Machine Learning

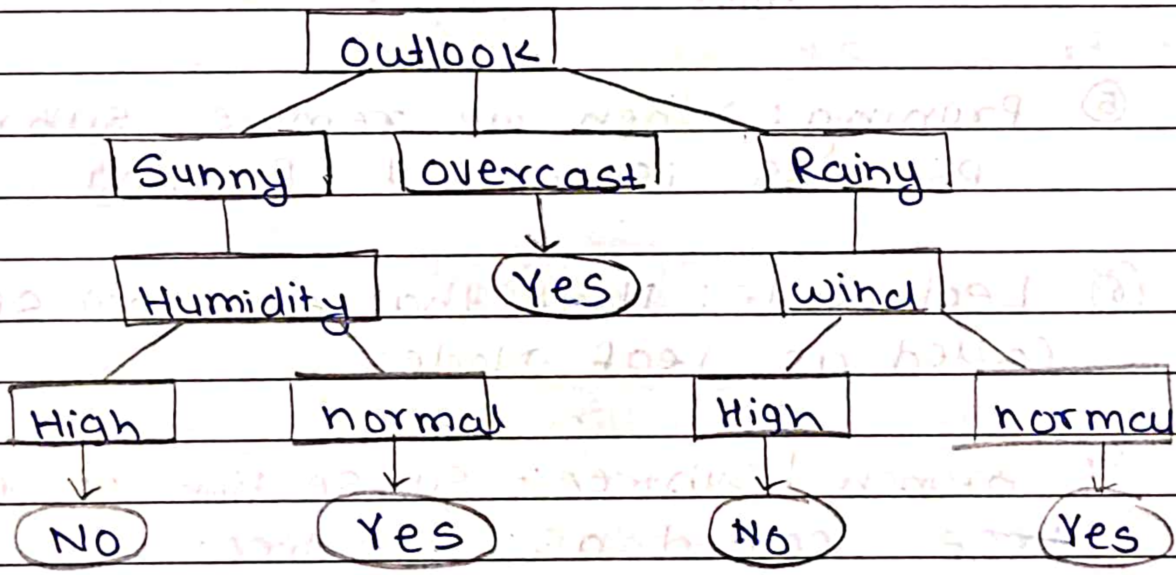
V-4- Desision tree & Probabilistic Models.

Q-1- What is Binary Decision tree ? — 8M

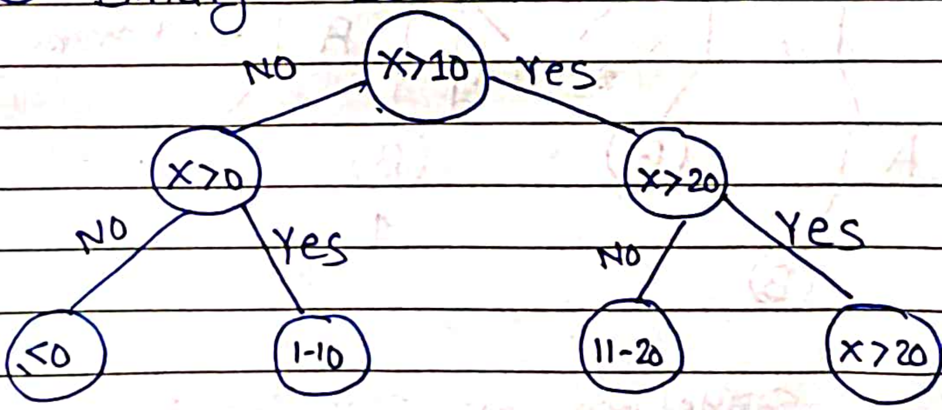
→ Decision tree is a supervised machine learning approch usualy used for classification but can also be used as regression

Binary decision tree is a decision tree which operates by every attribute series by binary desision i.e. Yes or No.

eg:- (i) Decision tree for Playing Tennis



(ii) Binary decision tree for Numbers.



• Key terms

① Root : The Decision tree beigns with root node further divided into subnodes. childs

② Splitting : The Process of dividing Parent node into subnode called as splitting

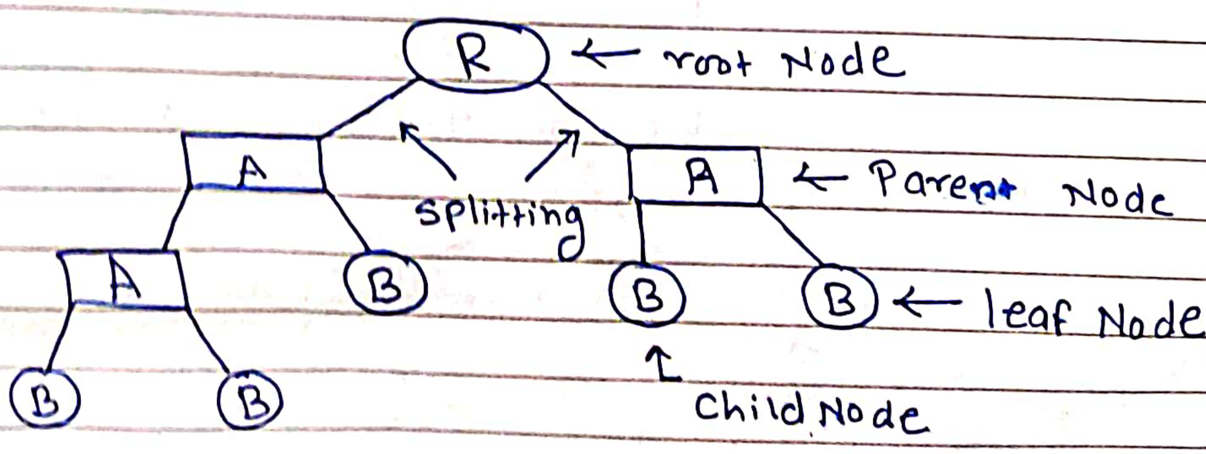
③ Parent Node: A Node which is divided into sub node called as parent node

④ Child Node :- All subnodes called as child node

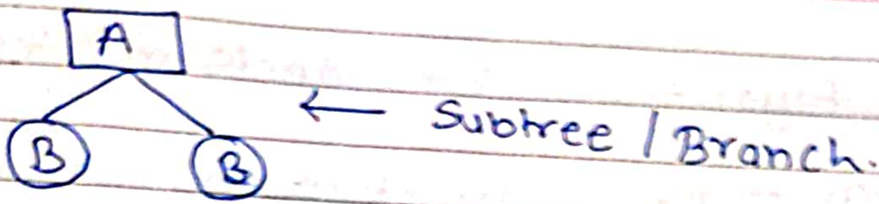
⑤ Pruning : When we remove subnode the Process is called Pruning.

⑥ Leaf Node : Node that does not split further called as leaf Node

⑦ Branch | Subtree :- Subsection of entire tree called as sub tree.



Structure of Decision tree

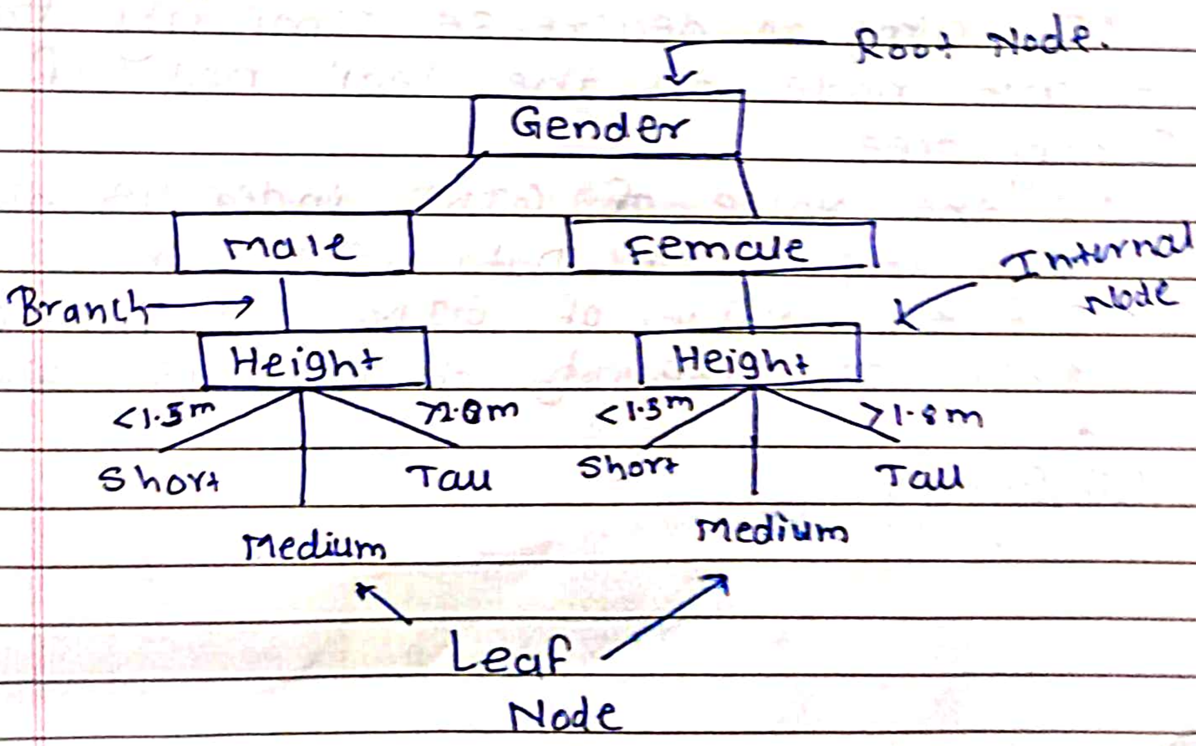


R : Root Node of the tree

A : Parent Node / Decision Node

B : Leaf Node of the tree.

eg:- A Example dataset of 30 students of gender boys & girls of class IX & X and hight 5ft to 6ft. 15 out of 30 play cricket in leisure time.



Date
Page

• Algorithms for decision tree

greedy ← ① ID3 (Iterative Dichotomiser 3)
↑acc/dynamic ← ② C4.5
↓ ③ CART (Classification & Regression Tree).
Binary tree

Q-2- Define : (i) GINI index —— | 2 M |
(ii) Entropy —— | 2 M |

→

(i) GINI index

: GINI index is a Powerful measure that measure randomness or impurity of the values of dataset

• It aims to deacrease Impurity from the root node to the leaf node of the Binary tree

• If the value of GINI index is higher then impurity of Data is higher

• If the value of GINI index is lower then The impurity of Data is lower.

•

(ii) Entropy

Q-3- Explain Naive Bayes Classifier working — 10m

→ Bayes Thm = $\frac{P(B|A)\,P(A)}{P(B)} = P(A|B)$

- Using Bayes Thm we can find the Probability of A happening given that B is occured here B is the evidance & A is the hypothe

- Assumption is made that feature is indepedent
- Presence of one feature does not affect other features

Naive bayes Model: It is a Probabilistic machine learning model that used for classification tasks.

let's consider a Example of 1000 fruits which could be either 'banana', 'orange' or 'other' there are three Possible classes of Variable 'Y' either 'banana', 'orange' or 'other'. given that it is long sweet & yellow

Type	Long	Sweet	Yellow	Total
Banana	400	350	450	500
orange	0	150	300	300
other	100	150	50	200
Total	500	650	800	1000

Step 1

$P(\text{Banana}) = \frac{500}{1000} = 0.50$

$P(\text{orange}) = \frac{300}{1000} = 0.30$

$P(\text{other}) = \frac{200}{1000} = 0.20$

Step 2

$P(\text{long}) = \frac{500}{1000} = 0.50$

$P(\text{sweet}) = \frac{650}{1000} = 0.65$

$P(\text{Yellow}) = \frac{800}{1000} = 0.80$

Step 3

Ⓐ For Banana.

$P(\text{long} \mid \text{Banana}) = \cancel{\frac{400}{500} = 0.80}$

$= \cancel{P(\text{Banana} \mid \text{orange})}\ P(c$

$$= \frac{P(\text{Banana} \mid \text{long}) \cdot P(\text{long})}{P(\text{Banana})}$$

$$= \frac{\frac{400}{500} \times \frac{500}{1000}}{\frac{500}{1000}} = 0.80$$

$$P(\text{Sweet} \mid \text{Banana}) = \frac{P(\text{Banana} \mid \text{Sweet}) \, P(\text{Sweet})}{P(\text{Banana})}$$

$$= \frac{\frac{350}{650} \times \frac{650}{1000}}{\frac{500}{1000}} = 0.7$$

$$P(\text{Yellow} \mid \text{Banana}) = \frac{P(\text{Banana} \mid \text{Yellow}) \, P(\text{Yellow})}{P(\text{Banana})}$$

$$= \frac{\frac{450}{800} \times \frac{800}{1000}}{\frac{500}{1000}} = 0.9$$

Step 4

$P(\text{fruit} \mid \text{Banana}) = 0.8 \times 0.7 \times 0.9 = 0.504$

Q-4- Write Short Note on (16)

(i) Multinominal Naive Bayes

→

If feature of Bayes distrubution used in text classification for multinomially distrubuted data.

- **Discreate count**
- multinominal distrubution
- Document classification

BG	O	A	B	AB
P	0.44	0.42	0.10	0.04

$$\boxed{P(X = x_1, \dots x_k) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}}$$

$6 \longrightarrow$ Person.

$1 \rightarrow$ O $\quad 2 \rightarrow$ A

$2 \rightarrow$ B $\quad 1 \rightarrow$ AB

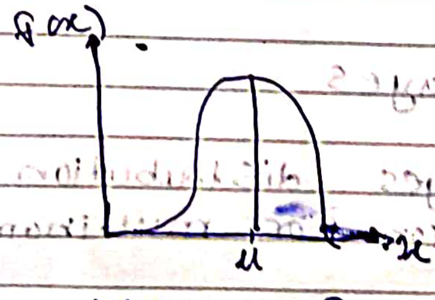
$$P(x_1 = 1, x_2 = 2, x_3 = 2, x_4 = 1)$$

$$= \frac{6!}{1!\,2!\,2!\,1!} \times (0.44)^1 \times (0.42)^2 \times (0.10)^2 (0.04)^1$$

180 x $\qquad = \underline{0.0065}$

(ii) Gaussian Naive Bayes

- used for classification
- Gaussian Distribution
- used for feature are contious nature.



$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

PDF

- Normal Distribution

(ii) Bernoulli Naive Bayes

- If features are in Binary nature then we use Bernoulli naive Bayes

- P (success) = P
  P(faliure) = q = 1 - P

If random variable value is 1 → success
If random variable value is 0 → failure

'X has a Bernoulli distribution'

$$P(X = x) = p^{x} \cdot (1-p)^{1-x}$$

$$P(X = x) = \begin{cases} p & \text{if } \cancel{x} = 1 \\ q & \text{if } x = 0 \end{cases}$$