

## importing libraries

In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

## Importing dataset

In [2]:

```
df=pd.read_csv('Combined_Cycle_Power_Plant.csv')
```

In [19]:

```
X=df.iloc[:, :-1].values
Y=df.iloc[:, -1].values
```

## Splitting the dataset into the Training & Test set

In [20]:

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(X,Y,test_size=0.2,random_state=0)
```

## Random Forest Regression model Training on the Training set

In [21]:

```
from sklearn.ensemble import RandomForestRegressor
regressor=RandomForestRegressor(n_estimators = 10, random_state = 0)
regressor.fit(x_train,y_train)
```

Out[21]:

```
RandomForestRegressor(n_estimators=10, random_state=0)
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.

On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

## Test set results Predictions

In [22]:

```
y_pred=regressor.predict(x_test)
```

In [23]:

```
np.set_printoptions(precision=2)
```

In [25]:

```
print(np.concatenate((y_pred.reshape(len(y_pred),1), y_test.reshape(len(y_test),1)),1))
```

```
[[ 434.05  431.23]
 [ 458.79  460.01]
 [ 463.02  461.14]
 ...
 [ 469.48  473.26]
 [ 439.57  438.  ]
 [ 460.38  463.28]]
```

## Model Performance Evaluation

In [30]:

```
from sklearn.metrics import r2_score,mean_squared_error as mse,mean_absolute_error as mae
```

In [31]:

```
print(r2_score(y_pred,y_test))
print(mae(y_pred,y_test))
np.sqrt(mse(y_pred,y_test))
```

```
0.9599399648940179
2.452366771159876
```

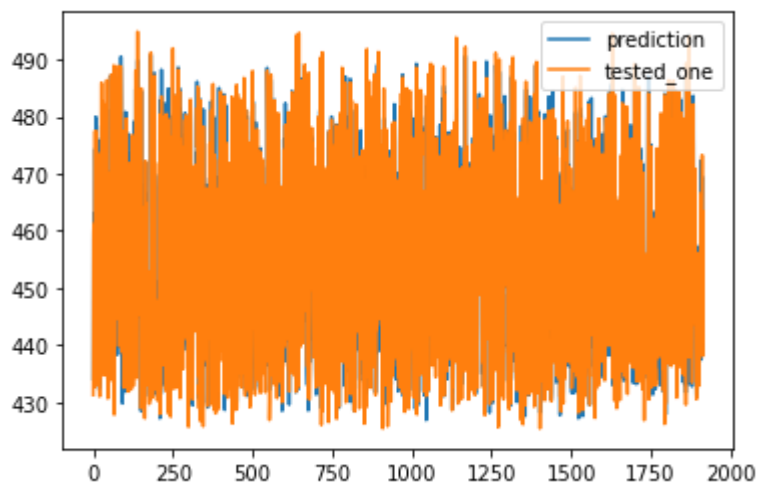
Out[31]:

```
3.3517479009675766
```

## Visualization

In [34]:

```
plt.plot(y_pred)
plt.plot(y_test)
plt.legend(['prediction', 'tested_one'])
plt.show()
```



In [ ]: