UNIT 4: LEARNING WITH TREES

1. Explain with an example the steps involved in Iterative Dichotomiser 3(ID3) Algorithm. What are the features of ID3 Algorithm?
2. Consider the following set of training examples:

Instance	Classification	a_1	a_2
1	+	T	T
2	+	T	T
3	−	T	F
4	+	F	F
5	−	F	T
6	−	F	T

   (a) What is the entropy of this collection of training examples with respect to the target function classification?

   (b) What is the information gain of *a2* relative to these training examples?

3. Write and explain the steps involved in Classification and Regression Trees (CART).
4. What is a decision tree? Discuss with an example. Which are the appropriate problems for decision tree learning?
5. What are the similarities and differences between ID3 and CART?
6. Consider the training examples in Fig below. Compute the information gain for each of the attributes. Construct a 2 level decision tree.

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

7. Compare ID3 and CART with respect to learning rules.

UNIT 4: PROBABILISTIC LEARNING

8. List and briefly explain a few applications of the Expectation Maximisation Algorithm.
9. What are the two distinctive steps of the Expectation Maximisation Algorithm? Briefly explain how it this algorithm fits into the Gaussian Mixture Model(GMM).
10. What is the difference between K-means Clustering and GMM? Why do we need GMM?
11. With any example of your choice illustrate the use of k-Nearest Neighbour(KNN) as a probabilistic learning algorithm.
12. All numericals to solve KNN is the slide 'Probabilistic Learning_KNN.pdf'

UNIT 4: SUPPORT VECTOR MACHINES (SVM)

13. Numericals for Linearly Separable and Non-Linearly separable data in the slides 'Support Vector Machines_MKN.pdf' and 'SVM.example_PB.pdf'
14. Explain how extensions of SVM is applied to ***any one*** of the below scenarios:
    (i) Multi-class classification
    (ii) Regression
15. With suitable example, illustrate the Optimal separators in Classification problem.
16. Suppose that the following are a set of points in two classes:

$$\text{class 1} \quad : \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\text{class 2} \quad : \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Plot them and find the optimal separating line. What are the support vectors, and what is the margin?

17. Suppose that the following are the given positively labeled data points in $\mathbb{R}2$:

$$\left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 6 \\ 1 \end{pmatrix}, \begin{pmatrix} 6 \\ -1 \end{pmatrix} \right\}$$

and the following negatively labeled data points in $\mathbb{R}2$:

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$$

Obtain a separating hyperplane that accurately discriminates the two classes using SVM.

18. Suppose that the points are now:

$$\text{class 1} \quad : \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\text{class 2} \quad : \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Try out the different basis functions that were given in the chapter to see which separate this data and which do not.