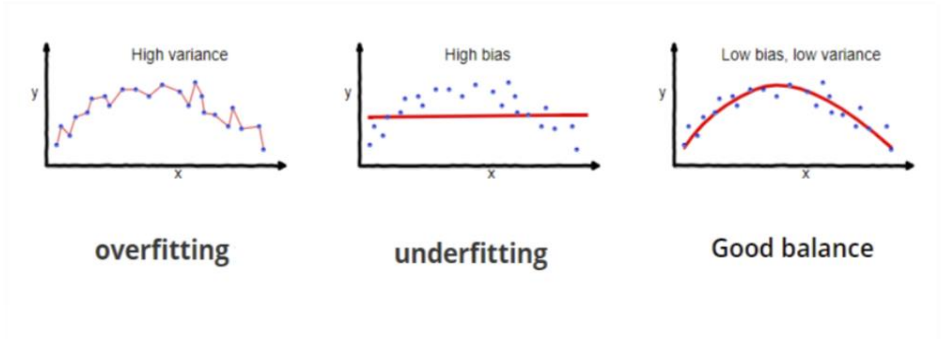
Ridge & Lasso Regression

Bias & Variance

Bias : Biases are the underlying assumptions that are made by data to simplify the target function. Bias does help us generalize the data better and make the model less sensitive to single data point

Variance : Variance is a type of error that occurs due to a model's sensitivity to small fluctuations in the dataset.



Problem with Linear Regression

- Linear Regression can cause overfitting.
- Suppose you have 2 data pts in training example and we try to fit a line passing through both the pts, then the error is 0
- If the testing data has pts far from the training examples then our linear regression model cannot generalize the data.

  Even though the line was the best fit line it may lead to be an overfitting condition and thus not generalizable.

  To Solve this problem we use Ridge and Lasso Regression.

Ridge & Lasso Regression

- In Ridge regression, we add a penalty term which is equal to the square of the coefficient.
- In Lasso regression, we add a penalty term which is equal to the absolute value of the coefficient.
- We also add a coefficient **α** to control that penalty term.

$$New\ Loss\ 1(y, y_{pred}) = MSE(y, y_{pred}) + \alpha \sum_{i=1}^{m} \theta_i^2$$

$$New\ Loss\ 2(y, y_{pred}) = MSE(y, y_{pred}) + \alpha \sum_{i=1} |\theta_i|$$

Ridge & Lasso Regression

- Ridge Regression regularizes the loss function by adding a large penalty for the error.
- Ridge regression are sensitive to noise as the errors are squared.
- Lasso Regression not only regularizes the loss function but also helps in feature selection as some of the coefficients equals zero.
- Lasso regression are less sensitive to noise as only the magnitude of the errors are used.