MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent unit of MAHE, Manipal)

DEPARTMENT OF MECHATRONICS ENGINEERING

II SEMESTER M.TECH. (INDUSTRIAL AUTOMATION AND ROBOTICS)
MAKE UP EXAMINATION- 24th JUNE 2024

SUBJECT: Data Analytics for Automation Subject Code: MTE 5404

Time: 3 Hour **Exam time: 9:30-12:30 am** **MAX. MARKS: 50**

Instructions to Candidates:

- Answer **ALL** the questions.
- Assume suitablt any data that is not provided.

Q. No.		M	CO	PO	LO	BL
1.	Explain th TF-IDF feature extractors in case of Text Processing.	**2**	**1**	1,3	1,2	3
2	Explain the concept of a binomial distribution and its characteristics. Provide an example to illustrate the binomial distribution.	**4**	**3**	5	4	3
3.	Differentiate between probability and non-probability sampling techniques. Discuss the advantages and disadvantages of each type of sampling technique	**4**	**2**	1,3	1,2	5
4.	A market research firm wants to conduct a survey to determine the preference for a new product in a city with a population of 50,000. If the desired level of precision is ±5% and the estimated proportion of the population interested in the new product is 0.6, calculate the required sample size using Cochran's formula. Assume a 95% confidence level.	**2**	**2**	1,3	1,2	3
5.	Analyze the concept of stratified sampling and its applications. Explain why stratified sampling is more accurate when dealing with skewed populations	**4**	**2**	1,3	1,2	4
6.	Evaluate the suitability of different non-probability sampling techniques, such as convenience sampling, purposive sampling, and snowball sampling, for various research scenarios. Provide examples to support your evaluation	**4**	**1**	1,3	1,2	4
7.	Describe the significance of Schema in a database.	**2**	**2**	1,3	1,2	4
8.	Explain the concept of hypothesis testing and differentiate between Type I and Type II errors. Discuss the importance of the level of significance in hypothesis testing	**4**	**2**	1,3	1,2	3

9.	Compare and contrast the assumptions and conditions for applying the Z-test and t-test in hypothesis testing. Provide examples to illustrate the use of each test.	**4**	**3**	5	4	4
10.	A company claims that the average lifespan of their batteries is 3 years. A random sample of 20 batteries showed a mean lifespan of 2.8 years with a standard deviation of 0.6 years. At a 5% significance level, test the claim of the company using an appropriate hypothesis test.	**2**	**3**	5	4	4
11.	Analyze the concept of p-values in hypothesis testing. Discuss the criteria for rejecting or failing to reject the null hypothesis based on p-values and the significance level.	**4**	**4**	1,3	1,2, 3	3
12.	Evaluate the suitability of the chi-square test as a non-parametric hypothesis testing technique. Explain the steps involved in conducting a chi-square test and interpret the results.	**4**	**4**	1,3	1,2, 3	4
13.	Enlist four characteritistics of a good sample.	**2**	**3**	5	4	3
14.	Explain the various types of NoSQL Databases that exist in the manner of their data being stored.	**4**	**4**	1,3	1,2, 3	5
15	Enlist the problems of traditional file systems that are overcome by Data based management systems	**4**	**5**	**3,5**	1,2, 3,17	5