**→ Linear regression**

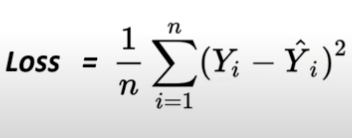
- Advantages:

1. Very simple to implements
2. Performs well on data with linear relationship

- Disadvantages:

1. Not suitable for data having non-linear relationship
2. Underfitting issue
3. Sensitive to outliers

- Loss/Cost Functions:

1. It measures how far an estimated value is from its true value
2. It is helpful to determine which model performs better & which parameter are better
3. 

Yi cap = model predicted value

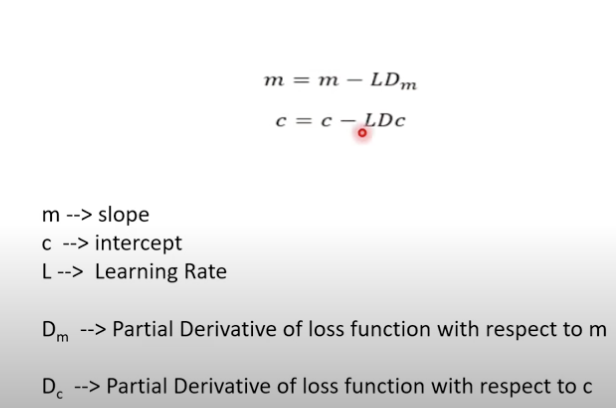
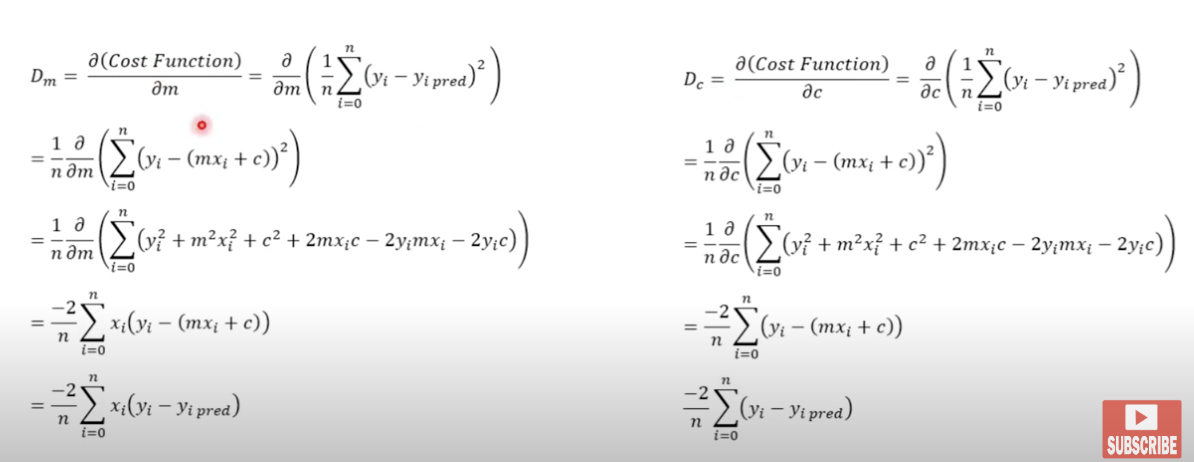
Y = real value

* Low cost/loss value → high accuracy

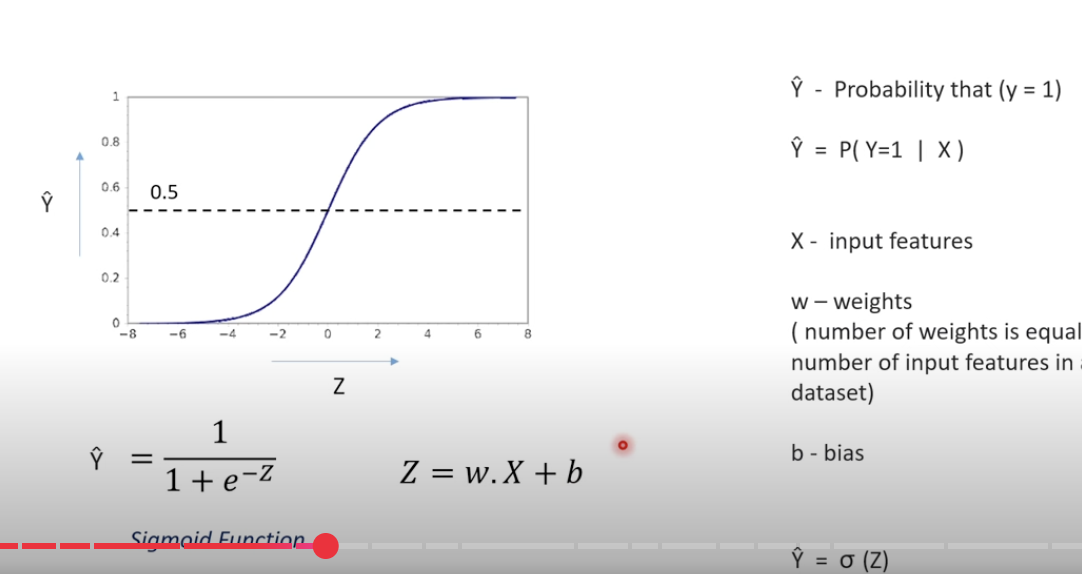
- Model Optimization:

1. Optimization refers to determining best parameters for a model, such that the loss function of model increases, as a result of which the model can predict more accurately

- Gradient Descent:

1. Is an optimization algorithm used for minimizing the loss function in various machine learning algorithms. It is used for updating the parameters of the learning model.
2. 
3. 

**→ Logistic Regression**

****

- Advantages:

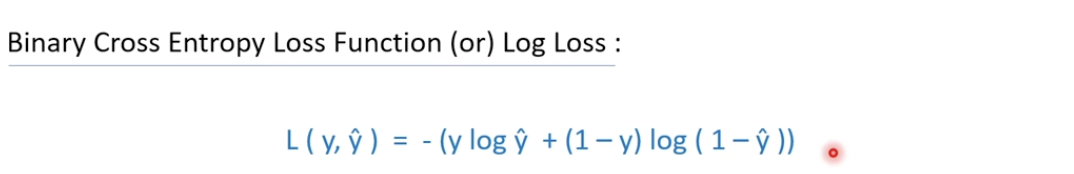
1. Easy to implement
2. Performs well on data with linear relationship
3. Less prone to overfitting for low dimensional dataset

-Disadvantages:

1. High dimensional dataset causes over-fitting
2. Difficult to capture complex relationships in a dataset.
3. Sensitive to outliers
4. Needs a larger dataset

-Loss Function:

1. Loss function measure show far an estimated value is from its true value



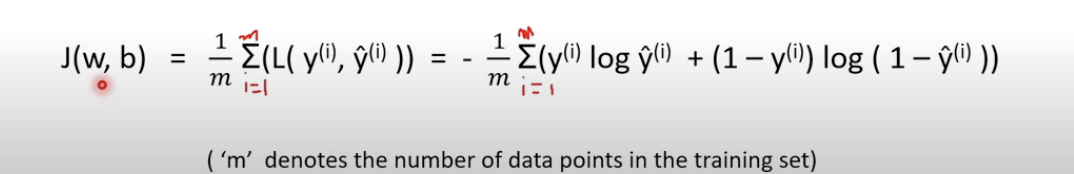
When y = 1 L(1, Y^cap) = -log Y^cap

We always want a smaller loss function value, hence y^cap should be very large, so that - log(1-y^cap) will be a large negative number.

When Y = 0 L(0, y^cap) = -log(1-y^cap)

We always want a smaller loss function value, hence y^cap should be very small, so that - log(1-y^cap) will be a large negative number.

→ Cost Function

1. Cost function deals with a penalty for a number of training sets or the complete batch.
2. Loss function mainly applies for a single training set 

**→ Lasso Regression**

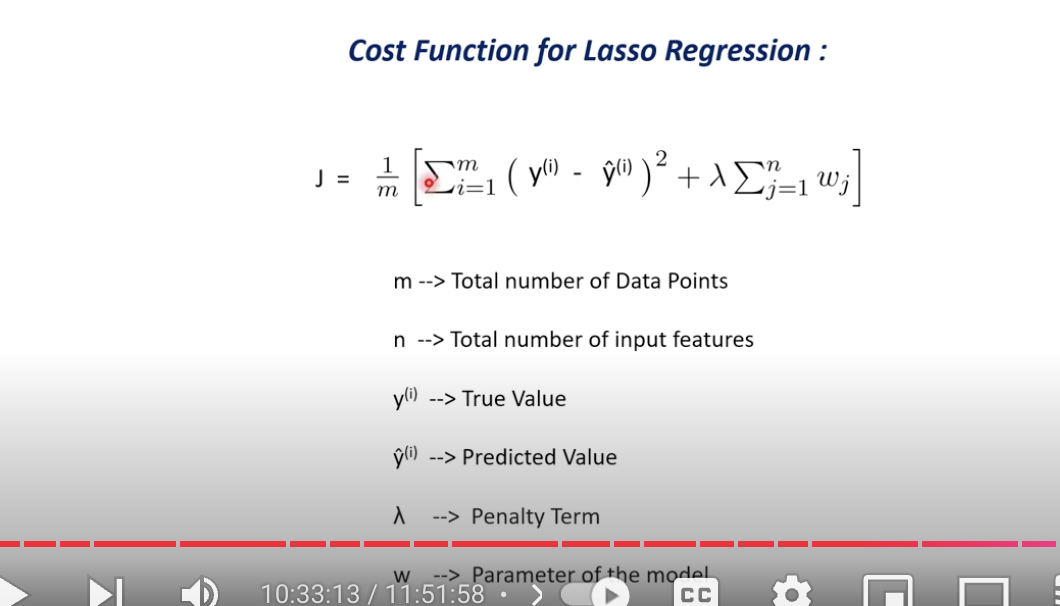
**→ About Lasso Regression**

1. Supervised learning model
2. Regression model
3. Least Absolute Shrinkage and Selection Operator
4. Built on top of linear regression model and mainly used to avoid overfitting using L1 regularization.

→ Regularization

1. It is used to reduce the overfitting of the model by adding a penalty term(lambda symbol) to the model. Lasso regression uses L1 regularization technique.
2. The ‘penalty’ term reduces the value of the coefficient or eliminates a few coefficient, so that the model has fewer coefficients. As a result, overfitting can be avoided.

→ Cost Function

1. 

→ Support Vector Model(SVM):

→About SVM

1. Supervised Learning model
2. Both classification and regression
3. Hyperplane(Is a line(in 2d space) or a plane that separates the data points into 2 classes.)
4. Support Vectors( are the data points which lie nearest to the hyperplane. If these data points change, the position of the hyperplane changes)

→ Advantages

1. Works well with smaller datasets
2. Works efficiently when there is a clear margin of separation
3. Works well with high dimensional data

→ Disadvantages

1. Not suitable for large datasets as the training time is higher
2. Not suitable for noisier datasets with overlapping classes.

