LGM - Data Science Internship - Dec 2023¶

Stock Market Prediction And Forecasting Using Stacked LSTM

In [23]:
```
# Importing the necessary Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.preprocessing import MinMaxScaler
```

In [2]:
```
# # Data Processing
data = pd.read_csv("C:/Users/MANOJ S/Downloads/NSE-TATAGLOBAL.csv")
data
```

Out[2]:

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
0	2018-09-28	234.05	235.95	230.20	233.50	233.75	3069914	7162.35
1	2018-09-27	234.55	236.80	231.10	233.80	233.25	5082859	11859.95
2	2018-09-26	240.00	240.00	232.50	235.00	234.25	2240909	5248.60
3	2018-09-25	233.30	236.75	232.00	236.25	236.10	2349368	5503.90
4	2018-09-24	233.55	239.20	230.75	234.00	233.30	3423509	7999.55
...	...	...	...	...	...	...	...	...
2030	2010-07-27	117.60	119.50	112.00	118.80	118.65	586100	694.98
2031	2010-07-26	120.10	121.00	117.10	117.10	117.60	658440	780.01
2032	2010-07-23	121.80	121.95	120.25	120.35	120.65	281312	340.31
2033	2010-07-22	120.30	122.00	120.25	120.75	120.90	293312	355.17
2034	2010-07-21	122.10	123.00	121.05	121.10	121.55	658666	803.56

2035 rows × 8 columns

In [3]:
```
data.head() # Shows first 5 rows of Dataset
```

Out[3]:

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
0	2018-09-28	234.05	235.95	230.20	233.50	233.75	3069914	7162.35
1	2018-09-27	234.55	236.80	231.10	233.80	233.25	5082859	11859.95
2	2018-09-26	240.00	240.00	232.50	235.00	234.25	2240909	5248.60
3	2018-09-25	233.30	236.75	232.00	236.25	236.10	2349368	5503.90
4	2018-09-24	233.55	239.20	230.75	234.00	233.30	3423509	7999.55

In [4]:
```
data.tail()
```

Out[4]:

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
2030	2010-07-27	117.6	119.50	112.00	118.80	118.65	586100	694.98
2031	2010-07-26	120.1	121.00	117.10	117.10	117.60	658440	780.01
2032	2010-07-23	121.8	121.95	120.25	120.35	120.65	281312	340.31
2033	2010-07-22	120.3	122.00	120.25	120.75	120.90	293312	355.17
2034	2010-07-21	122.1	123.00	121.05	121.10	121.55	658666	803.56

In [5]:
```
#Sort the dataset according to the date
data['Date'] = pd.to_datetime(data['Date'])
df = data.sort_values(by='Date')
df.head()
```

Out[5]:

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
2034	2010-07-21	122.1	123.00	121.05	121.10	121.55	658666	803.56
2033	2010-07-22	120.3	122.00	120.25	120.75	120.90	293312	355.17
2032	2010-07-23	121.8	121.95	120.25	120.35	120.65	281312	340.31
2031	2010-07-26	120.1	121.00	117.10	117.10	117.60	658440	780.01
2030	2010-07-27	117.6	119.50	112.00	118.80	118.65	586100	694.98

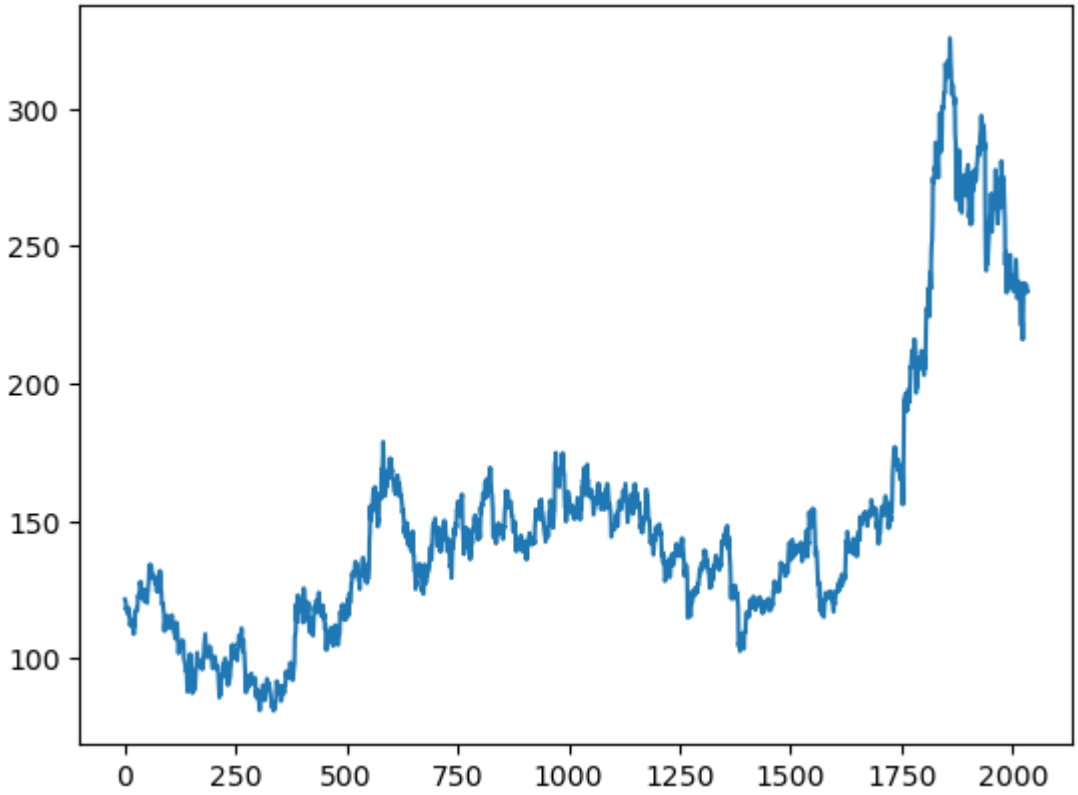
In [6]:
```
# # Index is not in order, Lets make it in order
df.reset_index(inplace=True)
df.head()
```

Out[6]:

	index	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
0	2034	2010-07-21	122.1	123.00	121.05	121.10	121.55	658666	803.56
1	2033	2010-07-22	120.3	122.00	120.25	120.75	120.90	293312	355.17
2	2032	2010-07-23	121.8	121.95	120.25	120.35	120.65	281312	340.31
3	2031	2010-07-26	120.1	121.00	117.10	117.10	117.60	658440	780.01
4	2030	2010-07-27	117.6	119.50	112.00	118.80	118.65	586100	694.98

In [7]:
```
plt.plot(df['Close'])
```

Out[7]:
```
[<matplotlib.lines.Line2D at 0x26fd925c640>]
```



In [8]:
```
df1 = df['Close']
df1
```

Out[8]:
```
0       121.55
1       120.90
2       120.65
3       117.60
4       118.65
         ...
2030    233.30
2031    236.10
2032    234.25
2033    233.25
2034    233.75
Name: Close, Length: 2035, dtype: float64
```

In [9]:
```
scaler = MinMaxScaler (feature_range=(0,1))
df1 = scaler.fit_transform(np.array(df1).reshape(-1,1))
df1
```

Out[9]:
```
array([[0.16584967],
       [0.16319444],
       [0.1621732 ],
       ...,
       [0.62622549],
       [0.62214052],
       [0.62418301]])
```

In [10]:
```
 # Splitting the data into training and test sets

training_size = int(len(df1)*0.70)
test_size = len(df1)-training_size

train_data = df1[0:training_size,:]
test_data = df1[training_size:len (df1), :1]
```

In [11]:
```
#convert an array of values into a dataset matrix
def datasetCrtd(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 [12]:
```
time_step = 100
X_train, y_train = datasetCrtd(train_data, time_step)
X_test, ytest = datasetCrtd(test_data, time_step)
```

In [13]:
```
print(X_train.shape), print(y_train.shape)
```

```
(1323, 100)
(1323,)
```

Out[13]:
```
(None, None)
```

In [14]:
```
#Reshape input which is required for LSTM
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 [15]:
```
# # Building the Model
!pip install tensorflow
```

```
Requirement already satisfied: tensorflow in c:\users\manoj s\anaconda3\li
b\site-packages (2.15.0)
Requirement already satisfied: tensorflow-intel==2.15.0 in c:\users\manoj
s\anaconda3\lib\site-packages (from tensorflow) (2.15.0)
Requirement already satisfied: numpy<2.0.0,>=1.23.5 in c:\users\manoj s\an
aconda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (1.2
6.2)
Requirement already satisfied: tensorflow-estimator<2.16,>=2.15.0 in c:\us
ers\manoj s\anaconda3\lib\site-packages (from tensorflow-intel==2.15.0->te
nsorflow) (2.15.0)
Requirement already satisfied: wrapt<1.15,>=1.11.0 in c:\users\manoj s\ana
conda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (1.1
4.1)
Requirement already satisfied: h5py>=2.9.0 in c:\users\manoj s\anaconda3\l
ib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (3.7.0)
Requirement already satisfied: keras<2.16,>=2.15.0 in c:\users\manoj s\ana
conda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (2.1
5.0)
Requirement already satisfied: six>=1.12.0 in c:\users\manoj s\anaconda3\l
ib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (1.16.0)
Requirement already satisfied: libclang>=13.0.0 in c:\users\manoj s\anacon
da3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (16.0.6)
Requirement already satisfied: tensorboard<2.16,>=2.15 in c:\users\manoj s
\anaconda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow)
(2.15.1)
Requirement already satisfied: google-pasta>=0.1.1 in c:\users\manoj s\ana
conda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (0.2.
0)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in
c:\users\manoj s\anaconda3\lib\site-packages (from tensorflow-intel==2.15.
0->tensorflow) (0.31.0)
Requirement already satisfied: flatbuffers>=23.5.26 in c:\users\manoj s\an
aconda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (23.
5.26)
Requirement already satisfied: typing-extensions>=3.6.6 in c:\users\manoj
s\anaconda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow)
(4.3.0)
Requirement already satisfied: gast!=0.5.0,!=0.5.1,!=0.5.2,>=0.2.1 in c:\u
sers\manoj s\anaconda3\lib\site-packages (from tensorflow-intel==2.15.0->t
ensorflow) (0.5.4)
Requirement already satisfied: ml-dtypes~=0.2.0 in c:\users\manoj s\anacon
da3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (0.2.0)
Requirement already satisfied: astunparse>=1.6.0 in c:\users\manoj s\anaco
nda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (1.6.3)
Requirement already satisfied: absl-py>=1.0.0 in c:\users\manoj s\anaconda
3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (2.0.0)
Requirement already satisfied: packaging in c:\users\manoj s\anaconda3\lib
\site-packages (from tensorflow-intel==2.15.0->tensorflow) (21.3)
Requirement already satisfied: termcolor>=1.1.0 in c:\users\manoj s\anacon
da3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (2.4.0)
Requirement already satisfied: opt-einsum>=2.3.2 in c:\users\manoj s\anaco
nda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (3.3.0)
Requirement already satisfied: setuptools in c:\users\manoj s\anaconda3\li
b\site-packages (from tensorflow-intel==2.15.0->tensorflow) (63.4.1)
Requirement already satisfied: protobuf!=4.21.0,!=4.21.1,!=4.21.2,!=4.21.
3,!=4.21.4,!=4.21.5,<5.0.0dev,>=3.20.3 in c:\users\manoj s\anaconda3\lib\s
ite-packages (from tensorflow-intel==2.15.0->tensorflow) (4.23.4)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in c:\users\manoj s\ana
conda3\lib\site-packages (from tensorflow-intel==2.15.0->tensorflow) (1.6
0.0)
Requirement already satisfied: wheel<1.0,>=0.23.0 in c:\users\manoj s\anac
```

```
onda3\lib\site-packages (from astunparse>=1.6.0->tensorflow-intel==2.15.0-
>tensorflow) (0.37.1)
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in
c:\users\manoj s\anaconda3\lib\site-packages (from tensorboard<2.16,>=2.15
->tensorflow-intel==2.15.0->tensorflow) (0.7.2)
Requirement already satisfied: google-auth-oauthlib<2,>=0.5 in c:\users\ma
noj s\anaconda3\lib\site-packages (from tensorboard<2.16,>=2.15->tensorflo
w-intel==2.15.0->tensorflow) (1.2.0)
Requirement already satisfied: requests<3,>=2.21.0 in c:\users\manoj s\ana
conda3\lib\site-packages (from tensorboard<2.16,>=2.15->tensorflow-intel==
2.15.0->tensorflow) (2.28.1)
Requirement already satisfied: werkzeug>=1.0.1 in c:\users\manoj s\anacond
a3\lib\site-packages (from tensorboard<2.16,>=2.15->tensorflow-intel==2.1
5.0->tensorflow) (2.0.3)
Requirement already satisfied: google-auth<3,>=1.6.3 in c:\users\manoj s\a
naconda3\lib\site-packages (from tensorboard<2.16,>=2.15->tensorflow-intel
==2.15.0->tensorflow) (2.25.2)
Requirement already satisfied: markdown>=2.6.8 in c:\users\manoj s\anacond
a3\lib\site-packages (from tensorboard<2.16,>=2.15->tensorflow-intel==2.1
5.0->tensorflow) (3.3.4)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in c:\users\manoj
s\anaconda3\lib\site-packages (from packaging->tensorflow-intel==2.15.0->t
ensorflow) (3.0.9)
Requirement already satisfied: pyasn1-modules>=0.2.1 in c:\users\manoj s\a
naconda3\lib\site-packages (from google-auth<3,>=1.6.3->tensorboard<2.16,>
=2.15->tensorflow-intel==2.15.0->tensorflow) (0.2.8)
Requirement already satisfied: rsa<5,>=3.1.4 in c:\users\manoj s\anaconda3
\lib\site-packages (from google-auth<3,>=1.6.3->tensorboard<2.16,>=2.15->t
ensorflow-intel==2.15.0->tensorflow) (4.9)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in c:\users\manoj s
\anaconda3\lib\site-packages (from google-auth<3,>=1.6.3->tensorboard<2.1
6,>=2.15->tensorflow-intel==2.15.0->tensorflow) (5.3.2)
Requirement already satisfied: requests-oauthlib>=0.7.0 in c:\users\manoj
s\anaconda3\lib\site-packages (from google-auth-oauthlib<2,>=0.5->tensorbo
ard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow) (1.3.1)
Requirement already satisfied: idna<4,>=2.5 in c:\users\manoj s\anaconda3
\lib\site-packages (from requests<3,>=2.21.0->tensorboard<2.16,>=2.15->ten
sorflow-intel==2.15.0->tensorflow) (3.3)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\manoj s\anac
onda3\lib\site-packages (from requests<3,>=2.21.0->tensorboard<2.16,>=2.15
->tensorflow-intel==2.15.0->tensorflow) (2022.9.14)
Requirement already satisfied: charset-normalizer<3,>=2 in c:\users\manoj
s\anaconda3\lib\site-packages (from requests<3,>=2.21.0->tensorboard<2.16,
>=2.15->tensorflow-intel==2.15.0->tensorflow) (2.0.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\manoj s\a
naconda3\lib\site-packages (from requests<3,>=2.21.0->tensorboard<2.16,>=
2.15->tensorflow-intel==2.15.0->tensorflow) (1.26.11)
Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in c:\users\manoj s\an
aconda3\lib\site-packages (from pyasn1-modules>=0.2.1->google-auth<3,>=1.
6.3->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow) (0.4.
8)
Requirement already satisfied: oauthlib>=3.0.0 in c:\users\manoj s\anacond
a3\lib\site-packages (from requests-oauthlib>=0.7.0->google-auth-oauthlib<
2,>=0.5->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow)
(3.2.2)
```

In [16]:
```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, LSTM
```

```
WARNING:tensorflow:From C:\Users\MANOJ S\anaconda3\lib\site-packages\keras
\src\losses.py:2976: The name tf.losses.sparse_softmax_cross_entropy is de
precated. Please use tf.compat.v1.losses.sparse_softmax_cross_entropy inst
ead.
```

In [17]:
```
model = Sequential()
model.add(LSTM(50, return_sequences=True, input_shape=(100,1)))
model.add(LSTM(50, return_sequences=True))
model.add(LSTM(50))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')
model.summary()
Model: "sequential"
```

```
WARNING:tensorflow:From C:\Users\MANOJ S\anaconda3\lib\site-packages\keras
\src\backend.py:873: The name tf.get_default_graph is deprecated. Please u
se tf.compat.v1.get_default_graph instead.

WARNING:tensorflow:From C:\Users\MANOJ S\anaconda3\lib\site-packages\keras
\src\optimizers\__init__.py:309: The name tf.train.Optimizer is deprecate
d. Please use tf.compat.v1.train.Optimizer instead.

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 100, 50)           10400

 lstm_1 (LSTM)               (None, 100, 50)           20200

 lstm_2 (LSTM)               (None, 50)                20200

 dense (Dense)               (None, 1)                 51

=================================================================
Total params: 50851 (198.64 KB)
Trainable params: 50851 (198.64 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

In [18]:
```
model.fit(X_train, y_train, validation_split=0.1, epochs=60, batch_size=64,
```

```
Epoch 1/60
WARNING:tensorflow:From C:\Users\MANOJ S\anaconda3\lib\site-packages\ke
ras\src\utils\tf_utils.py:492: The name tf.ragged.RaggedTensorValue is
deprecated. Please use tf.compat.v1.ragged.RaggedTensorValue instead.

19/19 [==============================] - 16s 331ms/step - loss: 0.0080
- val_loss: 0.0025
Epoch 2/60
19/19 [==============================] - 4s 186ms/step - loss: 0.0018 -
val_loss: 0.0016
Epoch 3/60
19/19 [==============================] - 4s 193ms/step - loss: 9.3081e-
04 - val_loss: 0.0012
Epoch 4/60
19/19 [==============================] - 4s 196ms/step - loss: 8.4935e-
04 - val_loss: 0.0011
Epoch 5/60
19/19 [==============================] - 4s 188ms/step - loss: 8.3440e-
04 - val_loss: 0.0011
```

In [19]:
```
# Prediction of scaled test set
test_predict = model.predict(X_test)
```

```
16/16 [==============================] - 3s 54ms/step
```

In [20]:
```
# Transform to original test set
test_predict1 = scaler.inverse_transform(test_predict)
```

In [21]:
```
test_predict1
```

Out[21]:
```
array([[143.22514 ],
       [143.38019 ],
       [142.3973  ],
       [139.57753 ],
       [137.95882 ],
       [137.60919 ],
       [139.0274  ],
       [140.71329 ],
       [140.98082 ],
       [140.39905 ],
       [140.04247 ],
       [141.31248 ],
       [142.16986 ],
       [143.79814 ],
       [146.05193 ],
       [143.74869 ],
       [141.09236 ],
       [140.87192 ],
       [141.78345 ],
```

In [22]:
```
# # Performance Metrics
# # Mean Square error

import math
from sklearn.metrics import mean_squared_error
import sklearn.metrics as metrics
math.sqrt(mean_squared_error(ytest, test_predict)) #MSE is Low
```

Out[22]: 0.03793273114459891

Conclusion : Hence, We Can Conclude that the mean squared error is low and is 0.032445

In [ ]:
```

```