DATA PREPROCESSING
CREATING A DATASET WTH SLIDING WINDOWS

Team ID	PNT2022TMID06799
Project Name	Crude Oil Price Prediction

Creating A Dataset With Sliding Windows

In [18]:
```
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 [19]:
```
time_step = 10
X_train, y_train = create_dataset(train_data, time_step)
X_test, y_test = create_dataset(test_data, time_step)
```

In [20]:
```
print(X_train.shape), print(y_train.shape)
```

```
(5329, 10)
(5329,)
```

Out[20]: (None, None)

In [21]:
```
print(X_test.shape), print(y_test.shape)
```

```
(2865, 10)
(2865,)
```

Out[21]: (None, None)

In [22]:
```
X_train
```

Out[22]:
```
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]])
```