Sub Code: SDS401 **ROLL NO…….……..**

EVEN SEMESTER EXAMINATION, 2022 – 23

2nd yr. B. Tech- Computer Science & Engg

Scalable Data Science

Duration: 3:00 hrs **Max Marks: 100**

Note: - Attempt all questions. All Questions carry equal marks. In case of any ambiguity or missing data, the same may be assumed and state the assumption made in the answer.

Q 1.	Answer any four parts of the following. a) Discuss about Data Science? b) What do you understand by the term Normal Distribution? c) Explain overfitting with suitable diagram. d) Difference between convex and non-convex cost function; what does it mean when a cost function is non-convex? e) What do you understand by logistic regression? Explain with suitable diagram. f) Why is Naive Bayes referred to as Naive?	5x4=20
Q 2.	Answer any four parts of the following. a) What do you understand about Linear Regression? b) How will you explain linear regression to a non-tech person? Discuss. c) How is AUC different from ROC? d) What is a confusion matrix? e) Explain A/B testing. f) What is pickle module in Python?	5x4=20
Q 3.	Answer any two parts of the following. a) What is the difference between a box plot and a histogram? Discuss. b) Write short notes about kernel? Explain the kernel trick. c) You randomly draw a coin from 100 coins — 1 unfair coin (head-head), 99 fair coins (head-tail) and roll it 10 times. If the result is 10 heads, what is the probability that the coin is unfair?	10x2= 20
Q 4.	Answer any two parts of the following. a) How is Deep Learning different from Machine Learning? Discuss with example. b) What is meant by 'curse of dimensionality'? How can we solve it? c) Describe different regularization methods, such as L1 and L2 regularization?	10x2= 20
Q 5.	Answer any two parts of the following. a) If you are a data scientist, how will you collect the data. What will be your data acquisition and retention strategy? b) What is the Central Limit Theorem? Explain it. Why is it important? c) How can you tell if a given coin is biased? Make an unfair coin fair	10x2= 20

****