

ML ALGO REPORT

Analysis of Data_Set-1 :

prediction value by both model

own_model_prediction = [2.384060066057183,5.680787126761226] # slope and constant

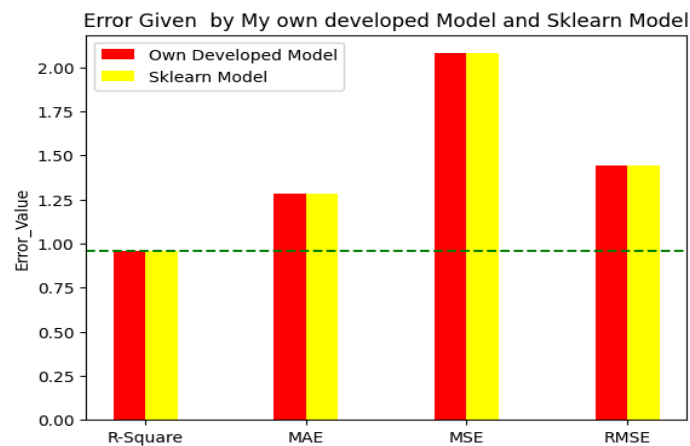
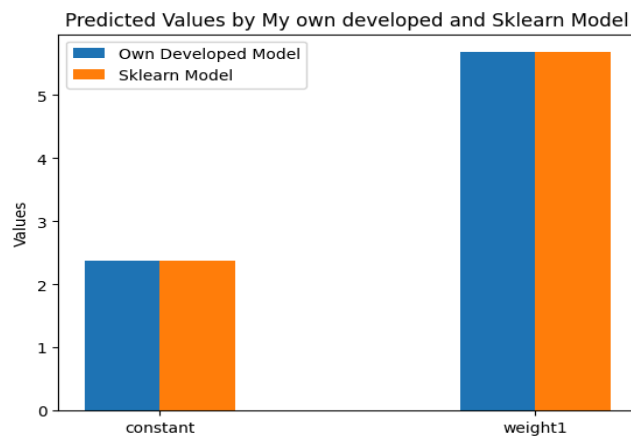
sk_model_prediction = [2.38406007,5.68078713]

error value by both model

error_own_model =[0.95795719055863581,1.2805559784291467,2.0785254017773265,1.4417091945941547] #R-square, MAE,MSE,RMSE

error_sk_model =

[0.9579571905586357,1.280555978429147,2.078525401777328,1.4417091945941551]



Observation :

After developing my own Linear regression model ,I verified it with inbuilt linear regression model(SKLearn) and applied on Dataset-1(Data.csv).I found that result predicted by both model is almost same.Also error(R-squared,MAE,MSE,RMSE) given by both model are closely same.Moreover I got $R^2(0.95)$ which is close to 1 indicates that our own developed model is up to the mark.

Best Fit Hyperplane $Y = 2.38X + 5.68$

Analysis of Data_Set - 2

prediction value by both model

own_model_prediction = [39.7306395177676,0.9729974518460589] # alpha and beta

sk_model_prediction = [39.73063952,0.97299745]

error value by both model

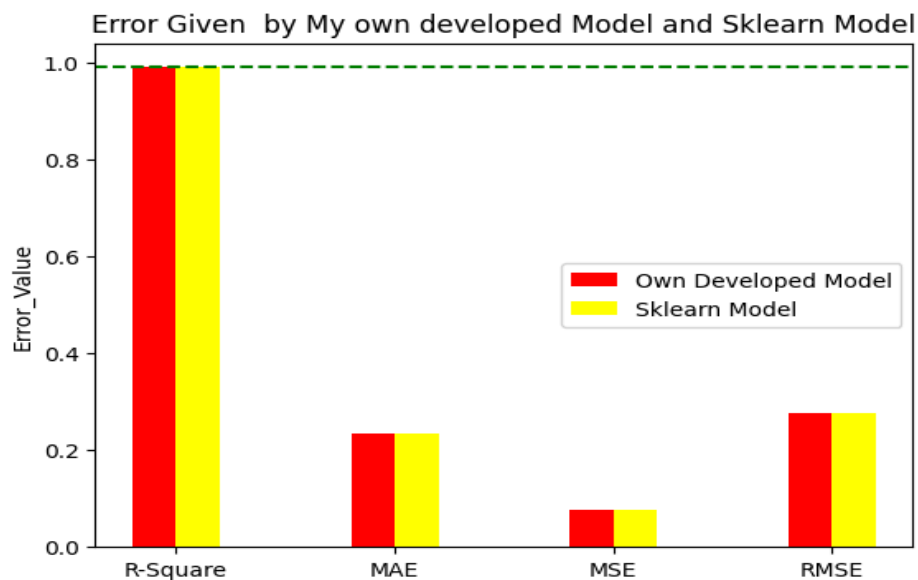
error_own_model

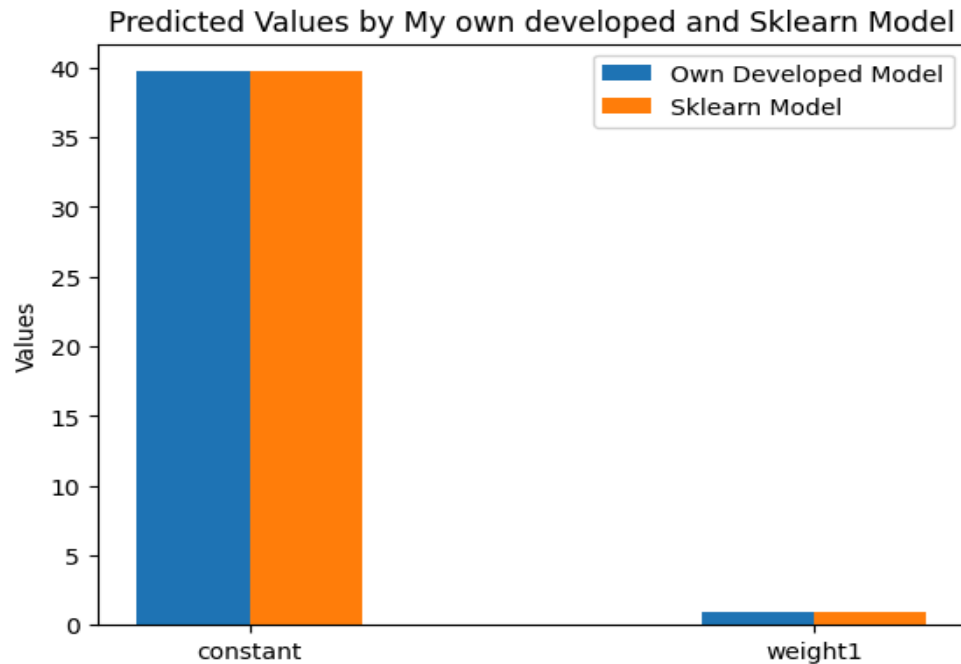
=[0.9904038522690993,0.2349883528902577,0.07643342704351971,0.27646596000867757]

#R-square, MAE,MSE,RMSE

error_sk_model =

[0.9904038522690993,0.23498835289025738,0.07643342704351966,0.27646596000867746]





Observation :

After developing my own Linear regression model ,I verified it with inbuilt linear regression model(SKLearn) and applied on Dataset-2(Data.csv).I found that result predicted by both model is almost same.Also error(R-squared,MAE,MSE,RMSE) given by both model are closely same.Moreover I got $R^2(0.99)$ which is close to 1 indicates that our own developed model is up to the mark.

Best Fit Hyperplane $Y = 0.97X + \log(39.73)$

Analysis of Data_Set-3 :

prediction value by both model

own_model_prediction = [0.09419021414817955,1.1770620783119932]

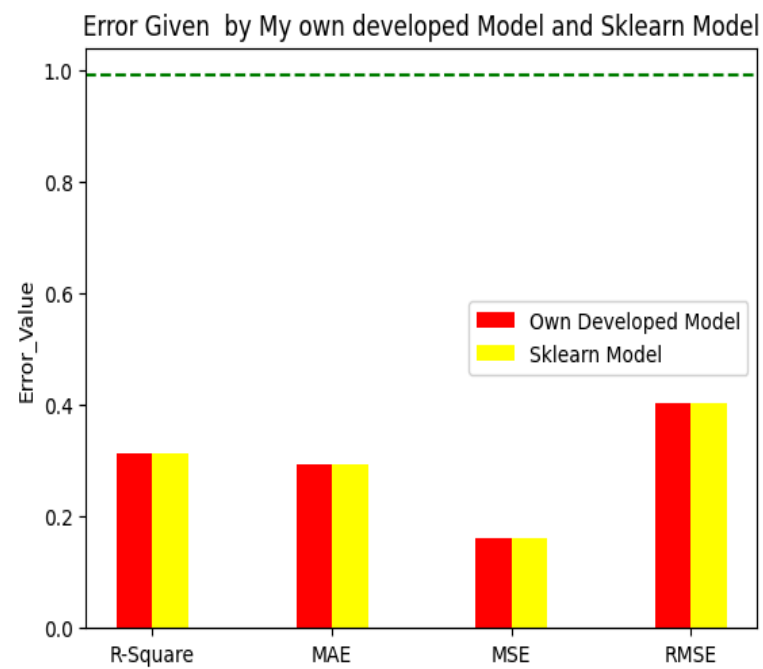
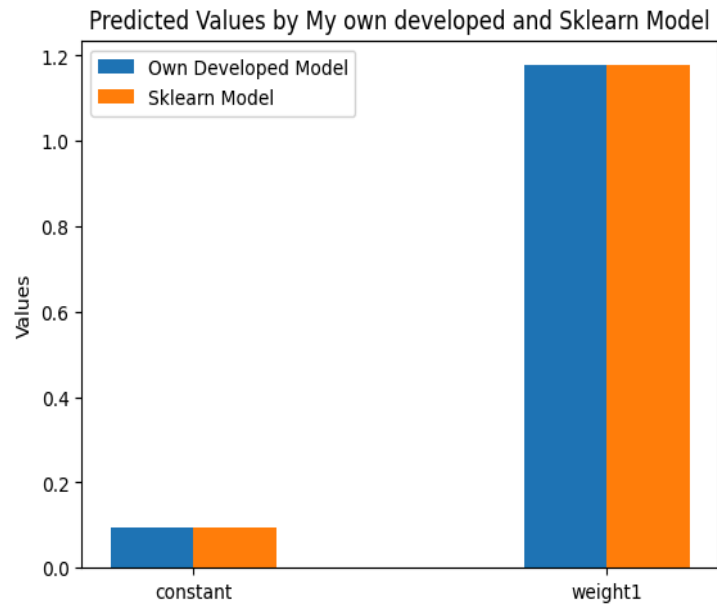
sk_model_prediction = [0.09419021,1.17706208] # slope ,intercept

error value by both model

**error_own_model =[0.3136973226728079,
0.29467793301310385,0.16173044143088552,0.4021572347116057] #R-square,
MAE,MSE,RMSE**

error_sk_model =

[0.3136973226728079,0.29467793301310374,0.16173044143088552,0.4021572347116057]



Observation :

After developing my own Linear regression model ,I verified it with inbuilt linear regression model(SKLearn) and applied on Dataset-3(Data.csv).I found that result predicted by both model is almost same.Also error(R-squared,MAE,MSE,RMSE) given by both model are closely same.Moreover I got $R^2(0.31)$ which is not close to 1 indicates that our own developed model is not best fit for this data.

$$\text{Best Fit Hyperplane } Y = 0.094X + 1.774$$

Data_Set -4 Analysis:

prediction value by both model

own_model_prediction = [13.23947782, 6.13243763,2.39226554,7.74681038] #

weights(0,1,2,3)

sk_model_prediction = [13.239477824445359,6.13243763,2.39226554,7.74681038]

error value by both model

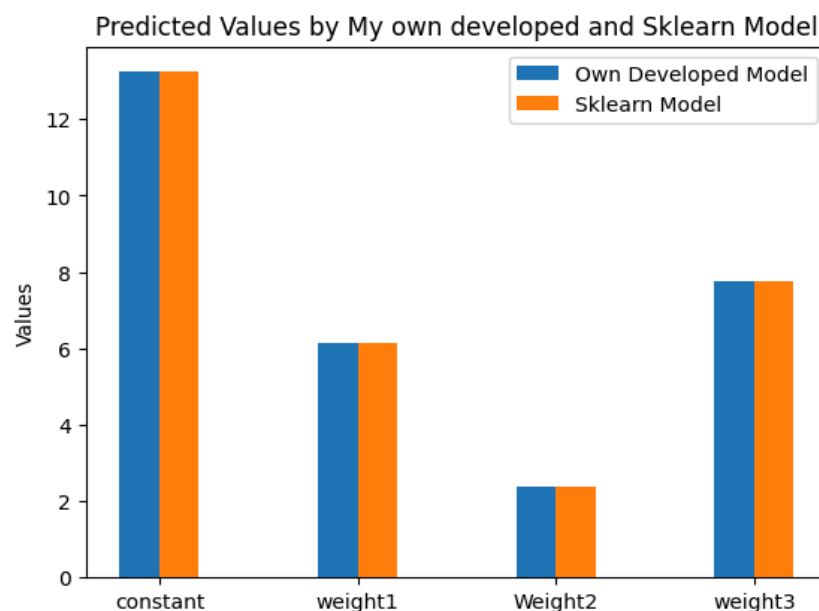
error_own_model

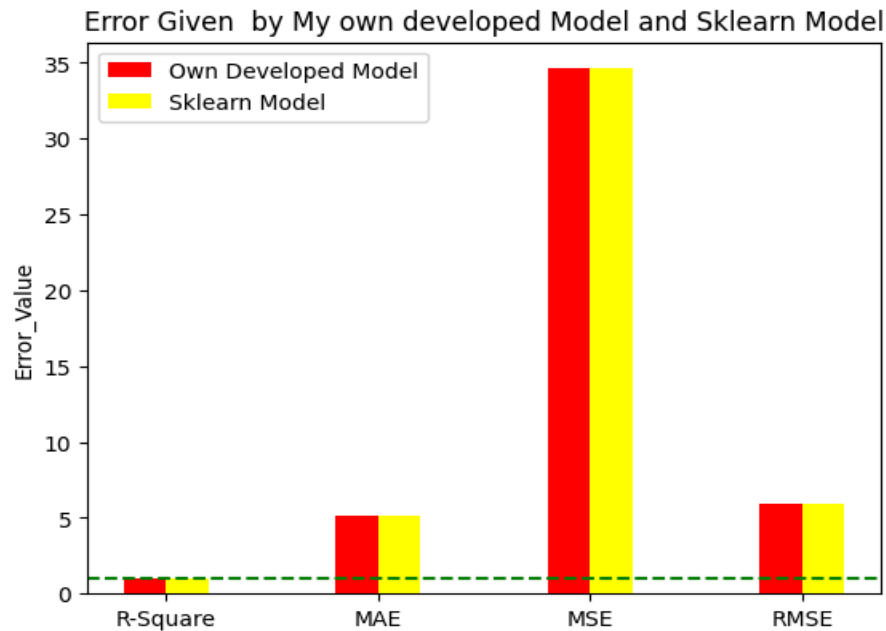
=[0.9841749058943147,5.155505630377769,34.62048082924355,5.8839171331047435]

#R-square, MAE,MSE,RMSE

error_sk_model =

[0.9841749058943147,5.15550562646378,34.62048082924356,5.883917133104745]





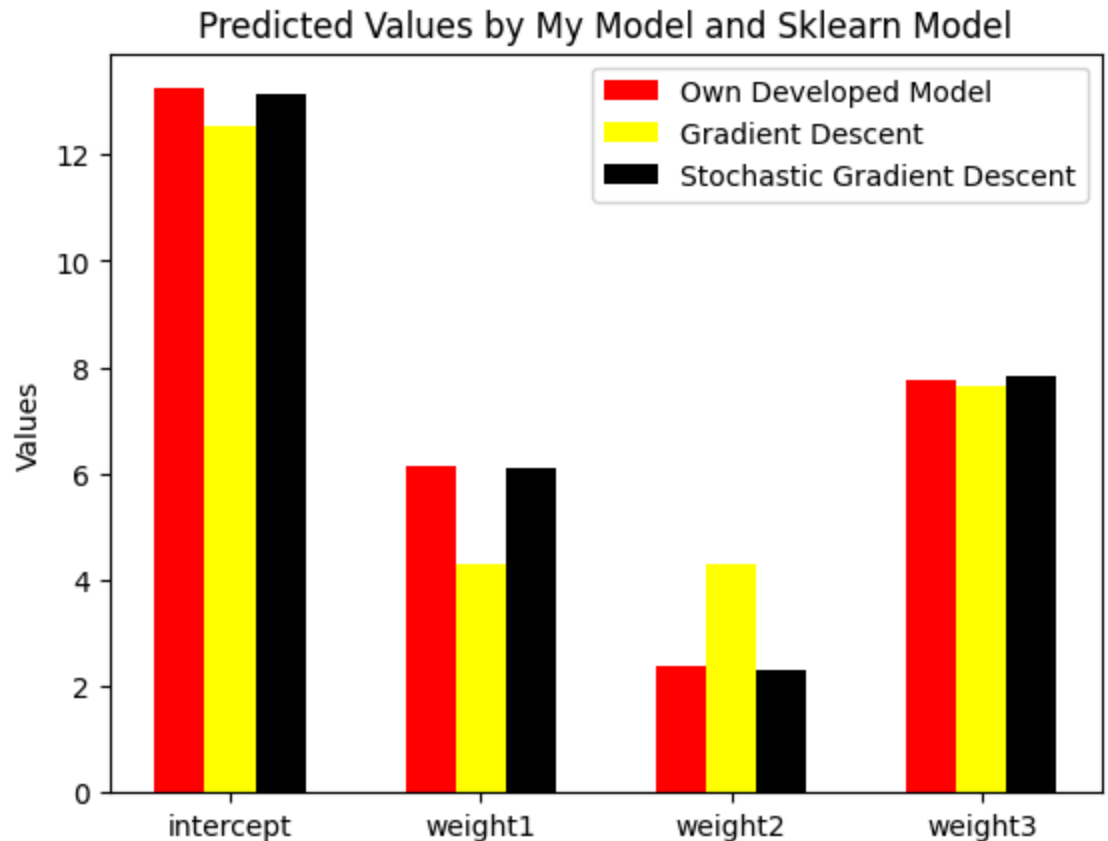
Observation :

After developing my own Multi_Linear regression model ,I verified it with inbuilt linear regression model(SKLearn) and applied on Dataset-4(Data.csv).I found that result predicted by both model is almost same.Also error(R-squared,MAE,MSE,RMSE) given by both model are closely same.Moreover I got $R^2(0.98)$ which is not close to 1 indicates that our own developed model is best fit for this data.

Best Fit Hyperplane $Y = 6.13x_1 + 2.39x_2 + 7.74x_3 + 13.23$

Gradient Descent v/s Stochastic Gradient Descent on Data_Set -4

```
own_model_prediction = [13.23947782, 6.13243763, 2.39226554, 7.74681038] #
weights(0,1,2,3)
GD_prediction =
[12.526059486410091, 4.293667771213319, 4.293667771213319, 7.658473605875606]
SGD_prediction = [13.154571488555078, 6.091250059460483, 2.3134524409841464,
7.827527898786717]
```



Observation :

After applying GD on dataset-4 ,weight(w_0, w_1, w_2, w_3) which is given by GD is little bit different as compared to Multi_Linear model.As You can see in the above graph weight and weight3 are almost same but weight and weight are different.In case of SGD weight was changing as I was running the code again and again.GD was giving more stable result as compared to SGD.