Regression

- In Simple Linear Regression, we assumed that the **data is linearly related to the target** and hence can be modelled with linear models such as linear regression.
- If we find data is more complex & non linear, we use linear model by adding a simple step of **polynomial regression**.
- The high degree polynomial comes with its own cost. The curve may fit the training data very precisely but it may start missing the estimation on test data and this scenario is called as **overfitting.**

- **Overfitting :**
- When the model is able to fit the train data perfectly and the train error is very small.
- but test error is huge which is indicating that the model is unable to predict on unseen data then we say the model has overfitted the data.
- And this happens due to something called as **Low bias (Training Error)** and **high variance(Testing Error)**

Low bias (Training Error) & **high variance(Testing Error)**

- **Underfitting :**

- When the model is unable to fit to the train data and the train error is hugh (**High bias**) & test error is huge **(High Variance)** then we say that the model is underfitting.

Best Fit : High bias (Training Error) & **High variance/Low variance(Testing Error)**

Y = 2.3X1+4X2+40X3

High val of slope increases complexity of model & computation power also.

the model is more bias means mor interest to.

error = actual value and predicted value then it is called loss.
loss single data point
cost for entire dataset.

Regularization : Its a solution to overfitting.

The solution to underfitting is to train more. Or improve the data.

Regularization basically modifies the loss function and ensures that overfitting is reduced.

In Regression problem regularization is applied using

1) **Ridge Regression**
2) **Lasso Regression**

lambda or alpha is hyperparameter.
internally lambda called as alpha
lambda value is 0 to any value

Ridge Regression:

LF/CF = MSE + lambda * (M1^2 + M2^2 + M3^2) squre of m

- It penalizes high value of slope.
- Uses when input features are highly correlated with each other.
- Slope val move towards zero.
- The val of lambda vary from (0 to ....)

m1m2m3 are slopes., here modify the value of slope.
and giving best fit line.

Lasso Regression:

LF/CF = MSE + lambda * |M1 + M2 + M3| absolute of m

- It penalizes high value of slope.
- Slope val reaches to zero for few features.
- Helps in feature selection.
- The val of lambda vary from (0 to ....)

What is L1 and L2 regularization?
L1 Regularization, also called a lasso regression.
L2 Regularization, also called a ridge regression.

Elastic Net Regression:

LF = MSE + lambda1 * |M1 + M2 + M3| + lambda2 *(M1^2 + M2^2 + M3^2)

- It combines benefits of ridge & lasso.
- Multicolinarity(High correlation val betn independent feature) from Ridge.
- Feature Selection from Lasso.
- The val of lambda vary from (0 to ....)

What is lasso regression in simple terms?
Lasso regression algorithm is defined as a regularization algorithm that assists in the elimination of irrelevant parameters, thus helping in the concentration of selection and regularizes the models. Lasso models can be evaluated using various metrics such as RMSE and R-Square.

What is Lasso Regression in machine learning?
Lasso regression makes coefficients to absolute zero; while ridge regression is a model turning method that is used for analyzing data suffering from multicollinearity

What is the difference between elastic net and lasso regression?
Lasso will eliminate many features, and reduce overfitting in your linear model. Ridge will reduce the impact of features that are not important in predicting your y values. Elastic Net combines feature elimination from Lasso and feature coefficient reduction from the Ridge model to improve your model's predictions.