

Feature Scaling

Date	19 September 2022
Team ID	PNT2022TMID25121
Project Name	Project - Crude Oil Price Prediction
Maximum Marks	4 Marks

In [19]:

```
#Importing Libraries
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from numpy import ndarray
from sklearn.preprocessing import StandardScaler
```

In [20]:

```
ds=pd.read_excel(r"C:\Users\Dhyalan\Desktop\Crude Oil Prices Daily.xlsx")
```

In [21]:

```
ds.head()
```

Out[21]:

	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

FEATURE SCALING

In [22]:

```
hd.describe()
```

Out[22]:

	Closing Value
count	8216.000000
mean	43.492139
std	29.616804
min	10.250000
25%	19.577500
50%	29.610000
75%	63.402500
max	145.310000

In [24]:

```

x = ds.iloc[:, 1:3].values
print ("\nOriginal data values : \n", x)

""" Handling the missing values """

from sklearn import preprocessing

""" MIN MAX SCALER """

min_max_scaler = preprocessing.MinMaxScaler(feature_range =(0, 1))

# Scaled feature
x_after_min_max_scaler = min_max_scaler.fit_transform(x)

print ("\nAfter min max Scaling : \n", x_after_min_max_scaler)

""" Standardisation """

Standardisation = preprocessing.StandardScaler()

# Scaled feature
x_after_Standardisation = Standardisation.fit_transform(x)

print ("\nAfter Standardisation : \n", x_after_Standardisation)

```

Original data values :

```

[[25.56]
 [26.  ]
 [26.53]
 ...
 [73.05]
 [73.78]
 [73.93]]

```

After min max Scaling :

```

[[0.11335703]
 [0.11661484]
 [0.12053902]
 ...
 [0.46497853]
 [0.47038353]
 [0.47149415]]

```

After Standardisation :

```

[[-0.60550861]
 [-0.59065128]
 [-0.57275494]
 ...
 [ 0.99807057]
 [ 1.02272024]
 [ 1.02778524]]

```

