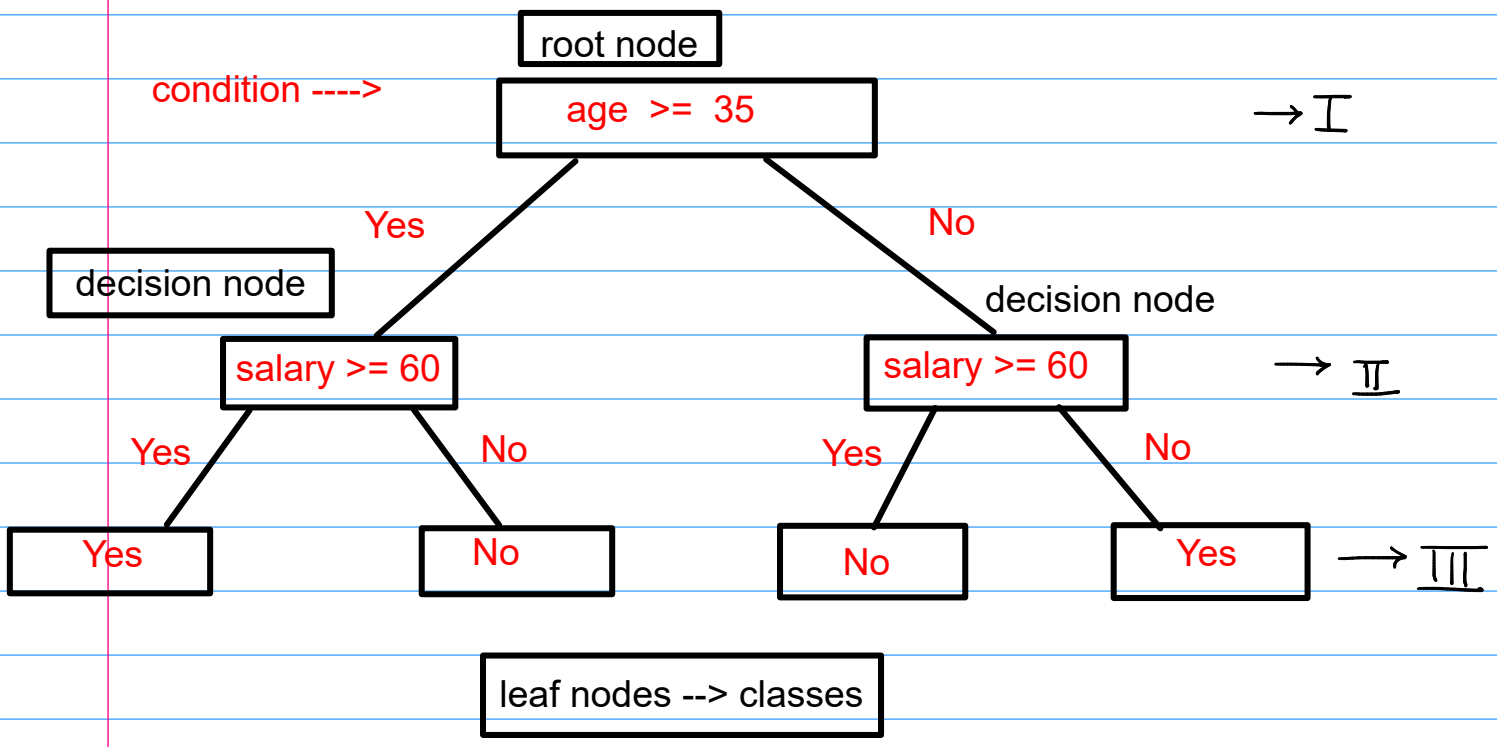
DECISION TREE

- Looks like a tree
- a feature which can take up as the root of the tree
- CART algo = Classification And Regresion Trees
-classification and regression related problems then always go for decision tree
- multiple number of check conditions in the data base then go for decision tree
- algo contains conditional control statements like we do with if-else statements

age	salary	purchase
20	50k	1
30	80k	0
23	60k	1
..	..	..
..	..	..
.	..	..
..	..	..
..	..	..

condition 1:
age = 10 - 50 years
salary = 50k - 100k

condition 2:
age = 30years
salary = 80k



Criteria-

These four terms are criteria used in machine learning
(especially decision trees and feature selection)

Entropy → measures impurity or randomness

Information Gain → reduction in entropy after a split

Gain Ratio → normalized information gain (used in C4.5)

Gini Index → measures impurity (used in CART)

Chi Square → is a statistical test of independence
(It measures how strongly a feature is associated with the target class)
- Used mainly in CHAID (Chi-square Automatic Interaction Detection) decision trees
- Determines whether a split is statistically significant
- Helps stop splitting if the feature and class are independent

entropy / Information Gain / Gain Ratio / Gini Index
→ impurity-based split criteria

Chi-square
→ statistical significance–based split criteria

When to use what ?

- Building decision trees → Gini or Entropy

- Fast & practical → Gini

- Theory & explanation(Theoretical clarity) → Entropy + information gain

- Selecting important features → Chi-square
(selecting features before training)

Gini Index and Entropy are both used in decision trees to measure node impurity,
Use Gini when -
- You want faster training
- is faster and favors balanced splits,
- good for large datasets,
.

while Entropy-
- is slower but more sensitive to class imbalances,
- providing deeper insight into disorder, often yielding slightly better results but at a higher computational cost

Choose Gini for speed/efficiency (CART algorithm default) and Entropy for more theoretical splits, though results are often similar.