1. What are Entropy, Information Gain, Gini Index? Write short notes on these.
2. Differentiate between Supervised and Unsupervised learning.
3. Write down the limitations, advantages and applications of Decision Tree and Naive Bayes Classifier.
4. What are the trade-offs of using decision trees vs. neural networks for certain machine learning applications?
5. What are some of the alternative algorithms to decision trees?
6. Consider the following dataset and build a decision tree for it.

	Animals	Size of Animal	Body Color	Can we Pet them
0	Dog	Medium	Black	Yes
1	Dog	Big	White	No
2	Rat	Small	White	Yes
3	Cow	Big	White	Yes
4	Cow	Small	Brown	No
5	Cow	Big	Black	Yes
6	Rat	Big	Brown	No
7	Dog	Small	Brown	Yes
8	Dog	Medium	Brown	Yes
9	Cow	Medium	White	No
10	Dog	Small	Black	Yes
11	Rat	Medium	Black	No
12	Rat	Small	Brown	No
13	Cow	Big	White	Yes

Target

7. Consider the given dataset in question 6 and predict the classification label for the following test data by using Naive Bayes Classifier.
   Test_set1 = {Cow, Big, Black}
   Test_set2 = {Rat, Small, Black}
   Test_set3 = {Dog, Medium, Brown}
   Test_set4 = {Dog, Big, White}
   Test_set5 = {Cow, Big, Black}
   Test_set6 = {Rat, Small, White}
8. Write down the limitations and significance of using Naive Bayes Classifier.
9. What is Clustering? How is clustering different from classification tasks?
10. Briefly describe Partitioning, Hierarchical, Density based methods of clustering.
11. Describe how the clustering is different from segmentation?
12. Draw the flowchart of K-means clustering and explain the steps. Try to design an algorithm implementing clustering by yourself.
13. Consider the following data table. Assume K=3 and perform K-means clustering based on the given data. Show all the necessary iterations.

ID	*Breadth (cm)*	*Height (cm)*	*Length (cm)*
1.	55	14	33

2.	33	45	66
3.	74	25	45
4.	55	52	44
5.	33	24	33
6.	45	45	55
7.	55	33	42

14. What is the Elbow Method? Describe how you will choose the elbow point.
15. What are the limitations of K-means clustering ?
16. What is the DBSCAN algorithm used for?
17. Describe the min points, epsilon, core point, border point, outlier.
18. What are these terms:
    i) Directly Density Reachable
    ii) Density Connected
    iii) Density Reachable
19. How do you implement the DBSCAN algorithm? Show the steps for the following dataset and find out the Core Points, Border Points and Outliers. Select the clusters as well.

Points	X	Y
P1	2	-2
P2	1	2
P3	-5	-10
P4	1	15
P5	-10	5
P6	16	6
P7	21	12
P8	55	55