



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2025-2026

Programme Name & Branch :B.Tech CSE (All Branches)
Course Code and Course Name :BCSE318L Data Privacy
Faculty Name(s) :M.Diviya, GLADYS GNANA KIRUBA B,NAVEE
Class Number(s) : VL2025260505869,2063,2228
Date of Examination :30.1.2026
Exam Duration :90 minutes
Maximum Marks :100

General instruction(s):

- Answer All Questions
- CO1: Characterize basic rules, principles for protecting privacy and pers identifiable information.
- CO2 : Formulate data that supports useful statistical inference while min disclosure of sensitive information.

Q. No	Question	Mod
1.	The City of Neotropolis launches a "Smart Health Corridor" project. They collect anonymized mobility and health service access data from participating citizens' apps (with explicit consent) to improve public health planning. How can privacy preserving tools be used to mediate the different, and sometimes competing, needs of various stakeholders. Explain in detail the various stake holders involved in the process.	1
2.	FinGlobal Inc. is building a new platform to analyze cross-border transaction patterns. They must protect sensitive customer data while enabling fraud detection analytics. Before choosing tools, their security team must answer a foundational question. Based on the analytical needs explain the methods of protecting the data applied purposefully rather than arbitrarily?	1
3.	The given dataset consists of both continuous and categorical attributes along with a sensitive attribute. The data must be anonymized without adding noise or modifying original values. Based on the nature of the attributes and the privacy requirement, identify and justify the most appropriate non perturbative methods that can be applied and how?	2

ID	Age	Annual Income (₹)	Gender	Education	Region	Disease
1	20	180000	F	UG	South	Flu
2	27	240000	M	UG	North	Diabetes
3	31	320000	F	PG	South	Flu
4	36	410000	M	PG	East	Hypertension
5	41	480000	M	PG	West	Diabetes
6	45	610000	F	PhD	South	Cancer
7	50	600000	M	DSc	North	Hypertension



VIT

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.N

SLOT:

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2025-2026

8	58	390000	F	UG	East	Flu
9	29	280000	M	UG	West	Diabetes
10	55	450000	M	PG	South	Cancer

4. Consider the following dataset

ID	Age	Monthly_Income (₹)	Gender	Education	Risk_Score
1	23	18000	F	UG	0.42
2	35	42000	M	PG	0.31
3	29	32000	F	UG	0.38
4	41	55000	M	PhD	0.27
5	52	60000	M	PG	0.22
6	26	25000	F	UG	0.40
7	33	48000	M	PG	0.34
8	45	52000	F	PhD	0.29
9	38	45000	M	UG	0.36
10	28	30000	F	PG	0.39

Based on the given data

(a) Apply additive noise to the data using a zero-mean distribution with $\sigma_{\text{Age}} = 2$ and $\sigma_{\text{Income}} = 3000$, restricting noise values to 2σ for both Age and for Monthly Income.

(b) Analyse the changes imparted in the original data by using the following transformation function to age and income attributes.

$T(x) = Ax + b$, Where $A = 1.2$ and $b = 3$

5. Explain the concept of information loss in data privacy and compute it using Mean Square Error (MSE), Mean Absolute Error (MAE), Mean Variation, and Direct comparison of categorical values between the original dataset and the masked dataset generated in Question 4.
