

Code: 19IT4601E

**III B.Tech - II Semester – Regular Examinations – JUNE 2022****MACHINE LEARNING  
(INFORMATION TECHNOLOGY)**

Duration: 3 hours

Max. Marks: 70

- 
- Note: 1. This question paper contains two Parts A and B.  
2. Part-A contains 5 short answer questions. Each Question carries 2 Marks.  
3. Part-B contains 5 essay questions with an internal choice from each unit. Each question carries 12 marks.  
4. All parts of Question paper must be answered in one place.
- 

**PART – A**

1. a) Define Machine Learning. Give an example.
- b) Why do we perform ‘Normalization’ for the features of the target dataset?
- c) What is ‘Naïve’ assumption in Naïve Bayes Algorithm?
- d) Discuss the advantages and Disadvantages of Eager Learners.
- e) How to evaluate the rules generated by any rule based approach?

**PART – B****UNIT – I**

2. a) How Find-S algorithm identifies the most specific hypothesis to satisfy all positive instances? Justify with an example. 8 M
- b) Analyze Find-S algorithm to identify the setbacks to apply it in various use cases. 4 M

**OR**

3. Consider the following dataset identify the root node of the decision tree using Information Gain and Gini Impurity.

A	B	Class Label
T	F	+
T	T	+
T	T	+
T	F	-
T	T	+
F	F	-
F	F	-
F	F	-
T	T	-
T	F	-

12 M

### UNIT – II

4. a) Discuss various approaches to estimate confidence intervals from a given sample. 6 M
- b) How to compare two learning algorithms based on various evaluation metrics of confusion matrix? 6 M

OR

5. a) Discuss about Binomial Distribution and how it is used in Error Estimation? 6 M
- b) Assume that 15 % of the students in your class use Mac OS. If you select 20 students at random then
- i) What is the probability that exactly 4 of them are using Mac OS?
- ii) What is the probability that between 3 to 7 are of them are using Mac OS? 6 M

### UNIT-III

6. a) Discuss Brute force MAP learning with an example. 6 M  
b) How gradient search helps in maximizing the likelihood of the model predictions? 6 M

OR

7.

<i>Owns Home</i>	<i>Married</i>	<i>Gender</i>	<i>Employed</i>	<i>Credit rating</i>	<i>Class</i>
Yes	Yes	Male	Yes	A	II
No	No	Female	Yes	A	I
Yes	Yes	Female	Yes	B	III
Yes	No	Male	No	B	II
No	Yes	Female	Yes	B	III
No	No	Female	Yes	B	I
No	No	Male	No	B	II
Yes	No	Female	Yes	A	I
No	Yes	Female	Yes	A	III
Yes	Yes	Female	Yes	A	III

Apply naïve bayes and identify the class label for the following test row:

<i>Owns Home</i>	<i>Married</i>	<i>Gender</i>	<i>Employed</i>	<i>Credit rating</i>	<i>Class</i>
Yes	No	Male	Yes	A	?

12 M

### UNIT – IV

8. a) Find the Euclidian distance between points A(13, 11) and B(21, 31). 6 M  
b) Discuss the heuristics to decide optimal 'k' in KNN classification algorithm. 6 M

OR

9. a) Prove that points A(0, 4), B(6, 2), and C(9, 1) are collinear. Note that, to prove the given three points to be collinear, it is sufficient to prove that the sum of the distances between two pairs of points is equal to the distance between the third pair. 6 M
- b) Compare Case based reasoning and rule based reasoning. 6 M

**UNIT – V**

10. a) Discuss learn one rule function in sequential covering algorithm. 6 M
- b) Interpret the role of class prevalence in sequential covering algorithm. 6 M

OR

11. a) Discuss the FOIL algorithm with an example. 6 M
- b) How search is optimized in FOIL to find the most promising literal? 6 M