

CO-327 MACHINE LEARNING

Time: 3:00 Hours

Max Marks: 40

NOTE: Answer any FIVE questions. Assume suitable missing data, if any.

Q1. Answer all the following questions:

- [a] Consider the problem of sorting 'n' numbers. Is it wise to apply machine learning to solve this problem? Justify your answer.
- [b] Illustrate difference between hard margin and soft margin SVM.
- [c] Give two real life examples of unsupervised learning.
- [d] Write the applications of reinforcement learning. [4 X 2=8]

Q2.[a] Machine learning algorithms are used for model prediction. How do we build model using machine learning algorithms. Explain stepwise procedure with schematic diagram. [4]

[b] You are given the project for faculty appraisal in university based on academics as well as research records. Analyse and suggest the various dependent and independent attributes which could be present in the training data provided by the university. Categorize the variables according to their type. [4]

Q3 [a] Consider the following dataset. Construct a decision tree using information gain as attribute selection measure. [4]

Price	Maintenance	Capacity	Airbag	Profitable
Low	Low	2	No	Yes
Low	Med	4	Yes	No
Low	Low	4	No	Yes
Low	High	4	No	No
Med	Med	4	No	No
Med	Med	4	Yes	Yes
Med	High	2	Yes	No
Med	High	5	No	Yes
High	Med	4	Yes	Yes
High	High	2	Yes	No
High	High	5	Yes	Yes

[b] Explain splitting of decision tree based on different types of attributes with suitable examples. Also, tell the significance of Occam's Razor in decision trees. [4]

P.T.O

Q4. [a] Compute eigenvalues and eigenvectors for given 2 X 2 matrix.

$$\begin{bmatrix} 5 & -3 \\ -6 & 2 \end{bmatrix}$$

How do you retrieve the original data from principal components in PCA method? [4]

[b] Explain taxonomy of dimensionality reduction methods used in machine learning. Briefly differentiate between wrapper classes and filter classes. [4]

Q5. [a] What is need of Reinforcement learning? Compare supervised machine learning and Reinforcement learning. Also highlight the challenges that we face in reinforcement learning. [4]

[b] Differentiate between leave one out and hold out cross validation with the help of an example. [4]

Q6..[a] Take the following training samples

X	Y	Z	Fruit
4	6	7	Apple
2	4	9	Mango
4	2	4	Orange
5	5	7	Apple
4	3	9	Orange

Predict the label of test sample (X, Y, Z) =(3, 9, 3) with the KNN classifier (K=3) using Manhattan distance. [4]

[b] Explain role of prior probability, likelihood, evidence, Bayes rule and posterior probability in Naïve Bayes classifiers. [4]

Q7. [a] Performance analysis of classifier is better practice. Do you agree with this statement? What can be the different performance evaluation metrics for any classifier? Explain their significance and how do you calculate them? [4]

[b] How do you deal with multicollinearity and imbalanced data issue in machine learning? [4]