**Description**

The code provided reads data from the fifa19 dataset, Applies some basic preprocessing like standard scaling and removing null entries if any exists. Feature selection is done based on correlation with the Label feature. The data Is split into training and testing sets (80% - 20%) then polynomial linear regression is applied in two different ways:

1. A From scratch method that makes polynomial features from the existing features by multiplying all combinations of features up to a user specified degree and returning the new features, then fitting these features into a from scratch linear regression model that utilizes Gradient Descent algorithm to minimize the loss function and tune the parameters (Theta) which are then used to predict Y.
2. Built in methods provided by SkLearn to make both polynomial features and to fit them into a linear regression model. This is used as a benchmark to compare it to the from scratch method.

**Evaluation**

Finally, the two approaches are compared using the MSE and R2 Score of each. The results are very close.

|  |  |  |
| --- | --- | --- |
|  | Built In Method | From Scratch Method |
| MSE | 1.5792838689816608 | 1.8194559472945897 |
| R2 Score | 0.9537178167806503 | 0.9466793809737813 |

Note: the from scratch method approaches the values of the built in method at higher epochs and learning rates.