# [5461] - 602

# B.E. (Information Technology)

# MACHINE LEARNING & APPLICATIONS

## (2015 Pattern) (414454) (End Semester)

*Time : 2½ Hours]*

*[Max. Marks : 70*

*Instructions to the candidates:*

1) *Answer Q1 or Q2, Q3 or Q4, Q5 or Q6, Q7 or Q8, and Q9 or Q10.*
2) *Neat diagrams must be drawn wherever necessary.*
3) *Figures to the right side indicate full marks.*
4) *Use of calculator is allowed.*
5) *Assume suitable data if necessary.*

**Q1)** a) Explain with example k-fold cross validation. **[5]**

b) Write short note on Vapnik-Chervonenkis dimension. **[5]**

OR

**Q2)** a) Explain two methods for reducing dimensionality. **[5]**

b) Write short note on Gram Matrix with an example. **[5]**

**Q3)** a) What are support vector margins and also explain soft margin. **[5]**

b) Explain the term bias-Variance dilemma. **[5]**

OR

**Q4)** a) Explain Predictive and descriptive task. **[5]**

b) Explain Perceptron training algorithm for linear classification. **[5]**

**Q5)** a) Consider following five data points: **[12]**

(0,3), (3,3) (3,0), (-2,-4) and (-4,-2)

Clusters are formed as follows:

*P.T.O.*

Case 1:

A. First two points together in one cluster.

B. Remaining three in another cluster.

Case 2:

A. First three points together in one cluster.

B. Remaining two in another cluster.

Find Out:

i) Within-cluster scatters for both cases.

ii) Between-cluster scatters for both cases.

iii) Also comment which clustering produces tighter cluster whose centroids are further apart.

b) Define and explain following terms.                                        **[6]**

i) Minority Class.

ii) Gini Index.

iii) Entropy.

**OR**

*Q6)* a) Find all association rules in the following database in the following database with minimum support = 2 and minimum confidence = 65%.**[10]**

| Transactions | Data Items |
|---|---|
| T1 | Milk, Bread, Cornflakes |
| T1 | Bread, Jam |
| T1 | Milk, Bread, cornflakes, Jam |
| T1 | Milk, cornflakes |
| T1 | Bread, Butter, Jam |
| T1 | Bread, Butter |
| T1 | Milk, Bread, Butter |

b) Consider following splits having four features:                          **[8]**

Length = [3,4,5]          [2+,0-][1+, 3-] [2+, 2-]

Gills = [Yes, No]         [0+, 4-] [5+, 1-]

Beak = [Yes, No]          [5+, 3-] [0+2-]

Teeth = [many, few]       [3+, 4-] [2+, 1-]

Find

Total weighted Entropy & Gini-index of all Features.

**Q7)** a) Define Bayes Rule and solve following example. **[8]**

Example:

5% of people in a city having cancer. In that city 10% people are smoker, Also 20% of people with cancer and smoker.

Find out the probability of people who are smoker possess cancer.

b) Define **[8]**

i) Bernoulli's distribution

ii) Binomial distribution.

iii) MAP decision rule.

iv) Maximum likelihood function.

OR

**Q8)** a) For the given dataset apply Naïve Bayes algorithm and predict the outcome for the car={Red, Domestic, SUV}. **[8]**

| Color | Type | Origin | Stolen |
|-------|------|--------|--------|
| Red | Sports | Domestic | Yes |
| Red | Sports | Domestic | No |
| Red | Sports | Domestic | Yes |
| Yellow | Sports | Domestic | No |
| Yellow | Sports | Imported | Yes |
| Yellow | SUV | Imported | No |
| Yellow | SUV | Imported | Yes |
| Yellow | SUV | Domestic | No |
| Red | SUV | Imported | No |
| Red | Sports | Imported | Yes |

b) Write short note on GMM. **[8]**

**Q9)** a) Write short note on Feed - forward Neural Network. **[8]**

b) Write short note on Ensemble learning. **[8]**

OR

**Q10)** a) Explain why we use non-linearity function? States & explain 3 types of neurons that add non-linearity in their computations. **[8]**

b) Write short note on Reinforcement learning. **[8]**

● ● ●