PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc(SS) DEGREE EXAMINATION MAY 2024
(Eighth Semester)

Branch – **SOFTWARE SYSTEMS (five year integrated)**

MACHINE LEARNING

Time: Three Hours Maximum: 75 Marks

SECTION-A (10 Marks)

Answer **ALL** questions
ALL questions carry **EQUAL** marks (10 x 1 = 10)

1. What is the application of machine learning methods to a large database called?
   i) Big data computing ii) Data mining
   iii) Internet of things iv) Artificial intelligence

2. Identify the type of learning in which label training data is used.
   i) Semi unsupervised learning ii) Supervised learning
   iii) Reinforcement learning iv) Unsupervised learning

3. Artificial neural network is used for
   i) Classification ii) Pattern recognition
   iii) All of the above iv) Clustering

4. Pattern recall takes more time for
   i) MLFNN ii) Equal for both MLFNN and basis function
   iii) None of the mentioned iv) Basis function

5. Which of the following statements is not an advantage of Reduced error pruning?
   i) Linear computational complexity ii) Over pruning
   iii) Simplicity iv) Speed

6. What are different types of nodes a decision tree has?
   i) Root node ii) Internal nodes
   iii) Leaf nodes iv) All of the above

7. In terms of the bias-variance trade-off, which of the following is substantially more harmful to the test error than the training error?
   i) Bias ii) Loss
   iii) Variance iv) Risk

8. Which of the following learning algorithms will return a classifier if the training data is not linearly separable?
   i) Hard margin SVM ii) Soft margin SVM
   iii) Perceptron iv) Naïve bayes

9. K-Means Clustering comes under
   i) Supervised learning Algorithm ii) Unsupervised Learning Algorithm
   iii) Reinforcement Learning iv) None of the above

10. Which of the following optimized techniques are used in K-Means Clustering Algorithm?
    i) K-Means++ ii) Elbow plot
    iii) Only 2 iv) Both (i) and (ii)

Cont...

SECTION - B (25 Marks)

Answer **ALL** questions
ALL questions carry **EQUAL** Marks (5 x 5 = 25)

11. a) Explain the concept of supervised learning.
(OR)
b) Describe the key differences between classification and regression in supervised learning.

12. a) Explain the Radial Basis Networks.
(OR)
b) Differentiate between Linear Regression and Logistic Regression.

13. a) Differentiate between Parametric Methods and Non-Parametric Methods.
(OR)
b) How to Evaluate Machine Learning Algorithms?

14. a) Illustrate the VC dimension.
(OR)
b) Discuss the concept of PAC learning and its connection to generalization in machine learning.

15. a) Elucidate the K-Means clustering Algorithm with its Initialization Methods and Convergence Criteria.
(OR)
b) Discuss the importance of Dimensionality Reduction.

SECTION -C (40 Marks)

Answer **ALL** questions
ALL questions carry **EQUAL** Marks (5 x 8 = 40)
Question No. 16 is Compulsory

16. Discuss any four Supervised learning algorithms.

17. a) Discuss about Naïve Bayes with examples.
(OR)
b) Write about a K-Nearest Neighbour Algorithms.

18. a) Construct the Maximum likelihood estimation in machine learning.
(OR)
b) Determine the Model Selection in Machine Learning.

19. a) Evaluating Bias Variance Trade Off – Machine Learning.
(OR)
b) Infer the principles of online learning.

20. a) Explain the use of hierarchical clustering in real=world scenarios.
(OR)
b) Describe the application of principal components analysis in dimensionality reduction.

Z-Z-Z END