

Assignment

on

Classification & Decision Trees

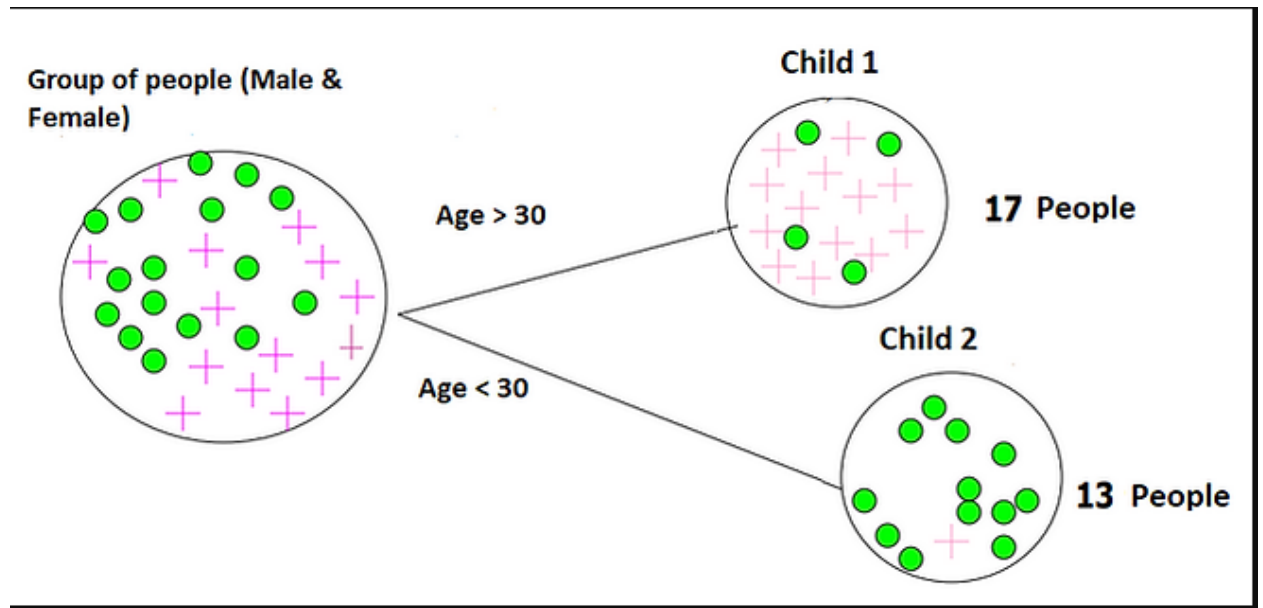
1. What is Classification?
2. Why Classification is used?
3. What are the different methods of Classification?
4. Enlist five different Examples of Classification
5. What is an attribute? What are the different types of attributes?
6. What is Dataset?
7. What is a training Dataset?
8. What is testing Dataset?
9. What is a validation Dataset?
10. What is Entropy?
11. What is Conditional Entropy ?
12. What is the formula used to compute the Entropy?
13. What is Information Gain? What is the formula used to compute the Information Gain?
14. What is a Decision Trees?
15. Define the following
 - Root node
 - Leaf Node
 - Decision Node
16. What is Contingency Table?
17. If we had a total of 100 data points in our dataset with 30 belonging to the positive class and 70 belonging to the negative class, then compute the Entropy of this dataset.
18. Consider the following Table

Credit Rating	Liability		
	Normal	High	Total
Excellent	3	1	4
Good	4	2	6
Poor	0	4	4
Total	7	7	14

Then compute

- Entropy (Liability)
- Entropy (Liability | Credit Rating = Good)
- Entropy (Liability | Credit Rating = Poor)
- Entropy (Liability | Credit Rating = Excellent)
- Find Information Gain (Liability | Credit Rating)

19. Consider the following dataset details



Compute the entropies of

- Parent Node and
- Child (1 & 2) nodes
- Information Gain