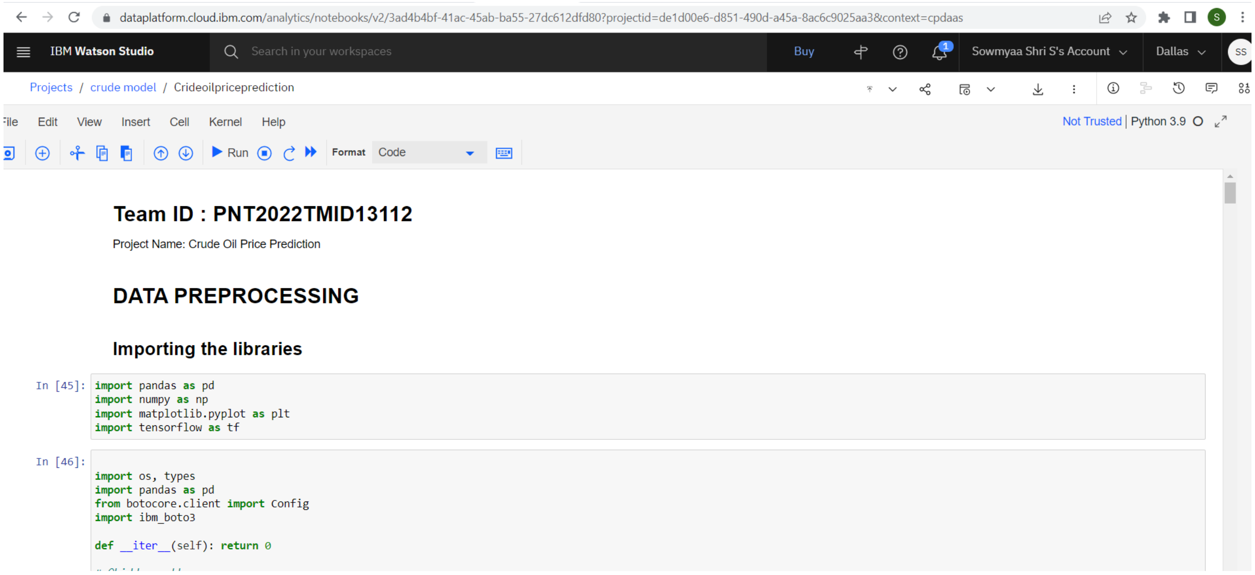
SPRINT 4
Training the model on IBM cloud

Date	19 november 2022
Team ID	PNT2022TMID13112
Project Name	Project - Crude Oil Price Prediction

1. IBM Watson studio

   Crude oil price prediction environment:


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format Code

Team ID : PNT2022TMID13112

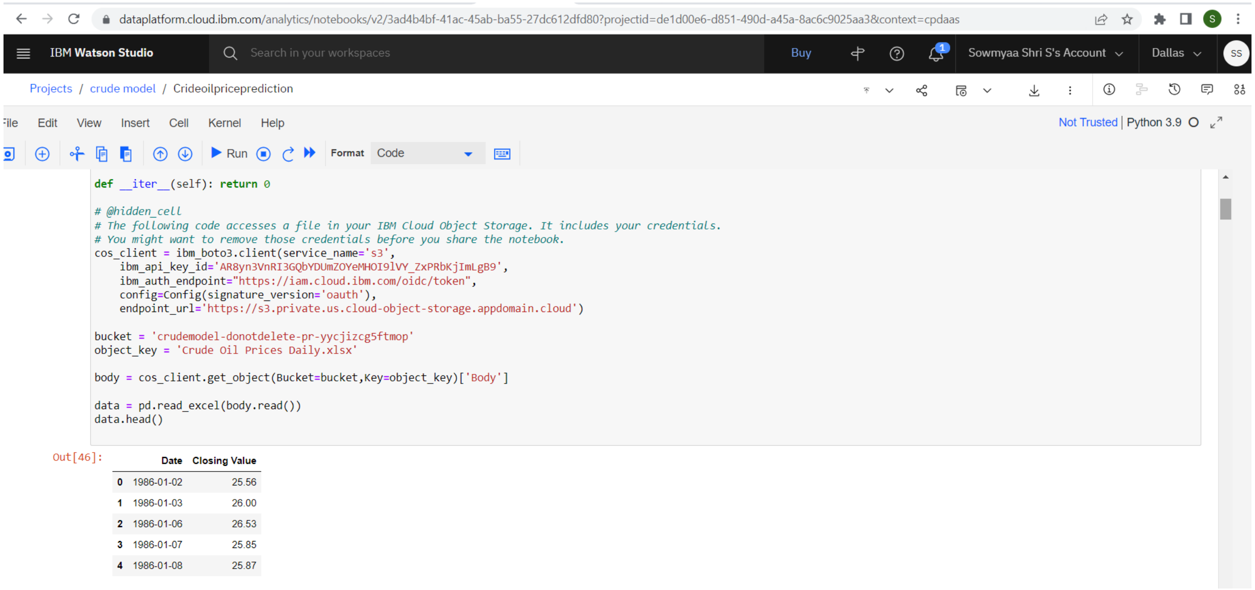
Project Name: Crude Oil Price Prediction

DATA PREPROCESSING

Importing the libraries

```
In [45]: import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import tensorflow as tf
```

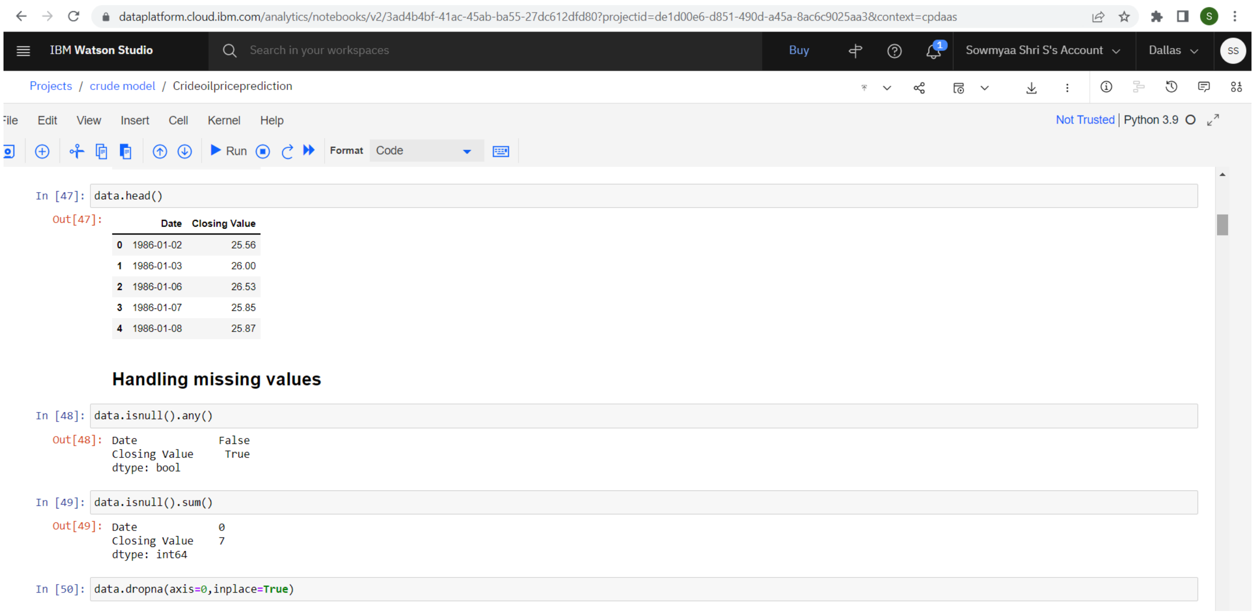
```
In [46]:
         import os, types
         import pandas as pd
         from botocore.client import Config
         import ibm_boto3

         def __iter__(self): return 0
```


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format
Code

```
def __iter__(self): return 0

# @hidden_cell
# The following code accesses a file in your IBM Cloud Object Storage. It includes your credentials.
# You might want to remove those credentials before you share the notebook.
cos_client = ibm_boto3.client(service_name='s3',
    ibm_api_key_id='AR8yn3VnRI3GQbYDUmZOYeMHOI9lVY_ZxPRbKjImLgB9',
    ibm_auth_endpoint="https://iam.cloud.ibm.com/oidc/token",
    config=Config(signature_version='oauth'),
    endpoint_url='https://s3.private.us.cloud-object-storage.appdomain.cloud')

bucket = 'crudemodel-donotdelete-pr-yycjizcg5ftmop'
object_key = 'Crude Oil Prices Daily.xlsx'

body = cos_client.get_object(Bucket=bucket,Key=object_key)['Body']

data = pd.read_excel(body.read())
data.head()
```

Out[46]:

	Date	Closing Value
0	1986-01-02	25.56
1	1986-01-03	26.00
2	1986-01-06	26.53
3	1986-01-07	25.85
4	1986-01-08	25.87


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format
Code

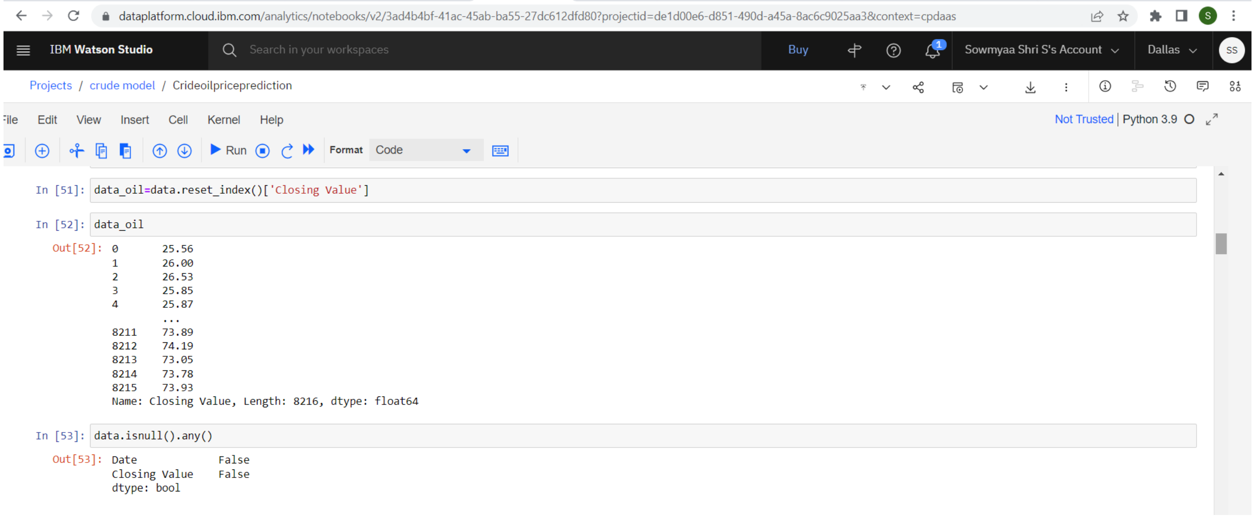
```
In [47]: data.head()
```

Out[47]:

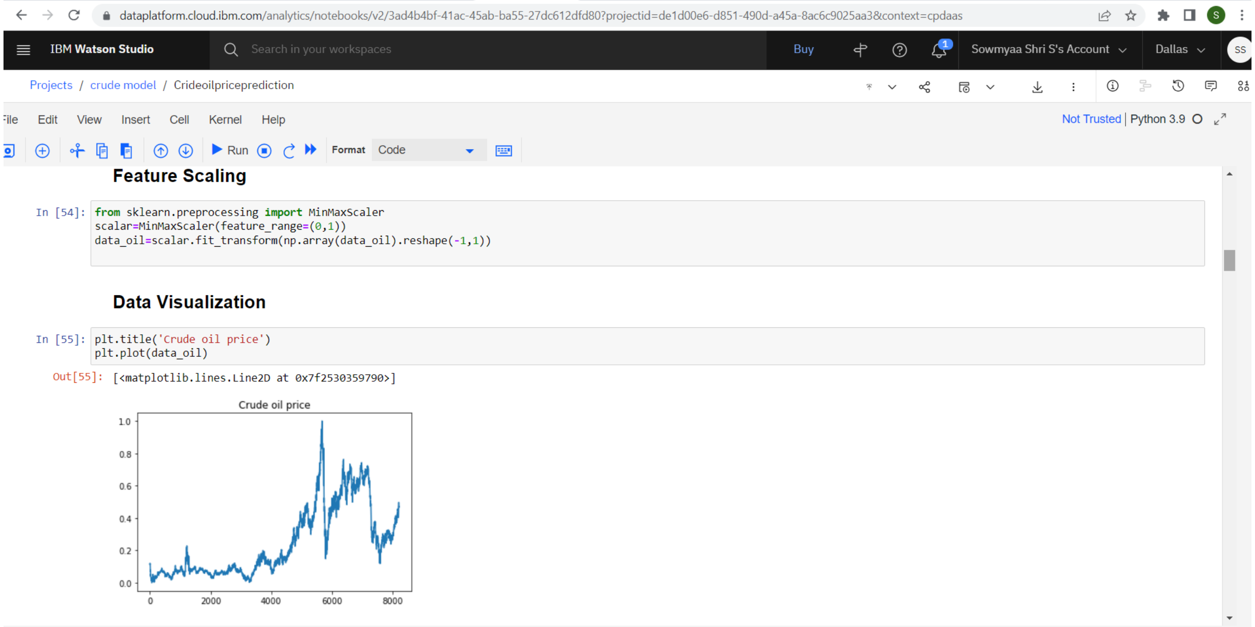
	Date	Closing Value
0	1986-01-02	25.56
1	1986-01-03	26.00
2	1986-01-06	26.53
3	1986-01-07	25.85
4	1986-01-08	25.87

Handling missing values

```
In [48]: data.isnull().any()

Out[48]: Date             False
         Closing Value     True
         dtype: bool

In [49]: data.isnull().sum()

Out[49]: Date             0
         Closing Value    7
         dtype: int64

In [50]: data.dropna(axis=0,inplace=True)
```


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format
Code

```
In [51]: data_oil=data.reset_index()['Closing Value']

In [52]: data_oil

Out[52]: 0       25.56
         1       26.00
         2       26.53
         3       25.85
         4       25.87
                 ...
         8211    73.89
         8212    74.19
         8213    73.05
         8214    73.78
         8215    73.93
         Name: Closing Value, Length: 8216, dtype: float64

In [53]: data.isnull().any()

Out[53]: Date             False
         Closing Value    False
         dtype: bool
```


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format
Code

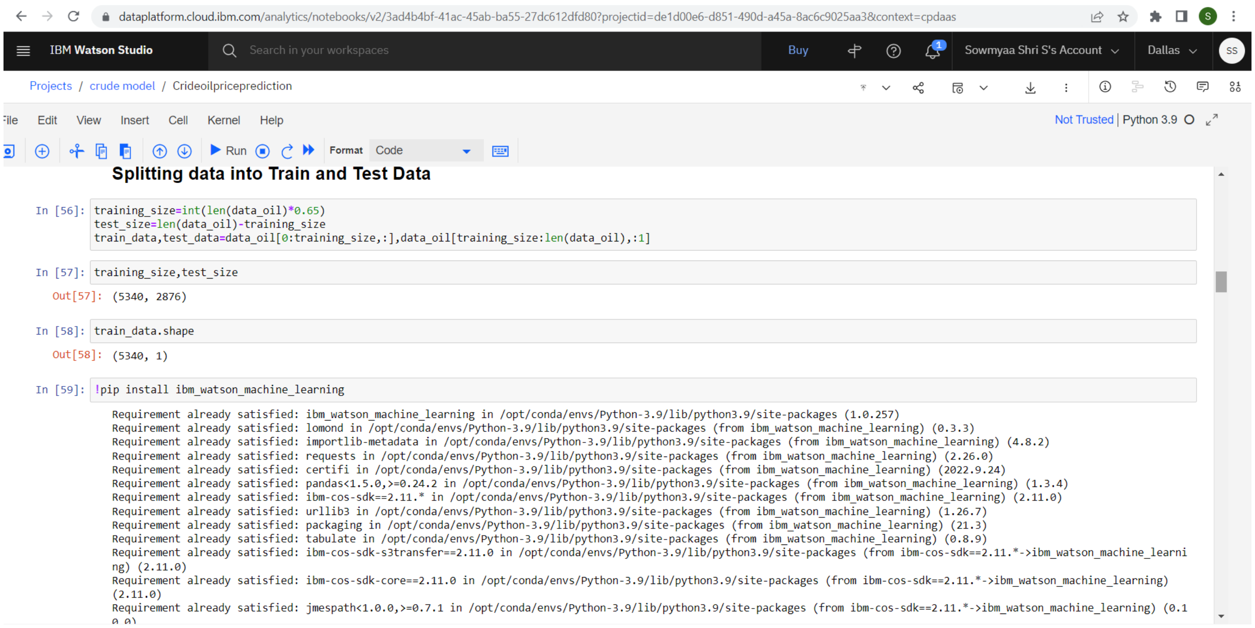
Feature Scaling

```
In [54]: from sklearn.preprocessing import MinMaxScaler
         scalar=MinMaxScaler(feature_range=(0,1))
         data_oil=scalar.fit_transform(np.array(data_oil).reshape(-1,1))
```

Data Visualization

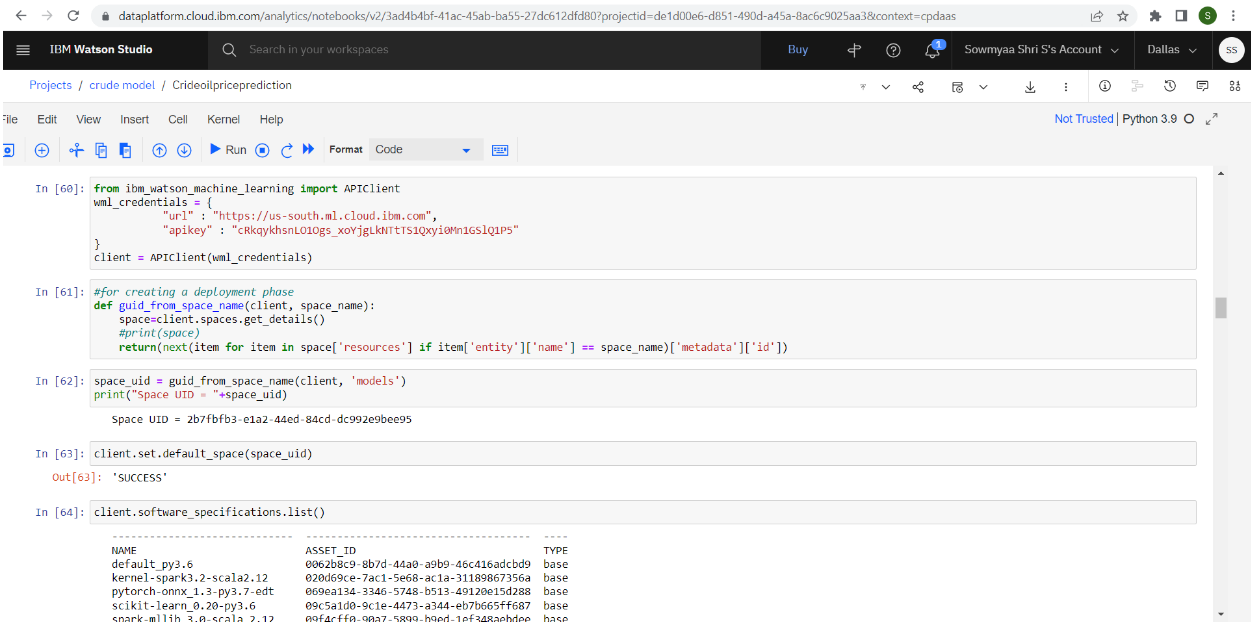
```
In [55]: plt.title('Crude oil price')
         plt.plot(data_oil)

Out[55]: [<matplotlib.lines.Line2D at 0x7f2530359790>]
```

Crude oil price
1.0
0.8
0.6
0.4
0.2
0.0
0
2000
4000
6000
8000

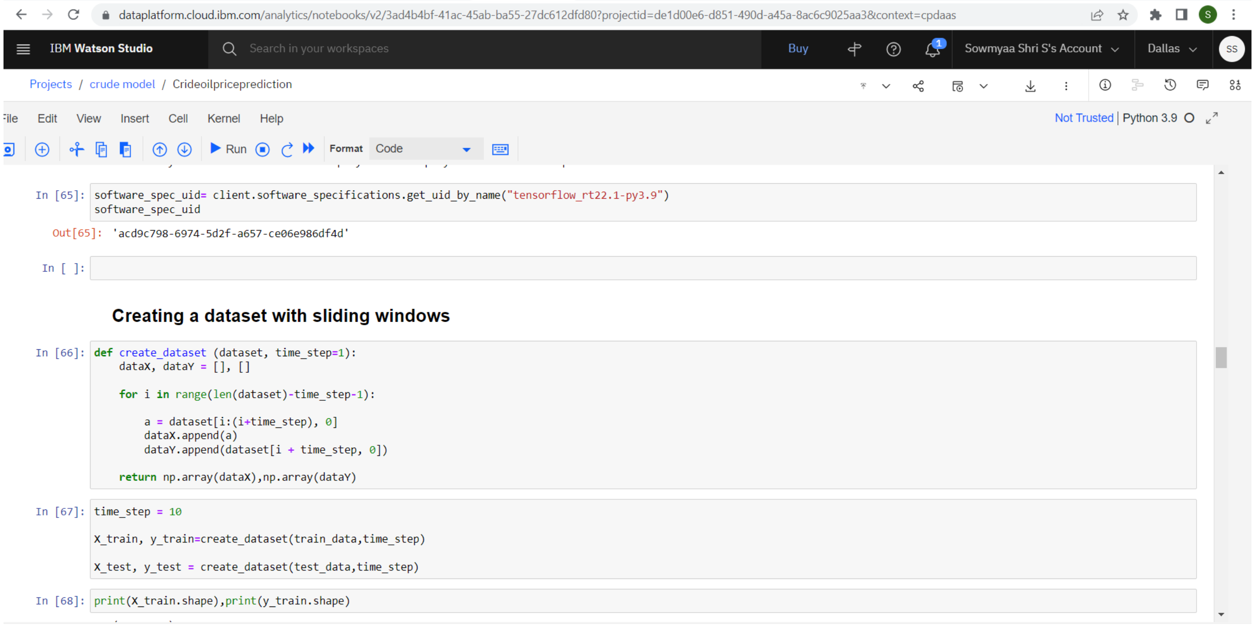

dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format
Code

Splitting data into Train and Test Data

```
In [56]: training_size=int(len(data_oil)*0.65)
         test_size=len(data_oil)-training_size
         train_data,test_data=data_oil[0:training_size,:],data_oil[training_size:len(data_oil),:1]

In [57]: training_size,test_size

Out[57]: (5340, 2876)

In [58]: train_data.shape

Out[58]: (5340, 1)

In [59]: !pip install ibm_watson_machine_learning

  Requirement already satisfied: ibm_watson_machine_learning in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (1.0.257)
  Requirement already satisfied: lomond in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (0.3.3)
  Requirement already satisfied: importlib-metadata in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (4.8.2)
  Requirement already satisfied: requests in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (2.26.0)
  Requirement already satisfied: certifi in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (2022.9.24)
  Requirement already satisfied: pandas<1.5.0,>=0.24.2 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (1.3.4)
  Requirement already satisfied: ibm-cos-sdk==2.11.* in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (2.11.0)
  Requirement already satisfied: urllib3 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (1.26.7)
  Requirement already satisfied: packaging in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (21.3)
  Requirement already satisfied: tabulate in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm_watson_machine_learning) (0.8.9)
  Requirement already satisfied: ibm-cos-sdk-s3transfer==2.11.0 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm_watson_machine_learni
  ng) (2.11.0)
  Requirement already satisfied: ibm-cos-sdk-core==2.11.0 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm_watson_machine_learning)
  (2.11.0)
  Requirement already satisfied: jmespath<1.0.0,>=0.7.1 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm_watson_machine_learning) (0.1
  0.0)
```


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format
Code

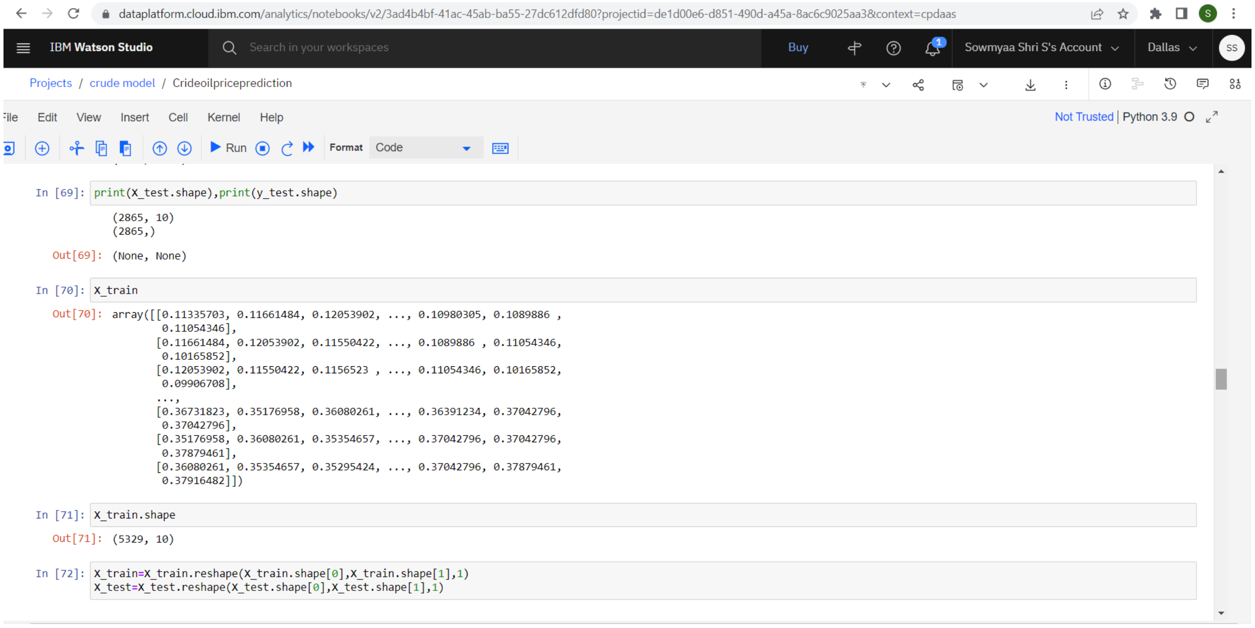
```
In [60]: from ibm_watson_machine_learning import APIClient
         wml_credentials = {
                   "url" : "https://us-south.ml.cloud.ibm.com",
                   "apikey" : "cRkqykhsnLO1Ogs_xoYjgLkNTtTS1Qxyi0Mn1GSlQ1P5"
         }
         client = APIClient(wml_credentials)

In [61]: #for creating a deployment phase
         def guid_from_space_name(client, space_name):
             space=client.spaces.get_details()
             #print(space)
             return(next(item for item in space['resources'] if item['entity']['name'] == space_name)['metadata']['id'])

In [62]: space_uid = guid_from_space_name(client, 'models')
         print("Space UID = "+space_uid)

  Space UID = 2b7fbfb3-e1a2-44ed-84cd-dc992e9bee95

In [63]: client.set.default_space(space_uid)

Out[63]: 'SUCCESS'

In [64]: client.software_specifications.list()

  ----------------------------  ------------------------------------  ----
  NAME                          ASSET_ID                              TYPE
  default_py3.6                 0062b8c9-8b7d-44a0-a9b9-46c416adcbd9  base
  kernel-spark3.2-scala2.12     020d69ce-7ac1-5e68-ac1a-31189867356a  base
  pytorch-onnx_1.3-py3.7-edt    069ea134-3346-5748-b513-49120e15d288  base
  scikit-learn_0.20-py3.6       09c5a1d0-9c1e-4473-a344-eb7b665ff687  base
  spark-mllib_3.0-scala_2.12    09f4cff0-90a7-5899-b9ed-1ef348aebdee  base
```


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format Code

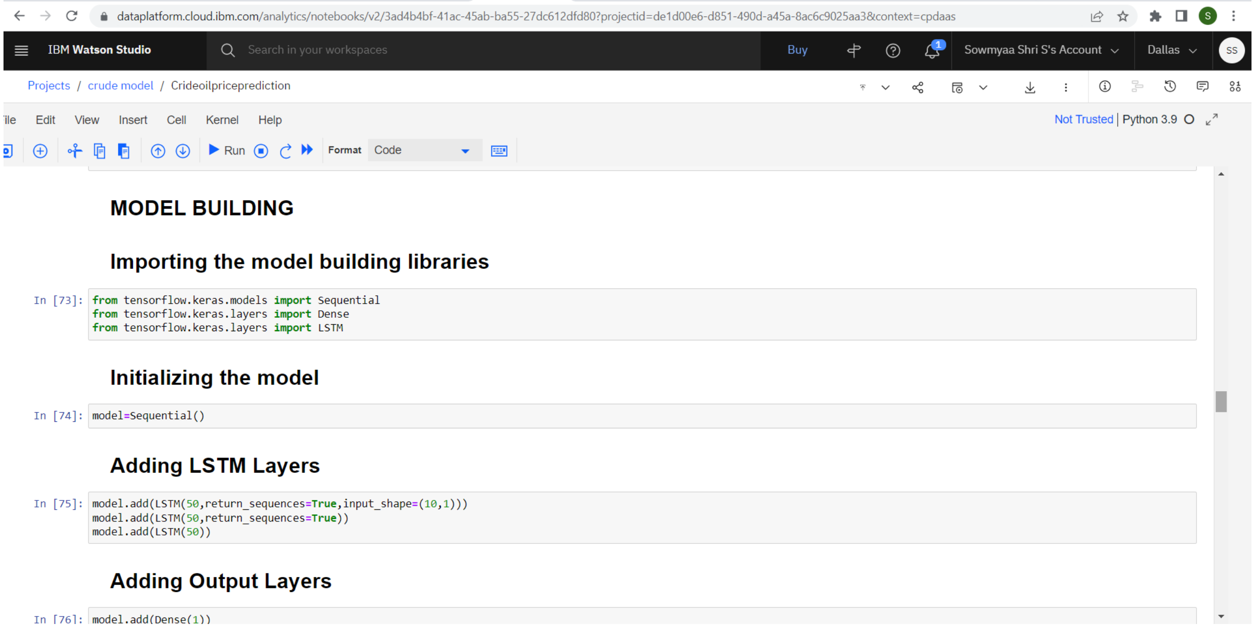
```
In [65]: software_spec_uid= client.software_specifications.get_uid_by_name("tensorflow_rt22.1-py3.9")
         software_spec_uid

Out[65]: 'acd9c798-6974-5d2f-a657-ce06e986df4d'

In [ ]:
```

Creating a dataset with sliding windows

```
In [66]: def create_dataset (dataset, time_step=1):
             dataX, dataY = [], []

             for i in range(len(dataset)-time_step-1):

                 a = dataset[i:(i+time_step), 0]
                 dataX.append(a)
                 dataY.append(dataset[i + time_step, 0])

             return np.array(dataX),np.array(dataY)

In [67]: time_step = 10

         X_train, y_train=create_dataset(train_data,time_step)

         X_test, y_test = create_dataset(test_data,time_step)

In [68]: print(X_train.shape),print(y_train.shape)
```


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format Code

```
In [69]: print(X_test.shape),print(y_test.shape)

           (2865, 10)
           (2865,)

Out[69]: (None, None)

In [70]: X_train

Out[70]: array([[0.11335703, 0.11661484, 0.12053902, ..., 0.10980305, 0.1089886 ,
                 0.11054346],
                [0.11661484, 0.12053902, 0.11550422, ..., 0.1089886 , 0.11054346,
                 0.10165852],
                [0.12053902, 0.11550422, 0.1156523 , ..., 0.11054346, 0.10165852,
                 0.09906708],
                ...,
                [0.36731823, 0.35176958, 0.36080261, ..., 0.36391234, 0.37042796,
                 0.37042796],
                [0.35176958, 0.36080261, 0.35354657, ..., 0.37042796, 0.37042796,
                 0.37879461],
                [0.36080261, 0.35354657, 0.35295424, ..., 0.37042796, 0.37879461,
                 0.37916482]])

In [71]: X_train.shape

Out[71]: (5329, 10)

In [72]: X_train=X_train.reshape(X_train.shape[0],X_train.shape[1],1)
         X_test=X_test.reshape(X_test.shape[0],X_test.shape[1],1)
```


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format
Code

MODEL BUILDING

Importing the model building libraries

```
In [73]: from tensorflow.keras.models import Sequential
         from tensorflow.keras.layers import Dense
         from tensorflow.keras.layers import LSTM
```

Initializing the model

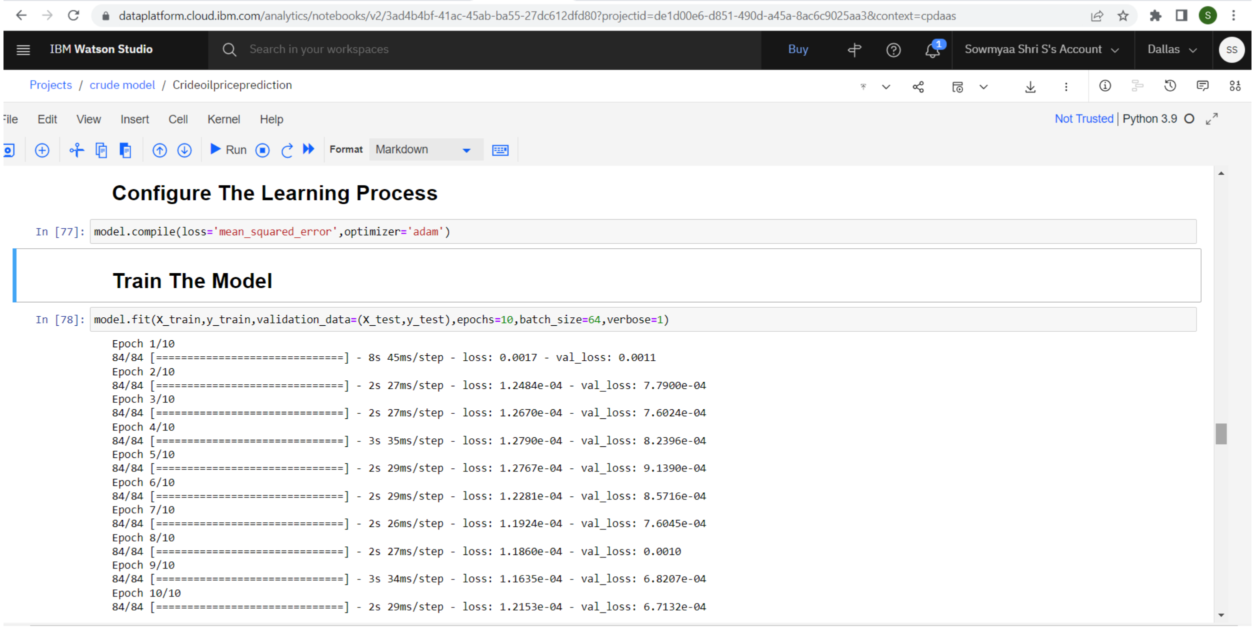
```
In [74]: model=Sequential()
```

Adding LSTM Layers

```
In [75]: model.add(LSTM(50,return_sequences=True,input_shape=(10,1)))
         model.add(LSTM(50,return_sequences=True))
         model.add(LSTM(50))
```

Adding Output Layers

```
In [76]: model.add(Dense(1))
```

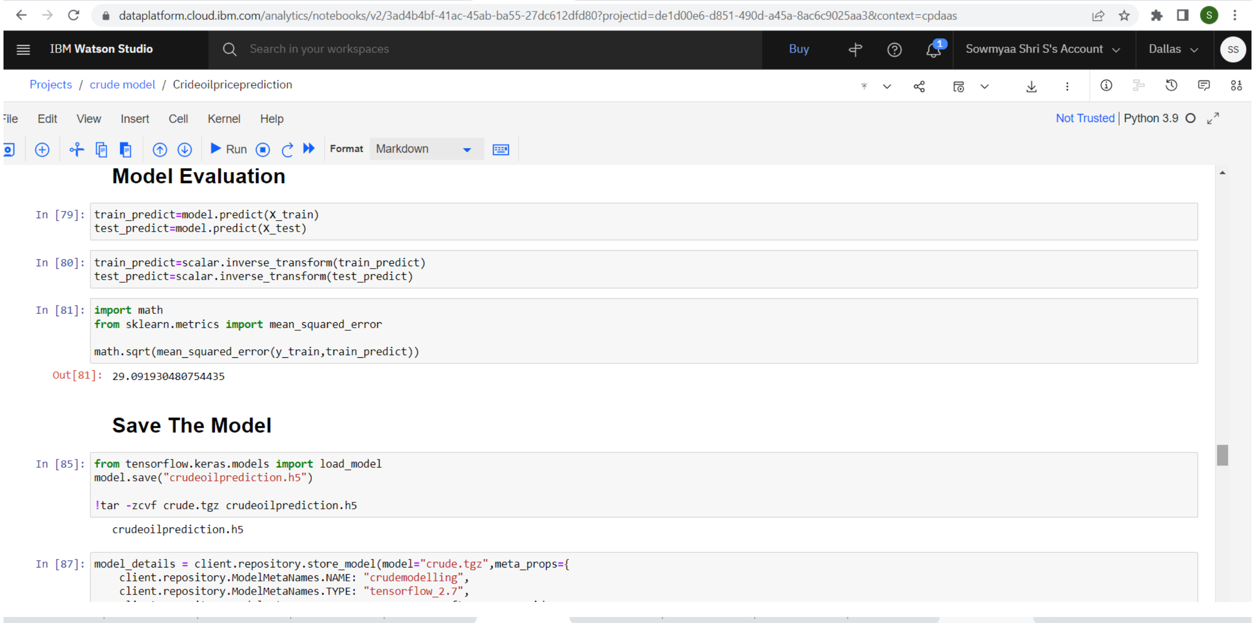

dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
SS
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format
Markdown

Configure The Learning Process

```
In [77]: model.compile(loss='mean_squared_error',optimizer='adam')
```

Train The Model

```
In [78]: model.fit(X_train,y_train,validation_data=(X_test,y_test),epochs=10,batch_size=64,verbose=1)

   Epoch 1/10
   84/84 [==============================] - 8s 45ms/step - loss: 0.0017 - val_loss: 0.0011
   Epoch 2/10
   84/84 [==============================] - 2s 27ms/step - loss: 1.2484e-04 - val_loss: 7.7900e-04
   Epoch 3/10
   84/84 [==============================] - 2s 27ms/step - loss: 1.2670e-04 - val_loss: 7.6024e-04
   Epoch 4/10
   84/84 [==============================] - 3s 35ms/step - loss: 1.2790e-04 - val_loss: 8.2396e-04
   Epoch 5/10
   84/84 [==============================] - 2s 29ms/step - loss: 1.2767e-04 - val_loss: 9.1390e-04
   Epoch 6/10
   84/84 [==============================] - 2s 29ms/step - loss: 1.2281e-04 - val_loss: 8.5716e-04
   Epoch 7/10
   84/84 [==============================] - 2s 26ms/step - loss: 1.1924e-04 - val_loss: 7.6045e-04
   Epoch 8/10
   84/84 [==============================] - 2s 27ms/step - loss: 1.1860e-04 - val_loss: 0.0010
   Epoch 9/10
   84/84 [==============================] - 3s 34ms/step - loss: 1.1635e-04 - val_loss: 6.8207e-04
   Epoch 10/10
   84/84 [==============================] - 2s 29ms/step - loss: 1.2153e-04 - val_loss: 6.7132e-04
```

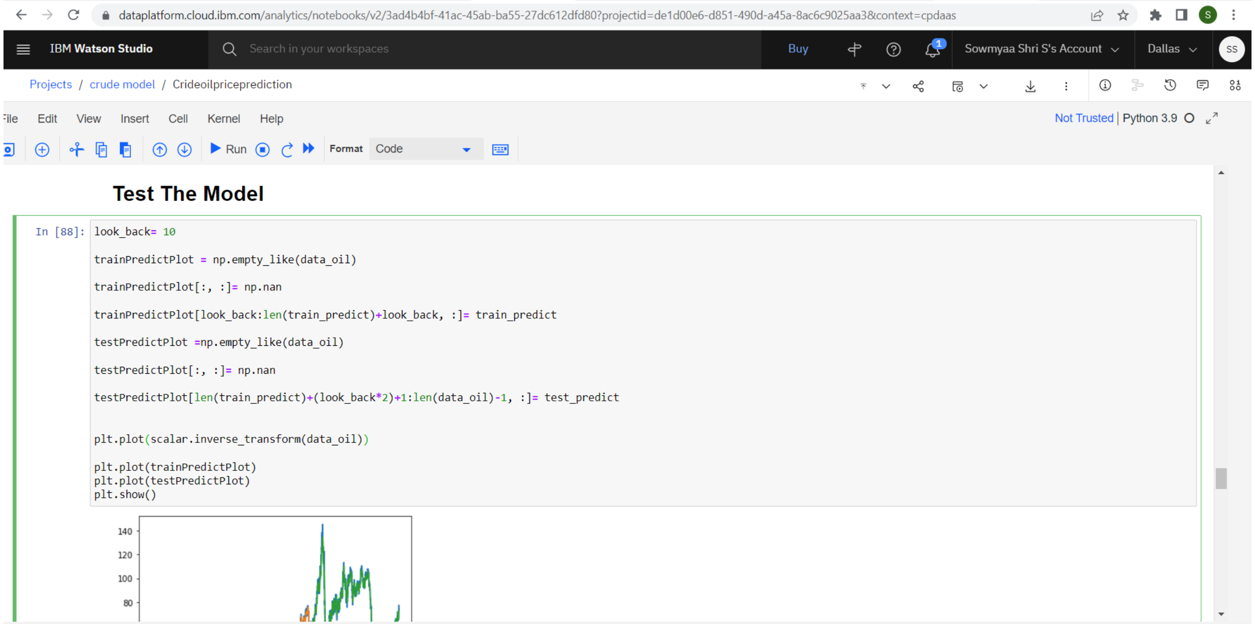

dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format Markdown

Model Evaluation

```
In [79]: train_predict=model.predict(X_train)
         test_predict=model.predict(X_test)

In [80]: train_predict=scalar.inverse_transform(train_predict)
         test_predict=scalar.inverse_transform(test_predict)

In [81]: import math
         from sklearn.metrics import mean_squared_error

         math.sqrt(mean_squared_error(y_train,train_predict))

Out[81]: 29.091930480754435
```

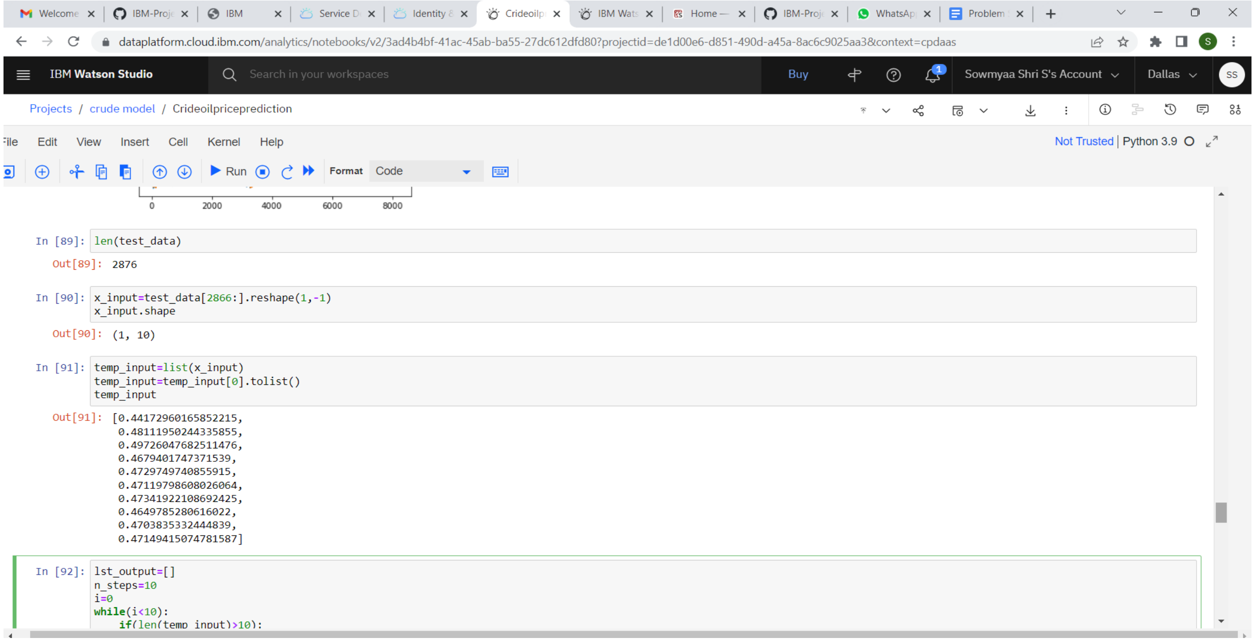
Save The Model

```
In [85]: from tensorflow.keras.models import load_model
         model.save("crudeoilprediction.h5")

         !tar -zcvf crude.tgz crudeoilprediction.h5

         crudeoilprediction.h5

In [87]: model_details = client.repository.store_model(model="crude.tgz",meta_props={
             client.repository.ModelMetaNames.NAME: "crudemodelling",
             client.repository.ModelMetaNames.TYPE: "tensorflow_2.7",
```


dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas
IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Not Trusted | Python 3.9
Run
Format Code
140
120
100
80

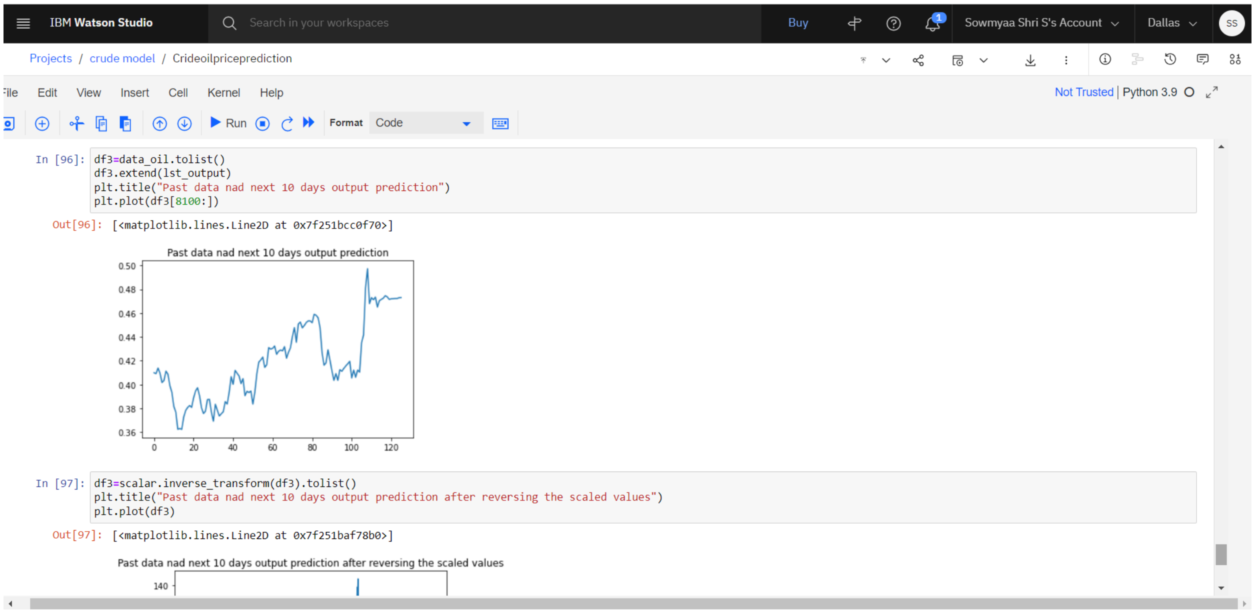
Test The Model

```
In [88]: look_back= 10

         trainPredictPlot = np.empty_like(data_oil)

         trainPredictPlot[:, :]= np.nan

         trainPredictPlot[look_back:len(train_predict)+look_back, :]= train_predict

         testPredictPlot =np.empty_like(data_oil)

         testPredictPlot[:, :]= np.nan

         testPredictPlot[len(train_predict)+(look_back*2)+1:len(data_oil)-1, :]= test_predict


         plt.plot(scalar.inverse_transform(data_oil))

         plt.plot(trainPredictPlot)
         plt.plot(testPredictPlot)
         plt.show()
```


IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Run
Format Code
Not Trusted | Python 3.9
0
2000
4000
6000
8000

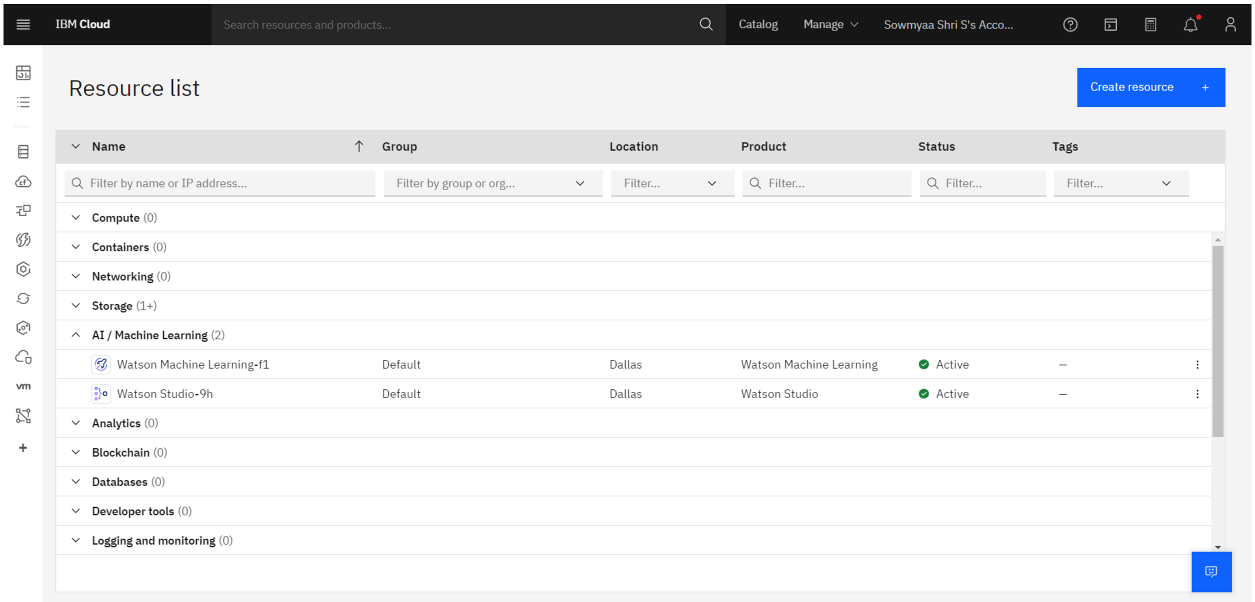
dataplatform.cloud.ibm.com/analytics/notebooks/v2/3ad4b4bf-41ac-45ab-ba55-27dc612dfd80?projectid=de1d00e6-d851-490d-a45a-8ac6c9025aa3&context=cpdaas

```
In [89]: len(test_data)

Out[89]: 2876

In [90]: x_input=test_data[2866:].reshape(1,-1)
         x_input.shape

Out[90]: (1, 10)

In [91]: temp_input=list(x_input)
         temp_input=temp_input[0].tolist()
         temp_input

Out[91]: [0.44172960165852215,
          0.48111950244335855,
          0.49726047682511476,
          0.4679401747371539,
          0.4729749740855915,
          0.47119798608026064,
          0.47341922108692425,
          0.4649785280616022,
          0.4703835332444839,
          0.47149415074781587]

In [92]: lst_output=[]
         n_steps=10
         i=0
         while(i<10):
             if(len(temp_input)>10):
```


IBM Watson Studio
Search in your workspaces
Buy
Sowmyaa Shri S's Account
Dallas
Projects / crude model / Crideoilpriceprediction
File Edit View Insert Cell Kernel Help
Run
Format Code
Not Trusted | Python 3.9
Past data nad next 10 days output prediction
0.50
0.48
0.46
0.44
0.42
0.40
0.38
0.36
0
20
40
60
80
100
120
Past data nad next 10 days output prediction after reversing the scaled values
140

```
In [96]: df3=data_oil.tolist()
         df3.extend(lst_output)
         plt.title("Past data nad next 10 days output prediction")
         plt.plot(df3[8100:])

Out[96]: [<matplotlib.lines.Line2D at 0x7f251bcc0f70>]

In [97]: df3=scalar.inverse_transform(df3).tolist()
         plt.title("Past data nad next 10 days output prediction after reversing the scaled values")
         plt.plot(df3)

Out[97]: [<matplotlib.lines.Line2D at 0x7f251baf78b0>]
```

2. Resources list



Resource list

Name	Group	Location	Product	Status	Tags
Compute (0)					
Containers (0)					
Networking (0)					
Storage (1+)					
AI / Machine Learning (2)					
Watson Machine Learning-f1	Default	Dallas	Watson Machine Learning	● Active	–
Watson Studio-9h	Default	Dallas	Watson Studio	● Active	–
Analytics (0)					
Blockchain (0)					
Databases (0)					
Developer tools (0)					
Logging and monitoring (0)					