

U.G. 5th Semester Examination - 2021

COMPUTER SCIENCE

[HONOURS]

Discipline Specific Elective (DSE)

Course Code : COM.SC-H-DSE-L-502

Full Marks : 60

Time : 2½ Hours

The figures in the right-hand margin indicate marks.

Candidates are required to give their answers in their own words as far as practicable.

GROUP - A

1. Answer any **ten** questions : 2×10=20
- a) Define supervised learning.
 - b) Write down and explain the standard expression of multiple linear regression.
 - c) “We can get multiple local optimum solutions if we solve a linear regression problem by minimizing the sum of squared errors using gradient descent”- justify whether this statement is true or false.
 - d) Differentiate between classification and regression.
 - e) What is model underfitting?

[Turn Over]

- f) What is prior probability?
- g) Find the inverse of the following matrix
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$
- h) What is histogram plot?
- i) Define tensor.
- j) What is learning in the field of machine learning?
- k) Why does oscillation effect occur in gradient decent technique?
- l) State and explain the vector notation of linearly seperable functions.
- m) What is threshold logic unit?
- n) How does the bias change with respect to the change in number of training samples?
- o) What is max norm?

GROUP - B

Answer any **four** questions: 5×4=20

- 2. What are different ways of data visualization? Is data visualization necessary for machine learning? 3+2
- 3. “XNOR function is not linearly separable by a single decision boundary line.” Explain. 5
- 4. With the help of a neat diagram explain NAND function implementation using Perceptron. 5

5. What is regularization? Explain the need of it. 2+3=5
6. Define the expected squared loss function for regression problem $y = f(x)$ where $y \in R$ and $x \in R^d$. 5
7. With a suitable diagram explain a three input artificial neuron and chart all the input combinations. 5

GROUP - C

Answer any **two** questions: 10×2=20

8. What is feature selection in machine learning? With suitable examples classify and explain different feature selection techniques. 2+8
9. The sales of a company (in million dollars) for each year are shown in the table below.

$x(\text{year})$	$y(\text{sales})$
2005	12
2006	19
2007	29
2008	37
2009	45

- a) Find the least square regression line $y = ax + b$.
- b) Use the least square regression line as a model to estimate the sales of the company in 2012.

6+4

10. a) State and explain Bayes' theorem.
- b) Consider all the people of a city ABC all are tested for COVID19. Out of a population of 100000 only 2000 are actually infected but the test shows 5000 positive results.
- i) Create the confusion matrix for the test results of COVID19 for the city ABC.
- ii) Consider that you are a member of that city and you are tested positive. What is the probability that you have COVID 19? 4+(2+4)
11. Consider the following two class problem table:

Outlook	Temperature	Humidity	Wind	Play Tennis
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

Using Naive Bayes classifier find out if a person will play tennis when the attributes of a day are (Sunny, Cool, High, Strong). 10