

**VIT**

Vellore Institute of Technology

(Approved by the Council for Higher Education, India, 2014)

School of Computer Science and Engineering**Fall Semester 2023-24****Continuous Assessment Test – II****SLOT: B1+TB1****Programme Name & Branch: M.Tech (Integrated) & Data Science****Course Name & Code: MDI3002 – Foundations of Data Science****Class Number (s): VL2023240104666, VL2023240104667****Faculty Name (s): Dr. S. Malini, Dr. P. Kauser Ahmed****Exam Duration: 90 Min.****Maximum Marks: 50**

Q. No.	Question	Max Marks										
1.	Assume that, you are working in a final year project in data science field, you are going to do the project by your own. What are the various steps/phases needed for your project completion in a successful manner. Explain the phases in detail with proper example and diagrams. (10)	10										
2.	Use Newton's Method to determine x_3 (upto 5 precisions) for $f(x)=x^3-7x^2+8x-3$, if $x_0=5$	10										
3.	a) Find the critical points of the function $f(x,y)=-x^2-5y^2+8xy-10y-13$. Identify the maximized or minimized points. Check the presence of any saddle point. (7 Marks) b) Differentiate between the batch gradient descent and stochastic gradient descent (3 Marks)	10										
4.	a) Assume that, you are working in a company. You have received the complete details of various applicants. In which way, you are going to short-list the candidates and Explain the various ways to verify the quality of the data received with one simple example. (7 Marks) b) Identify the normalized values for the following set of values <table><thead><tr><th>Employee Name</th><th>Years of Experience</th></tr></thead><tbody><tr><td>ABC</td><td>8</td></tr><tr><td>XYZ</td><td>20</td></tr><tr><td>PQR</td><td>10</td></tr><tr><td>MNO</td><td>15</td></tr></tbody></table> Using i) min-max normalization ii) z-score normalization (3 marks)	Employee Name	Years of Experience	ABC	8	XYZ	20	PQR	10	MNO	15	10
Employee Name	Years of Experience											
ABC	8											
XYZ	20											
PQR	10											
MNO	15											
5.	Assume that, you are working in a Bioinformatics based IT concern. You are going to present your work to the end users and to your higher officials. You need to show the demonstration and visualization of your work. You are insisted that your work should be more attractive. Explain the various steps followed to achieve the proper presentation (visualization) using one simple dataset. Explain the various steps from the way of fixing the dataset to present the dataset in which form to others using the programming language (R/Python).	10										