

U.G. 5th Semester Examination - 2020

COMPUTER SCIENCE

[HONOURS]

Discipline Specific Elective (DSE)

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

Full Marks : 60

Time : $2\frac{1}{2}$ 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 : $2 \times 10 = 20$
- a) Briefly describe machine learning.
 - b) What do you mean by Posterior Probability?
 - 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) "Low dissimilarity between each member of a class and low similarity between members of different class are followed in supervised task of data mining"—exemplify the sentence.

- e) Differentiate between supervised and unsupervised learning.
- f) "In a linear least-squares regression problem, adding regularization can decrease the error of the solution on the training data." Whether this statement is true or false—justify your answer.
- g) Write down the Bayes' Theorem.
- h) What is the oscillation effect in Gradient Descent technique and what is the reason behind the oscillation effect?
- i) What does it mean by linearly separable problem? Give a suitable example.
- j) Differentiate between Intra cluster similarity and Inter cluster similarity with suitable example.
- k) Write down two drawbacks of single-layer perceptron network.
- l) How are artificial neurons different from biological neurons?
- m) How is residual sum of square method useful for fitting a linear model?
- n) Differentiate between Logistic Regression and Linear Regression.
- o) Why is it called the 'Naive' Bayes classifier?

GROUP - B

Answer any **four** of the following : 5×4=20

2. Write a stochastic gradient decent algorithm, for optimizing the following function

$$L(\theta) = \frac{1}{N} \sum_{n=1}^N l(x_n, y_n, \theta)$$

3. "XOR function is not linearly separable by a single decision boundary line." Explain.
4. Explain the concept of a Perceptron with a neat diagram and represent the Boolean functions of AND using Perceptron. 2+3=5
5. Explain the Minkowski distance and relate it with Manhattan and Euclidean distance. What is clustering? 3+2=5
6. Illustrate the candidate elimination algorithm with a suitable example.
7. Explain 'Naive' Bayes Classifier with an example.

GROUP - C

Answer any **two** of the following : 10×2=20

8. In the context of logistic regression, define the prediction and loss function. Show that if the class conditional densities, $p(c_k)$ $k=1,2$ are Gaussian, with equal co-variance, the posterior distribution $p(x)$ is a logistic. 5+5=10

9. What are the shortcomings of back propagation algorithm? Explain how can the shortcomings be overcome by Kohonen Self-Organizing Neural Network. 2+8=10

10. Define the expected squared loss function for regression problem $y = f(x)$ where $y \in \mathbb{R}$ and $x \in \mathbb{R}^d$. Derive the bias-variance decomposition of the expected squared loss function from first principles. 5+5=10

11. The values of independent variable x and dependent variable y is given below:

x	y
0	2
1	3
2	4
3	5
4	6

Calculate the least square regression line $y=ax+b$. Also estimate the value of y when x is 10. 4+6=10
