BCSE318L	DATA PRIVACY	L	T	P	C
		3	0	0	3
Pre-requisite	**NIL**	**Syllabus version**			
		1.0			

Course Objectives

1. To impart the need of data privacy.
2. To categorize the statistical and computational techniques required to share data, with a primary focus on the social, and health sciences.
3. To formulate architectural, algorithmic, and technological foundations for the maintaining the data privacy.

Course Outcomes

After completion of this course, the student shall be able to:

1. Characterize basic rules, principles for protecting privacy and personally identifiable information.
2. Formulate data that supports useful statistical inference while minimizing the disclosure of sensitive information.
3. Identify the list of threats on the various types of anonymized data.
4. Classify and analyze the methods of test data generation with Privacy and utility.

Module:1	**Data privacy and Importance**	**5 hours**

Need for Sharing Data - Methods of Protecting Data - Importance of Balancing Data Privacy and Utility – Disclosure - Tabular Data - Micro data - Approaches to Statistical disclosure control – Ethics – principles - guidelines and regulations.

Module:2	**Microdata**	**7 hours**

Disclosure - Disclosure risk - Estimating re-identification risk - Non-Perturbative Micro data masking - Perturbative Micro data masking - Information loss in Micro data.

Module:3	**Static Data Anonymization on Multidimensional Data**	**7 hours**

Privacy – Preserving Methods - Classification of Data in a Multidimensional Dataset - Group-based Anonymization: k-Anonymity, l-Diversity, t-Closeness.

Module:4	**Anonymization on Complex Data Structures**	**8 hours**

Privacy-Preserving Graph Data, Privacy-Preserving Time Series Data, Time Series Data Protection Methods, Privacy Preservation of Longitudinal Data, Privacy Preservation of Transaction Data.

Module:5	**Threats to Anonymized Data**	**6 hours**

Threats to Anonymized Data, Threats to Data Structures, Threats by Anonymization Techniques: Randomization, k-Anonymization, l-Diversity, t-Closeness.

Module:6	**Dynamic Data Protection**	**5 hours**

Dynamic Data Protection: Tokenization, Understanding Tokenization, Use Cases for Dynamic Data Protection, Benefits of Tokenization Compared to Other Methods, Components for Tokenization.

Module:7	**Privacy-Preserving Test Data Generation and Privacy Regulations**	**5 hours**

Test Data Fundamentals - Insufficiencies of Anonymized Test Data. Privacy regulations: UK Data Protection Act, Swiss Data Protection Act, HIPPA, General Data Protection Regulation.

Module:8	**Contemporary Issues**	**2 hours**
	Total Lecture hours:	**45 hours**

Text Book

1. NatarajVenkataramanan, AshwinShriram, Data Privacy: Principles and Practice, 2016, 1st Edition, Taylor & Francis. (ISBN No.: 978-1-49-872104-2), United Kingdom.

Reference Books			
1.	AncoHundepool, Josep Domingo-Ferrer, Luisa Franconi, Sarah Giessing, Eric Schulte Nordholt, Keith Spicer, Peter-Paul de Wolf, Statistical Disclosure Control, 2012, 1st Edition Wiley. (ISBN No.: 978-1-11-997815-2), United States.		
2.	George T. Duncan. Mark Elliot, Juan-Jose Salazar-GonZalez, Statistical Confidentiality: Principle and Practice. 2011, 1st Edition, Springer. (ISBN No.: 978-1-44-197801-1).		
Mode of Evaluation: CAT / written assignment / Quiz / FAT			
Recommended by Board of Studies	04-03-2022		
Approved by Academic Council	No.65	Date	17-03-2022