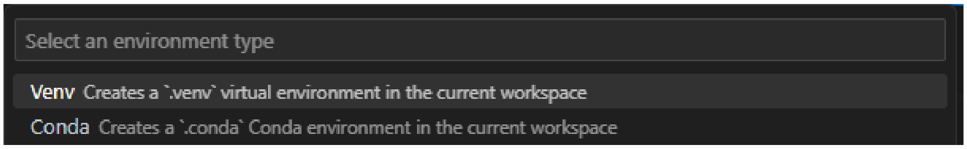
Tasks B.1 - Setup

Jiin Wen Tan 102846565

Environment setup:

I started off with Visual studio Code as my IDE and created a virtual environment with command palette

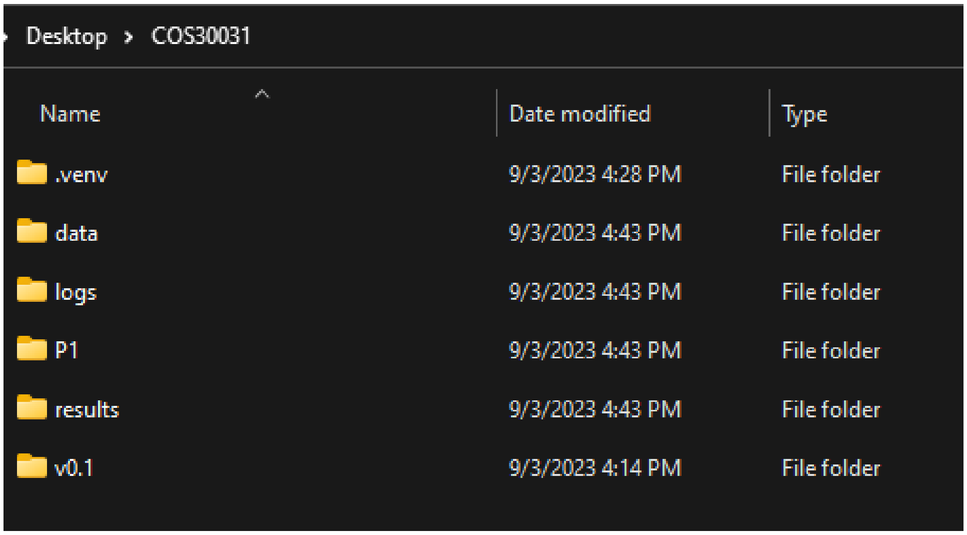


Select an environment type

Venv Creates a `.venv` virtual environment in the current workspace

Conda Creates a `.conda` Conda environment in the current workspace

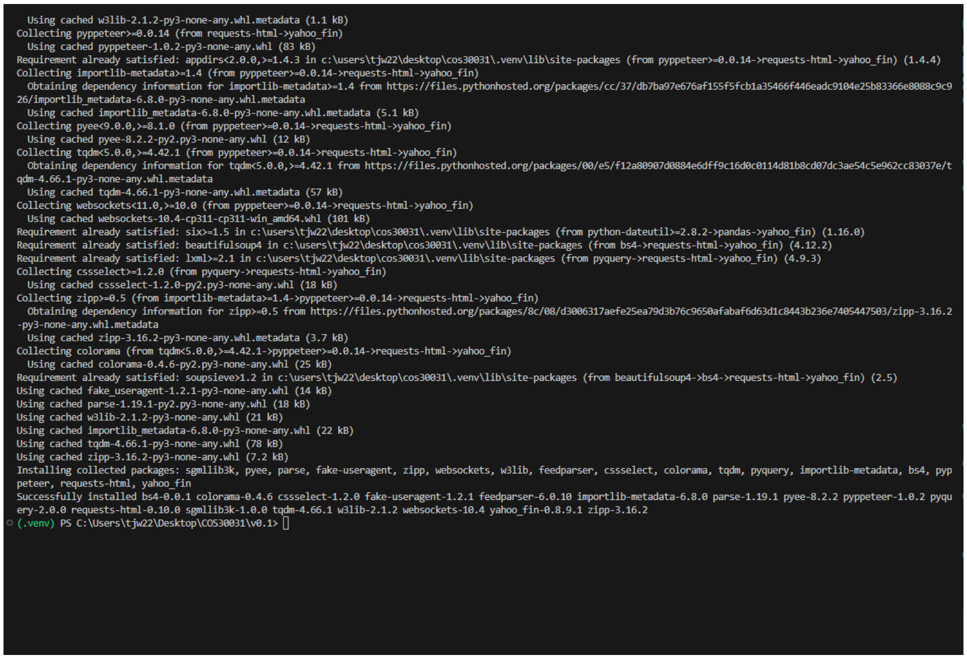
Then I downloaded both source code from Canvas (v0.1) and github repositories (P1) and relocate them under the same folder named COS30031



Desktop › COS30031

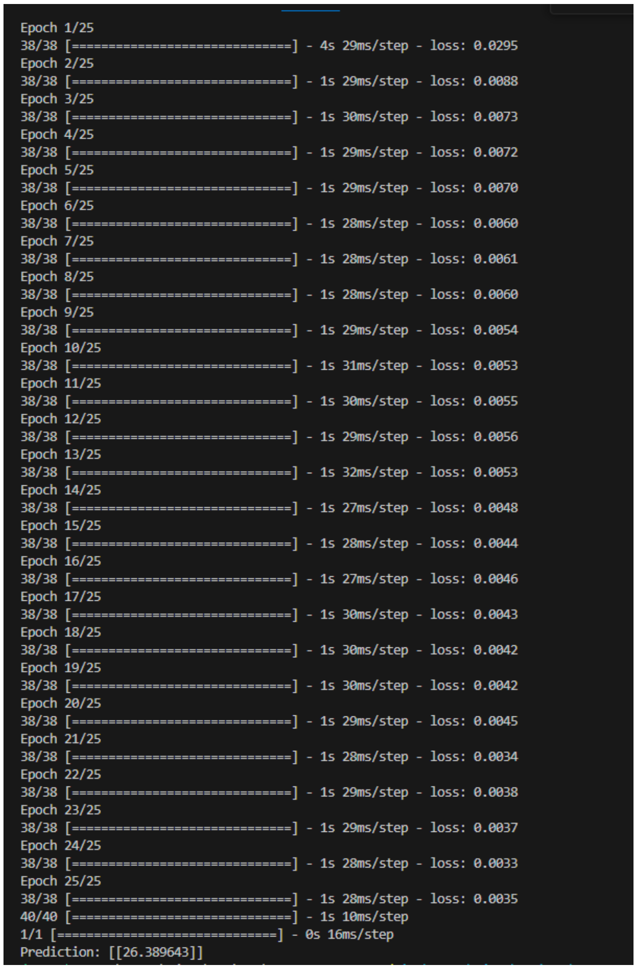
Name	Date modified	Type
.venv	9/3/2023 4:28 PM	File folder
data	9/3/2023 4:43 PM	File folder
logs	9/3/2023 4:43 PM	File folder
P1	9/3/2023 4:43 PM	File folder
results	9/3/2023 4:43 PM	File folder
v0.1	9/3/2023 4:14 PM	File folder

Right after, pip installed all necessary libraries such as numpy, matplotlib, pandas, tensorflow e.g

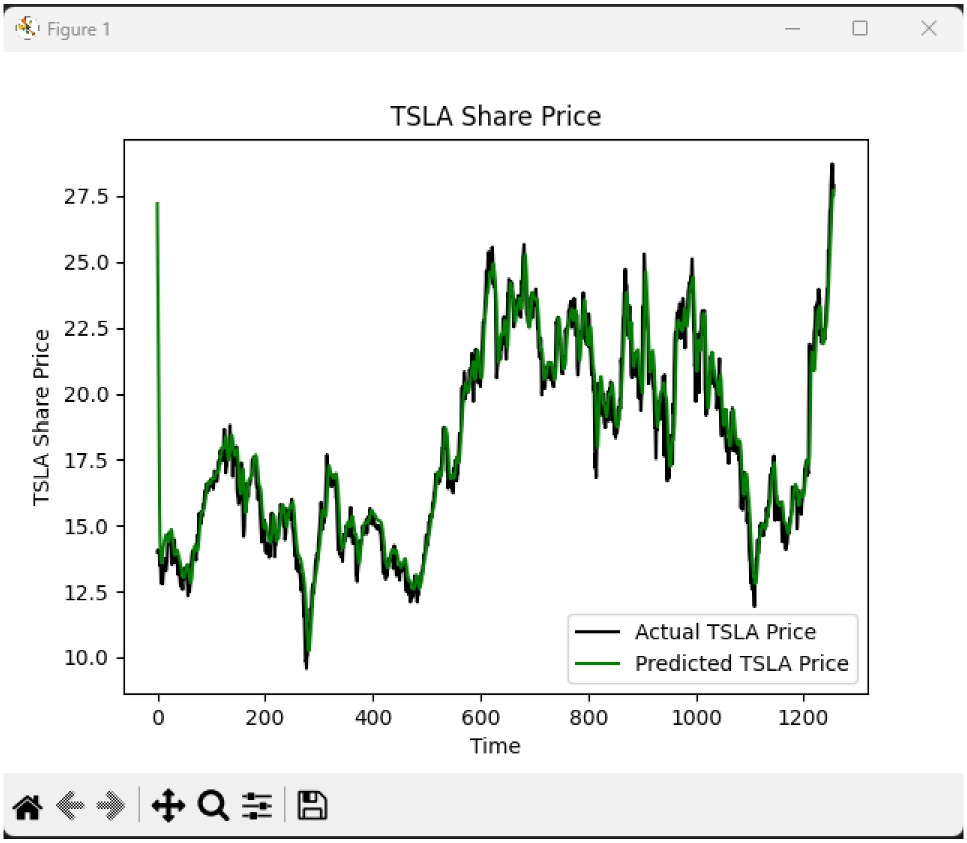


```
  Using cached w3lib-2.1.2-py3-none-any.whl.metadata (1.1 kB)
Collecting pyppeteer>=0.0.14 (from requests-html->yahoo_fin)
  Using cached pyppeteer-1.0.2-py3-none-any.whl (83 kB)
Requirement already satisfied: appdirs<2.0.0,>=1.4.3 in c:\users\tjw22\desktop\cos30031\.venv\lib\site-packages (from pyppeteer>=0.0.14->requests-html->yahoo_fin) (1.4.4)
Collecting importlib-metadata>=1.4 (from pyppeteer>=0.0.14->requests-html->yahoo_fin)
  Obtaining dependency information for importlib-metadata>=1.4 from https://files.pythonhosted.org/packages/cc/37/db7ba97e676af155f5fcb1a35466f446eadc9104e25b83366e8088c9c9
26/importlib_metadata-6.8.0-py3-none-any.whl.metadata
  Using cached importlib_metadata-6.8.0-py3-none-any.whl.metadata (5.1 kB)
Collecting pyee<9.0.0,>=8.1.0 (from pyppeteer>=0.0.14->requests-html->yahoo_fin)
  Using cached pyee-8.2.2-py2.py3-none-any.whl (12 kB)
Collecting tqdm<5.0.0,>=4.42.1 (from pyppeteer>=0.0.14->requests-html->yahoo_fin)
  Obtaining dependency information for tqdm<5.0.0,>=4.42.1 from https://files.pythonhosted.org/packages/00/e5/f12a80907d0884e6dff9c16d0c0114d81b8cd07dc3ae54c5e962cc83037e/t
qdm-4.66.1-py3-none-any.whl.metadata
  Using cached tqdm-4.66.1-py3-none-any.whl.metadata (57 kB)
Collecting websockets<11.0,>=10.0 (from pyppeteer>=0.0.14->requests-html->yahoo_fin)
  Using cached websockets-10.4-cp311-cp311-win_amd64.whl (101 kB)
Requirement already satisfied: six>=1.5 in c:\users\tjw22\desktop\cos30031\.venv\lib\site-packages (from python-dateutil>=2.8.2->pandas->yahoo_fin) (1.16.0)
Requirement already satisfied: beautifulsoup4 in c:\users\tjw22\desktop\cos30031\.venv\lib\site-packages (from bs4->requests-html->yahoo_fin) (4.12.2)
Requirement already satisfied: lxml>=2.1 in c:\users\tjw22\desktop\cos30031\.venv\lib\site-packages (from pyquery->requests-html->yahoo_fin) (4.9.3)
Collecting cssselect>=1.2.0 (from pyquery->requests-html->yahoo_fin)
  Using cached cssselect-1.2.0-py2.py3-none-any.whl (18 kB)
Collecting zipp>=0.5 (from importlib-metadata>=1.4->pyppeteer>=0.0.14->requests-html->yahoo_fin)
  Obtaining dependency information for zipp>=0.5 from https://files.pythonhosted.org/packages/8c/08/d3006317aefe25ea79d3b76c9650afabaf6d63d1c8443b236e7405447503/zipp-3.16.2
-py3-none-any.whl.metadata
  Using cached zipp-3.16.2-py3-none-any.whl.metadata (3.7 kB)
Collecting colorama (from tqdm<5.0.0,>=4.42.1->pyppeteer>=0.0.14->requests-html->yahoo_fin)
  Using cached colorama-0.4.6-py2.py3-none-any.whl (25 kB)
Requirement already satisfied: soupsieve>1.2 in c:\users\tjw22\desktop\cos30031\.venv\lib\site-packages (from beautifulsoup4->bs4->requests-html->yahoo_fin) (2.5)
Using cached fake_useragent-1.2.1-py3-none-any.whl (14 kB)
Using cached parse-1.19.1-py2.py3-none-any.whl (18 kB)
Using cached w3lib-2.1.2-py3-none-any.whl (21 kB)
Using cached importlib_metadata-6.8.0-py3-none-any.whl (22 kB)
Using cached tqdm-4.66.1-py3-none-any.whl (78 kB)
Using cached zipp-3.16.2-py3-none-any.whl (7.2 kB)
Installing collected packages: sgmllib3k, pyee, parse, fake-useragent, zipp, websockets, w3lib, feedparser, cssselect, colorama, tqdm, pyquery, importlib-metadata, bs4, pyp
peteer, requests-html, yahoo_fin
Successfully installed bs4-0.0.1 colorama-0.4.6 cssselect-1.2.0 fake-useragent-1.2.1 feedparser-6.0.10 importlib-metadata-6.8.0 parse-1.19.1 pyee-8.2.2 pyppeteer-1.0.2 pyqu
ery-2.0.0 requests-html-0.10.0 sgmllib3k-1.0.0 tqdm-4.66.1 w3lib-2.1.2 websockets-10.4 yahoo_fin-0.8.9.1 zipp-3.16.2
(.venv) PS C:\Users\tjw22\Desktop\COS30031\v0.1> 
```

Running v0.1



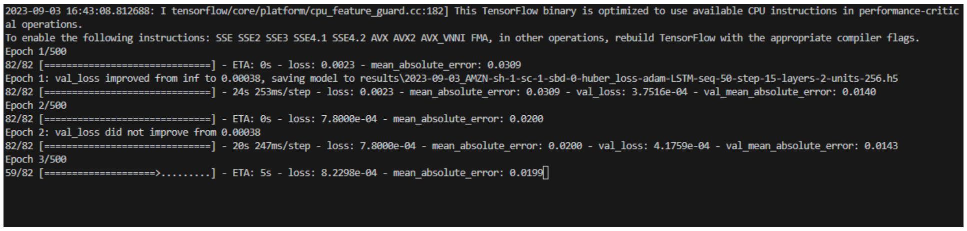
```
Epoch 1/25
38/38 [==============================] - 4s 29ms/step - loss: 0.0295
Epoch 2/25
38/38 [==============================] - 1s 29ms/step - loss: 0.0088
Epoch 3/25
38/38 [==============================] - 1s 30ms/step - loss: 0.0073
Epoch 4/25
38/38 [==============================] - 1s 29ms/step - loss: 0.0072
Epoch 5/25
38/38 [==============================] - 1s 29ms/step - loss: 0.0070
Epoch 6/25
38/38 [==============================] - 1s 28ms/step - loss: 0.0060
Epoch 7/25
38/38 [==============================] - 1s 28ms/step - loss: 0.0061
Epoch 8/25
38/38 [==============================] - 1s 28ms/step - loss: 0.0060
Epoch 9/25
38/38 [==============================] - 1s 29ms/step - loss: 0.0054
Epoch 10/25
38/38 [==============================] - 1s 31ms/step - loss: 0.0053
Epoch 11/25
38/38 [==============================] - 1s 30ms/step - loss: 0.0055
Epoch 12/25
38/38 [==============================] - 1s 29ms/step - loss: 0.0056
Epoch 13/25
38/38 [==============================] - 1s 32ms/step - loss: 0.0053
Epoch 14/25
38/38 [==============================] - 1s 27ms/step - loss: 0.0048
Epoch 15/25
38/38 [==============================] - 1s 28ms/step - loss: 0.0044
Epoch 16/25
38/38 [==============================] - 1s 27ms/step - loss: 0.0046
Epoch 17/25
38/38 [==============================] - 1s 30ms/step - loss: 0.0043
Epoch 18/25
38/38 [==============================] - 1s 30ms/step - loss: 0.0042
Epoch 19/25
38/38 [==============================] - 1s 30ms/step - loss: 0.0042
Epoch 20/25
38/38 [==============================] - 1s 29ms/step - loss: 0.0045
Epoch 21/25
38/38 [==============================] - 1s 28ms/step - loss: 0.0034
Epoch 22/25
38/38 [==============================] - 1s 29ms/step - loss: 0.0038
Epoch 23/25
38/38 [==============================] - 1s 29ms/step - loss: 0.0037
Epoch 24/25
38/38 [==============================] - 1s 28ms/step - loss: 0.0033
Epoch 25/25
38/38 [==============================] - 1s 28ms/step - loss: 0.0035
40/40 [==============================] - 1s 10ms/step
1/1 [==============================] - 0s 16ms/step
Prediction: [[26.389643]]
```



Running P1

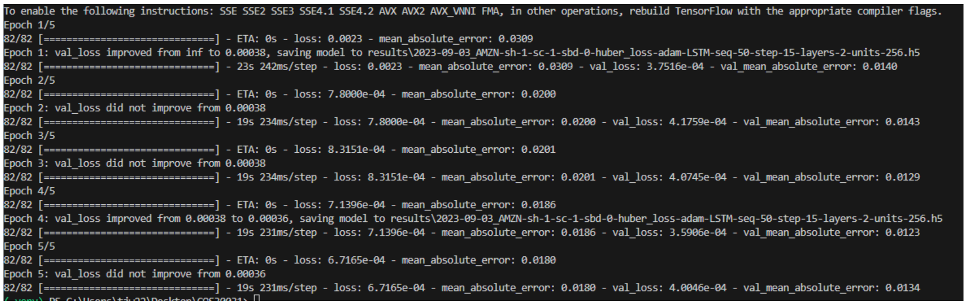
Epoch was initially set to 500, due to long hours and processor limitation I changed it to 5 just to have a quick view and ensure everything was able to run smoothly.

Before change:

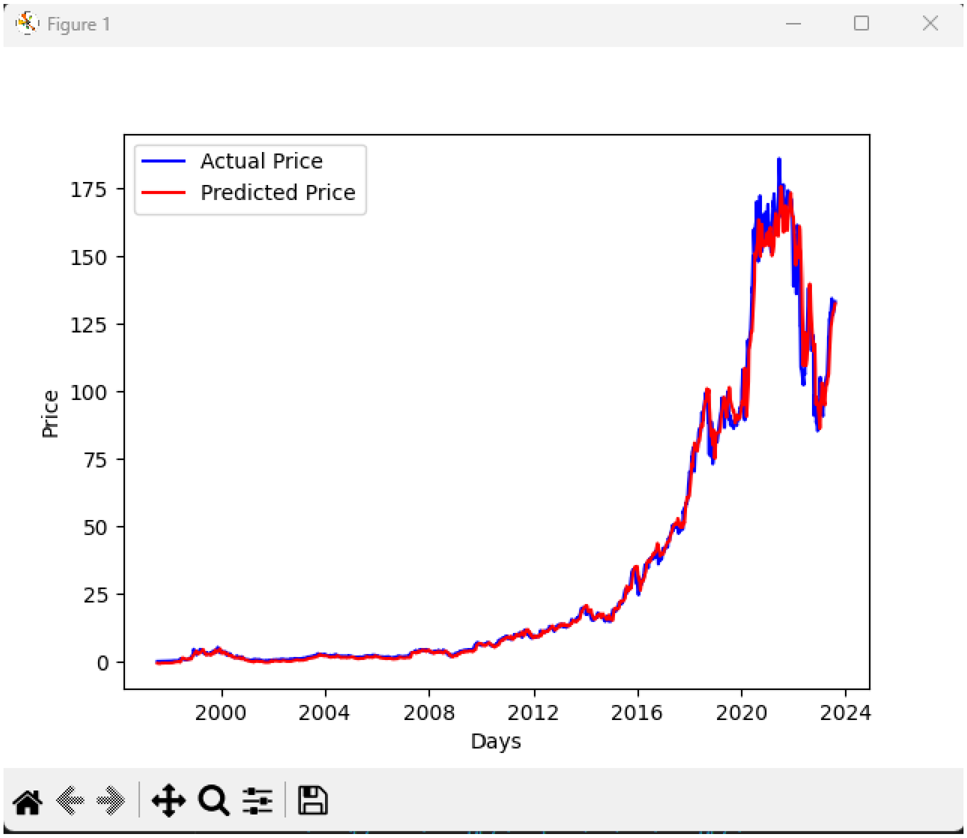


```
2023-09-03 16:43:08.812688: I tensorflow/core/platform/cpu_feature_guard.cc:182] This TensorFlow binary is optimized to use available CPU instructions in performance-critic
al operations.
To enable the following instructions: SSE SSE2 SSE3 SSE4.1 SSE4.2 AVX AVX2 AVX_VNNI FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
Epoch 1/500
82/82 [==============================] - ETA: 0s - loss: 0.0023 - mean_absolute_error: 0.0309
Epoch 1: val_loss improved from inf to 0.00038, saving model to results\2023-09-03_AMZN-sh-1-sc-1-sbd-0-huber_loss-adam-LSTM-seq-50-step-15-layers-2-units-256.h5
82/82 [==============================] - 24s 253ms/step - loss: 0.0023 - mean_absolute_error: 0.0309 - val_loss: 3.7516e-04 - val_mean_absolute_error: 0.0140
Epoch 2/500
82/82 [==============================] - ETA: 0s - loss: 7.8000e-04 - mean_absolute_error: 0.0200
Epoch 2: val_loss did not improve from 0.00038
82/82 [==============================] - 20s 247ms/step - loss: 7.8000e-04 - mean_absolute_error: 0.0200 - val_loss: 4.1759e-04 - val_mean_absolute_error: 0.0143
Epoch 3/500
59/82 [====================>.........] - ETA: 5s - loss: 8.2298e-04 - mean_absolute_error: 0.0199
```

After change:



```
To enable the following instructions: SSE SSE2 SSE3 SSE4.1 SSE4.2 AVX AVX2 AVX_VNNI FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
Epoch 1/5
82/82 [==============================] - ETA: 0s - loss: 0.0023 - mean_absolute_error: 0.0309
Epoch 1: val_loss improved from inf to 0.00038, saving model to results\2023-09-03_AMZN-sh-1-sc-1-sbd-0-huber_loss-adam-LSTM-seq-50-step-15-layers-2-units-256.h5
82/82 [==============================] - 23s 242ms/step - loss: 0.0023 - mean_absolute_error: 0.0309 - val_loss: 3.7516e-04 - val_mean_absolute_error: 0.0140
Epoch 2/5
82/82 [==============================] - ETA: 0s - loss: 7.8000e-04 - mean_absolute_error: 0.0200
Epoch 2: val_loss did not improve from 0.00038
82/82 [==============================] - 19s 234ms/step - loss: 7.8000e-04 - mean_absolute_error: 0.0200 - val_loss: 4.1759e-04 - val_mean_absolute_error: 0.0143
Epoch 3/5
82/82 [==============================] - ETA: 0s - loss: 8.3151e-04 - mean_absolute_error: 0.0201
Epoch 3: val_loss did not improve from 0.00038
82/82 [==============================] - 19s 234ms/step - loss: 8.3151e-04 - mean_absolute_error: 0.0201 - val_loss: 4.0745e-04 - val_mean_absolute_error: 0.0129
Epoch 4/5
82/82 [==============================] - ETA: 0s - loss: 7.1396e-04 - mean_absolute_error: 0.0186
Epoch 4: val_loss improved from 0.00038 to 0.00036, saving model to results\2023-09-03_AMZN-sh-1-sc-1-sbd-0-huber_loss-adam-LSTM-seq-50-step-15-layers-2-units-256.h5
82/82 [==============================] - 19s 231ms/step - loss: 7.1396e-04 - mean_absolute_error: 0.0186 - val_loss: 3.5906e-04 - val_mean_absolute_error: 0.0123
Epoch 5/5
82/82 [==============================] - ETA: 0s - loss: 6.7165e-04 - mean_absolute_error: 0.0180
Epoch 5: val_loss did not improve from 0.00036
82/82 [==============================] - 19s 231ms/step - loss: 6.7165e-04 - mean_absolute_error: 0.0180 - val_loss: 4.0046e-04 - val_mean_absolute_error: 0.0134
```



Summary of v0.1 code base understanding:

The purpose of this program is to predict the stock prices of Tesla by fetching data from Yahoo Finance. The data will be normalized to a range of 0 to 1 with the MinMaxScaler then the program takes 60 days (about 2 months) of historical data and produces the 61st day's profit/loss. The model chosen is a sequential neural network that uses LTSM layers which consist of three layers in this case: `model.add(LSTM(units=50, return_sequences=True, input_shape=(x_train.shape[1], 1)))`, `model.add(LSTM(units=50, return_sequences=True))`, `model.add(LSTM(units=50))` (Dense layer). In between layers were separated by a dropout layer to avoid overfitting. This model also uses Adam optimizer and Mean squared error as its loss function. The model is then trained using the prepared training dataset for 25 epochs with a batch size of 32. Test start and end date is 2020 – 2022 by merging training and datasets for more accurate predictions. After combining the test's initial 60 days that was drawn from the last 60 days of the training dataset will undergoes scaling and segmented into 60 days sequence. With the use of matplotlib to generate a graph that visualizes the datasets.