

MAHATMA GANDHI UNIVERSITY, KOTTAYAM

MGU-UGP (HONOURS)

END SEMESTER EXAMINATION

SEMESTER-V

DSE – MG5DSEMATA00 – SUPERVISED MACHINE

LEARNING:ALGORITHMS AND APPLICATIONS

(2024 Admission onwards)

MODEL QUESTION PAPER

Duration: 2 Hours

Maximum Marks: 70

*Remember(K), Understand(U), Apply(A), Analyse(An), Evaluate(E),
Create(C),Skill(S),Interest(I),Appreciation(Ap)*

Students should attempt at least one question from each course outcome.

Part A

One or two sentences

Answer any **5** questions

Each question carries **2** marks

1. State any two advantages of the KNN algorithm. [U/CO:1]
2. What is conditional probability? [U/CO:1]
3. Express the main steps of decision tree algorithm in flow chart form using an example. [U/CO:2]
4. Define Entropy. [K/CO:2]
5. What is information gain? [K/CO:2]
6. Define a Support Vector Machine. [K/CO:3]
7. Name any two kernel functions used in SVM. [K/CO:3]
8. Define Bagging Algorithm. [U/CO:4]

[2 × 5 = 10]

Part B

Short answer

Answer any **5** questions

Each question carries **6** marks

9. Explain the KNN algorithm and describe its main steps. [U/CO:1]
10. Describe the role of Bayes' theorem in Naïve Bayes classification. [U/CO:1]
11. What is the purpose of the function SplitDataset()? [E/CO:2]
12. Write a short note on the ID3 algorithm and recursive construction. [E/CO:2]
13. Explain the concepts of entropy using an example. [A/CO:2]
14. Explain the kernel trick used in SVM. [U/CO:3]
15. Describe the difference between linear and non-linear SVM. [U/CO:3]
16. What are the advantages of bagging learning technique over single decision tree algorithm? [A/CO:4]

[6 × 5 = 30]

Part C

Essays

Answer any **3** questions

Each question carries **10** marks

17. Compare Naïve Bayes with the K-Nearest Neighbour algorithm based on working principle, advantages, and limitations. [U/CO:1]
18. Consider the following dataset containing the age and annual income of five customers.

Customer	Age	Annual Income ()	Class
A	22	25000	No
B	25	30000	No
C	28	35000	Yes
D	35	50000	Yes
E	40	60000	Yes

Write a Python program using the K-Nearest Neighbours (KNN) algorithm to predict the class of a new customer with age = 30 years and annual income = 40000, taking $K = 3$. [A/CO:1]

Table 1: Fish Data Set

No Surface	Flippers	Fish?
1	1	Maybe
1	1	Yes
1	0	No
0	1	No
0	1	No

19. Write the Python program for calculating Shannon Entropy of the following data
[U/CO:2]
20. Write a python program to implement the Splitdataset() function used in the ID3 decision tree algorithm.
[E/CO:2]
21. Consider the following two-dimensional training data:

Class	Points
+1	(0, 2), (1, 3), (3, 6)
-1	(0, 0), (1, 1), (2, -1)

The separating hyperplane is given by

$$x_2 - x_1 - 1 = 0.$$

- (a) Determine the equations of the two margin boundaries.
- (b) Calculate the width of the margin.
- (c) Identify the support vectors from the given data points.
- (d) Classify the point (2, 4) using the given SVM decision function.

[A/CO:3]

22. Explain the application, optimization, and interpretability of Random Forest Regression in machine learning.
[A/CO:4]

[10 × 3 = 30]