Decision Tree →

it Can be used in the Concept of classification and the Concept of regression.

As a classifier (Tree Structured)

Mastly Decision Tree is used for classification and The role of Decision Tree is → it is a classifier.

Data Set → Algorithm → Model classifier

When this classifier is in Tree Structure then We Can say that it is a classifier and Tree Structured classifier.

In the Decision Tree (two Nodes)

(1) Decision Node (Test)

(2) Leaf Node (classification / value)

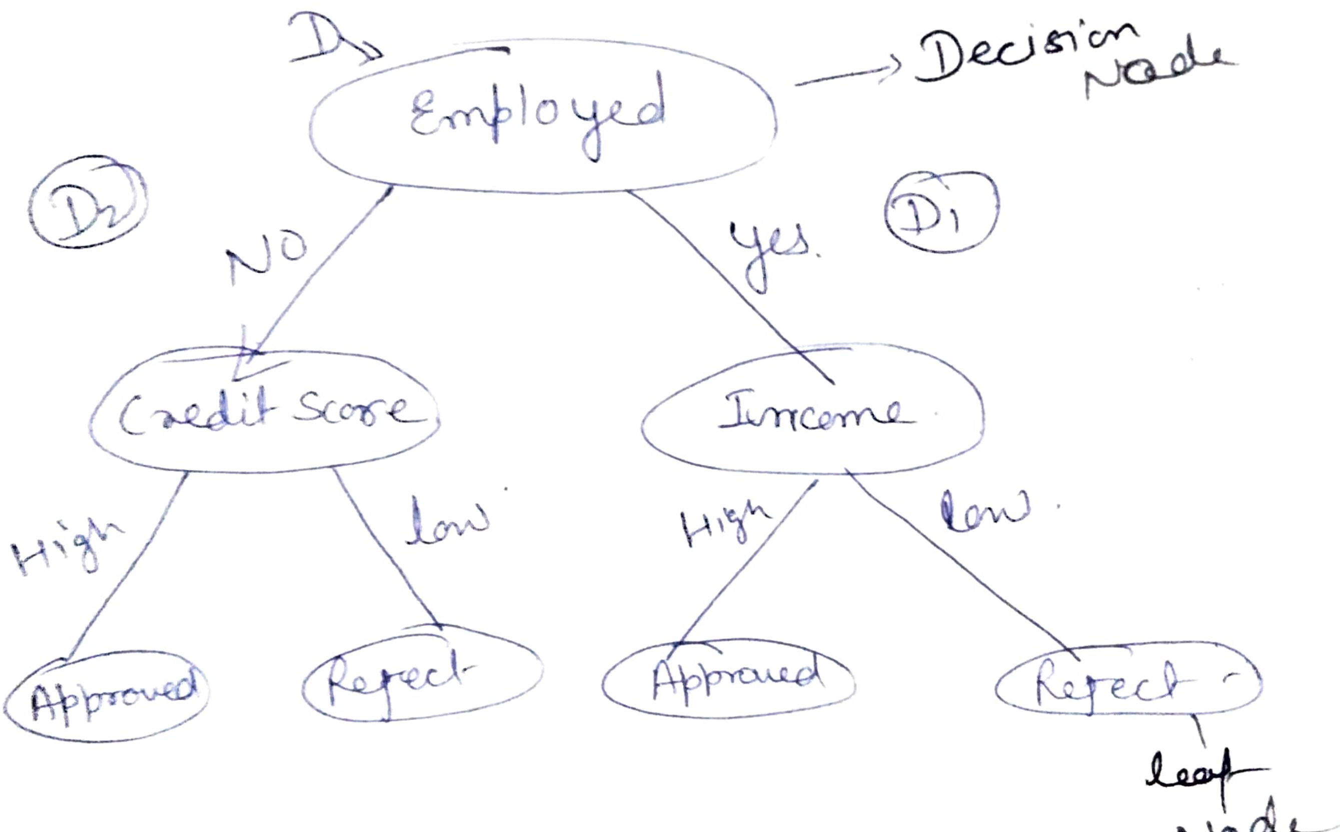
* Nodes that have branches are called Decision Node.

* Nodes that don't have any branch are called leaf Node.

Test is Performed on the feature / Attribute (value).

Dataset

Employed	Credit Score	Income	
Y			D1
Y			
Y			
N			D2
N			



This is the Model and classifier and

(i) We will give whole data set to Decision Node.

(ii) Depending on this test data set will split (this is called spliting action.

(iii) Then this spliting action is perfomed up to the leaf Node.

~~Person is Employed~~

There is a Person Who is an Employee, but if his income is low, then his loan will be approved or Not.

* There is a Person Who is an Employee and having high Credit Score then his loan will be approved or Not.

For Example →

Age	Competition	Type	Profit
old	Yes	S/W	Down
old	NO	S/W	Down
old	NO	H/W	Down
Mid	Yes	S/W	Down
Mid	yes	H/W	Down
Mid	NO	H/W	up
Mid	NO	S/W	up
New	yes	S/W	up
New	NO	H/W	up
New	NO	S/W	up

(i) Selection of Target Attribute
so Profit is target Attribute.

(ii) selection of Decision Node.

Information Gain

$$IG = -\frac{P}{P+N}\log_2\left(\frac{P}{P+N}\right) - \frac{N}{P+N}\log_2\left(\frac{N}{P+N}\right)$$

Entropy

$$E(A) = \sum_{i=1}^{N} \frac{P_i + N_i}{P+N} I(P_i N_i)$$

Information gain of that attribute X Probability of that Attribute

First we have to find Information gain of target attributes :- (4)

$$IG = -\left[\frac{5}{10}\log_2\left(\frac{5}{10}\right) + \frac{5}{10}\log_2\left(\frac{5}{10}\right)\right]$$

$$\Rightarrow -\left[0.5 \times \log_2 2^{-1} + 0.5 \log_2 2^{-1}\right]$$

$$\Rightarrow -\left[0.5 \times \left(-1 \log_2 2\right) + 0.5\left(-1 \log_2 2\right)\right]$$

$$\Rightarrow -\left[-0.5 - 0.5\right]$$

$$\Rightarrow -\left[-1\right]$$

$$\boxed{I.G. = 1}$$

Now we have to find out Entropy of remaining attributes.

Entropy for Age :- E(A)

Age :→

	Down	up
old	3	0
Mid	2	2
New	0	3

How Many times the value of old is Down in the Data Set.

$$I(\text{Old}) = -\left[\frac{3}{3}\log_2\left(\frac{3}{3}\right) + \frac{0}{3}\log_2\left(\frac{0}{3}\right)\right] = 0 \times \frac{3}{10} \Rightarrow 0$$

$$I(\text{Mid}) = -\left[\frac{2}{4}\log_2\left(\frac{2}{4}\right) + \frac{2}{4}\log_2\left(\frac{2}{4}\right)\right] = 1 \times \frac{4}{10} \Rightarrow 0.4$$

$$I(\text{New}) = -\left[\frac{0}{3}\log_2\left(\frac{0}{3}\right) + \frac{3}{3}\log_2\left(\frac{3}{3}\right)\right] = 0 \times \frac{3}{10} \Rightarrow 0$$

$$E(\text{Age}) = 0 + 0.4 + 0 \Rightarrow 0.4$$

$$\boxed{E(\text{Age}) \Rightarrow 0.4}$$

$$\text{Gain} \Rightarrow IG - E(A)$$

$$\text{Gain}(\text{Age}) \Rightarrow 1 - 0.4 = 0.6$$

$$\text{Gain}(\text{Competition}) = 0.124$$

$$\text{Gain}(\text{Type}) = 0$$

$$\boxed{I.G. = 1}$$

Attribute that have highest gain will be a Root Node.

There are 3 variant of Age ~~attribute~~