|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Roll No |  |  |  |  |  |  |  |  |  |  |  |

PRESIDENCY UNIVERSITY BENGALURU

SCHOOL INFORMATION SCIENCE

MAKEUP EXAMINATION – JULY 2024

|  |  |
| --- | --- |
| **Semester :I** | **Date :08.07.2024** |
| **Course Code :CSA1003** | **Time :9.30 am to 12.30pm** |
| **Course Name :Fundamentals of Data Science** | **Max Marks :100** |
| **Program :BCA** | **Weightage :50%** |

**Instructions:**

1. *Read all questions carefully and answer accordingly.*
2. *Question paper consists of 3 parts.*
3. *Scientific and non-programmable calculator are permitted.*
4. *Do not write any information on the question paper other than Roll Number.*

|  |  |  |  |
| --- | --- | --- | --- |
| **PART A** | | | |
| **ANSWER ANY 5 QUESTIONS 5Q X 2M=10M** | | | |
| 1 | Define Data Science | (CO 1) | [Knowledge] |
|  | | | |
| 2 | What is meant by Data Mining? | (CO 1) | [Knowledge] |
|  | | | |
| 3 | What is the use of Pie Chart? | (CO 2) | [Knowledge] |
|  | | | |
| 4 | Identify the mean of 1st 10 Odd integers | (CO 2) | [Knowledge] |
|  | | | |
| 5 | Define Correlation Coefficient and write the formulae to find r. | (CO 3) | [Knowledge] |
|  | | | |
| 6 | What is the use of shuffle and sort technique in map reduce process | (CO 4) | [Knowledge] |
|  |  |  |  |
| 7 | Write the formulae to find Euclidean Distance | (CO 4) | [Knowledge] |
|  | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **PART B** | | | |
| **ANSWER ANY 5 QUESTIONS 5Q X 10M=50M** | | | |
| 8 | Illustrate Key Aspects of Data Science Process | (CO 1) | [Comprehension] |
|  | | | |
| 9 | Explain statistical descriptions of Data[Mean, Median, Mode, Variance, Standard Deviation] | (CO 1) | [Comprehension] |
|  | | | |
| 10 | Demonstrate Different types of Charts with Example. | (CO 2) | [Comprehension] |
|  | | | |
| 11 | 1. What is meant by Percentile? Give Example 2. A garden contains 39 plants. The following plants were chosen at random, and their heights were recorded in cm: 38, 51, 46, 79, and 57. Calculate their heights’ standard deviation. | (CO 2) | [Comprehension] |
|  | | | |
| 12 | |  | | --- | | Solve the prediction to obtain mean, variance and standard deviation for the following observation and compare which one is best. | | x                                     y | | 30                                    40 | | 60                                    50 | | 90                                    70 | | 85                                    61 | | 72                                    87 | | 63                                   79 | | 45                                  12 | | 22                                  14 | | 13.5                              16 | | 14                                  18 | | (CO 3) | [Comprehension] |
|  | | | |
| 13 | Compare Supervised Learning with Unsupervised Learning | (CO 4) | [Comprehension] |
|  |  |  |  |
| 14 | Explain KNN Algorithm in detail. | (CO 4) | [Comprehension] |
|  | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **PART C** | | | |
| **ANSWER ANY 2 QUESTIONS 2Q X 20M=40M** | | | |
| 14 | Explain different types of facets[Types of V’s] of data | (CO1) | [Application] |
|  | | | |
| 15 | Let us assume that you start up a company that has developed a drug that is supposed to increase IQ. You know that the standard deviation of IQ in the general population is 15. You test your drug on 36 patients and obtain a mean IQ of 97.65. Using an alpha value of 0.05, is this IQ signifcantly different than the population mean of 100?  Z Area between mean and Z Area beyond Z  …. ……. ………….  1.94 0.4738 0.0262  1.95 0.4744 0.0256  1.96 0.475 0.025  1.97 0.4756 0.0244  1.98 0.4761 0.0239  …… ……….. ……….    Z Area between mean and Z Area beyond z  …. ….. …..  -0.96 -0.3315 0.8315  -0.95 -0.3289 0.8289  -0.94 -0.3264 0.8264  -0.93 -0.3238 0.8238  -0.92 -0.3212 0.8212 | (CO2) | [Application] |
|  | | | |
| 16 | Perform KNN Algorithm on following data. Predict the class for New Entry. [Where K = 3]     | **BRIGHTNESS** | **SATURATION** | **CLASS** | | --- | --- | --- | | 20 | 35 | ? |  | **BRIGHTNESS** | **SATURATION** | **CLASS** | | --- | --- | --- | | 40 | 20 | Red | | 50 | 50 | Blue | | 60 | 90 | Blue | | 10 | 25 | Red | | 70 | 70 | Blue | | 60 | 10 | Red | | 25 | 80 | Blue | | (CO4) | [Application] |
|  | | | |
|  | | | |