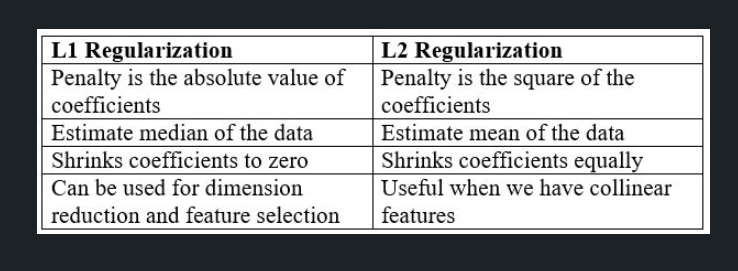
𝘿𝙖𝙮 3:- 𝙇𝙞𝙣𝙚𝙖𝙧 𝙍𝙚𝙜𝙧𝙚𝙨𝙨𝙞𝙤𝙣 𝙢𝙤𝙙𝙚𝙡 𝙚𝙫𝙖𝙡𝙪𝙖𝙩𝙞𝙤𝙣 𝙢𝙚𝙩𝙧𝙞𝙘𝙨  
  
There are several evaluation metrics that can be used to evaluate the performance of a linear regression model.  
Some of the most common metrics include:  
  
1. 𝙈𝙚𝙖𝙣 𝙎𝙦𝙪𝙖𝙧𝙚𝙙 𝙀𝙧𝙧𝙤𝙧 (𝙈𝙎𝙀):This is the average squared difference between the predicted values and the true values. It is calculated as: MSE = (1/n) ∑(ŷ - y)^2, where ŷ is the predicted value and y is the true value.  
  
2. 𝙍𝙤𝙤𝙩 𝙈𝙚𝙖𝙣 𝙎𝙦𝙪𝙖𝙧𝙚𝙙 𝙀𝙧𝙧𝙤𝙧 (𝙍𝙈𝙎𝙀): This is the square root of the MSE. It is often used because it is in the same units as the original data and is easier to interpret.  
  
3. 𝙈𝙚𝙖𝙣 𝘼𝙗𝙨𝙤𝙡𝙪𝙩𝙚 𝙀𝙧𝙧𝙤𝙧 (𝙈𝘼𝙀): This is the average absolute difference between the predicted values and the true values. It is calculated as: MAE = (1/n) ∑|ŷ - y|.  
  
4. 𝙍-𝙨𝙦𝙪𝙖𝙧𝙚𝙙: This metric measures the proportion of the variance in the dependent variable that is explained by the model. It ranges from 0 to 1, with higher values indicating a better fit.  
  
5. 𝘼𝙙𝙟𝙪𝙨𝙩𝙚𝙙 𝙍-𝙨𝙦𝙪𝙖𝙧𝙚𝙙:This is a modified version of R-squared that adjusts for the number of predictors in the model. It is typically used when comparing models with different numbers of predictors.  
  
6. 𝙁-𝙨𝙩𝙖𝙩𝙞𝙨𝙩𝙞𝙘: This is a measure of the overall significance of the model. It compares the fit of the model to the fit of a model with no predictors, and is calculated as: F = (R-squared / (1 - R-squared)) \* ((n - k - 1) / k), where n is the number of observations, and k is the number of predictors.  
  
7. 𝘼𝙠𝙖𝙞𝙠𝙚 𝙄𝙣𝙛𝙤𝙧𝙢𝙖𝙩𝙞𝙤𝙣 𝘾𝙧𝙞𝙩𝙚𝙧𝙞𝙤𝙣 (𝘼𝙄𝘾) 𝙖𝙣𝙙 𝘽𝙖𝙮𝙚𝙨𝙞𝙖𝙣 𝙄𝙣𝙛𝙤𝙧𝙢𝙖𝙩𝙞𝙤𝙣 𝘾𝙧𝙞𝙩𝙚𝙧𝙞𝙤𝙣 (𝘽𝙄𝘾): These are both measures of the relative quality of a statistical model. They take into account the complexity of the model and the number of parameters, and are often used to compare different models.

𝘿𝙖𝙮: 4 : 𝙇𝙞𝙣𝙚𝙖𝙧 𝙍𝙚𝙜𝙧𝙚𝙨𝙨𝙞𝙤𝙣 : - 𝙍𝙚𝙜𝙪𝙡𝙖𝙧𝙞𝙯𝙖𝙩𝙞𝙤𝙣 𝙏𝙚𝙘𝙝𝙣𝙞𝙦𝙪𝙚𝙨  
  
Regularization is a technique used to prevent overfitting. Overfitting occurs when a model is trained too well on the training data and performs poorly on new, unseen data. Regularization helps to constrain, or regularize, the model so that it is not too complex and is better able to generalize to new data.  
  
There are two main types of regularization techniques for linear regression: 𝙍𝙞𝙙𝙜𝙚 𝙍𝙚𝙜𝙧𝙚𝙨𝙨𝙞𝙤𝙣 𝙖𝙣𝙙 𝙇𝙖𝙨𝙨𝙤 𝙍𝙚𝙜𝙧𝙚𝙨𝙨𝙞𝙤𝙣.  
  
𝙍𝙞𝙙𝙜𝙚 𝙍𝙚𝙜𝙧𝙚𝙨𝙨𝙞𝙤𝙣 is a regularization technique that adds a penalty term to the objective function. The penalty term is the sum of the squares of the coefficients of the features in the model. This has the effect of decreasing the magnitude of the coefficients, which can help to reduce overfitting. 𝙄𝙩 𝙞𝙨 𝙖𝙡𝙨𝙤 𝙠𝙣𝙤𝙬𝙣 𝙖𝙨 𝙇2 𝙧𝙚𝙜𝙪𝙡𝙖𝙧𝙞𝙯𝙖𝙩𝙞𝙤𝙣  
  
𝙇𝙖𝙨𝙨𝙤 𝙍𝙚𝙜𝙧𝙚𝙨𝙨𝙞𝙤𝙣 is another regularization technique that adds a penalty term to the objective function. The penalty term is the sum of the absolute values of the coefficients of the features in the model. This has the effect of decreasing the magnitude of the coefficients, but it can also set some of the coefficients to zero, which can be useful for feature selection.𝙄𝙩 𝙞𝙨 𝙖𝙡𝙨𝙤 𝙠𝙣𝙤𝙬𝙣 𝙖𝙨 𝙇1 𝙧𝙚𝙜𝙪𝙡𝙖𝙧𝙞𝙯𝙖𝙩𝙞𝙤𝙣  
  
Regularization is used to tune the complexity of the model and can be controlled by a hyperparameter, which is a parameter that is set before training the model and is not adjusted during training.



𝐌𝐮𝐥𝐭𝐢𝐩𝐥𝐞 𝐋𝐢𝐧𝐞𝐚𝐫 𝐑𝐞𝐠𝐫𝐞𝐬𝐬𝐢𝐨𝐧  
  
• It is a statistical method used to model the relationship between multiple independent variables and a single dependent variable by fitting a linear equation to the observed data.  
  
• The goal of multiple linear regression is to find the values of coefficients that minimise the residual sum of squares, using a method called OLS (Ordinary least squares)  
  
• Only difference between Linear regression and this is the number of independent variables taken into consideration, the rest of everything is the same.   
  
𝐏𝐨𝐥𝐲𝐧𝐨𝐦𝐢𝐚𝐥 𝐑𝐞𝐠𝐫𝐞𝐬𝐬𝐢𝐨𝐧  
  
• Polynomial regression is a form of non-linear regression which models the relationship between the independent variable x and the dependent variable y as an n-th degree polynomial. An n-th degree polynomial equation is of the form:  
y = b0 + b1x + b2x^2 + ... + bn\*x^n  
  
• This method can model various non-linear relationships, such as quadratic, cubic, or higher-order relationships.  
  
• The difference between this Polynomial regression and multiple linear regs lies in the type of relationship they model, multiple linear regression is restricted to Linear mod whereas polynomial regression can model non-linear relationships.   
  
• Additionally, multiple linear regression can model a relationship between more than one independent variable, while polynomial regression uses only one independent variable and models its relationship with the dependent variable in a non-linear manner.  
  
🎯  The higher the degree of the polynomial equation that is used, the more likely it is to overfit the data. Therefore, it is important to use cross-validation techniques to choose the degree of the polynomial equation that generalizes well to unseen data.  
  
Below is the implementation of Multiple Linear Regression from scratch using the Gradient descent method and also Polynomial regression for non-linear data.