

IBM EPBL/IBM-Project-15304-16 Crude Oil Price Prediction(Final D...

dataplatform.cloud.ibm.com/analytics/notebooks/v2/556c44a3-e8ea-4343-b32c-33494198be0b/view?projectid=3d8925a2-0bfe-4f56-b114-0412f53a77d7&c...

IBM Watson Studio

Projects / Crude Oil Price Prediction / Crude Oil Price Prediction(Final D...

```
In [2]: # Import statements
import numpy as np
import pandas as pd
import datetime
import matplotlib.pyplot as plt
import warnings
import itertools
import statsmodels.api as sm
import seaborn as sns
import math
from pylab import rcParams
from sklearn.preprocessing import MinMaxScaler

#Jupyter Notebook Specific
sns.set_context("paper", font_scale=1.3)
sns.set_style('white')
warnings.filterwarnings("ignore")
plt.style.use('fivethirtyeight')

In [3]: import os, types
import pandas as pd
from boto3.client import Config
import boto3

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 = boto3.client(service_name='s3',
    ibm_api_key_id='1xEp1bzVw0XmoQ0nKcfCFMjUvezzzWxvUmsPIDP9VSF',
    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 = 'crudeoilpriceprediction-donotdelete-pr-ekltddoz8wv161'
```

https://dataplatform.cloud.ibm.com/analytics/notebooks/v2/556c44a3-e8ea-4343-b32c-33494198be0b/view?projectid=3d8925a2-0bfe-4f56-b114-0412f53a77d7&context=cpdaas#

IBM EPBL/IBM-Project-15304-16 Crude Oil Price Prediction(Final D...

dataplatform.cloud.ibm.com/analytics/notebooks/v2/556c44a3-e8ea-4343-b32c-33494198be0b/view?projectid=3d8925a2-0bfe-4f56-b114-0412f53a77d7&c...

IBM Watson Studio

Projects / Crude Oil Price Prediction / Crude Oil Price Prediction(Final D...

```
# 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 = boto3.client(service_name='s3',
    ibm_api_key_id='1xEp1bzVw0XmoQ0nKcfCFMjUvezzzWxvUmsPIDP9VSF',
    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 = 'crudeoilpriceprediction-donotdelete-pr-ekltddoz8wv161'
object_key = 'Crude Oil Prices Daily.xlsx'

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

df_data_0 = pd.read_excel(body.read())
df_data_0['Date'] = pd.to_datetime(df_data_0['Date'])

In [4]: df_data_0.isnull().any()

Out[4]: Date False
Closing Value True
dtype: bool

In [5]: df_data_0.dropna(axis=0, inplace=True)

In [6]: data_oil = df_data_0.reset_index()[['Closing Value']]

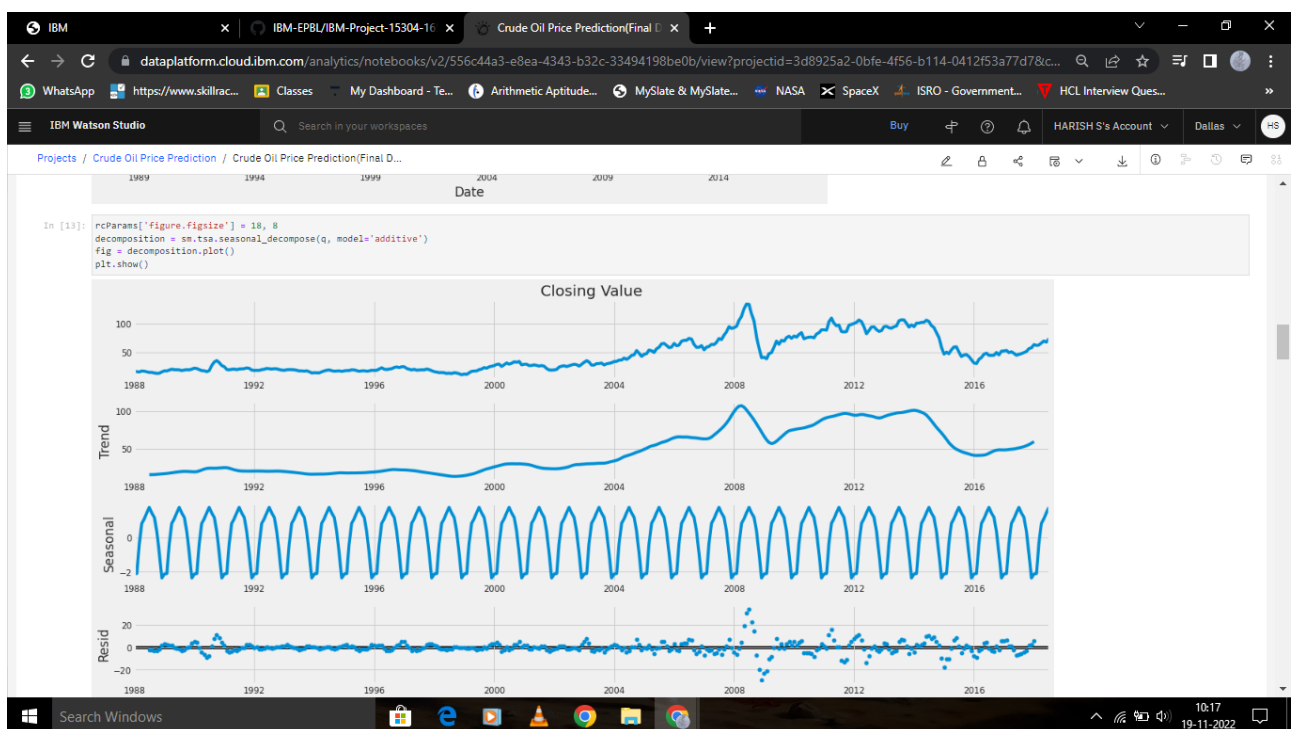
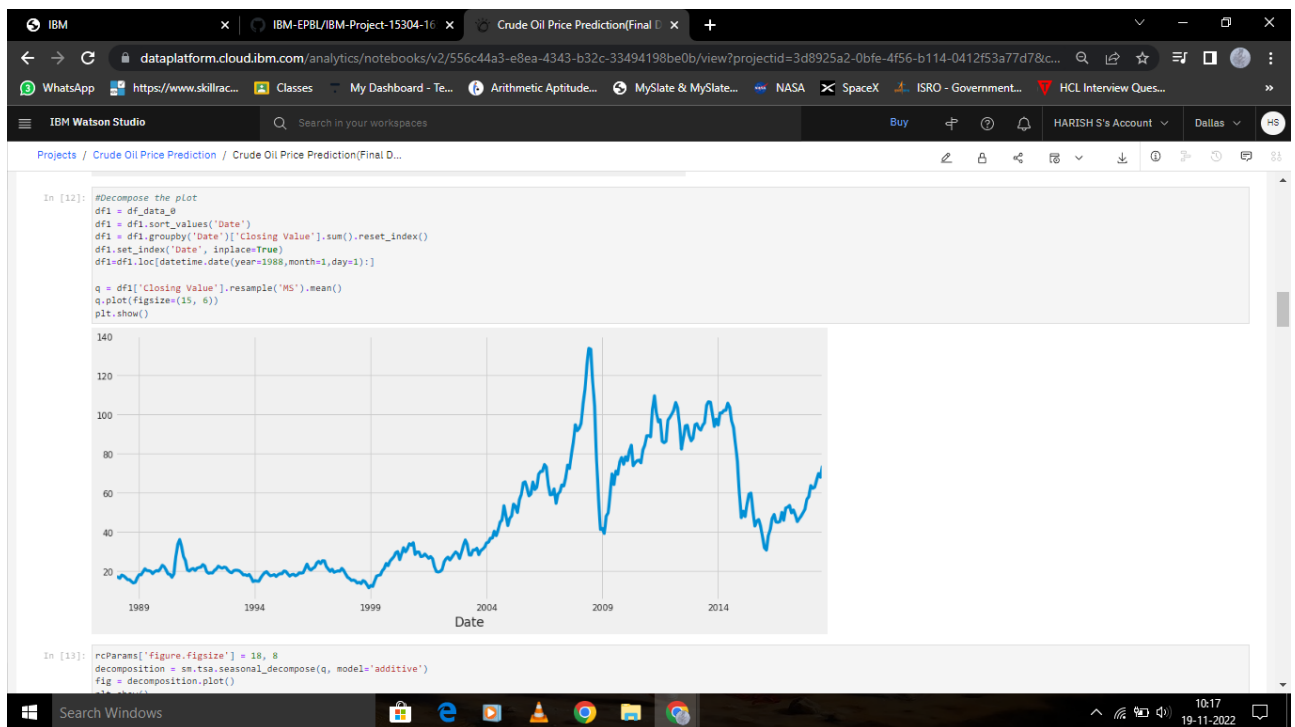
In [7]: len(data_oil)

Out[7]: 8216

In [9]: plt.1 = plt.figure(figsize=(15, 6))
time = pd.to_datetime(df_data_0['Date'])
data = list(df_data_0['Closing Value'])
copdata = pd.Series(data, time)
plt.plot(copdata)

Out[9]: [matplotlib.lines.Line2D at 0x7f27702a8bb0]
```





```
IBM | IBM-EPBL/IBM-Project-15304-16 | Crude Oil Price Prediction(Final D... | +
dataplatform.cloud.ibm.com/analytics/notebooks/v2/556c44a3-e8ea-4343-b32c-33494198be0b/view?projectid=3d8925a2-0bfe-4f56-b114-0412f53a77d7&c...
WhatsApp | https://www.skillrac... | Classes | My Dashboard - Te... | Arithmetic Aptitude... | MySlate & MySlate... | NASA | SpaceX | ISRO - Government... | HCL Interview Ques...
IBM Watson Studio | Search in your workspaces | Buy | HARISH S's Account | Dallas | HS

Projects / Crude Oil Price Prediction / Crude Oil Price Prediction(Final D...

In [14]: sc = MinMaxScaler(feature_range = (0, 1))
data_oil = sc.fit_transform(np.array(data_oil).reshape(-1,1))

In [15]: len(data_oil)
Out[15]: 8216

In [16]: train_size = int(len(data_oil) * 0.65)
test_size = len(data_oil) - train_size
train, test = data_oil[0:train_size, :], data_oil[train_size:len(data_oil), :]

In [17]: len(test)
Out[17]: 2876

In [18]: def create_dataset(_data_set, _look_back=1):
data_x, data_y = [], []
for i in range(len(_data_set) - _look_back - 1):
a = _data_set[i:(i + _look_back), 0]
data_x.append(a)
data_y.append(_data_set[i + _look_back, 0])
return np.array(data_x), np.array(data_y)

time_step = 10
X_train, Y_train = create_dataset(train, time_step)
X_test, Y_test = create_dataset(test, time_step)

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)

In [19]: X_train.shape
Out[19]: (5329, 10, 1)

In [20]: from keras.models import Sequential
from keras.layers import Dense
from keras.layers import LSTM
```

```
IBM | IBM-EPBL/IBM-Project-15304-16 | Crude Oil Price Prediction(Final D... | +
dataplatform.cloud.ibm.com/analytics/notebooks/v2/556c44a3-e8ea-4343-b32c-33494198be0b/view?projectid=3d8925a2-0bfe-4f56-b114-0412f53a77d7&c...
WhatsApp | https://www.skillrac... | Classes | My Dashboard - Te... | Arithmetic Aptitude... | MySlate & MySlate... | NASA | SpaceX | ISRO - Government... | HCL Interview Ques...
IBM Watson Studio | Search in your workspaces | Buy | HARISH S's Account | Dallas | HS

Projects / Crude Oil Price Prediction / Crude Oil Price Prediction(Final D...

Out[19]: (5329, 10, 1)

In [20]: from keras.models import Sequential
from keras.layers import Dense
from keras.layers import LSTM
from keras.layers import Dropout
from sklearn.metrics import mean_squared_error
from keras.callbacks import ReduceLROnPlateau, EarlyStopping, ModelCheckpoint
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error

In [21]: regressor = Sequential()
regressor.add(LSTM(units = 60, return_sequences = True, input_shape = (X_train.shape[1], 1)))
regressor.add(Dropout(0.1))

regressor.add(LSTM(units = 60, return_sequences = True))
regressor.add(Dropout(0.1))

regressor.add(LSTM(units = 60))
regressor.add(Dropout(0.1))

regressor.add(Dense(units = 1))

regressor.compile(optimizer = 'adam', loss = 'mean_squared_error')
# early_stopping = EarlyStopping(monitor='val_loss', patience=20)
history = regressor.fit(X_train, Y_train, epochs = 20, batch_size = 64, validation_data=(X_test, Y_test), verbose=1)
# history = regressor.fit(X_train, Y_train, epochs = 100, batch_size = 64, validation_data=(X_test, Y_test))

Epoch 1/20
84/84 [=====] - 5s 22ms/step - loss: 0.0016 - val_loss: 0.0018
Epoch 2/20
84/84 [=====] - 1s 14ms/step - loss: 2.2518e-04 - val_loss: 7.3368e-04
Epoch 3/20
84/84 [=====] - 1s 16ms/step - loss: 2.0074e-04 - val_loss: 7.3821e-04
Epoch 4/20
84/84 [=====] - 1s 14ms/step - loss: 1.9867e-04 - val_loss: 7.6358e-04
Epoch 5/20
84/84 [=====] - 1s 14ms/step - loss: 1.9252e-04 - val_loss: 0.0010
Epoch 6/20
```

IBM Watson Studio interface showing a Jupyter Notebook titled "Crude Oil Price Prediction(Final D...". The notebook contains Python code for training a neural network model using Keras and TensorFlow.

```

regressor.add(Dropout(0.1))

regressor.add(Dense(units = 1))

regressor.compile(optimizer = 'adam', loss = 'mean_squared_error')
# early_stopping = EarlyStopping(monitor='val_loss',patience=20)
history = regressor.fit(X_train, Y_train, epochs = 20, batch_size = 64, validation_data=(X_test, Y_test), verbose=1)
# history = regressor.fit(X_train, Y_train, epochs = 100, batch_size = 64, validation_data=(X_test, Y_test))

Epoch 1/20
84/84 [=====] - 5s 22ms/step - loss: 0.0016 - val_loss: 0.0018
Epoch 2/20
84/84 [=====] - 1s 14ms/step - loss: 2.2518e-04 - val_loss: 7.3368e-04
Epoch 3/20
84/84 [=====] - 1s 16ms/step - loss: 2.0074e-04 - val_loss: 7.3821e-04
Epoch 4/20
84/84 [=====] - 1s 14ms/step - loss: 1.9867e-04 - val_loss: 7.6358e-04
Epoch 5/20
84/84 [=====] - 1s 14ms/step - loss: 1.9252e-04 - val_loss: 0.0010
Epoch 6/20
84/84 [=====] - 1s 14ms/step - loss: 1.9391e-04 - val_loss: 7.7598e-04
Epoch 7/20
84/84 [=====] - 1s 15ms/step - loss: 2.0522e-04 - val_loss: 6.7668e-04
Epoch 8/20
84/84 [=====] - 2s 20ms/step - loss: 1.9437e-04 - val_loss: 6.8141e-04
Epoch 9/20
84/84 [=====] - 1s 16ms/step - loss: 1.7695e-04 - val_loss: 8.4088e-04
Epoch 10/20
84/84 [=====] - 2s 18ms/step - loss: 1.7679e-04 - val_loss: 6.4298e-04
Epoch 11/20
84/84 [=====] - 1s 14ms/step - loss: 1.6428e-04 - val_loss: 6.2306e-04
Epoch 12/20
84/84 [=====] - 1s 16ms/step - loss: 1.5652e-04 - val_loss: 8.3206e-04
Epoch 13/20
84/84 [=====] - 1s 17ms/step - loss: 1.5308e-04 - val_loss: 5.8922e-04
Epoch 14/20
84/84 [=====] - 2s 18ms/step - loss: 1.5349e-04 - val_loss: 5.4442e-04
Epoch 15/20
84/84 [=====] - 2s 21ms/step - loss: 1.4928e-04 - val_loss: 9.9499e-04
Epoch 16/20
84/84 [=====] - 1s 15ms/step - loss: 1.5385e-04 - val_loss: 5.2495e-04

```

IBM Watson Studio interface showing the same Jupyter Notebook, now displaying the results of the model evaluation and a plot of the training and testing loss over epochs.

```

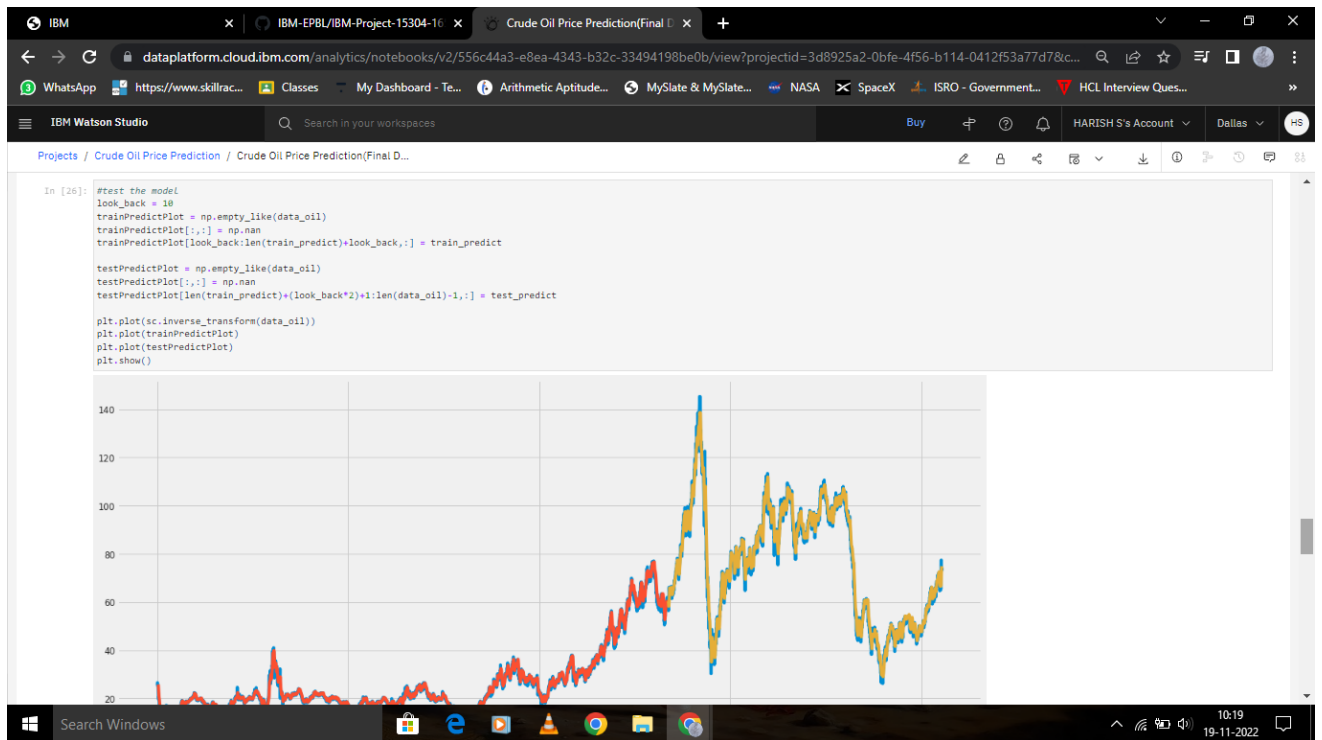
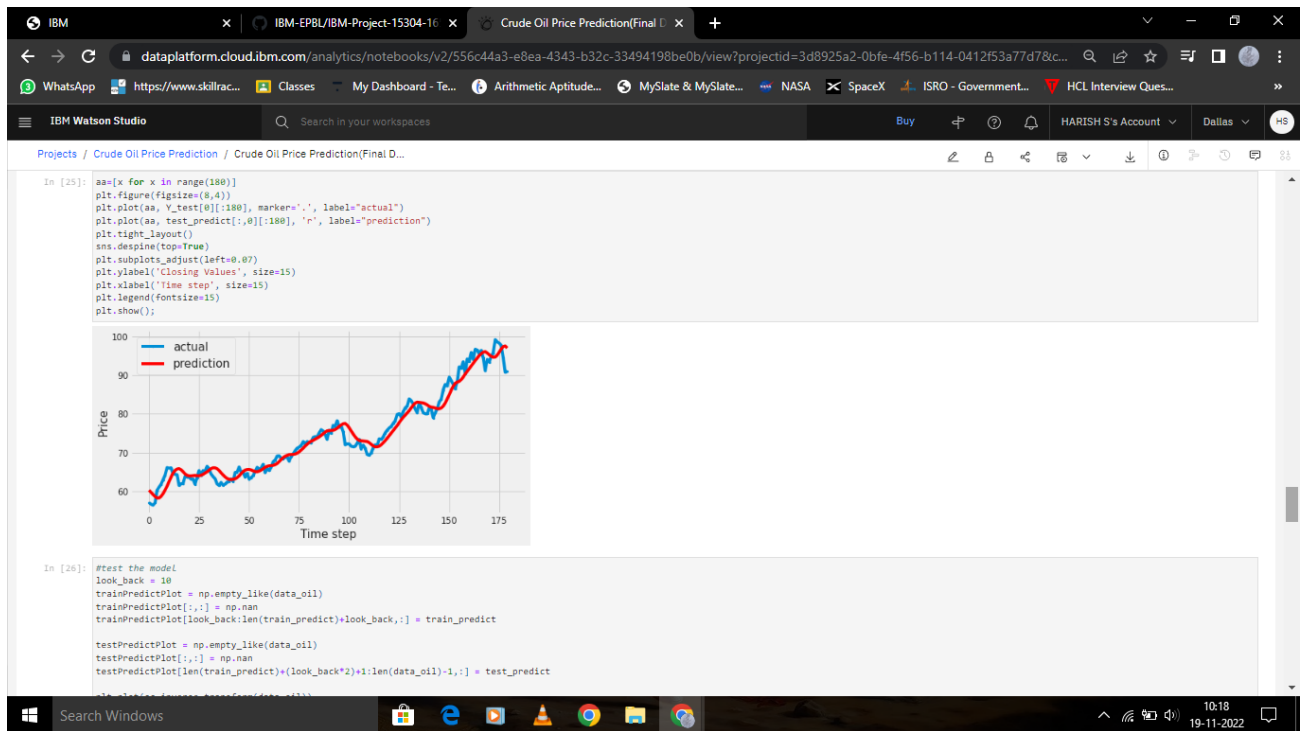
In [23]: # inverse predictions
train_predict = sc.inverse_transform(train_predict)
Y_train = sc.inverse_transform(Y_train)
test_predict = sc.inverse_transform(test_predict)
Y_test = sc.inverse_transform(Y_test)

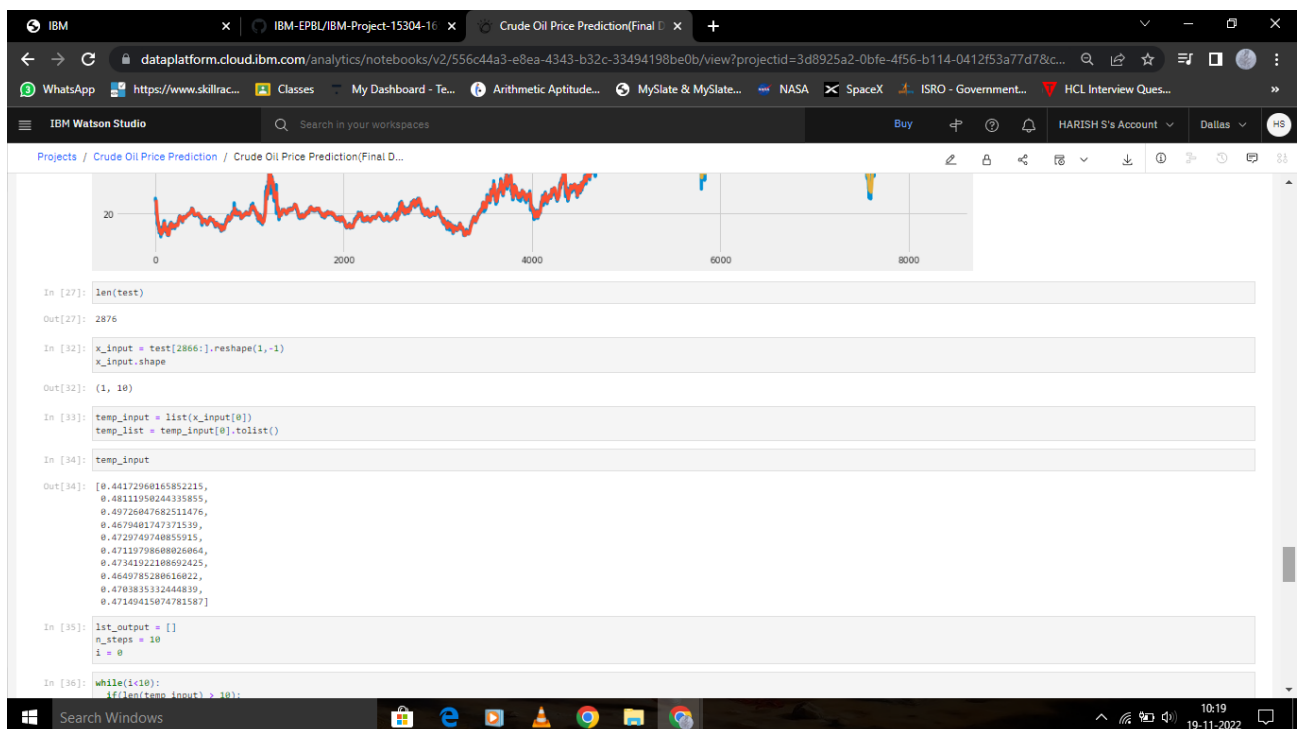
In [24]: print('Train Mean Absolute Error:', mean_absolute_error(Y_train[0], train_predict[:,0]))
print('Train Root Mean Squared Error:', np.sqrt(mean_squared_error(Y_train[0], train_predict[:,0])))
print('Test Mean Absolute Error:', mean_absolute_error(Y_test[0], test_predict[:,0]))
print('Test Root Mean Squared Error:', np.sqrt(mean_squared_error(Y_test[0], test_predict[:,0])))

plt.figure(figsize=(8,4))
plt.plot(history.history['loss'], label='Train Loss')
plt.plot(history.history['val_loss'], label='Test Loss')
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epochs')
plt.legend(loc='upper right')
plt.show()

Train Mean Absolute Error: 0.0002561245450949
Train Root Mean Squared Error: 1.295364160453239
Test Mean Absolute Error: 2.1252786170200526
Test Root Mean Squared Error: 2.873785434361001

```





```
IBM | IBM-EPBL/IBM-Project-15304-16 | Crude Oil Price Prediction(Final I | +
dataplatform.cloud.ibm.com/analytics/notebooks/v2/556c44a3-e8ea-4343-b32c-33494198be0b/view?projectid=3d8925a2-0bfe-4f56-b114-0412f53a77d7&c...
WhatsApp | https://www.skillrac... | Classes | My Dashboard - Te... | Arithmetic Aptitude... | MySlate & MySlate... | NASA | SpaceX | ISRO - Government... | HCL Interview Ques...
IBM Watson Studio | Search in your workspaces | Buy | HARISH S's Account | Dallas | HS

Projects / Crude Oil Price Prediction / Crude Oil Price Prediction(Final D...

In [36]: while(i<10):
          if(len(temp_input) > 10):
              x_input = np.array(temp_input[1:])
              print("{} day input {}".format(i,x_input),end="\n")
              x_input = x_input.reshape(1,-1)
              x_input = x_input.reshape((1,n_steps,1))

              yhat = regressor.predict(x_input,verbose=0)
              print("{} day output {}".format(i,yhat),end="\n")
              temp_input.extend(yhat[0].tolist())
              temp_input = temp_input[1:]
              print("-----",end="\n")
              list_output.extend(yhat.tolist())
              i = i+1
          else:
              x_input = x_input.reshape((1,n_steps,1))
              yhat = regressor.predict(x_input,verbose=0)
              print("{} day output {}".format(i,yhat),end="\n")
              temp_input.extend(yhat[0].tolist())
              list_output.extend(yhat.tolist())
              i = i+1

0 day output [[0.47814104]]
1 day input [[0.4811195 0.49726048 0.46794017 0.47297497 0.47119799 0.47341922
0.46497853 0.47038353 0.47149415 0.47814104]]
1 day output [[0.47829878]]
-----
2 day input [[0.49726048 0.46794017 0.47297497 0.47119799 0.47341922 0.46497853
0.47038353 0.47149415 0.47814104 0.47829878]]
2 day output [[0.47898063]]
-----
3 day input [[0.46794017 0.47297497 0.47119799 0.47341922 0.46497853 0.47038353
0.47149415 0.47814104 0.47829878 0.47898063]]
3 day output [[0.48020533]]
-----
4 day input [[0.47297497 0.47119799 0.47341922 0.46497853 0.47038353 0.47149415
0.47814104 0.47829878 0.47898063 0.48020533]]
4 day output [[0.48179677]]
```

```
IBM | IBM-EPBL/IBM-Project-15304-16 | Crude Oil Price Prediction(Final I | +
dataplatform.cloud.ibm.com/analytics/notebooks/v2/556c44a3-e8ea-4343-b32c-33494198be0b/view?projectid=3d8925a2-0bfe-4f56-b114-0412f53a77d7&c...
WhatsApp | https://www.skillrac... | Classes | My Dashboard - Te... | Arithmetic Aptitude... | MySlate & MySlate... | NASA | SpaceX | ISRO - Government... | HCL Interview Ques...
IBM Watson Studio | Search in your workspaces | Buy | HARISH S's Account | Dallas | HS

Projects / Crude Oil Price Prediction / Crude Oil Price Prediction(Final D...

list_output.extend(yhat.tolist())
i = i+1

0 day output [[0.47814104]]
1 day input [[0.4811195 0.49726048 0.46794017 0.47297497 0.47119799 0.47341922
0.46497853 0.47038353 0.47149415 0.47814104]]
1 day output [[0.47829878]]
-----
2 day input [[0.49726048 0.46794017 0.47297497 0.47119799 0.47341922 0.46497853
0.47038353 0.47149415 0.47814104 0.47829878]]
2 day output [[0.47898063]]
-----
3 day input [[0.46794017 0.47297497 0.47119799 0.47341922 0.46497853 0.47038353
0.47149415 0.47814104 0.47829878 0.47898063]]
3 day output [[0.48020533]]
-----
4 day input [[0.47297497 0.47119799 0.47341922 0.46497853 0.47038353 0.47149415
0.47814104 0.47829878 0.47898063 0.48020533]]
4 day output [[0.48179677]]
-----
5 day input [[0.47119799 0.47341922 0.46497853 0.47038353 0.47149415 0.47814104
0.47829878 0.47898063 0.48020533 0.48179677]]
5 day output [[0.48339531]]
-----
6 day input [[0.47341922 0.46497853 0.47038353 0.47149415 0.47814104 0.47829878
0.47898063 0.48020533 0.48179677 0.48339531]]
6 day output [[0.48500329]]
-----
7 day input [[0.46497853 0.47038353 0.47149415 0.47814104 0.47829878 0.47898063
0.48020533 0.48179677 0.48339531 0.48500329]]
7 day output [[0.48659298]]
-----
8 day input [[0.47038353 0.47149415 0.47814104 0.47829878 0.47898063 0.48020533
0.48179677 0.48339531 0.48500329 0.48659298]]
8 day output [[0.48817706]]
-----
9 day input [[0.47149415 0.47814104 0.47829878 0.47898063 0.48020533 0.48179677
0.48339531 0.48500329 0.48659298 0.48817706]]
9 day output [[0.48970482]]
```