Assignment 1:

Enhanced Gradient Descent

(Implementing SGD with mini batch using Boston Housing Prices Dataset)

Link to code and readme file: https://github.com/nimratbedi/Machine-Learning/tree/master/Assignment1

Submitted To :
Prof. Anurag Nagar

Submitted By:
Nimrat Bedi
NXB200004
CS6375.002 2208

Implementing Stochastic Gradient Descent

The loss function for the Linear Model that we need to minimize.

$$L(w,b) = min_{w,b} + \sum_{i=1}^{n} (y_i - W^T x_i - b)^2$$

Now, we calculate the gradients for our loss function L w.r.t Weights(W) and Intercept (b). Following is the equations for calculating the gradients,

$$\frac{\partial L}{\partial W} = \sum_{i=1}^{n} (-2\, x_i)\, (y_i - W^T x_i - b)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{n} (-2)\, (y_i - W^T x_i - b)$$

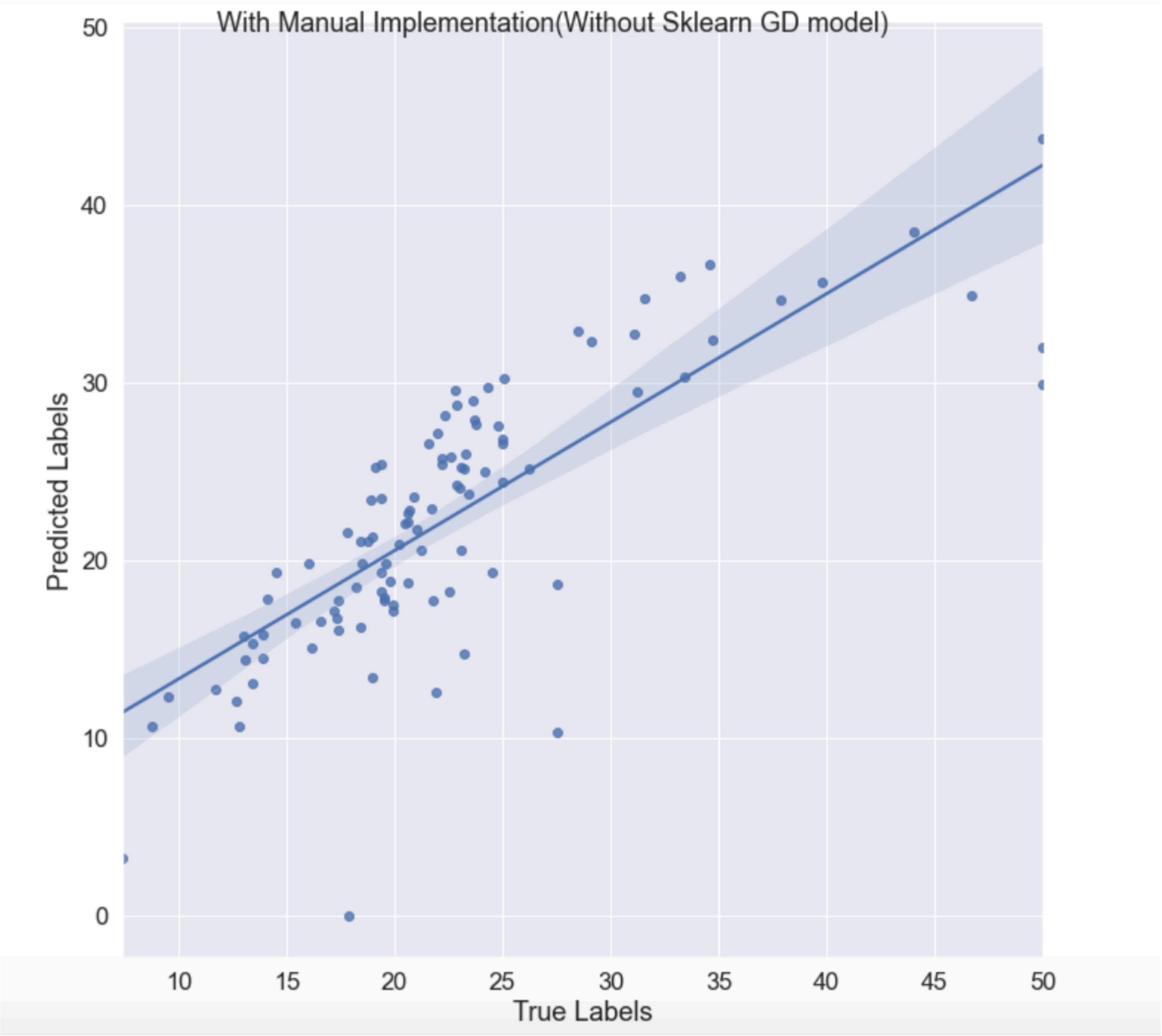
After calculating the gradients we keep changing our weights and intercept value with each iteration.

$$W_{i+1} = W_i - r\, (\frac{\partial L}{\partial W})_{w_i}$$

$$b_{i+1} = b_i - r\, (\frac{\partial L}{\partial b})_{b_i}$$

Here r in the equation is learning rate but in code used as learningRate

Best fit line with manual implementation:

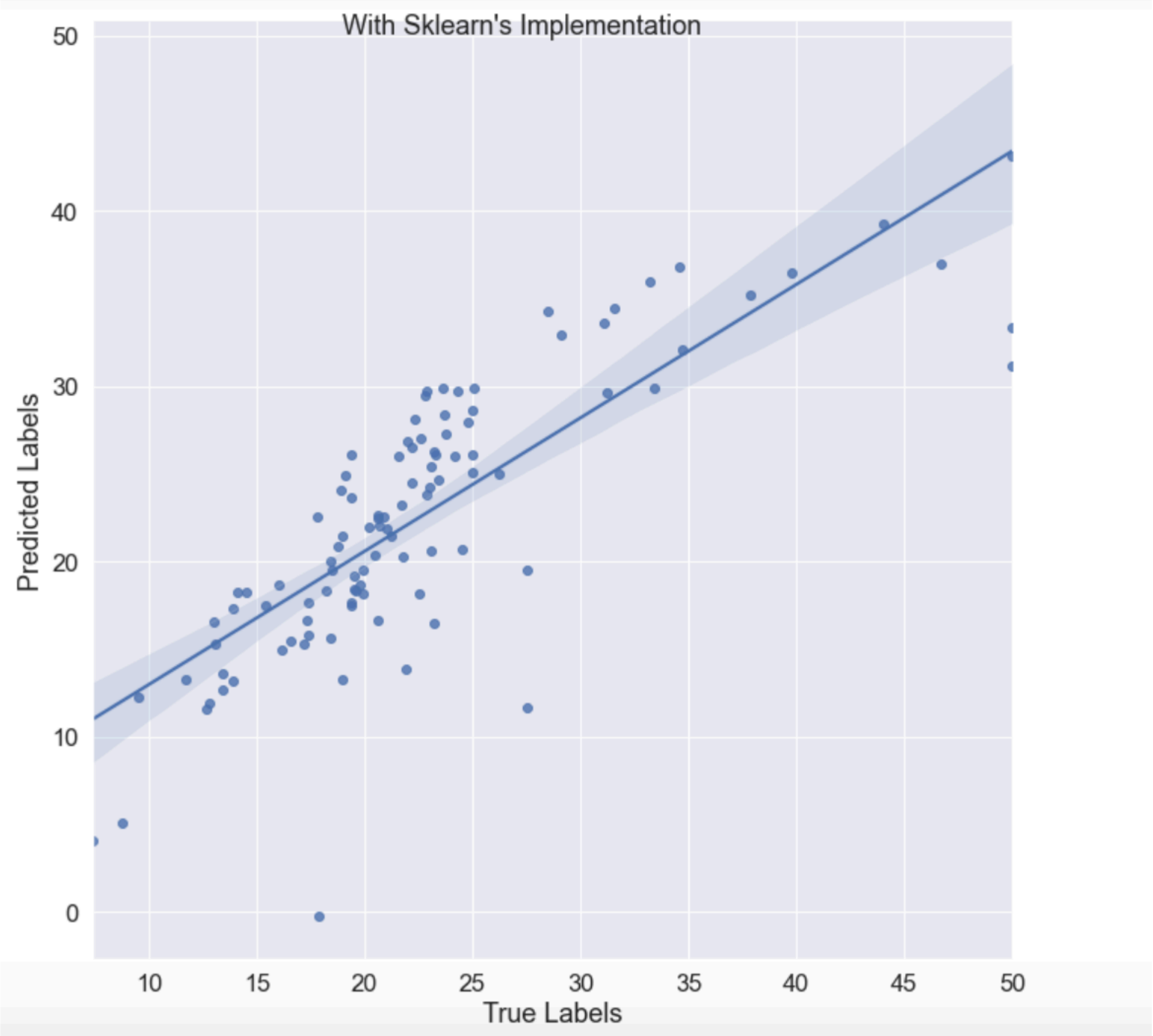


MSE(Mean Square Error) : 25.9

Weights are: [-0.21720708 0.90008082 0.27137001 0.6531707 -1.40960215 2.85138145   0.58618187 -1.73632078 1.17029756 -2.08251421 -2.06551   0.4840585 -3.55489378]

Y intercept(b) = [22.37548042]

Best fit line with sklearn SGDRegressor implementation:



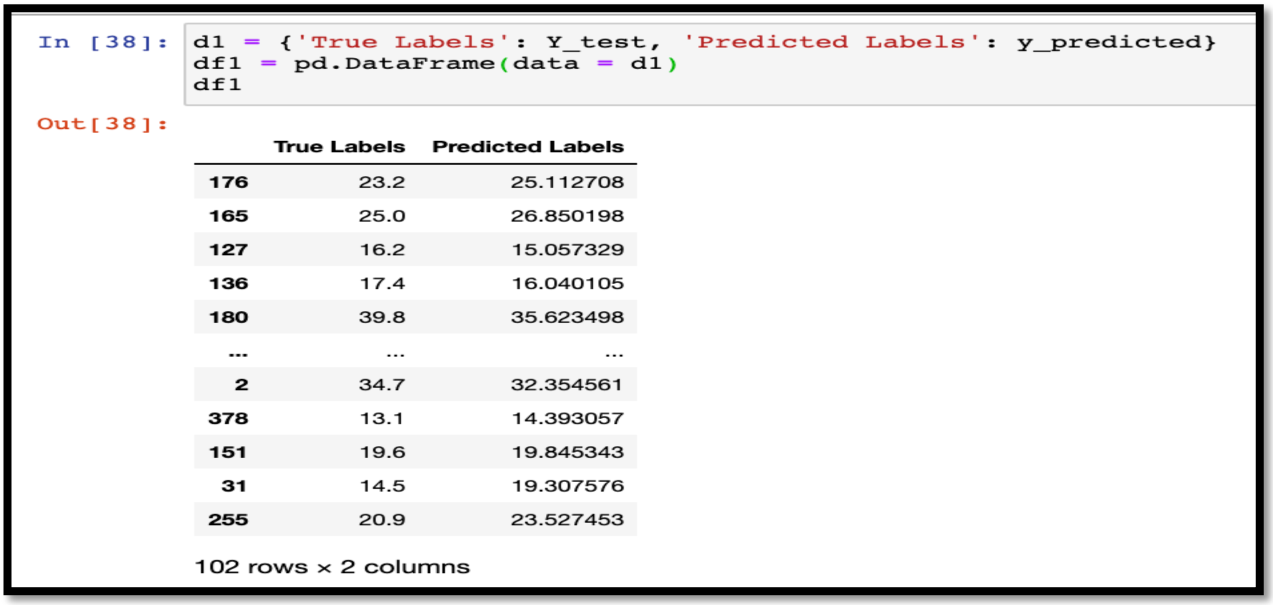
MSE(Mean Square Error) : 24.16

Weights are:[-0.89533898 1.06664128 0.14920204 0.43730232 -2.08909754 2.91771551   0.00530878 -3.01068891 2.30509997 -2.08863393 -2.18640495 0.97598674 -3.42241723]

Y Intercept = [22.51347788]

For Manual Implementation:

1.

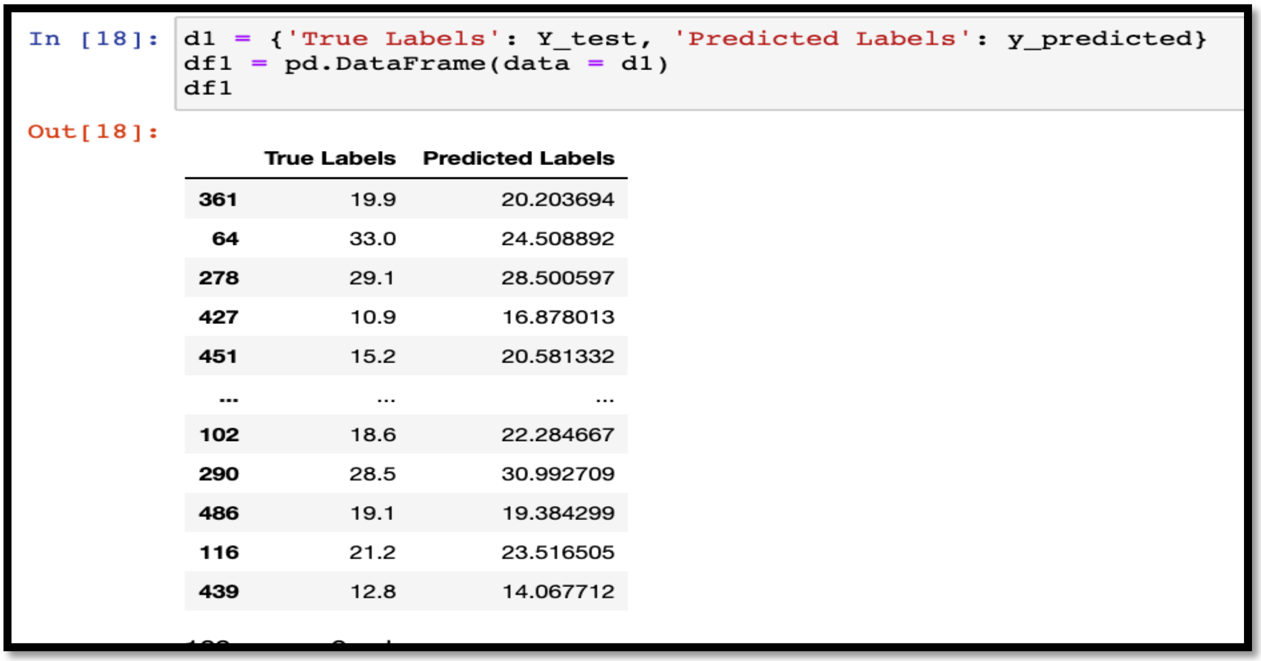


```
In [38]: d1 = {'True Labels': Y_test, 'Predicted Labels': y_predicted}
         df1 = pd.DataFrame(data = d1)
         df1
```

```
Out[38]:
            True Labels   Predicted Labels
      176          23.2          25.112708
      165          25.0          26.850198
      127          16.2          15.057329
      136          17.4          16.040105
      180          39.8          35.623498
      ...           ...                ...
        2          34.7          32.354561
      378          13.1          14.393057
      151          19.6          19.845343
       31          14.5          19.307576
      255          20.9          23.527453

     102 rows × 2 columns
```

2.

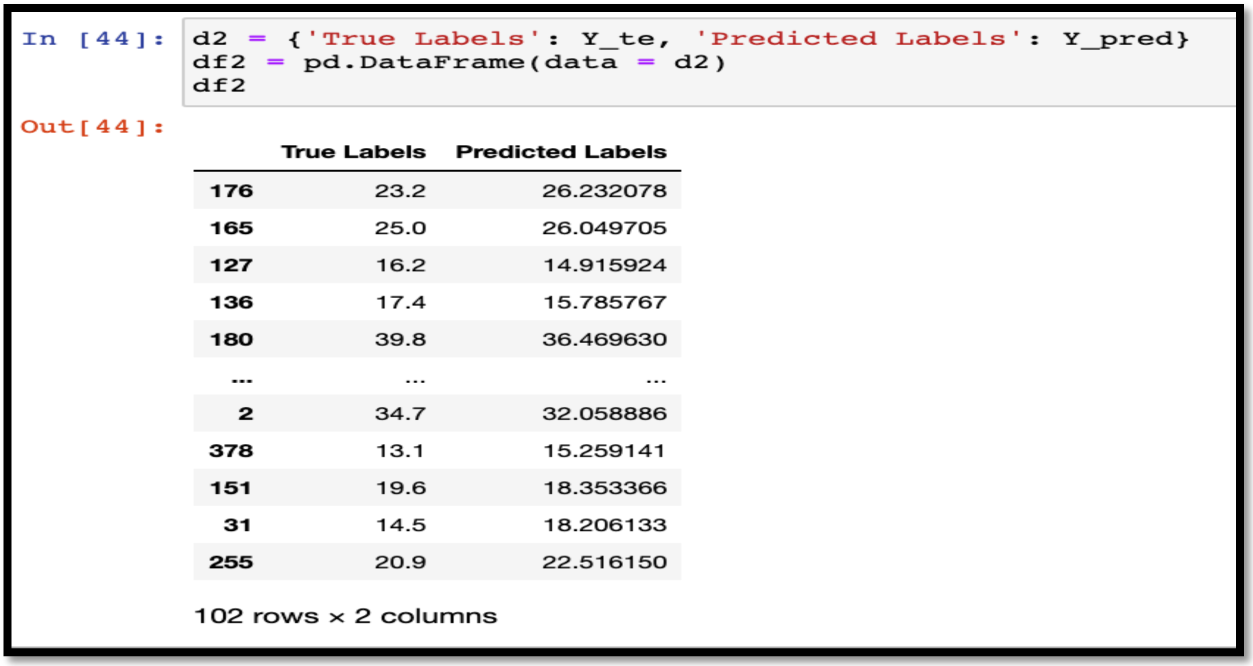


```
In [18]: d1 = {'True Labels': Y_test, 'Predicted Labels': y_predicted}
         df1 = pd.DataFrame(data = d1)
         df1
```

```
Out[18]:
            True Labels   Predicted Labels
      361          19.9          20.203694
       64          33.0          24.508892
      278          29.1          28.500597
      427          10.9          16.878013
      451          15.2          20.581332
      ...           ...                ...
      102          18.6          22.284667
      290          28.5          30.992709
      486          19.1          19.384299
      116          21.2          23.516505
      439          12.8          14.067712
```

For sklearn Implementation:

1.



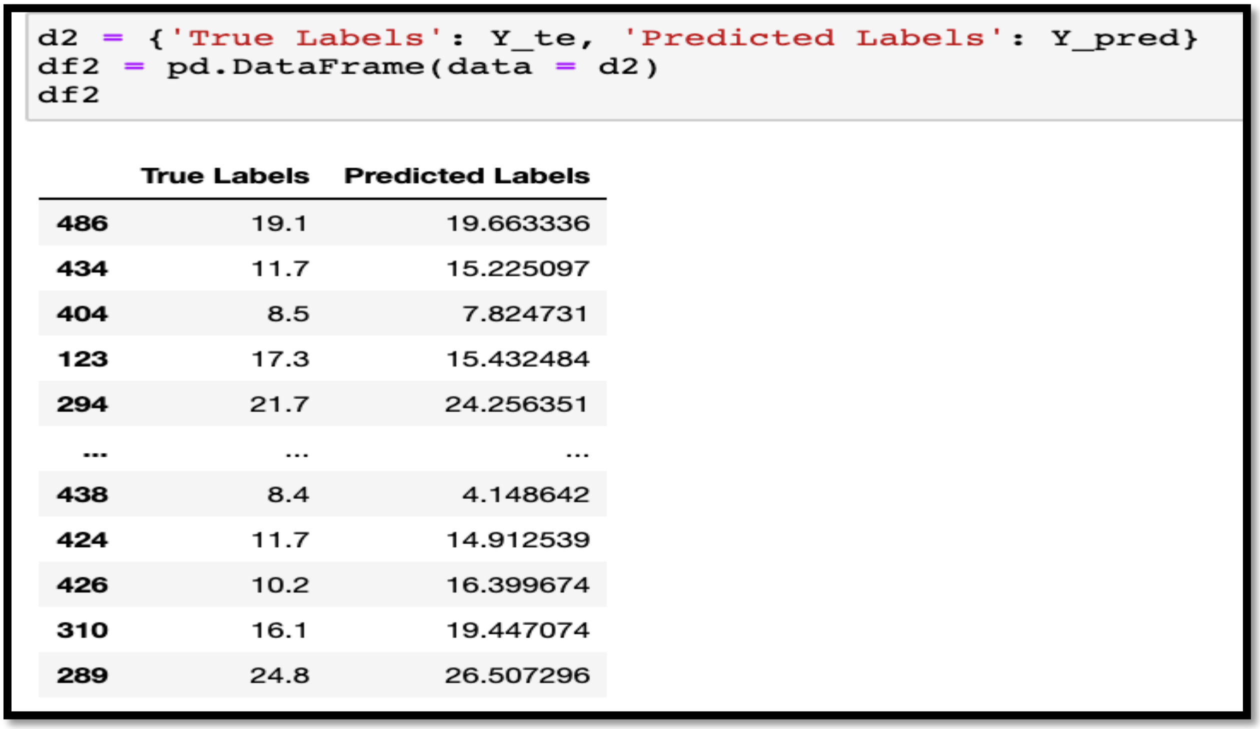
```
In [44]: d2 = {'True Labels': Y_te, 'Predicted Labels': Y_pred}
         df2 = pd.DataFrame(data = d2)
         df2
```

Out[44]:

	True Labels	Predicted Labels
176	23.2	26.232078
165	25.0	26.049705
127	16.2	14.915924
136	17.4	15.785767
180	39.8	36.469630
...	...	...
2	34.7	32.058886
378	13.1	15.259141
151	19.6	18.353366
31	14.5	18.206133
255	20.9	22.516150

102 rows × 2 columns

2.



```
d2 = {'True Labels': Y_te, 'Predicted Labels': Y_pred}
df2 = pd.DataFrame(data = d2)
df2
```

	True Labels	Predicted Labels
486	19.1	19.663336
434	11.7	15.225097
404	8.5	7.824731
123	17.3	15.432484
294	21.7	24.256351
...	...	...
438	8.4	4.148642
424	11.7	14.912539
426	10.2	16.399674
310	16.1	19.447074
289	24.8	26.507296

Final Verdit:

Mean Square Error:
(for 1st trail)
Manual Implementation = 25.9
Sklearn Implementation = 24.16
(for 2nd trail)
Manual Implementation = 18.99
Sklearn Implementation = 24.4