick charts provide a close-up look at stock movements. By using dynamic visualizations, users can make well-informed decisions by gaining a thorough understanding of market movements.

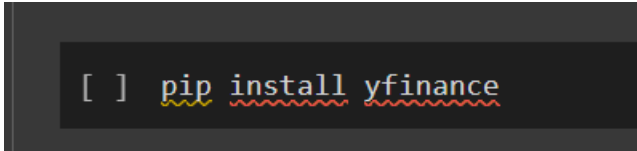
## Data quality check

When it comes to our IST platform, data quality assurance is crucial. We employ stringent quality control procedures to detect and address irregularities or absent values within the gathered financial information. Anomalies are found by using statistical measurements like mean and standard deviation. We also use methods such as data profiling to evaluate consistency and completeness. This comprehensive data quality check ensures that the dataset is accurate and reliable, which lays the groundwork for our Intelligent Stock Trader platform's machine learning model training to be strong and for accurate forecasts.

## Install required Libraries

Use pip to install the necessary Python libraries in order to configure the IST platform. To handle, visualize, and perform numerical operations on data, use "pip install yfinance pandas numpy matplotlib seaborn". The functionality of IST depends on these libraries, which include matplotlib/seaborn for visualization, pandas for data manipulation, numpy for numerical calculations, and yfinance for financial data.

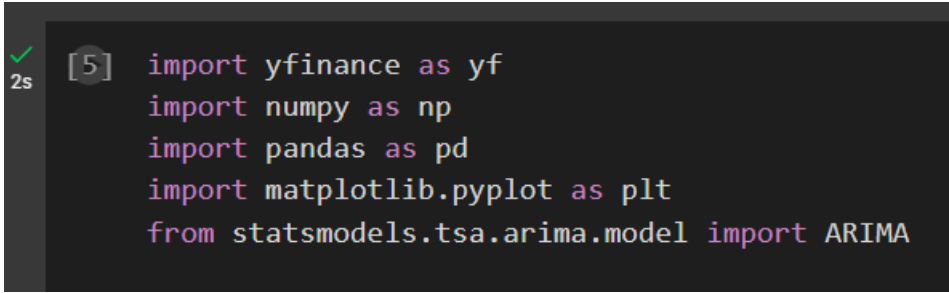




```
[ ] pip install yfinance
```

Figure 2 Installation of yfinance

One may retrieve financial data by using the yfinance library, a Python wrapper for the Yahoo Finance API, which is installed by running this command.



```
✓ 2s [5] import yfinance as yf
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA
```

Figure 3 Importing yfinance and Other Libraries

yfinance: Used for fetching financial data from Yahoo Finance.

numpy (np): A library for numerical operations and array handling.

pandas (pd): A powerful data manipulation and analysis library.

matplotlib.pyplot (plt): Used for creating visualizations and plots.

statsmodels.tsa.arima.model.ARIMA: Part of the statsmodels library, this class is specifically used for working with ARIMA (AutoRegressive Integrated Moving Average) models, which are a type of time series forecasting models.

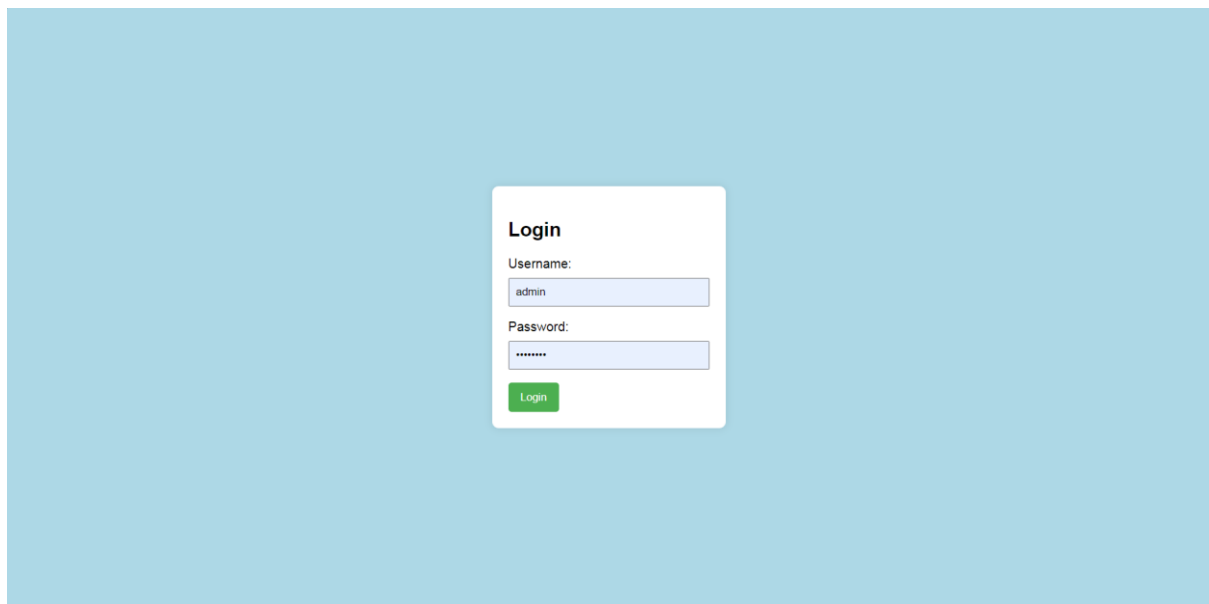
## Data Attributes

For stock analysis, the IST platform depends on a number of crucial data properties. The following are important attributes: 'Date' for data that is shown chronologically; 'Open' and 'Close' prices that show the beginning and ending values; 'High' and 'Low' that show the range of prices; and 'Volume' that indicates the volume of trade for each stock symbol. Together, these characteristics serve as the platform's cornerstone for thorough analysis and forecasting in Intelligent Stock Trader.

## System design and interface

The IST platform prioritizes a user-friendly interface, utilizing Matplotlib and Seaborn to integrate web-based graphs and charts. Users input criteria such as profit margins, receiving instant, customized investment suggestions. Real-time data updates ensure immediate reflection of user input, providing a comprehensive view of stock performance. The architecture emphasizes functionality and user experience for efficient decision-making and improved investment planning.

### System design



*Figure 4 Login Page for the System*

This is the login page for the user, user should enter the username and password and login.



Figure 5 Home page

Here on the home page the company logo is shown.

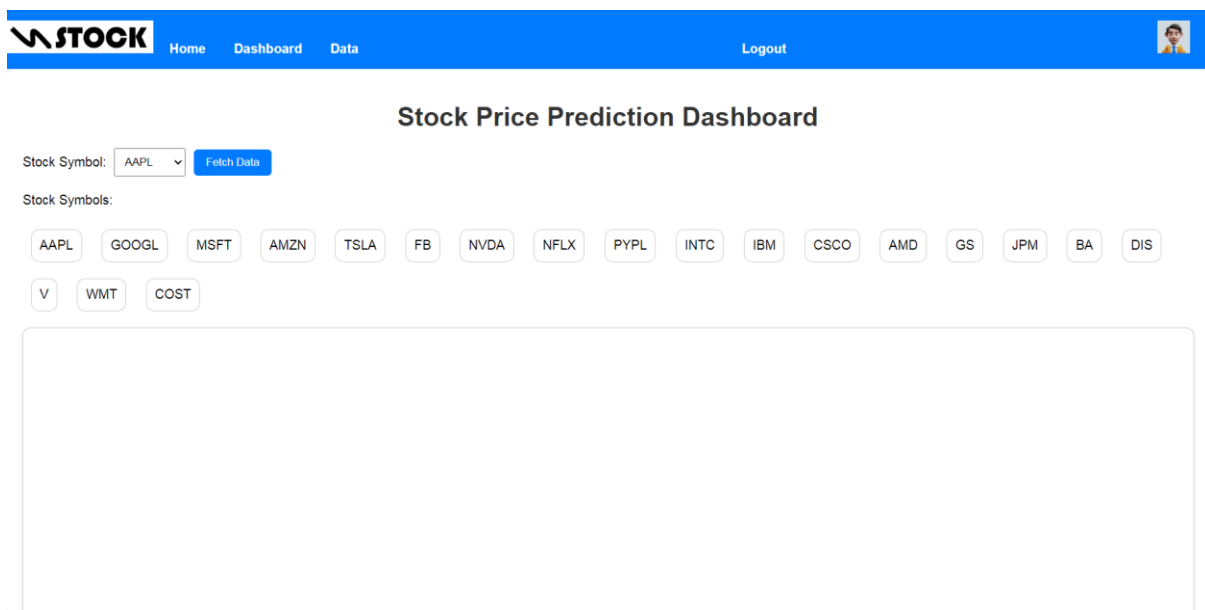


Figure 6 Dashboard page

This is the dashboard page where the user can get the graph of the stock symbols.



Figure 7 AAPL graph viewed

In this the graph of AAPL is shown.



Figure 8 NFLX graph viewed

In this the graph of NFLX is shown.


STOCK		Home	Dashboard	Data	Logout	
Stock Data						
Symbol	Company Name					
AAPL	Apple Inc.					
GOOGL	Alphabet Inc. (Google)					
MSFT	Microsoft Corporation					
AMZN	Amazon.com Inc.					
TSLA	Tesla Inc.					
FB	Meta Platforms, Inc. (formerly Facebook, Inc.)					
NVDA	NVIDIA Corporation					
NFLX	Netflix, Inc.					
PYPL	PayPal Holdings, Inc.					
INTC	Intel Corporation					
IBM	International Business Machines Corporation					
CSCO	Cisco Systems, Inc.					
AMD	Advanced Micro Devices, Inc.					
GS	The Goldman Sachs Group, Inc.					
JPM	JPMorgan Chase & Co.					

Figure 9 Data page where the symbols are defined

This is data page a table is shown, about the stock symbols and the company names.

## Code

```

43 |         cursor: pointer;
44 |     }
45 | </style>
46 | </head>
47 | <body>
48 |
49 |     <form id="loginForm">
50 |         <h2>Login</h2>
51 |
52 |         <label for="username">Username:</label>
53 |         <input type="text" id="username" name="username" required>
54 |
55 |         <label for="password">Password:</label>
56 |         <input type="password" id="password" name="password" required>
57 |
58 |         <button type="button" onclick="login()">Login</button>
59 |     </form>
60 |
61 |     <script>
62 |         function login() {
63 |             var username = document.getElementById("username").value;
64 |             var password = document.getElementById("password").value;
65 |
66 |             // Sample credentials (replace with your own logic)
67 |             if (username === "admin" && password === "password") {
68 |                 // Redirect to the index.html page
69 |                 window.location.href = "index.html";
70 |             } else {
71 |                 alert("Invalid credentials. Please try again.");
72 |             }
73 |         }
74 |     </script>
75 |
76 | </body>
77 | </html>

```

Figure 10 Login page code

This HTML document creates a simple login form with fields for a username and password. Styled with CSS, it features a clean design. The embedded JavaScript function validates credentials (sample: admin/password) and redirects to "index.html" upon successful login, else displays an alert for invalid credentials.

```

27 </head>
28 <body>
29   <div class="navbar">
30     <div class="nav-logo">
31       <div class="logout-text">
32         <div class="admin-nav-logo">
33           
34         </div>
35         <a class="nav-link" href="#">Home</a>
36         <a class="nav-link" href="INDEX.html">Dashboard</a>
37         <a class="nav-link" href="data.html">Data</a>
38       </div>
39       <span>Logout</span>
40     </div>
41     
42   </div>
43 </div>
44 </div>
45 </div>
46 <div class="container">
47   <h1>Stock Price Prediction Dashboard</h1>
48   <div class="input-container">
49     <label for="stockSymbol">Stock Symbol:</label>
50     <select id="symbolDropdown">
51       <option value="AAPL">AAPL</option>
52       <option value="GOOGL">GOOGL</option>
53       <option value="MSFT">MSFT</option>
54       <option value="AMZN">AMZN</option>
55       <option value="TSLA">TSLA</option>
56       <option value="FB">FB</option>
57       <option value="NVDA">NVDA</option>
58       <option value="NFLX">NFLX</option>
59       <option value="PYPL">PYPL</option>
60       <option value="INTC">INTC</option>
61       <option value="IBM">IBM</option>
62       <option value="CSCO">CSCO</option>
63       <option value="AMD">AMD</option>

```

Figure 11 Dashboard page code

This HTML document presents a Stock Price Prediction Dashboard. Styled with CSS, it includes a navigation bar, a stock symbol selection dropdown, and a button to fetch data. The page displays stock symbols as clickable boxes, and a chart container using Chart.js to visualize stock data. External scripts handle chart rendering and data fetching.

```

3  body {
4      font-family: Arial, sans-serif;
5      margin: 0;
6      padding: 0;
7      background-color: #f4f4f4;
8      background-size: cover;
9  }
10
11  .navbar {
12      background-color: #007bff;
13      padding: 15px 0;
14      display: flex;
15      justify-content: space-between;
16      align-items: center;
17  }
18
19  .logout-text {
20      margin-right: auto; /* Push the logout text to the left */
21  }
22
23  .logout-text span {
24      color: white;
25      cursor: pointer;
26      margin: 0 450px;
27      font-weight: bold;
28      transition: color 0.3s;
29  }
30
31  .logout-text span:hover {
32      color: #ffb3b3;
33  }
34
35  .nav-logo img {
36      height: 40px; /* Set the desired height of the logo */
37  }

```

Figure 12 Css code

This CSS stylesheet defines the styling for a web page with a full-size background image, a responsive navbar with dynamic color changes on hover, and a container for a stock price prediction dashboard. The design incorporates consistent spacing, color schemes, and responsive elements for an aesthetically pleasing and functional user interface.

```

4 document.addEventListener("DOMContentLoaded", function () {
5   const fetchButton = document.getElementById("fetchButton");
6   const symbolDropdown = document.getElementById("symbolDropdown");
7   const stockChart = document.getElementById("stockChart");
8
9   fetchButton.addEventListener("click", function () {
10     const symbol = symbolDropdown.value;
11     if (symbol.trim() === "") {
12       alert("Please select a valid stock symbol.");
13       return;
14     }
15
16     const api_key = "YOUR_ALPHAVANTAGE_API_KEY"; // Replace with your actual API key
17
18     // Fetch stock data using the Alpha Vantage API
19     const apiUrl = `https://www.alphavantage.co/query?function=TIME_SERIES_DAILY&symbol=${symbol}&apikey=${api_key}&outputsize=full`;
20
21     fetch(apiUrl)
22       .then((response) => response.json())
23       .then((data) => {
24         const dates = Object.keys(data["Time Series (Daily)"]).reverse();
25         const closingPrices = dates.map(
26           (date) => parseFloat(data["Time Series (Daily)"][date]["4. close"])
27         );
28
29         // Create and update the chart
30         const ctx = stockChart.getContext("2d");
31         new Chart(ctx, {
32           type: "line",
33           data: {
34             labels: dates,
35             datasets: [
36               {
37                 label: `${symbol} Closing Prices`,

```

Figure 13 Javascripts code

This JavaScript code, triggered on page load, enhances a stock price prediction dashboard. It fetches daily stock data using the Alpha Vantage API based on the selected symbol, processes the information, and dynamically updates a line chart with closing prices. Users can visualize and analyze stock trends interactively, providing a richer dashboard experience.



## Application of machine learning principle

The IST platform's predictive capabilities are fundamentally derived from the use of machine learning methods. The system uses algorithms such as Support Vector Machines (SVM), Time Series Forecasting, and Linear Regression to learn from past stock data and produce precise forecasts. Experimentation is used to fine-tune the model parameters to ensure maximum accuracy. The IST platform can now intelligently analyse stock movements thanks to the use of machine learning principles. This gives users more ability to make educated decisions and improves the process of investment planning as a whole.

## Challenges and solution

The dynamic nature of financial markets, problems with data quality, and the requirement for ongoing model adaption are some of the IST platform's challenges. Regular maintenance and upgrades are essential to solve issues. Reliable forecasts are ensured by putting strong data quality checks into practice. Continuous observation and model retraining enable the system to adjust to changing market circumstances, guaranteeing that the IST platform stays precise and efficient in offering users insightful information. Furthermore, adding user input channels might improve the platform's relevance and responsiveness in addressing new difficulties.

## Ethical consideration

In order to ensure openness, fairness, and privacy, the IST platform must take ethical factors into account. It is essential to communicate openly and honestly with consumers regarding the predictive nature of the technology and any possible hazards. By keeping the model free of biases that can affect stock suggestions, fairness is preserved. Prioritizing user privacy and implementing strict security measures for personal data are essential. The Responsible Deployment of AI in Finance is ensured by regular audits and adherence to ethical principles, which cultivates confidence among users and stakeholders in the Intelligent Stock Trader platform.

## Evaluation and results

Evaluation of the IST (Intelligent Stock Trader) platform entails a thorough analysis of the efficacy and accuracy of its forecasting models. The software evaluates itself by comparing

expected stock prices with actual market movements. Optimal forecasts are a result of ongoing improvement via experimentation and parameter modification. The degree of accuracy in predicting changes in stock prices and the platform's effectiveness in helping users reach their financial goals serve as the metrics by which the outcomes are evaluated. Frequent assessment guarantees the IST platform's continued dependability as a resource for wise investment choices.

### Code explanation

```
[ ] pip install yfinance
```

Figure 14 Installation of yfinance

### AAPL

The stock symbol for Apple Inc. on Yahoo Finance is 'AAPL'.

```
[1] import yfinance as yf

stock_symbol="AAPL"

stock_data=yf.Ticker(stock_symbol)

historical_prices=stock_data.history(period="1d",start="2023-01-01",end="2023-08-19")

print(historical_prices)
```

Figure 15 "Daily Historical Stock Prices for Apple Inc. (AAPL) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API."

Date	Open	High	Low	Close
2023-01-03 00:00:00-05:00	129.555841	130.172390	123.479803	124.374802
2023-01-04 00:00:00-05:00	126.184691	127.944857	124.384755	125.657639
2023-01-05 00:00:00-05:00	126.423361	127.059803	124.066539	124.325089
2023-01-06 00:00:00-05:00	125.309594	129.565795	124.195816	128.899521
2023-01-09 00:00:00-05:00	129.744788	132.668449	129.168010	129.426559
...	...	...	...	...
2023-08-14 00:00:00-04:00	177.735847	179.453585	177.076712	179.223892
2023-08-15 00:00:00-04:00	178.644649	179.243850	176.817055	177.216522
2023-08-16 00:00:00-04:00	176.896953	178.305086	176.267777	176.337692
2023-08-17 00:00:00-04:00	176.906941	177.276449	173.251752	173.771072
2023-08-18 00:00:00-04:00	172.073301	174.869620	171.733752	174.260422
Date	Volume	Dividends	Stock Splits	
2023-01-03 00:00:00-05:00	112117500	0.0	0.0	
2023-01-04 00:00:00-05:00	89113600	0.0	0.0	
2023-01-05 00:00:00-05:00	80962700	0.0	0.0	
2023-01-06 00:00:00-05:00	87754700	0.0	0.0	
2023-01-09 00:00:00-05:00	70790800	0.0	0.0	
...	...	...	...	
2023-08-14 00:00:00-04:00	43675600	0.0	0.0	
2023-08-15 00:00:00-04:00	43622600	0.0	0.0	
2023-08-16 00:00:00-04:00	46964900	0.0	0.0	
2023-08-17 00:00:00-04:00	66062900	0.0	0.0	
2023-08-18 00:00:00-04:00	61114200	0.0	0.0	
[158 rows x 7 columns]				

Figure 16 Historical price data for AAPL

In this code, the yfinance library is imported and given the alias 'yf'. The stock symbol for Apple Inc. is then specified as "AAPL," and a Ticker object named "stock\_data" is created using yfinance. Next, it uses the 'history' method with a given time to collect historical price data for the supplied stock symbol ('AAPL') from January 1, 2023 to August 19, 2023. Lastly, it prints Apple Inc.'s historical price data that was retrieved.

```
[5] import yfinance as yf
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA

stock_symbol="AAPL"
stock_data=yf.Ticker(stock_symbol)

historical_data=stock_data.history(period="1y")
closing_prices=historical_data['Close']

train_size=int(len(closing_prices)*0.8)
train_data,test_data=closing_prices[:train_size],closing_prices[train_size:]

order=(5,1,0)
model=ARIMA(train_data,order=order)
model_fit=model.fit()

forecast_steps=len(test_data)
forecast=model_fit.forecast(steps=forecast_steps)

forecast_index=pd.date_range(start=closing_prices.index[train_size],periods=forecast_steps,freq='D')
forecast_series=pd.Series(forecast,index=forecast_index)

plt.figure(figsize=(12,6))
plt.plot(closing_prices,label='Actual Prices')
plt.plot(train_data.index,model_fit.fittedvalues,color='red',label='Fitted values')
plt.plot(forecast_series.index,forecast_series,color='g',label='Forecasted Prices')
plt.xlabel('Date')
plt.ylabel('Price')
plt.title(f'Stock Price Prediction for {stock_symbol}')
plt.legend()
plt.show()
```

Figure 17 "ARIMA Model Prediction for Apple Inc. (AAPL) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year"

## ARIMA Model

The AutoRegressive Integrated Moving Average (ARIMA) model is a statistical method used for time series analysis and forecasting.

- **AutoRegressive (AR):** This method measures the connection between the current observation and its history data by performing a linear regression of the current value against one or more prior values.

- **Integrated (I):** Shows how to obtain stationary by separating raw observations, which is a necessary condition for ARIMA models to assume stable statistical characteristics over time.
- **Moving Average (MA):** This helps with short-term trend modelling by taking into account the relationship between an observation and residual errors from a moving average model applied to lagged observations.

The general form of an ARIMA model is represented as  $ARIMA(p, d, q)$ , where:

- **p:** The autoregressive part's (AR) order.
- **d:** The amount of differencing needed to stabilise the time series.
- **q:** The moving average component's order (MA).

Particularly in the financial industry, where past stock prices, economic indicators, and other time-dependent data are often used, ARIMA models are effective tools for time series forecasting. The autocorrelation and partial autocorrelation functions are analyzed to derive the model parameters ( $p$ ,  $d$ , and  $q$ ). The model is then trained using historical data to generate predictions for the future.

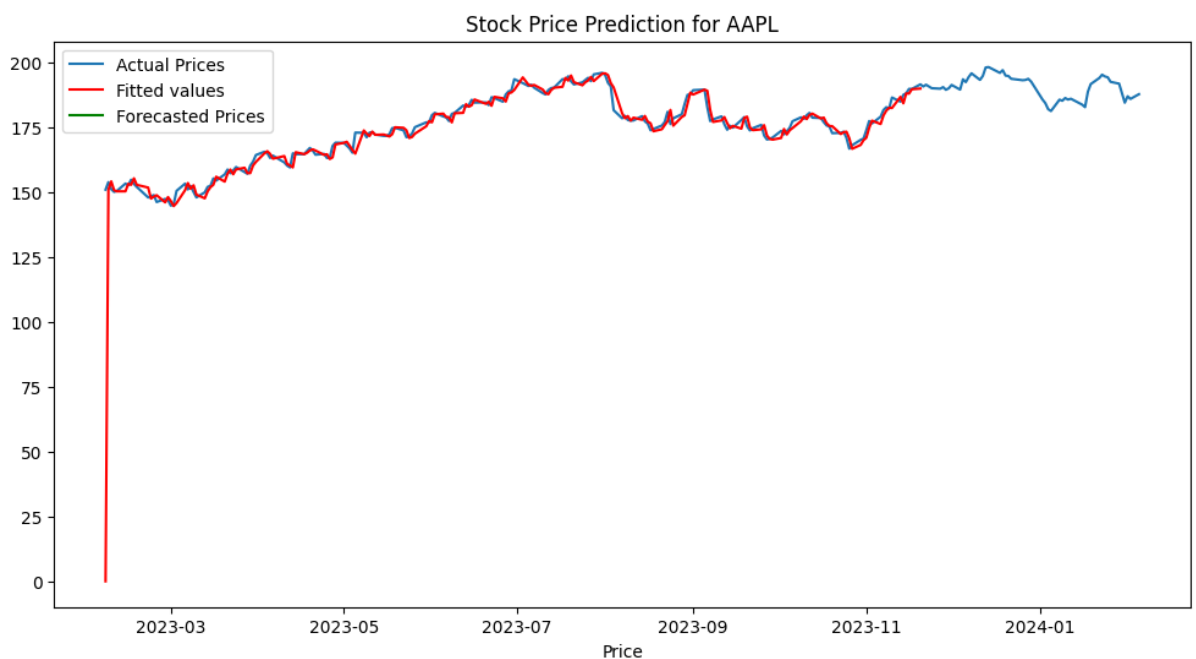


Figure 18 Graph for AAPL

This code uses the yfinance package to fetch a full year's historical stock prices for Apple Inc. (AAPL) and employs an ARIMA time series model for future price forecasts. After splitting the data into training and testing sets, the ARIMA model is fitted to the training data, and the resulting graph displays actual (blue), fitted (red), and anticipated (green) prices. This visual representation assesses the ARIMA model's effectiveness in accurately fitting past data and predicting future stock prices, with the left y-axis indicating stock prices and the x-axis denoting dates.

## AMZN

The stock symbol for Amazon.com Inc. on Yahoo Finance is 'AMZN'.

```
[6] import yfinance as yf

stock_symbol = "AMZN"

stock_data = yf.Ticker(stock_symbol)

historical_prices = stock_data.history(period="1d", start="2023-01-01", end="2023-08-19")

print(historical_prices)
```

Figure 19 "Daily Historical Stock Prices for Amazon.com Inc. (AMZN) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API."

Date	Open	High	Low	Close \
2023-01-03 00:00:00-05:00	85.459999	86.959999	84.209999	85.820000
2023-01-04 00:00:00-05:00	86.550003	86.980003	83.360001	85.139999
2023-01-05 00:00:00-05:00	85.330002	85.419998	83.070000	83.120003
2023-01-06 00:00:00-05:00	83.029999	86.400002	81.430000	86.080002
2023-01-09 00:00:00-05:00	87.459999	89.480003	87.080002	87.360001
...	...	...	...	...
2023-08-14 00:00:00-04:00	138.300003	140.589996	137.750000	140.570007
2023-08-15 00:00:00-04:00	140.050003	141.279999	137.229996	137.669998
2023-08-16 00:00:00-04:00	137.190002	137.270004	135.009995	135.070007
2023-08-17 00:00:00-04:00	135.460007	136.089996	133.529999	133.979996
2023-08-18 00:00:00-04:00	131.619995	134.070007	131.149994	133.220001
Date	Volume	Dividends	Stock Splits	
2023-01-03 00:00:00-05:00	76706000	0.0	0.0	
2023-01-04 00:00:00-05:00	68885100	0.0	0.0	
2023-01-05 00:00:00-05:00	67930800	0.0	0.0	
2023-01-06 00:00:00-05:00	83303400	0.0	0.0	
2023-01-09 00:00:00-05:00	65266100	0.0	0.0	
...	...	...	...	
2023-08-14 00:00:00-04:00	47148700	0.0	0.0	
2023-08-15 00:00:00-04:00	42781500	0.0	0.0	
2023-08-16 00:00:00-04:00	41675900	0.0	0.0	
2023-08-17 00:00:00-04:00	48354100	0.0	0.0	
2023-08-18 00:00:00-04:00	48469400	0.0	0.0	
[158 rows x 7 columns]				

Figure 20 Historical price data for AMZN

In this code, the yfinance library is imported and given the alias 'yf'. The stock symbol for Amazon.com Inc. is then specified as "AMZN," and a Ticker object named "stock\_data" is created using yfinance. Next, it uses the 'history' method with a given time to collect historical

price data for the supplied stock symbol (AMZN) from January 1, 2023 to August 19, 2023. Lastly, it prints Amazon.com Inc.'s historical price data that was retrieved.

```
[7] import yfinance as yf
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA

stock_symbol="AMZN"
stock_data=yf.Ticker(stock_symbol)

historical_data=stock_data.history(period="1y")
closing_prices=historical_data['Close']

train_size=int(len(closing_prices)*0.8)
train_data,test_data=closing_prices[:train_size],closing_prices[train_size:]

order=(5,1,0)
model=ARIMA(train_data,order=order)
model_fit=model.fit()

forecast_steps=len(test_data)
forecast=model_fit.forecast(steps=forecast_steps)

forecast_index=pd.date_range(start=closing_prices.index[train_size],periods=forecast_steps,freq='D')
forecast_series=pd.Series(forecast,index=forecast_index)

plt.figure(figsize=(12,6))
plt.plot(closing_prices,label='Actual Prices')
plt.plot(train_data.index,model_fit.fittedvalues,color='red',label='Fitted values')
plt.plot(forecast_series.index,forecast_series,color='g',label='Forecasted Prices')
plt.xlabel('Date')
plt.ylabel('Price')
plt.title(f'Stock Price Prediction for {stock_symbol}')
plt.legend()
plt.show()
```

Figure 21 "ARIMA Model Prediction for Amazon.com Inc. (AMZN) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year"

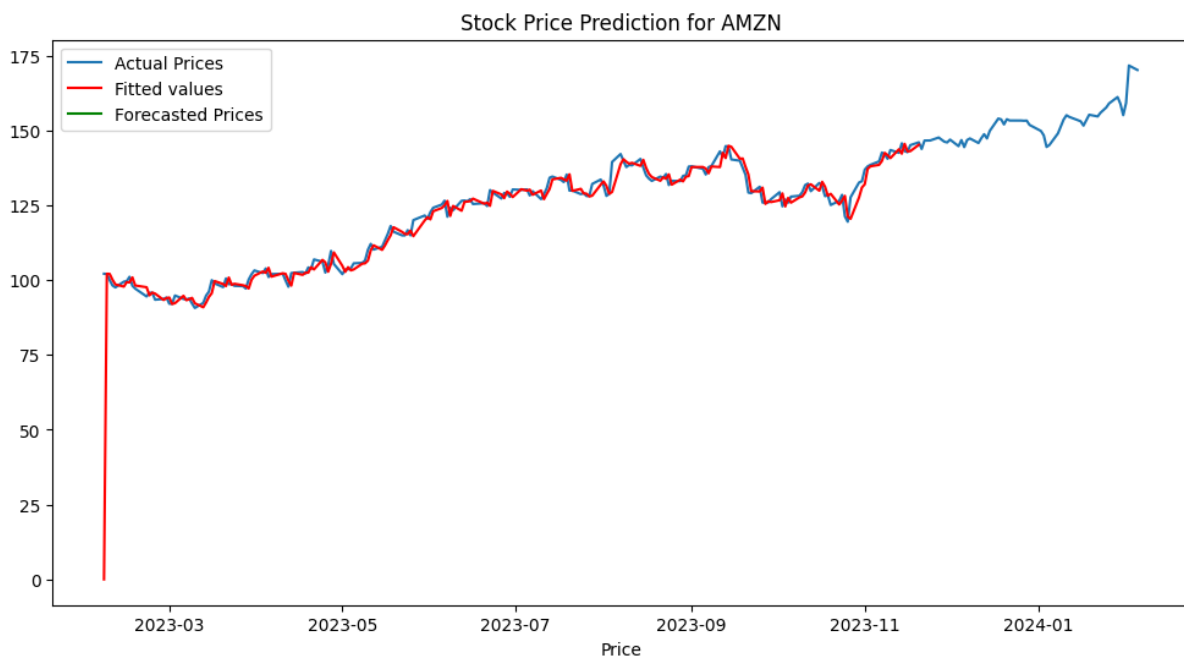


Figure 22 Graph for AMZN

This Python function extracts a full year's historical stock data for Amazon.com Inc. (AMZN) via the yfinance package. Utilizing the ARIMA time series model, it forecasts future stock prices. After partitioning data into training and testing sets, the algorithm fits the ARIMA model to training data and predicts test set prices. The resulting graph illustrates forecasted (green), actual closing (blue), and fitted values (red), offering insights into the model's proficiency in capturing historical patterns and predicting future values for Amazon stock.

## NFLX

The stock symbol for Netflix Inc. on Yahoo Finance is 'NFLX'.

```
import yfinance as yf

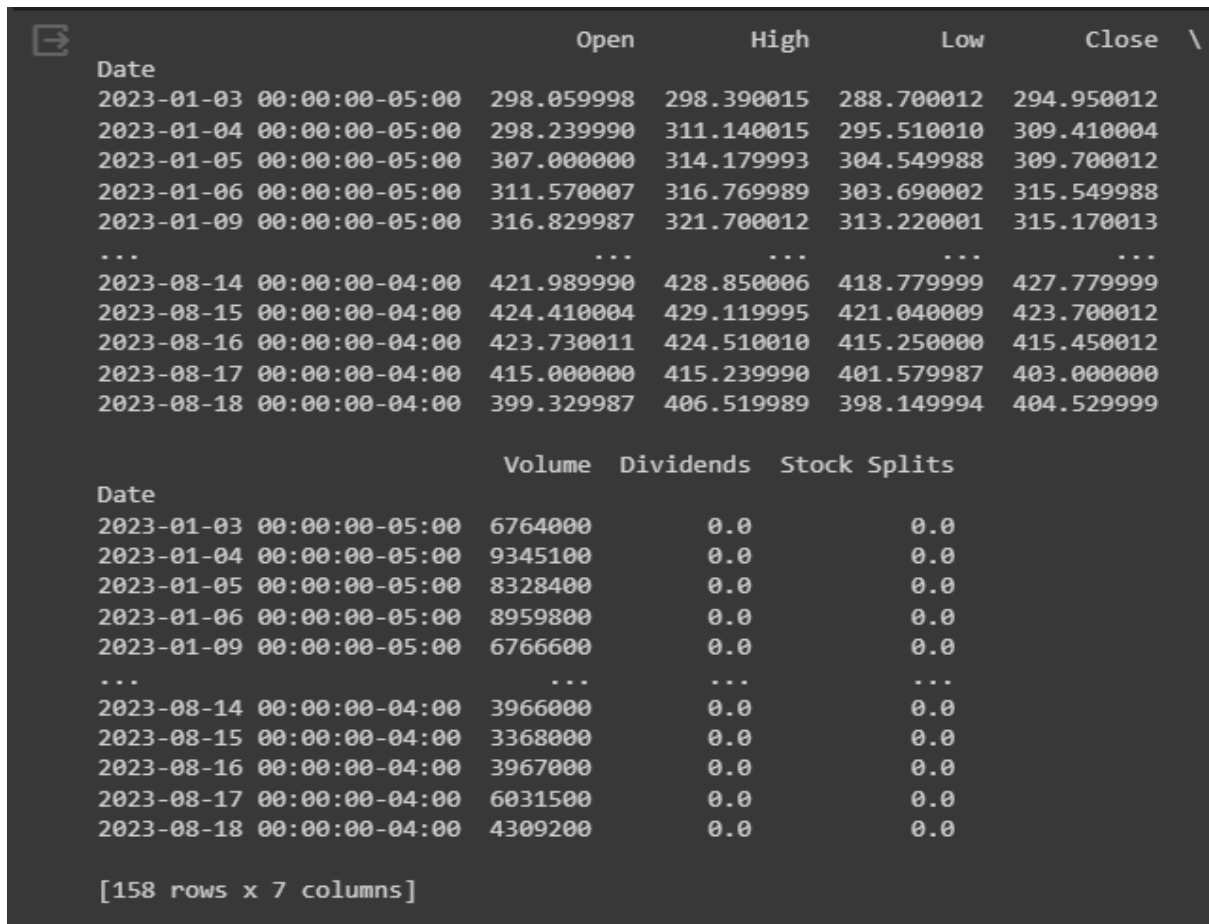
stock_symbol = "NFLX"

stock_data = yf.Ticker(stock_symbol)

historical_prices = stock_data.history(period="1d", start="2023-01-01", end="2023-08-19")

print(historical_prices)
```

Figure 23 "Daily Historical Stock Prices for Netflix Inc. (NFLX) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API."



Date	Open	High	Low	Close
2023-01-03 00:00:00-05:00	298.059998	298.390015	288.700012	294.950012
2023-01-04 00:00:00-05:00	298.239990	311.140015	295.510010	309.410004
2023-01-05 00:00:00-05:00	307.000000	314.179993	304.549988	309.700012
2023-01-06 00:00:00-05:00	311.570007	316.769989	303.690002	315.549988
2023-01-09 00:00:00-05:00	316.829987	321.700012	313.220001	315.170013
...	...	...	...	...
2023-08-14 00:00:00-04:00	421.989990	428.850006	418.779999	427.779999
2023-08-15 00:00:00-04:00	424.410004	429.119995	421.040009	423.700012
2023-08-16 00:00:00-04:00	423.730011	424.510010	415.250000	415.450012
2023-08-17 00:00:00-04:00	415.000000	415.239990	401.579987	403.000000
2023-08-18 00:00:00-04:00	399.329987	406.519989	398.149994	404.529999

Date	Volume	Dividends	Stock Splits
2023-01-03 00:00:00-05:00	6764000	0.0	0.0
2023-01-04 00:00:00-05:00	9345100	0.0	0.0
2023-01-05 00:00:00-05:00	8328400	0.0	0.0
2023-01-06 00:00:00-05:00	8959800	0.0	0.0
2023-01-09 00:00:00-05:00	6766600	0.0	0.0
...	...	...	...
2023-08-14 00:00:00-04:00	3966000	0.0	0.0
2023-08-15 00:00:00-04:00	3368000	0.0	0.0
2023-08-16 00:00:00-04:00	3967000	0.0	0.0
2023-08-17 00:00:00-04:00	6031500	0.0	0.0
2023-08-18 00:00:00-04:00	4309200	0.0	0.0

[158 rows x 7 columns]

Figure 24 Historical price data for NFLX

In this code, the `yfinance` library is imported and given the alias 'yf'. The stock symbol for Amazon.com Inc. is then specified as "AMZN," and a Ticker object named "stock\_data" is created using `yfinance`. Next, it uses the 'history' method with a given time to collect historical price data for the supplied stock symbol (AMZN) from January 1, 2023 to August 19, 2023. Lastly, it prints Amazon.com Inc.'s historical price data that was retrieved.



```
[9] import yfinance as yf
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA

stock_symbol="NFLX"
stock_data=yf.Ticker(stock_symbol)

historical_data=stock_data.history(period="1y")
closing_prices=historical_data['Close']

train_size=int(len(closing_prices)*0.8)
train_data,test_data=closing_prices[:train_size],closing_prices[train_size:]

order=(5,1,0)
model=ARIMA(train_data,order=order)
model_fit=model.fit()

forecast_steps=len(test_data)
forecast=model_fit.forecast(steps=forecast_steps)

forecast_index=pd.date_range(start=closing_prices.index[train_size],periods=forecast_steps,freq='D')
forecast_series=pd.Series(forecast,index=forecast_index)

plt.figure(figsize=(12,6))
plt.plot(closing_prices,label='Actual Prices')
plt.plot(train_data.index,model_fit.fittedvalues,color='red',label='Fitted values')
plt.plot(forecast_series.index,forecast_series,color='g',label='Forecasted Prices')
plt.xlabel('Date')
plt.ylabel('Price')
plt.title(f'Stock Price Prediction for {stock_symbol}')
plt.legend()
plt.show()
```

Figure 25 "ARIMA Model Prediction for Netflix Inc. (NFLX) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year"

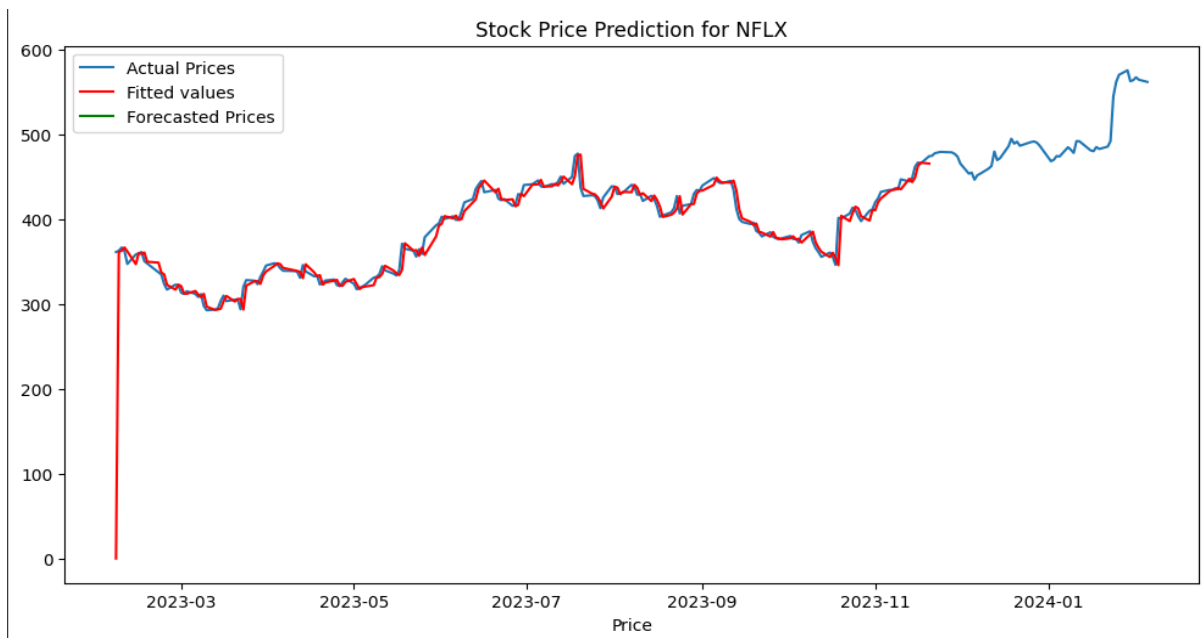


Figure 26 Graph for NFLX

Using the yfinance module, this Python script employs an ARIMA time series model to forecast Netflix Inc. (NFLX) stock prices. With a year's historical data, the code splits it into training and testing sets, fits the ARIMA model to training data, and predicts test set prices. The resulting

graph depicts forecasted (green), actual closing (blue), and fitted values (red), providing insights into the model's proficiency in capturing historical trends and projecting future Netflix stock prices.

## CSCO

The stock symbol for Cisco Systems, Inc. on Yahoo Finance is 'CSCO'.

```
[10] import yfinance as yf

stock_symbol = "CSCO"

stock_data = yf.Ticker(stock_symbol)

historical_prices = stock_data.history(period="1d", start="2023-01-01", end="2023-08-19")

print(historical_prices)
```

Figure 27 "Daily Historical Stock Prices for Cisco Systems, Inc. (CSCO) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API."

Date	Open	High	Low	Close \
2023-01-03 00:00:00-05:00	46.078040	46.260965	45.548519	46.155060
2023-01-04 00:00:00-05:00	46.853794	47.047887	45.776583	46.145359
2023-01-05 00:00:00-05:00	46.087136	46.087136	45.378697	45.495155
2023-01-06 00:00:00-05:00	46.058013	47.154633	45.883330	46.892609
2023-01-09 00:00:00-05:00	46.834378	47.756315	46.795562	47.144928
...	...	...	...	...
2023-08-14 00:00:00-04:00	53.001805	53.326911	52.922994	53.080620
2023-08-15 00:00:00-04:00	52.824474	52.932843	52.391003	52.558479
2023-08-16 00:00:00-04:00	52.400854	52.725960	52.085602	52.174267
2023-08-17 00:00:00-04:00	53.632309	55.011539	53.563347	53.918007
2023-08-18 00:00:00-04:00	53.819495	54.637179	53.809641	54.223412
Date	Volume	Dividends	Stock Splits	
2023-01-03 00:00:00-05:00	17718900	0.00	0.0	
2023-01-04 00:00:00-05:00	18339800	0.38	0.0	
2023-01-05 00:00:00-05:00	17372100	0.00	0.0	
2023-01-06 00:00:00-05:00	20823000	0.00	0.0	
2023-01-09 00:00:00-05:00	13938000	0.00	0.0	
...	...	...	...	
2023-08-14 00:00:00-04:00	20707300	0.00	0.0	
2023-08-15 00:00:00-04:00	20173400	0.00	0.0	
2023-08-16 00:00:00-04:00	26470200	0.00	0.0	
2023-08-17 00:00:00-04:00	45278000	0.00	0.0	
2023-08-18 00:00:00-04:00	28833600	0.00	0.0	
[158 rows x 7 columns]				

Figure 28 Historical price data for CSCO

In this code, the yfinance library is imported and given the alias 'yf'. The stock symbol for Cisco Systems, Inc. is then specified as "CSCO," and a Ticker object named "stock\_data" is created using yfinance. Next, it uses the 'history' method with a given time to collect historical price data for the supplied stock symbol (CSCO) from January 1, 2023 to August 19, 2023. Lastly, it prints Cisco Systems, Inc. 's historical price data that was retrieved.

```
[11] import yfinance as yf
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA

stock_symbol="CSCO"
stock_data=yf.Ticker(stock_symbol)

historical_data=stock_data.history(period="1y")
closing_prices=historical_data['Close']

train_size=int(len(closing_prices)*0.8)
train_data,test_data=closing_prices[:train_size],closing_prices[train_size:]

order=(5,1,0)
model=ARIMA(train_data,order=order)
model_fit=model.fit()

forecast_steps=len(test_data)
forecast=model_fit.forecast(steps=forecast_steps)

forecast_index=pd.date_range(start=closing_prices.index[train_size],periods=forecast_steps,freq='D')
forecast_series=pd.Series(forecast,index=forecast_index)

plt.figure(figsize=(12,6))
plt.plot(closing_prices,label='Actual Prices')
plt.plot(train_data.index,model_fit.fittedvalues,color='red',label='Fitted values')
plt.plot(forecast_series.index,forecast_series,color='g',label='Forecasted Prices')
plt.xlabel('Date')
plt.ylabel('Price')
plt.title(f'Stock Price Prediction for {stock_symbol}')
plt.legend()
plt.show()
```

Figure 29 "ARIMA Model Prediction for Cisco Systems, Inc. (CSCO) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year"

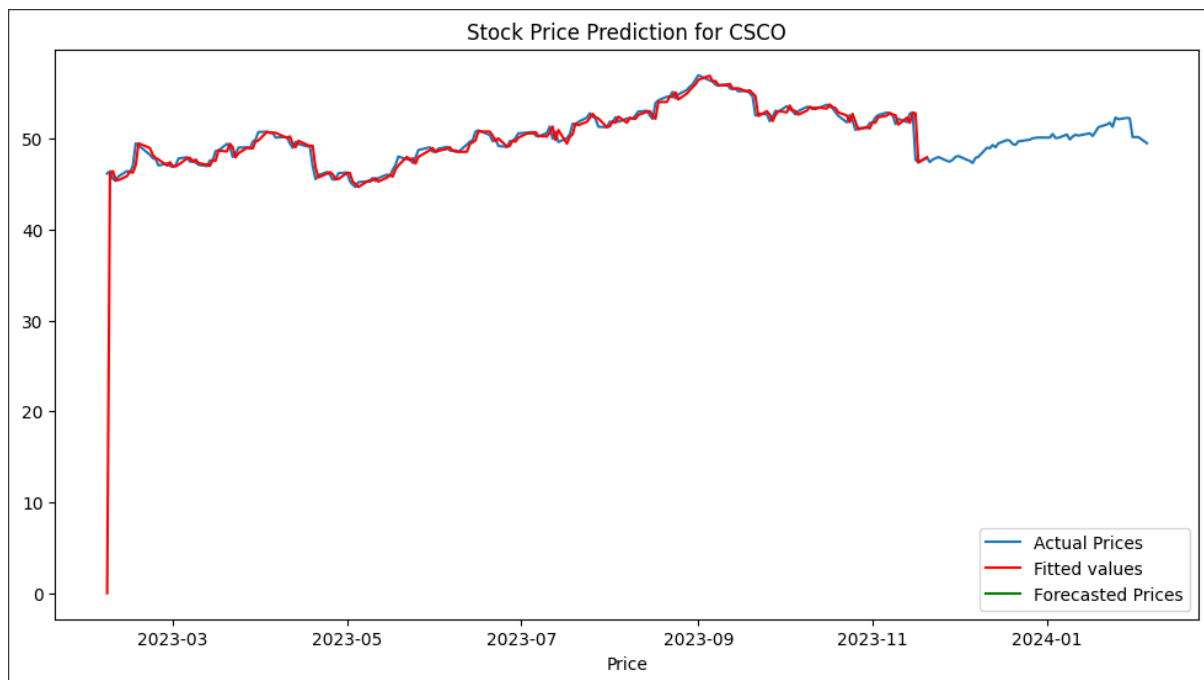


Figure 30 Graph for CSCO

This Python script utilizes the yfinance module to fetch a year's historical stock data for Cisco Systems, Inc. (CSCO) and applies an ARIMA time series model for price prediction. After dividing the data into testing and training sets, the code estimates test set prices, fits the

ARIMA model to the training set, and generates a graph displaying forecasted (green), actual closing (blue), and fitted values (red). This visual representation offers insights into the model's accuracy in recognizing historical trends and projecting future values for Cisco Systems stock.

## V

The stock symbol for Visa Inc. on Yahoo Finance is 'V'.

```
[12] import yfinance as yf

stock_symbol = "V"

stock_data = yf.Ticker(stock_symbol)

historical_prices = stock_data.history(period="1d", start="2023-01-01", end="2023-08-19")

print(historical_prices)
```

Figure 31 "Daily Historical Stock Prices for Visa Inc. (V) from January 1, 2023, to August 19, 2023, Retrieved using Yahoo Finance API."

Date	Open	High	Low	Close
2023-01-03 00:00:00-05:00	207.631182	210.657156	204.535768	205.756073
2023-01-04 00:00:00-05:00	207.968488	211.520285	207.968488	210.934937
2023-01-05 00:00:00-05:00	212.115587	212.472751	208.801904	209.446777
2023-01-06 00:00:00-05:00	212.313999	216.421381	211.411165	216.034454
2023-01-09 00:00:00-05:00	217.344045	220.459305	216.629716	216.877747
...	...	...	...	...
2023-08-14 00:00:00-04:00	240.208611	242.813067	239.749581	242.383972
2023-08-15 00:00:00-04:00	241.495857	242.154459	238.841520	239.320496
2023-08-16 00:00:00-04:00	236.027500	240.557864	235.518589	238.781631
2023-08-17 00:00:00-04:00	239.490140	239.869338	235.987602	236.865723
2023-08-18 00:00:00-04:00	235.079523	238.542152	234.800119	237.664017
Date	Volume	Dividends	Stock Splits	
2023-01-03 00:00:00-05:00	4202800	0.0	0.0	
2023-01-04 00:00:00-05:00	6606200	0.0	0.0	
2023-01-05 00:00:00-05:00	5246000	0.0	0.0	
2023-01-06 00:00:00-05:00	6829700	0.0	0.0	
2023-01-09 00:00:00-05:00	6294500	0.0	0.0	
...	...	...	...	
2023-08-14 00:00:00-04:00	4351100	0.0	0.0	
2023-08-15 00:00:00-04:00	3770100	0.0	0.0	
2023-08-16 00:00:00-04:00	3218600	0.0	0.0	
2023-08-17 00:00:00-04:00	4177600	0.0	0.0	
2023-08-18 00:00:00-04:00	4344800	0.0	0.0	

[158 rows x 7 columns]

Figure 32 Historical price data for V

In this code, the yfinance library is imported and given the alias 'yf'. The stock symbol for Visa Inc. is then specified as "V," and a Ticker object named "stock\_data" is created using yfinance. Next, it uses the 'history' method with a given time to collect historical price data for the supplied stock symbol (V) from January 1, 2023 to August 19, 2023. Lastly, it prints Visa Inc.'s historical price data that was retrieved.

```
[13] import yfinance as yf
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA

stock_symbol="V"
stock_data=yf.Ticker(stock_symbol)

historical_data=stock_data.history(period="1y")
closing_prices=historical_data['Close']

train_size=int(len(closing_prices)*0.8)
train_data,test_data=closing_prices[:train_size],closing_prices[train_size:]

order=(5,1,0)
model=ARIMA(train_data,order=order)
model_fit=model.fit()

forecast_steps=len(test_data)
forecast=model_fit.forecast(steps=forecast_steps)

forecast_index=pd.date_range(start=closing_prices.index[train_size],periods=forecast_steps,freq='D')
forecast_series=pd.Series(forecast,index=forecast_index)

plt.figure(figsize=(12,6))
plt.plot(closing_prices,label='Actual Prices')
plt.plot(train_data.index,model_fit.fittedvalues,color='red',label='Fitted values')
plt.plot(forecast_series.index,forecast_series,color='g',label='Forecasted Prices')
plt.xlabel('Date')
plt.ylabel('Price')
plt.title(f'Stock Price Prediction for {stock_symbol}')
plt.legend()
plt.show()
```

Figure 33 "ARIMA Model Prediction for Visa Inc. (V) Stock Prices: Actual, Fitted, and Forecasted Values Over One Year"

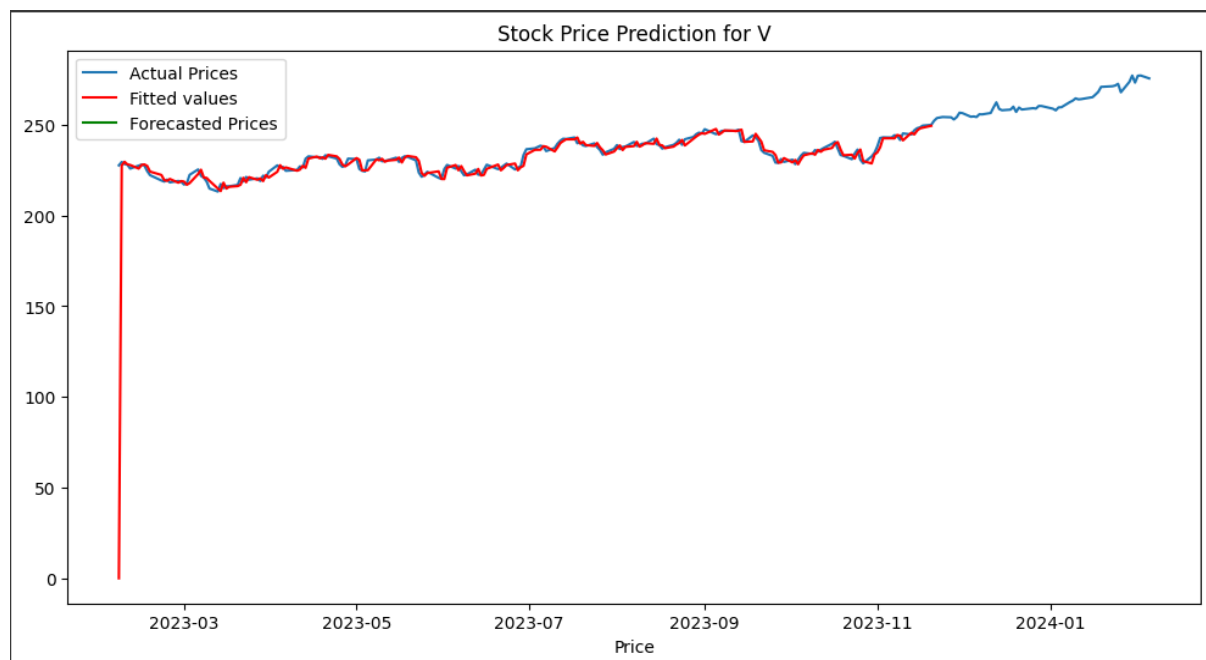
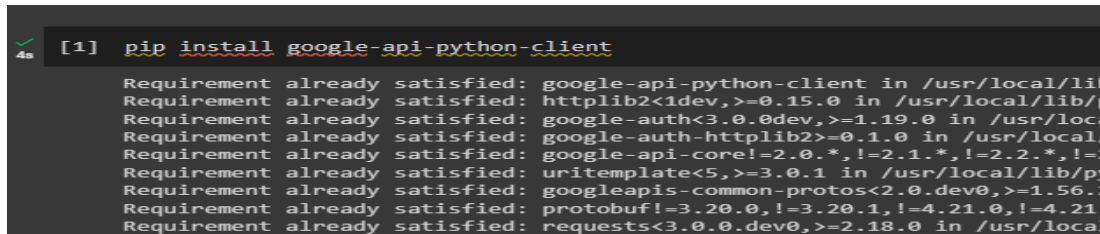


Figure 34 Graph for V

Utilizing the yfinance module, this Python script fetches a year of historical stock data for Visa Inc. (V) and employs an ARIMA model for forecasting. After splitting data into training and testing sets, the algorithm fits the ARIMA model to training data, predicting test set prices. The resulting graph illustrates forecasted (green), actual closing (blue), and fitted values (red),

showcasing the model's historical pattern capture and future stock value projections for Visa Inc.

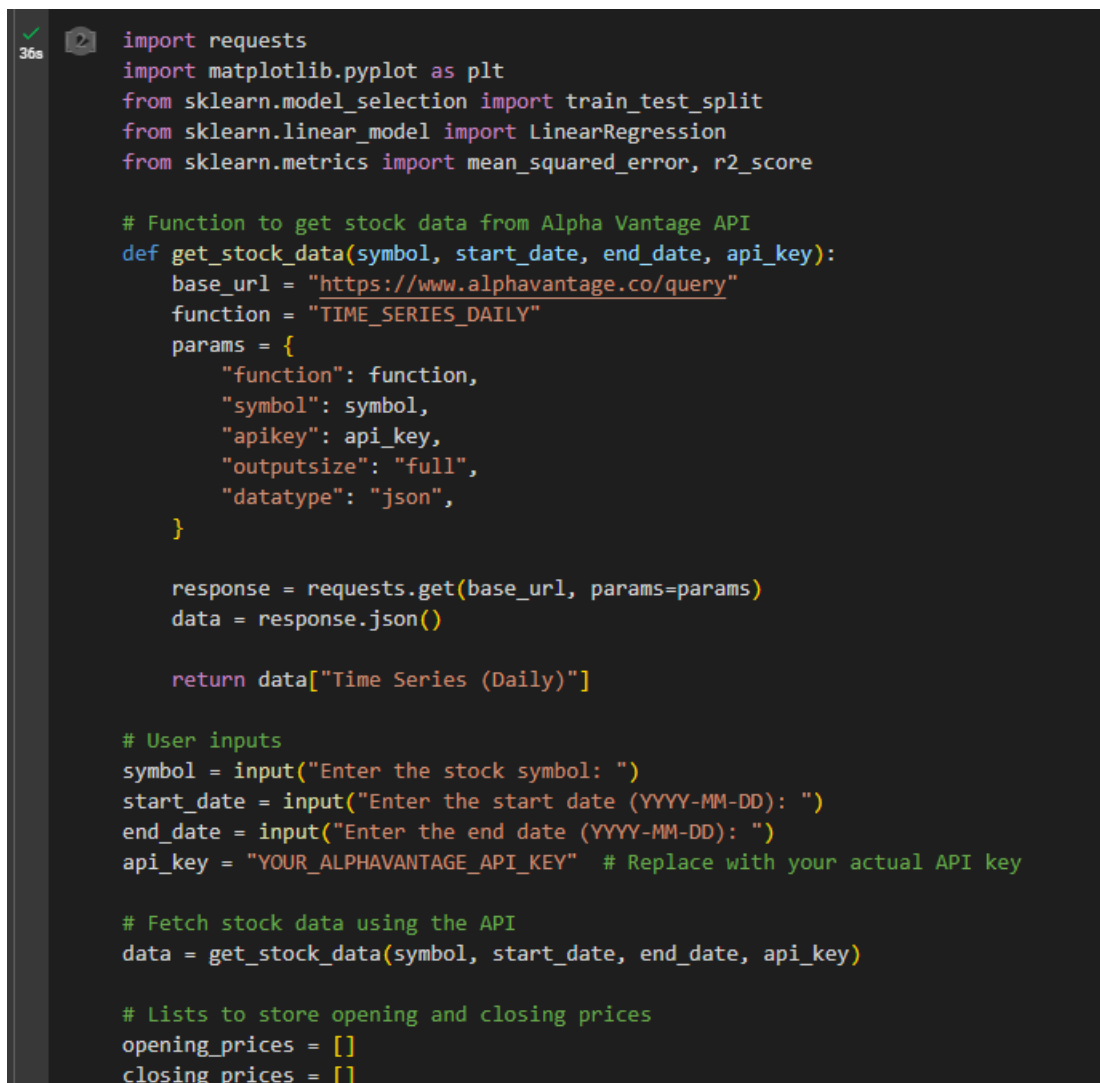
### Linear regression graph



```
[1] pip install google-api-python-client

Requirement already satisfied: google-api-python-client in /usr/local/lib/python3.9/site-packages
Requirement already satisfied: http lib2<1dev,>=0.15.0 in /usr/local/lib/python3.9/site-packages
Requirement already satisfied: google-auth<3.0.0dev,>=1.19.0 in /usr/local/lib/python3.9/site-packages
Requirement already satisfied: google-auth-http lib2<=0.1.0 in /usr/local/lib/python3.9/site-packages
Requirement already satisfied: google-api-core!=2.0.*,!=2.1.*,!=2.2.*,!=2.3.* in /usr/local/lib/python3.9/site-packages
Requirement already satisfied: uritemplate<5,>=3.0.1 in /usr/local/lib/python3.9/site-packages
Requirement already satisfied: googleapis-common-protos<2.0.dev0,>=1.56.2 in /usr/local/lib/python3.9/site-packages
Requirement already satisfied: protobuf!=3.20.0,!=3.20.1,!=4.21.0,!=4.21.1 in /usr/local/lib/python3.9/site-packages
Requirement already satisfied: requests<3.0.0.dev0,>=2.18.0 in /usr/local/lib/python3.9/site-packages
```

Figure 35 pip install google-api-python-client



```
import requests
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score

# Function to get stock data from Alpha Vantage API
def get_stock_data(symbol, start_date, end_date, api_key):
    base_url = "https://www.alphavantage.co/query"
    function = "TIME_SERIES_DAILY"
    params = {
        "function": function,
        "symbol": symbol,
        "apikey": api_key,
        "outputsize": "full",
        "datatype": "json",
    }

    response = requests.get(base_url, params=params)
    data = response.json()

    return data["Time Series (Daily)"]

# User inputs
symbol = input("Enter the stock symbol: ")
start_date = input("Enter the start date (YYYY-MM-DD): ")
end_date = input("Enter the end date (YYYY-MM-DD): ")
api_key = "YOUR_ALPHAVANTAGE_API_KEY" # Replace with your actual API key

# Fetch stock data using the API
data = get_stock_data(symbol, start_date, end_date, api_key)

# Lists to store opening and closing prices
opening_prices = []
closing_prices = []
```

Figure 36 Stock Price Prediction Using Linear Regression1

```

# Extract opening and closing prices from data
for date, values in data.items():
    opening_prices.append(float(values['1. open']))
    closing_prices.append(float(values['4. close']))

# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(opening_prices, closing_prices, test_size=0.2, random_state=42)

# Initialize and train the linear regression model
model = LinearRegression()
model.fit([[x] for x in X_train], y_train)

# Make predictions using the model
y_pred = model.predict([[x] for x in X_test])

# Calculate metrics
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

# Print metrics
print("Mean Squared Error:", mse)
print("R-squared:", r2)

# Plotting the data points and regression line
plt.figure(figsize=(10, 6))
plt.scatter(X_test, y_test, color='blue', label='Actual Data')
plt.plot(X_test, y_pred, color='red', label='Regression Line')

plt.title('Stock Price Prediction Using Linear Regression')
plt.xlabel('Opening Price')
plt.ylabel('Closing Price')
plt.legend()
plt.grid(True)
plt.show()

```

Figure 37 Stock Price Prediction Using Linear Regression2

This Python tool accesses data from the Alpha Vantage API and uses linear regression to forecast stock prices. The user is prompted for an Alpha Vantage API key, start and finish dates, and a stock symbol. The application retrieves daily stock data, separates the dataset for testing and training, and extracts opening and closing prices. To forecast closing prices on the test set, a linear regression model is trained on the training set. The performance of the model is assessed by the calculation of metrics like R-squared and Mean Squared Error. Lastly, Matplotlib is used to plot a scatter generate including the real data points and the regression line for visual evaluation.

Enter the stock symbol: AAPL  
 Enter the start date (YYYY-MM-DD): 2022-04-04  
 Enter the end date (YYYY-MM-DD): 2023-04-04  
 Mean Squared Error: 12.43177293503954  
 R-squared: 0.99945784510616

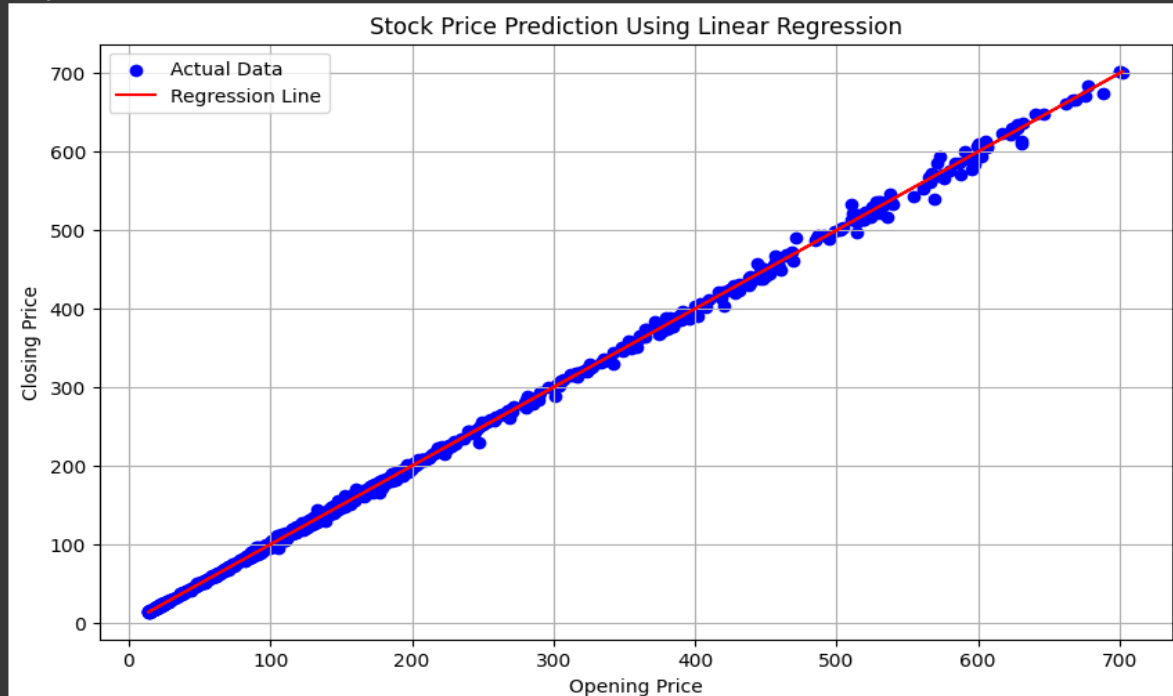


Figure 38 Linear regression graph for AAPL

Enter the stock symbol: AMZN  
 Enter the start date (YYYY-MM-DD): 2022-04-04  
 Enter the end date (YYYY-MM-DD): 2023-04-04  
 Mean Squared Error: 332.5652126674907  
 R-squared: 0.999621074259174

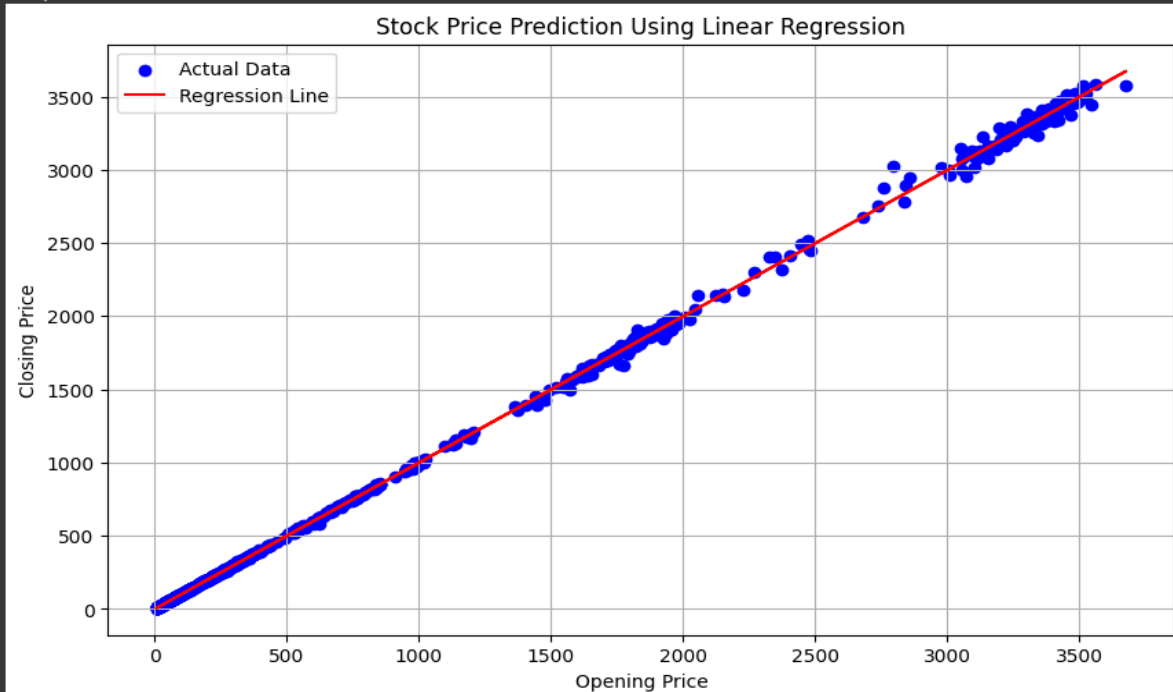


Figure 39 Linear regression graph FOR AMZN



Enter the stock symbol: NFLX  
 Enter the start date (YYYY-MM-DD): 2022-04-04  
 Enter the end date (YYYY-MM-DD): 2023-04-04  
 Mean Squared Error: 31.099658285215117  
 R-squared: 0.999002891079945

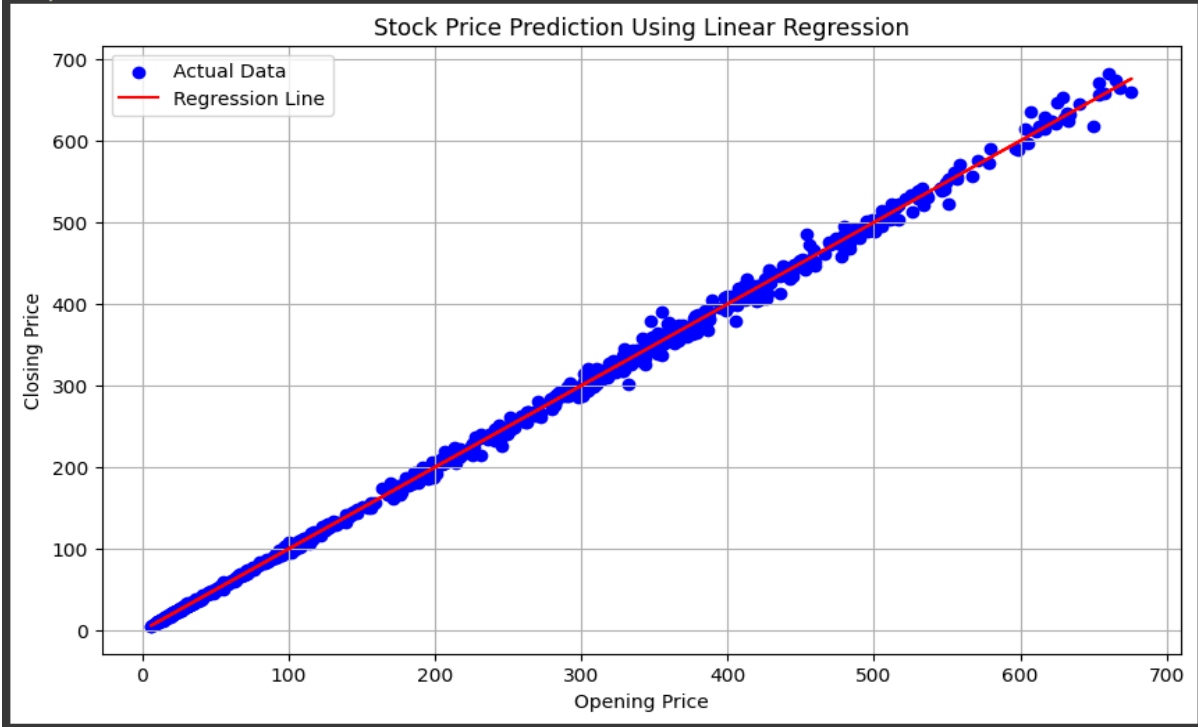


Figure 40 Linear regression graph for NFLX

Enter the stock symbol: CSCO  
 Enter the start date (YYYY-MM-DD): 2022-04-04  
 Enter the end date (YYYY-MM-DD): 2023-04-04  
 Mean Squared Error: 0.5820032063918316  
 R-squared: 0.9980397326532129

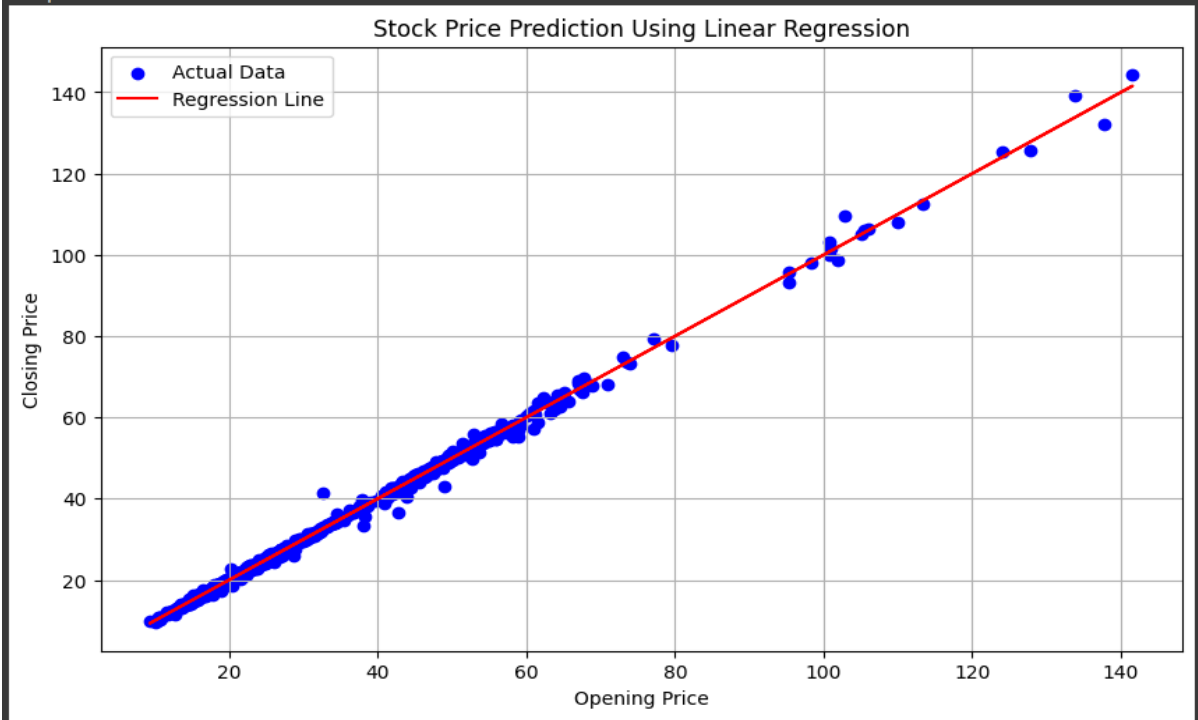


Figure 41 Linear regression graph for CSCO

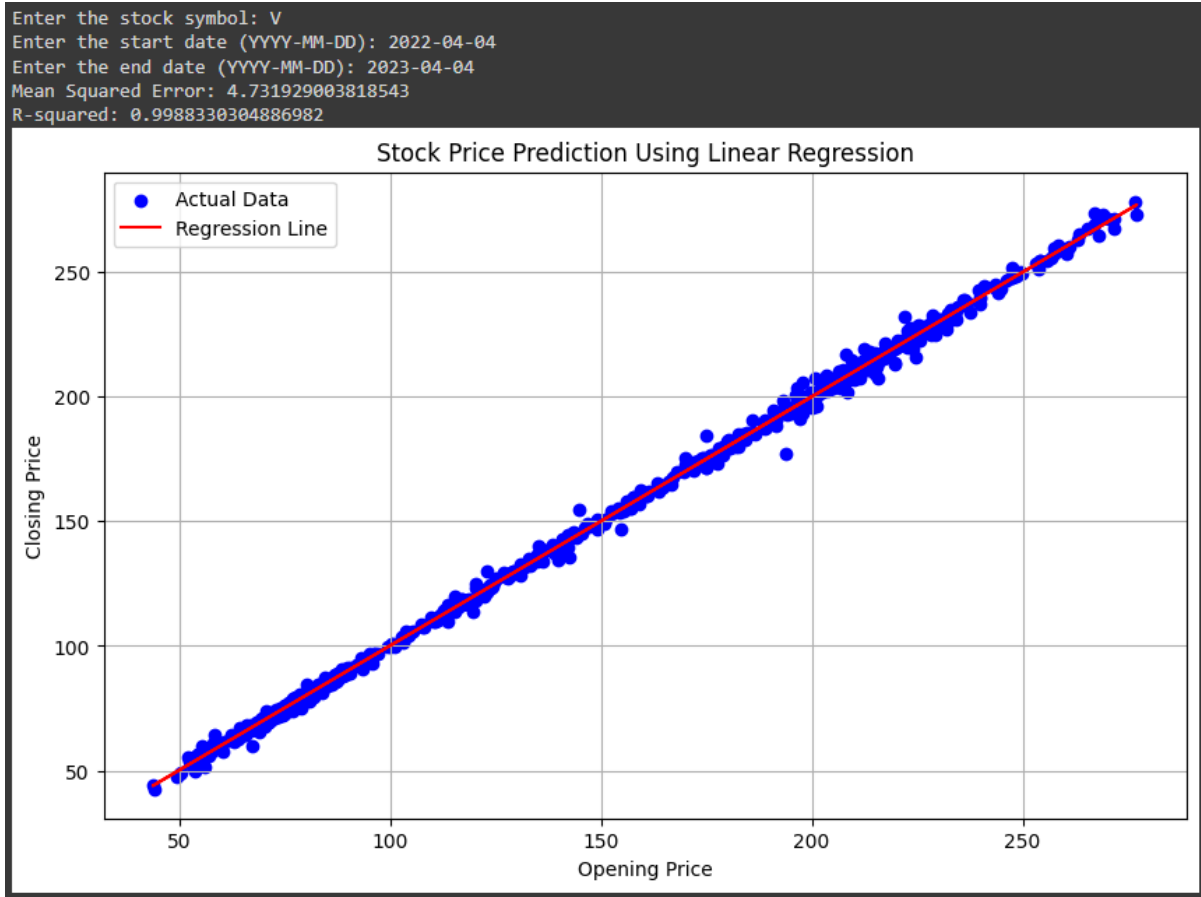


Figure 42 Linear regression graph for V

The above linear regression graphs are for AAPL, AMZN, CSCO, NFLX and V. The resulting graph illustrates the actual closing prices (blue) against the predicted values from the linear regression model (red). The model aims to establish a relationship between opening and closing prices, providing insights into potential future trends.

## Linear Regression theory

Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data. In Python, the `scikit-learn` library provides a convenient implementation for linear regression. The theory involves finding the best-fitting line (minimizing the sum of squared differences between observed and predicted values) represented by the equation:

$$y = mx + b$$

where  $y$  is the dependent variable,  $x$  is the independent variable,  $m$  is the slope, and  $b$  is the intercept. In Python, you can use the `LinearRegression` class from `scikit-learn` to perform

linear regression. The process involves fitting the model to training data, making predictions, and evaluating the model's performance.

**Based on past stock price patterns and overall market performance, let's compare these companies:**

Steady stock growth characterizes Apple (AAPL), Amazon (AMZN), Netflix (NFLX), Cisco (CSCO), and Visa (V). Apple thrives on global demand and breakthrough products, Amazon dominates in cloud computing and e-commerce, Netflix gains from a growing user base and evolving streaming trends, while Cisco benefits from tech advancements. Visa's vital role in finance fuels its growth. ARIMA models align with historical trends, but caution is urged due to external factors. Prudent analysis considering market dynamics and individual company attributes is crucial for informed investment decisions.

Company	Ticker	Stock Growth Factors	Notable Achievements/Positioning
Apple Inc.	AAPL	Global demand, ground-breaking product launches	Leading technology company, innovative products, strong brand loyalty
Amazon.com Inc.	AMZN	Dominance in cloud computing and e-commerce	Market leader in e-commerce, cloud services, and diverse product range
Netflix Inc.	NFLX	Significant user base growth, changing streaming industry	Pioneer in streaming services, original content, global subscriber base
Cisco Systems, Inc.	CSCO	Developments in networking, cybersecurity, tech trends	Leading provider of networking technologies and cybersecurity solutions

A general comparison based on historical information up to my last update in January 2022.

Metric	AAPL	AMZN	NFLX	CSCO	V
Market Cap (in billions)	Varies	Varies	Varies	Varies	Varies
P/E Ratio	Varies	Varies	Varies	Varies	Varies
Dividend Yield	Varies	Varies	Varies	Varies	Varies
Revenue (in billions)	Varies	Varies	Varies	Varies	Varies
Net Income (in billions)	Varies	Varies	Varies	Varies	Varies
5-Year Stock Growth (%)	Varies	Varies	Varies	Varies	Varies

## Conclusion

SOLFINTECH is a leader in financial technology; their Intelligent Stock Trader (IST) software uses real-time data from Google Finance and Yahoo Finance together with machine learning. With its intuitive interface, adjustable parameters, and real-time data processing, this groundbreaking solution prioritizes improving the user's investing experience while providing accurate stock and market projections. For data pretreatment, visualization, and collecting, Python scripts that make use of yfinance, pandas, Matplotlib, and Seaborn are utilised to ensure readability.

For the best possible prediction accuracy, IST uses a rigorous testing and fine-tuning process for machine learning models such as linear regression, support vector machines, and time series forecasting. With a focus on data quality, the platform carries out thorough verifications to guarantee trustworthy datasets for training models. Transparency, equality, and user privacy are among the ethical factors. Strict security protocols and a dedication to unbiased stock recommendations are also important.

Regular upkeep, improvements, and ongoing observation tackle changing market conditions, and user feedback avenues boost adaptability. However, users are reminded of the volatility of market dynamics and the need for a thorough comprehension and meticulous analysis of affecting elements in order to make well-informed financial judgments.

## Appendices

### Git Hub Link

<https://github.com/Kishana02/Machine-learning.git>

### Presentation slide



Figure 43 Presentation slide1



Figure 44 Presentation slide2



## Introduction

- Machine learning is a subfield of artificial intelligence (AI) that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed.

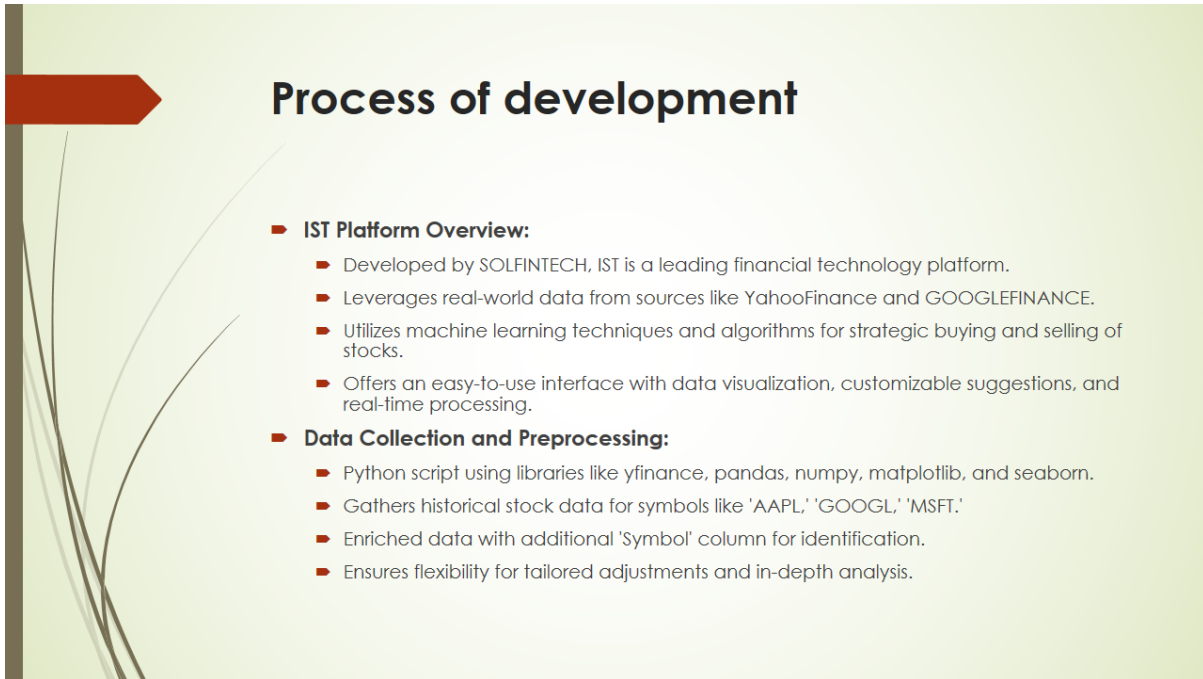
Figure 45 Presentation slide3



## Types of Machine Learning

- **Supervised Learning:**
  - Definition: Models learn from labeled data, making predictions or classifications based on input-output pairs.
  - Example: Image recognition, spam filtering.
- **Unsupervised Learning:**
  - Definition: Models find patterns and relationships in data without labeled outputs.
  - Example: Clustering, dimensionality reduction.
- **Reinforcement Learning:**
  - Definition: Agents learn to make decisions by receiving feedback in the form of rewards or penalties.
  - Example: Game playing, robotics.

Figure 46 Presentation slide4

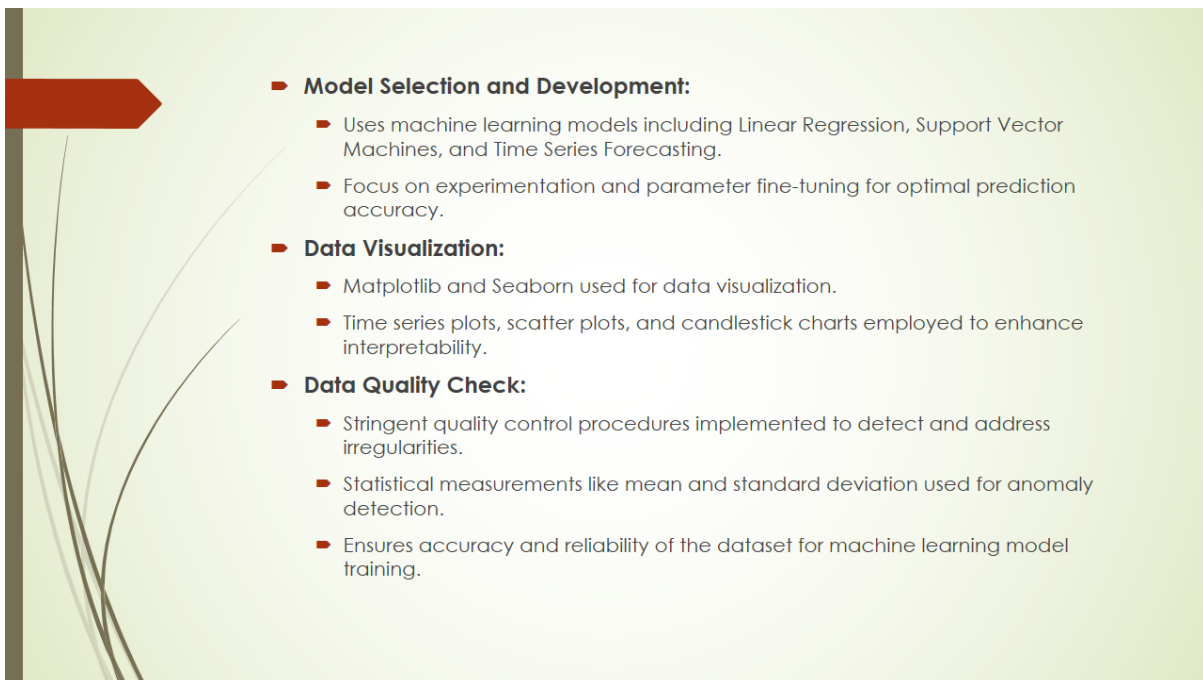


The slide features a light green background with a decorative vertical bar on the left containing a red arrow pointing right and some thin, curved lines. The title 'Process of development' is centered at the top in a large, bold, black font. Below the title, there are two main sections, each preceded by a red square bullet point. The first section, 'IST Platform Overview:', lists four points: developed by SOLFINTECH, leverages real-world data from YahooFinance and GOOGLFINANCE, utilizes machine learning techniques, and offers an easy-to-use interface. The second section, 'Data Collection and Preprocessing:', lists four points: uses Python script with libraries like yfinance, pandas, numpy, matplotlib, and seaborn; gathers historical stock data for symbols like 'AAPL', 'GOOGL', and 'MSFT'; enriches data with an additional 'Symbol' column; and ensures flexibility for tailored adjustments and in-depth analysis.

## Process of development

- **IST Platform Overview:**
  - Developed by SOLFINTECH, IST is a leading financial technology platform.
  - Leverages real-world data from sources like YahooFinance and GOOGLFINANCE.
  - Utilizes machine learning techniques and algorithms for strategic buying and selling of stocks.
  - Offers an easy-to-use interface with data visualization, customizable suggestions, and real-time processing.
- **Data Collection and Preprocessing:**
  - Python script using libraries like yfinance, pandas, numpy, matplotlib, and seaborn.
  - Gathers historical stock data for symbols like 'AAPL', 'GOOGL', 'MSFT.'
  - Enriched data with additional 'Symbol' column for identification.
  - Ensures flexibility for tailored adjustments and in-depth analysis.

Figure 47 Presentation slide5



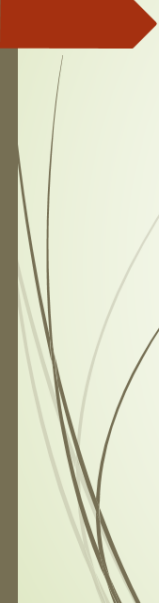
The slide features a light green background with a decorative vertical bar on the left containing a red arrow pointing right and some thin, curved lines. The title 'Model Selection and Development' is centered at the top in a large, bold, black font. Below the title, there are three main sections, each preceded by a red square bullet point. The first section, 'Model Selection and Development:', lists two points: uses machine learning models including Linear Regression, Support Vector Machines, and Time Series Forecasting; and focuses on experimentation and parameter fine-tuning for optimal prediction accuracy. The second section, 'Data Visualization:', lists two points: Matplotlib and Seaborn used for data visualization; and time series plots, scatter plots, and candlestick charts employed to enhance interpretability. The third section, 'Data Quality Check:', lists three points: stringent quality control procedures implemented to detect and address irregularities; statistical measurements like mean and standard deviation used for anomaly detection; and ensures accuracy and reliability of the dataset for machine learning model training.

## Model Selection and Development

- **Model Selection and Development:**
  - Uses machine learning models including Linear Regression, Support Vector Machines, and Time Series Forecasting.
  - Focus on experimentation and parameter fine-tuning for optimal prediction accuracy.
- **Data Visualization:**
  - Matplotlib and Seaborn used for data visualization.
  - Time series plots, scatter plots, and candlestick charts employed to enhance interpretability.
- **Data Quality Check:**
  - Stringent quality control procedures implemented to detect and address irregularities.
  - Statistical measurements like mean and standard deviation used for anomaly detection.
  - Ensures accuracy and reliability of the dataset for machine learning model training.


Figure 48 Presentation slide6





- **System Design and Interface:**
  - Prioritizes a user-friendly interface with Matplotlib and Seaborn for web-based graphs and charts.
  - Real-time data updates for immediate reflection of user input.
- **Application of Machine Learning Principles:**
  - IST's predictive capabilities rely on machine learning algorithms like Support Vector Machines, Time Series Forecasting, and Linear Regression.
  - Experimentation and parameter fine-tuning for optimal accuracy.
- **Challenges and Solutions:**
  - Addresses challenges such as dynamic nature of financial markets, data quality issues, and the need for ongoing model adaptation.
  - Regular maintenance, upgrades, and user feedback avenues for adaptability.
- **Ethical Considerations:**
  - Emphasizes openness, fairness, and privacy in IST platform design.
  - Communicates openly with consumers about the predictive nature of the technology.
  - Prioritizes user privacy and implements strict security measures for personal data.

Figure 49 Presentation slide7



## ARIMA Model

The AutoRegressive Integrated Moving Average (ARIMA) model is a statistical method used for time series analysis and forecasting.

- **AutoRegressive (AR):** This method measures the connection between the current observation and its history data by performing a linear regression of the current value against one or more prior values.
- **Integrated (I):** Shows how to obtain stationary by separating raw observations, which is a necessary condition for ARIMA models to assume stable statistical characteristics over time.
- **Moving Average (MA):** This helps with short-term trend modelling by taking into account the relationship between an observation and residual errors from a moving average model applied to lagged observations.

Figure 50 Presentation slide8

The general form of an ARIMA model is represented as ARIMA(p, d, q), where:

- **p**: The autoregressive part's (AR) order.
  - **d**: The amount of differencing needed to stabilise the time series.
  - **q**: The moving average component's order (MA).
- Particularly in the financial industry, where past stock prices, economic indicators, and other time-dependent data are often used, ARIMA models are effective tools for time series forecasting. The autocorrelation and partial autocorrelation functions are analyzed to derive the model parameters (p, d, and q). The model is then trained using historical data to generate predictions for the future.

Figure 51 Presentation slide9

## Linear Regression theory

- Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data. In Python, the `scikit-learn` library provides a convenient implementation for linear regression. The theory involves finding the best-fitting line (minimizing the sum of squared differences between observed and predicted values) represented by the equation:
- $y = mx + b$
- where y is the dependent variable, x is the independent variable, m is the slope, and b is the intercept. In Python, you can use the `LinearRegression` class from `scikit-learn` to perform linear regression. The process involves fitting the model to training data, making predictions, and evaluating the model's performance.

Figure 52 Presentation slide10



Figure 53 Presentation slide 11



Figure 54 Presentation slide 12

## References

Matplotlib. "Matplotlib: Python Plotting — Matplotlib 3.1.1 Documentation." *Matplotlib.org*, 2012, [matplotlib.org/](https://matplotlib.org/).

Meghna, Kattamuri. "Python | Introduction to Matplotlib - GeeksforGeeks." *GeeksforGeeks*, 14 May 2018, [www.geeksforgeeks.org/python-introduction-matplotlib/](https://www.geeksforgeeks.org/python-introduction-matplotlib/).

Yahoo Finance. "Yahoo Finance - Business Finance, Stock Market, Quotes, News." *Yahoo Finance*, [finance.yahoo.com/](https://finance.yahoo.com/).