Roll No.	2	0	2	2	2	4	4	8	7	

NIST INSTITUTE OF SCIENCE & TECHNOLOGY
(Autonomous)

B. Tech 4th Semester (2022 Batch)				Branch(s)	CSE/IT/CST/ECS
Subject Code	22CS4PE02T / 22IT4PE02T	Subject Name		DATA SCIENCE FOR ENGINEERS	
Time	*90 min*	Exam	Mid Semester	Max. Marks	50
Examination Superintendent		Prof. ChittaranjanBiswal			
Name of the Instructor(s)		Prof. P. K. Jena, Dr. A. Panigrahy, Prof. C. Palai, Prof. L. Behera, Prof. S. K. Kar, Dr. K. Prasad			
Date of Examination		18/04/2024	Sitting	2nd	

***Answer Question No.1from* PART-I *which is compulsory, any four from PART-II and any one from* PART-III.**

The figures in the right hand margin indicate marks.

PART-I

(Answer all the questions)

Q1.		CO	Level	**Level-1:** Knowledge **Level-2:** Comprehension **Level-3:** Application **Level-4:** Analysis **Level-5:** Synthesis **Level -6:** Evaluation	2 X 5
	(a)	1	1	Write four application areas of Data Science.	
	(b)	1	1	What is data pre-processing? Discuss about two different data pre-processing techniques.	
	(c)	2	2	Justify why qualitative and quantitative data collection is required for analysis?	
	(d)	2	1	Write the skills required for a Data Science Engineer.	
	(e)	1	1	What is conditional probability?	

Scanned with OKEN Scanner

PART-II

(Answer Any Four questions out of six)

Q2.		CO	Level	Level-1: Knowledge Level-2: Comprehension Level-3: Application Level-4: Analysis Level-5: Synthesis Level -6: Evaluation	4 X 6
	(a)	1	2	How Data science is multidisciplinary? Explain with Venn diagram.	
	(b)	1	2	How the discrete random variables are different from continuous random variables explain with example.	
	(c)	1	2	Describe the structured, give examples of four level of structured data	
	(d)	2	2	What is sampling? Write different types Probability and Non-probability sampling methods.	
	(e)	1	2	Explain the stages of data science project. Discuss the documentation methods.	
	(f)	1	2	Discuss different types of learning methods with example.	

PART-III

(Answer Any One question out of two)

<table>
<tr><th></th><th></th><th>CO</th><th>Level</th><th>Level-1: Knowledge Level-2: Comprehension Level-3: Application Level-4: Analysis Level-5: Synthesis Level -6: Evaluation</th><th>1 X 16</th></tr>
<tr><td>Q3.</td><td>(a)</td><td>1</td><td>2</td><td>Write the different parameters used for model evaluation?</td><td></td></tr>
<tr><td></td><td>(b)</td><td>2</td><td>3</td><td>The final exam scores of 20 students are (in increasing order): 48, 54, 58, 64, 67, 71, 74, 76, 80, 82, 83, 85, 87, 88, 90, 92, 93, 94, 95, 97. Calculate the values of median, and IQR.</td><td></td></tr>
<tr><td>Q4.</td><td>(a)</td><td>2</td><td>2</td><td>Differentiate between the Bayesian approach and Frequentist approach used for data analysis.</td><td></td></tr>
<tr><td></td><td>(b)</td><td>2</td><td>3</td><td>Discuss different types of measures used to understand the central tendency of data. Calculate the mode for the following data.
<table>
<tr><th>Mark obtained</th><th>No. of students</th></tr>
<tr><td>20-30</td><td>5</td></tr>
<tr><td>30-40</td><td>16</td></tr>
<tr><td>40-50</td><td>8</td></tr>
<tr><td>50-60</td><td>10</td></tr>
<tr><td>60-70</td><td>4</td></tr>
</table></td><td></td></tr>
</table>

Scanned with OKEN Scanner