SRI KRISHNA INSTITUTE OF TECHNOLOGY

(Accredited by NAAC, Approved by A.I.C.T.E. New Delhi, Recognised by Govt. of Karnataka & Affiliated to V.T U., Belagavi)
#29, Chimney Hills, Hesaraghatta Main Road, Chikkabanavara Post, Bengaluru- 560090

Department of Artificial Intelligence and Machine Learning

Subject Name: Data Science & Its Applications
Subject Code: 21AD62

SEM: 6th **DIV: A**

Faculty: Prof. Manzoor Ahmed

Module-1 Question Bank

<table>
<tr><th>SL#</th><th>Question</th><th>CO</th><th>Level</th><th>Marks</th></tr>
<tr><td>1.</td><td>What is Data Science? With example explain the role of a data scientist.</td><td>CO1</td><td>L2</td><td>8</td></tr>
<tr><td>2.</td><td>Define data visualization? With suitable example explain the Line and Bar Graph.</td><td>CO1</td><td>L2</td><td>8</td></tr>
<tr><td>3.</td><td>Explain with example the matplotlib library in python.</td><td>CO1</td><td>L2</td><td>8</td></tr>
<tr><td>4.</td><td>For the given data draw the Line chart with legend using plt.plot.
Variance = [1,2,4,8,16,32,64,128,256]
Bias_sqared = [256,128,64,32,16,8,4,2,1]
total_error = [x + y for x, y in zip(variance, bias_squared)]
xs = [i for i, _ in enumerate(variance)]</td><td>CO1</td><td>L2</td><td>6</td></tr>
<tr><td>5.</td><td>Draw the scatter plot to illustrate the relationship between number of friends and the number of minutes spend on every day.
friends = [70, 65, 72, 63, 71, 64, 60, 64, 67]
minutes = [175, 170, 205, 120, 220, 130, 105, 145, 190]
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i']</td><td>CO1</td><td>L2</td><td>6</td></tr>
<tr><td>6.</td><td>Define vectors. Explain different operations performed on vectors with example using python code (without using built-in functions).</td><td>CO1</td><td>L2</td><td>10</td></tr>
<tr><td>7.</td><td>For the given data find out the distance and magnitude of a vector.
List = [[1,2], [3,4], [5,6], [7,8]]</td><td>CO1</td><td>L2</td><td>8</td></tr>
<tr><td>8.</td><td>Define matrices. Explain different operations performed on matrices with example using python code (without using built-in functions).</td><td>CO1</td><td>L2</td><td>8</td></tr>
<tr><td>9.</td><td>Write a function to create a matrix given its shape and a function for generating its elements. Then use the function to generate to 5 x 5 identity matrix.</td><td>CO1</td><td>L2</td><td>6</td></tr>
<tr><td>10.</td><td>Explain standard deviation and interquartile range and write python code to compute standard deviation and interquartile range.</td><td>CO1</td><td>L2</td><td>6</td></tr>
<tr><td>11.</td><td>Consider the following employees data:
<table>
<tr><td>Name</td><td>John</td><td>Mike</td><td>Sara</td><td>Tom</td><td>Alex</td><td>Nina</td><td>David</td></tr>
<tr><td>Dept.</td><td>IT</td><td>Marketing</td><td>HR</td><td>IT</td><td>Finance</td><td>IT</td><td>Marketing</td></tr>
<tr><td>Salary</td><td>50000</td><td>60000</td><td>45000</td><td>55000</td><td>60000</td><td>52000</td><td>58000</td></tr>
<tr><td>Age</td><td>25</td><td>Nan</td><td>30</td><td>28</td><td>Nan</td><td>32</td><td>Nan</td></tr>
<tr><td>Hire_Date</td><td>01-01-2015</td><td>02-01-2016</td><td>01-01-2017</td><td>03-01-208</td><td>01-04-2019</td><td>01-01-2020</td><td>02-01-2021</td></tr>
</table></td><td>CO1</td><td>L2</td><td>8</td></tr>
</table>

	1. Find the standard deviation of salary of employees in each dept. of a company and identify the department with the highest standard deviation. 2. Find the mean and median salary of employees in each department of the company. Find out the above using python code without using built-in functions.			
12.	Explain dispersion and correlation with a suitable example.	**CO1**	**L2**	**8**
13.	Define probability. Explain with an example Dependent and Independent events.	**CO1**	**L2**	**8**
14.	Discuss Conditional probability with an example in detail.	**CO1**	**L2**	**8**
15.	Find the probability of the given events? a. A single letter is selected at random from the word 'MACHINE LEARNING'. The probability that it is a consonant. b. The probability of rolling 2 dice to get a sum of 4 or 7. c. Lottery tokens are numbered from 1to 25. What is the probability that a token drawn is multiple of 5 or 7? d. The probability of getting a face in 52 cards.	**CO1**	**L2**	**8**
16.	A car manufacturer purchases car batteries from two different suppliers. Supplier X provides 55% of the batteries and supplier Y provides the rest. If 5% of all batteries from supplier X are defective and 4% of the batteries from supplier Y are defective. Determine the probability that a randomly selected battery is defective.	**CO1**	**L2**	**6**
17.	A bag contains 5 red and 5 black balls. A ball is drawn at random, its colour is noted, and again the ball is returned to the bag. Also, 2 additional balls of the colour drawn are put in the bag. After that, the ball is drawn at random from the bag. What is the probability that the second ball drawn from the bag is red?	**CO1**	**L2**	**6**
18.	Of the students in the college, 60% of the students reside in the hostel and 40% of the students are day scholars. Previous year results report that 30% of all students who stay in the hostel scored A Grade and 20% of day scholars scored A grade. At the end of the year, one student is chosen at random and found that he/she has an A grade. What is the probability that the student is a hosteler?	**CO1**	**L2**	**6**
19.	From the pack of 52 cards, one card is lost. From the remaining cards of a pack, two cards are drawn and both are found to be diamond cards. What is the probability that the lost card is a diamond?	**CO1**	**L2**	**6**
20.	State and prove the Bayes Theorem using Total Probability and Conditional Probability.	**CO1**	**L2**	**6**
21.	State and illustrate the Central Limit Theorem with a python code using a suitable example.	**CO1**	**L2**	**6**
22.	Illustrate the difference between Normal and Binomial Distribution with a suitable example.	**CO1**	**L3**	**8**

Faculty Signature