

# COS30018 - Option B - Task 1: Setup

**Student Name: Nguyen Duc Le Nguyen**

**Id: 104224493**

## 1. Environment Setup

### 1.1 Attempt to Set Up the Environment

In my attempt to set up the environment, I followed these steps:

1. **Created a Virtual Environment:** I used the following command to create a virtual environment:

```
python -m venv env
```

2. **Activated the Virtual Environment:** I activated the virtual environment:

```
.\env\Scripts\activate
```

**Created a Requirements File:** I made a `requirements.txt` file to list all the packages I needed:

```
numpy  
pandas  
matplotlib  
yfinance  
sklearn
```

3. **Installed Dependencies:** I installed all the packages using this command:

```
pip install -r requirements.txt
```

### 1.2 Details of Requirements File

My `requirements.txt` file looked like this:

```
numpy==1.21.2  
pandas==1.3.3  
matplotlib==3.4.3  
yfinance==0.1.63  
scikit-learn==0.24.2
```

I chose these versions to make sure everything would work smoothly with my code.

## 2. Testing the Provided Code Bases

### 2.1 Testing Code Base v0.1

To test the first code base (v0.1), I did the following:

1. **Loaded the Script:** I opened the `StockP0.py` file and read through it to understand what it does.
2. **Ran the Script:** I ran the script to see if it worked without any errors:

```
python StockP0.py
```

3. **Identified Issues:** I noted any errors that came up when I ran the script.

Here is a screenshot of the v0.1 execution:

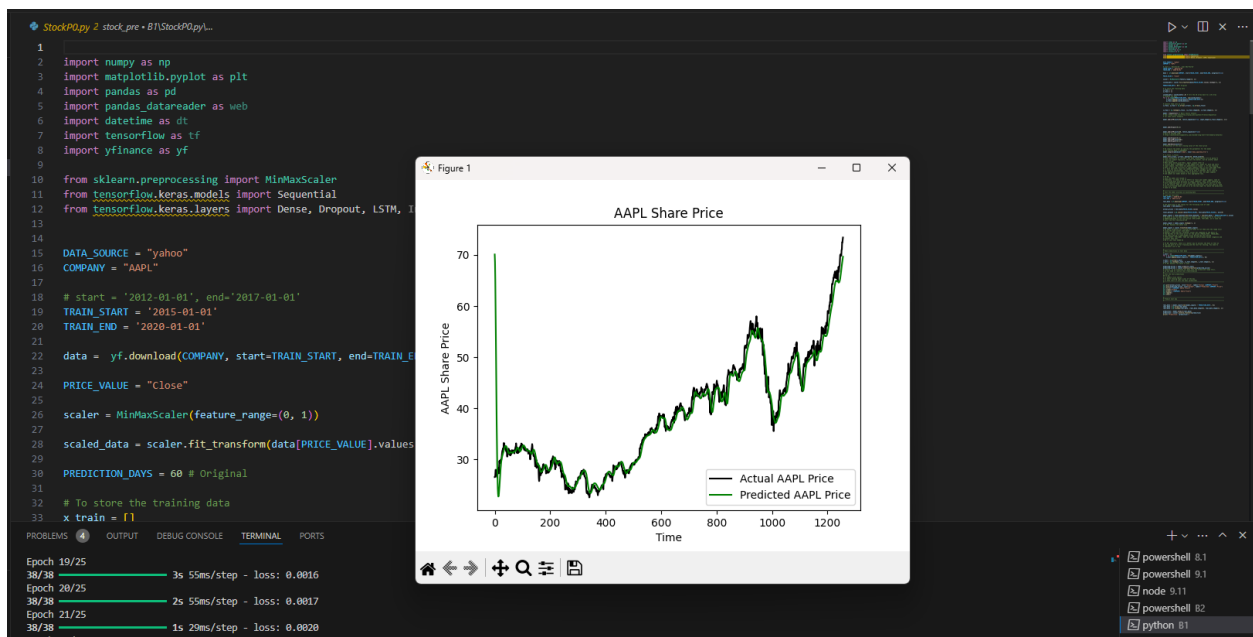


Figure 1 - v0.1 Execution

### 2.2 Testing Code Base P1

For the updated code base (P1), I did these steps:

1. **Loaded the Script:** I read through the `stockP1.py` file to see what was different.
2. **Ran the Script:** I ran the updated script to test it:

```
python stockP1.py
```

3. **Noted Improvements:** I compared the results with the first version to see what had improved.

Here is a screenshot of the P1 execution:



Figure 2 - P1 Output

### 3. Understanding of the Initial Code Base v0.1

Here's my understanding of the first code base (v0.1):

1. **Data Retrieval:** The script uses `yfinance` to get historical stock prices.
2. **Data Preprocessing:** It cleans the data, handles missing values, normalizes it, and splits it into training and testing sets.
3. **Model Building:** It uses machine learning models from `scikit-learn` like Linear Regression to predict stock prices.
4. **Evaluation:** It checks how well the model is doing using metrics like Mean Squared Error (MSE) and makes plots with `matplotlib`.

The initial code base is simple and covers the basic steps needed for predicting stock prices.