

Instructions to the Students:

1. Check that you have received a correct Question paper.
2. Assume suitable data if necessary and mention it clearly
3. Draw NEAT labeled diagrams wherever necessary

Q.1. Attempt any Six Questions

(1*6 = 6 Marks)

1. Define Hypothesis Space
2. What is information gain?
3. What is difference between supervised and unsupervised learning
4. What is cross validation?
5. Write Bayes theorem.
6. Define confusion matrix with suitable example.
7. What is SVM?

Q.2. Attempt any Two of the following

(2*3 =6 Marks)

1. Explain problem of overfitting in ML?
2. Differentiate Linear Vs Logistic regression. Give suitable example of each.
3. Apply KNN for following dataset and predict class of test example ($A1 = 3, A2 = 7$). Assume $K=3$

A1	A2	Class
7	7	True
7	4	True
3	4	False
1	4	True
5	3	False
6	3	True

Q.3. Attempt any One of the following

(1*8 = 8 Marks)

1. Explain Bayesian Learning and Naïve Bayes Classifier. Solve the following example.
A patient takes a lab test and the result comes back positive. The test returns a correct positive result in only 98% of the cases in which the disease is actually present, and a correct negative result in only 97% of the cases in which the disease is not present. Furthermore, .008 of the entire population have this cancer. Find $P(\text{cancer} | +)$ and $P(\neg \text{cancer} | +)$

2. Illustrate the operation of ID3 for the following training example. Consider Information Gain as Attribute Selection Measure.

PlayTennis: training examples

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