

## Multi class Regression:

- \* It will be used when Dependent variable consists of more than 2 classes.
- \* when there were 3 classes, the model will find best fit line for 3 Times.

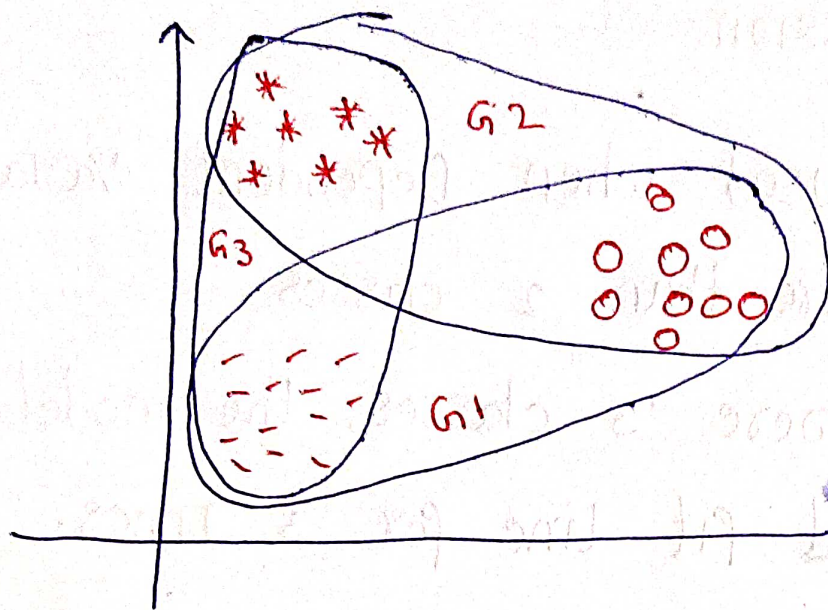
First it will group

- \* 1 and 2 classes and find line between this group and 3.

- \* Next, class 2 and 3 will find line between 1.

- \* Next class 1 and 3 groups and find best fit line with class 2.



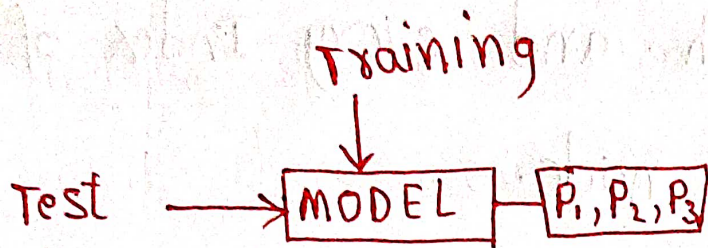


Let's see how it works with an example,

Age	Salary	class	one hot		
			$o_1$	$o_2$	$o_3$
18	25000	M.C	0	1	0
20	50000	Rich	0	0	1
25	45000	M.C	0	1	0
35	30000	Poor	1	0	0
50	55000	M.C	0	1	0
60	35000	Poor	1	0	0



If the model is trained and a new data is passed on it



\* Internally, It will give probabilities for each group of models out of 1 totally.

If resultant probability is

Resultant	one Hot	category
$[0.25, 0.25, \underline{0.5}]$	0, 0, 1	Rich
$[0.10, \underline{0.60}, 0.30]$	0, 1, 0	M.C
$[\underline{0.80}, 0.10, 0.10]$	1, 0, 0	poor



## Decision Trees:

\* Decision Tree splits the Data and will continue untill each and every Index of Data becomes Leaf node.

eg

Degree      company      Salary  $\geq 36,000$

B.Tech

GCS

yes

B.Tech

Info geek

yes

B.com

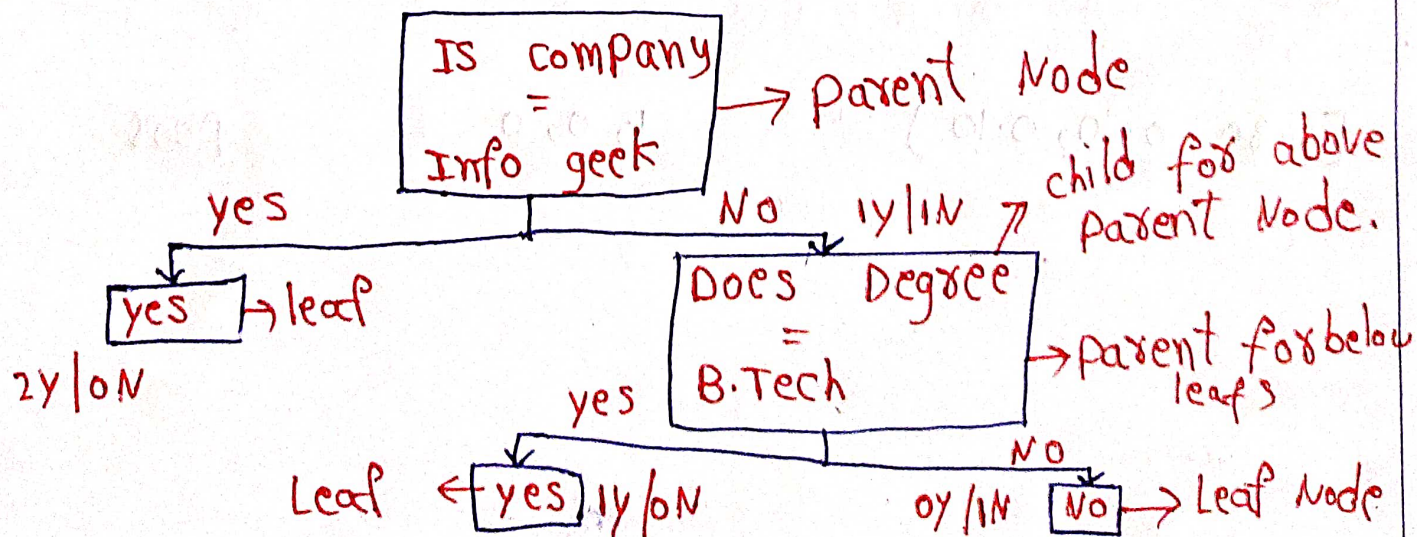
Info geek

yes

B.com

GCS

No





But how Decision tree knows about it's  
First split column:

It uses Information gain formulae to  
Find which column to split on order.

→ It is calculated using Entropy or Gini Impurity.

Entropy: used to check how pure is the  
split.

eg: let's take a split

$1y/1N \rightarrow$  Impure split  $\rightarrow$  Entropy = 1

$3y/0N \rightarrow$  pure split  $\rightarrow$  Entropy = 0

Gini Impurity:

$1y/1N \rightarrow$  Gini = 0.50

$0y/1N \rightarrow$  Gini = 0

## Post Pruning:

- \* used to reduce overfitting.
- \* By reducing the parameters, max-depth, we can reduce overfitting.

## Pre Pruning:

- \* Here, we use Hyper parameter tuning at the start only in order to select best parameters.

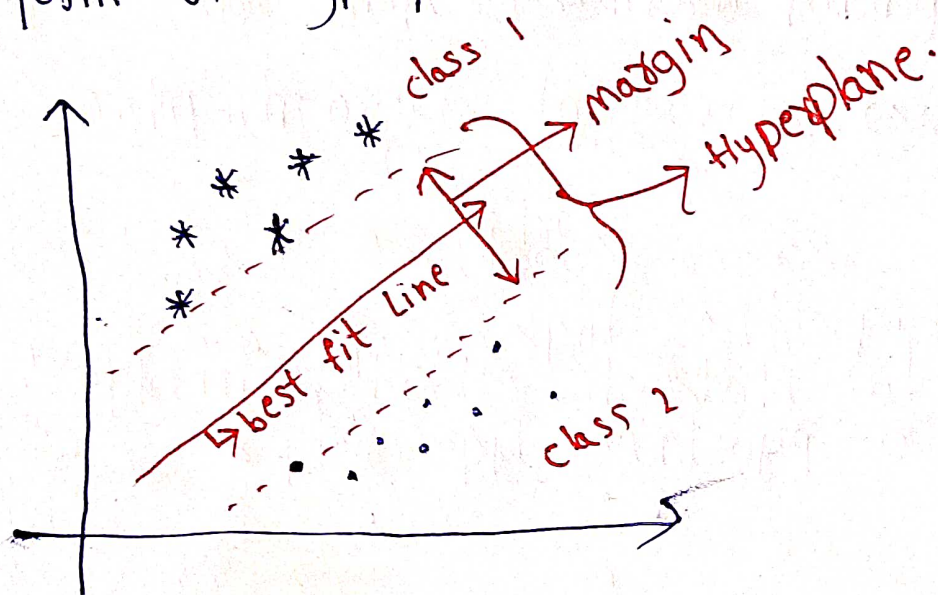
## Decision tree for continuous values:

- \* For splitting, better threshold value will be found
- \* Threshold value is the one that maximizes Reduction of Variance (or) Mean squared error.



## SVM

- \* SVM is majorly used while there is some uncertainty in predictions.
- \* It draws a best fit line first and then uses the support vectors to draw two more lines before and after the best fit line to form a Hyperplane.



- \* If predicted point is above hyperplane, then it's class A and vice versa is B.
- \* If it's exactly on Hyperplane, then there's been a uncertainty in that point, to decide class A (or) class B