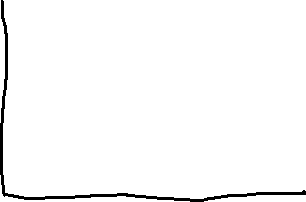
**Linear Regression**



Y=mx+c



Cost Function =



N=no. of points

Y= nearest point of the line



= point on the line

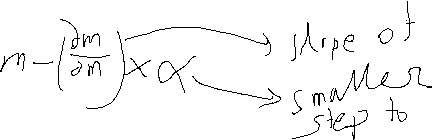
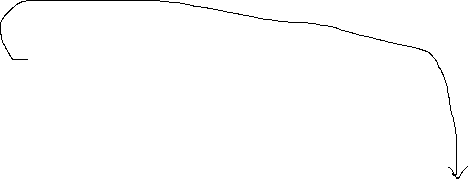
We have to minimize the error



When slope, m=1 and c=0

We get our C.F=0

\*When we get different m value we get different cost function value and by plotting those in graph we get a curvature-



Global Minima



Gradient Descent



\*At global minima point the slope will be 0, and at this point we can say that this will be the point of best fit line

Ridge and Lasso Regression

Reduce overfitting. That is the value of a point is higher from the best fit line.

In Ridge Regration =

**= 0 to any +ve value**

**We try to reduce**



Here is a huge path distance we see. This is steep slope

\*In the case of linear regression we stop if we get cf=0 which is = 0 . But in this case if it is 0 the cf will not be zero. So we will try to reduce that value



Here the slope will be reduce we will get a small value of this also, finally the cf will reduce

In this case this will be best fit line incase of another line

In this case the slope is moving towards 0.

**Lasso Regression:**

**Y= m1x+m2x+…+mnx**

**The magnitude of all the slope. It also works as feature selection. Here the slope is tending towards 0 and we don’t need the outliers of m we need just the value of m that’s minimize the value and the slope value is close to zero.**

# Bias And Variance



Degree of Polynomial is 1. It will act like linear reg.

But here the error is quite high, UNDERFITTING

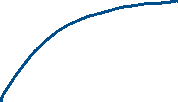
For the train data the acc Is down but for test data the acc. Is down

It gives us high BIAS, high VARIANCE



Degree of Polynomial is 2. For the train data the acc. Is high but for test data the acc. Is also high. This model is suitable.

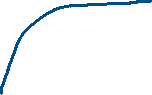
It gives us low BIAS, low VARIANCE



Degree of Polynomial is 4. But here the error is quite high, OVERFITTING

For the train data the acc. Is high but for test data the acc. Is low

It gives us low BIAS, high VARIANCE



|  |  |  |
| --- | --- | --- |
| Model1 | Model2 | Model3 |
| Train Error:1%  Test Error: 20%  Low Bias  High Variance  OVERFITTING | Train Error:25%  Test Error: 20%  high Bias  High Variance  UNDERFITTING | Train Error: <10%  Test Error: <10%  Low Bias  Low Variance  Generalized |

Based on what data sample u select the test variance doesn’t vary much this is low variance, when highly varies that’s high variance. So the variance is depeneded on the selection of test data set.

Bias is the measurement of how accurately a model can capture a pattern in training dataset.

**R square and Adjusted R square**

**= 1-**

**-> Sum of avg. totals**



**More nearer the value is one the line is best fitted. The more u add independent features the value of R2 gets increased.**

**Adjusted R2 = 1 - penalizes the features that are not correlated to the target feature**

**N =Total Number of samples**

**P= Number of predictors**

**If the value of independent p is not co related then the value of R2 will decrease, vice-versa will increase.**

**Hypothesis Test, Type 1 Error, Type 2 Error**

**Evalutate 2 mututal statement based on Population using Sample Data**

1. **Initial Assumption (H0) [H1 is Alternate assumption of H0]**
2. **Collect data based**
3. **Gather evidence to reject or not reject**

|  |  |  |
| --- | --- | --- |
|  | **H0** | **H1** |
| **Don’t Reject** | **Ok** | **Type-2 error** |
| **Reject** | **Type-1 error** | **OK** |

🡨 **Truth**

**Type-1 error 🡪 Rejecting a true null hypothesis (α)**

**Type-1 error 🡪 Accepting a false null hypothesis (β)**

**Rejecting a false null hypothesis is the goal (1- β) [Power of a test]**

**U like a girl**

**H0 🡪 She likes you back**

**Truth🡪 She likes you back**

**You didn’t invite for a date cause of situation u see. U miss a chance. This is type-1 error.**

**H0 🡪 She likes you back**

**Truth🡪 She doesn’t likes you back**

**You invite for a date .But she rejected because u r broke. This is type-1 error.**

**You didn’t invite for a date cause of situation u see. U miss a chance. This is type-1 error.**

**P-Value**

Probability of Null Hypothesis to be True

p = 0.1 , means 10 out of 100 times the Null hypothesis will be true

**Chi-Square Formula**

**χ2 =**

* Oi = observed value (actual value)
* Ei= expected value.

**E =**

**Confusion Matrix:**

Actual Value

|  |  |  |
| --- | --- | --- |
|  | **1** | **0** |
| **1** | **TP** | **FP** |
| **0** | **FN** | **TN** |

Predicted value

For balanced dataset, Accuracy =

For imbalanced dataset,

Recall = (Out of total actual positive values how many positive we predict correctly)

,Precision = (Out of total positive predicted results how many were actually positive)

**Logistic Regression**

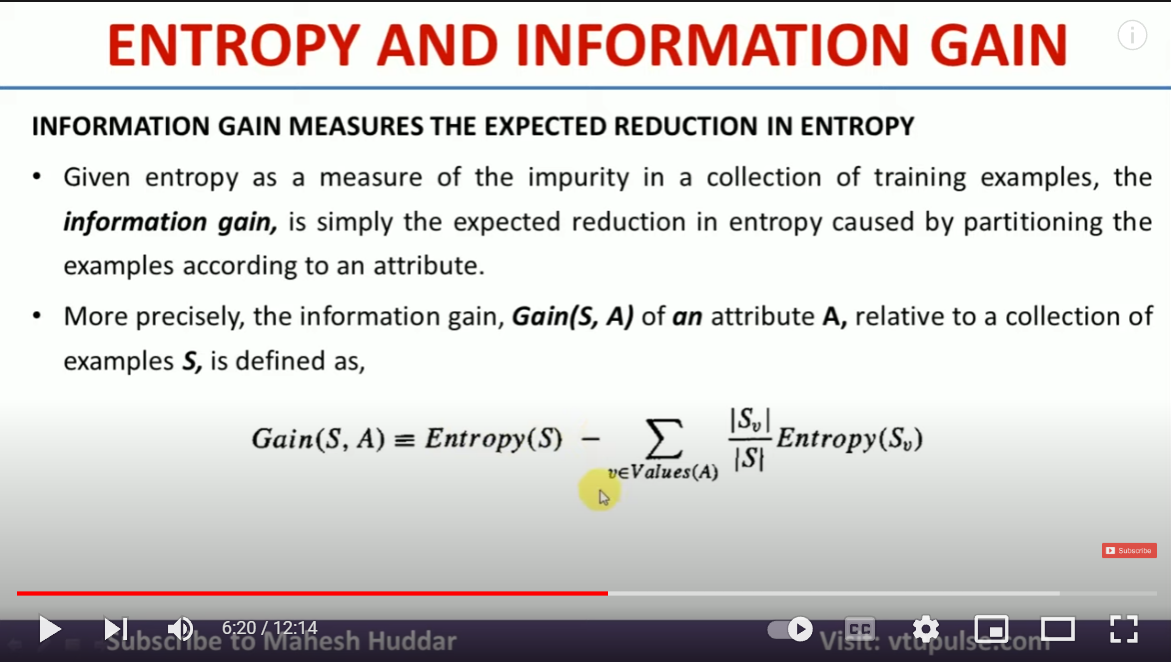
It predicts whether something true or false.

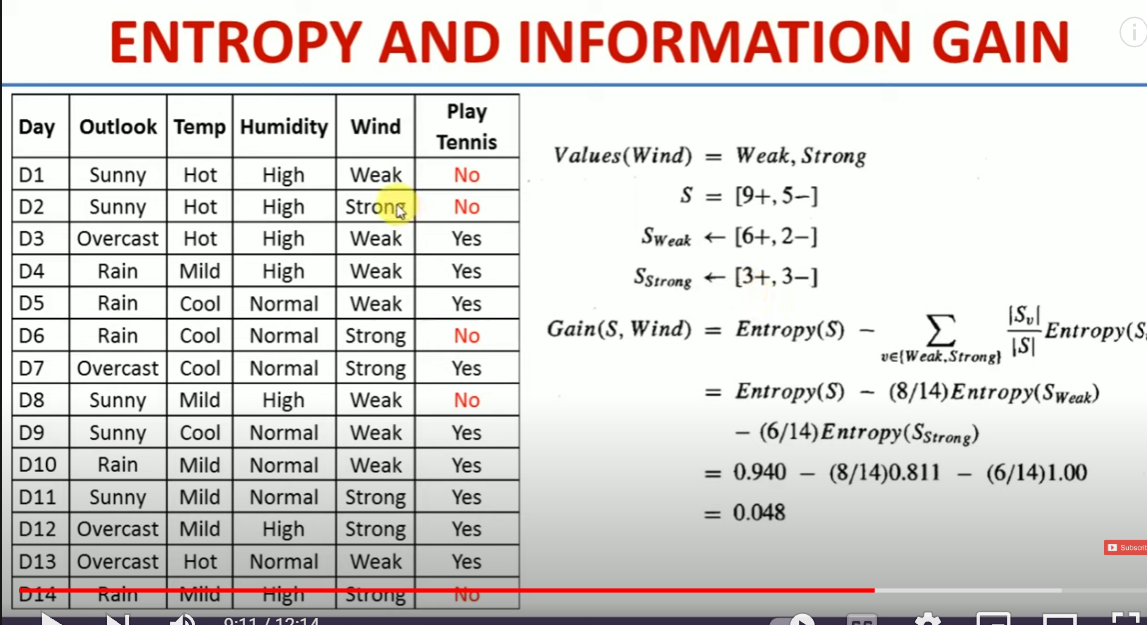
# Entropy In Decision Tree

Entropy(s) =

S= collection of training example

P=proportion of +ve example in S and -ve example in S





Which gain has higher value that will be the root node