3. Apply different outlier-detection methods on a noisy dataset and compare their effectiveness in terms of outliers reported.
4. Compute similarity between two documents after required document preparation.
5. Considering a time-stamped data (sales data/weather data), compare the aggregate values visually using different moving windows function.
6. Write a program to find the latent topics in a document using any topic modeling method and display top 5 terms that contribute to each topic along with their strength. Also, visualize the distribution of terms contributing to the topics.

Project: Students are encouraged to work on a good dataset in consultation with their faculty and apply the concepts learned in the course.

DISCIPLINE SPECIFIC ELECTIVE COURSE: Data Privacy

Credit distribution, Eligibility and Pre-requisites of the Course

Course title & Code	**Credits**	**Credit distribution of the course**			**Eligibility criteria**	**Pre-requisite of the course**
		Lecture	**Tutorial**	**Practical/ Practice**		
Data Privacy	**4**	**3**	**0**	**1**	Pass in Class XII	NIL

Learning Objectives

This course aims to provide students with an ability to identify privacy related aspects of data uses(including attacks on data privacy), evaluate proposed mechanisms for privacy protection and relate to ethical issues related to data privacy.

On successful completion of the course, students will be able to:

- Understand the basic principles of data privacy and the implications of data breaches.
- Identify and evaluate different methods of protecting sensitive data.
- Explain the role of privacy regulations in safeguarding personal information.
- Implement basic cryptographic techniques to secure data.
- Apply data anonymization techniques to protect personal information.
- Analyze the ethical considerations in data privacy

SYLLABUS OF DSE

Unit 1 (10 hours)

Introduction to Data Privacy and Privacy Regulations: Notion of data privacy, Historical context of data privacy, Types of sensitive data, Privacy laws and regulations.

Unit 2 (15 hours)

Data Privacy Attacks, Cryptography and Data Protection: Type of Attacks/ Data Breaches on Data Privacy, Impact of Data Breaches / Attacks, Introduction to cryptography, Symmetric and asymmetric encryption, Hashing and digital signatures.

Unit 3 (10 hours)

Data Collection, Use and Reuse: Harms Associated with Data collections, use and reuse, Introduction to data anonymization, Data Anonymization Techniques for anonymizing data, Challenges in anonymizing data

Unit 4 (10 hours)

Ethical considerations in Data Privacy: Privacy and Surveillance, Ethics of Data Collection and Use, Bias and discrimination in data analysis

Essential/recommended readings

1. Ronald Leenes, Rosamunde van Brakel, and Serge Gutwirth: *Data Protection and Privacy: The Age of Intelligent Machines*, Hart Publishing, 2017.
2. Naavi: *Personal Data Protection Act of India (PDPA 2020)*: Be Aware, Be Ready and Be Compliant, Notion Press, 2020.
3. Ravinder Kumar Gaurav Goyal, *The Right to Privacy in India: Concept and Evolution*, Publisher: Lightning Source, 2016.

Additional References

1. https://onlinecourses.nptel.ac.in/noc22_cs37/preview
2. https://www.coursera.org/learn/northeastern-data-privacy/home/info

Suggested Practical List : (30 Hours)

Practical exercises such as

1. Data Privacy Audit: Students can conduct a data privacy audit of a company or organization to identify potential vulnerabilities and risks in their data privacy practices.

2. Privacy Impact Assessment: Students can conduct a privacy impact assessment (PIA) of a new technology or system to identify potential privacy risks and develop strategies to mitigate them.
3. Regulation Compliance: Students can explore the requirements of the Data Protection Regulations and develop a plan for ensuring compliance with the regulation.
4. Cryptography: Students can learn about different cryptographic techniques and tools, such as encryption, hashing, and digital signatures, and implement them in practice.
5. Anonymization Techniques: Students can learn about data anonymization techniques, such as k-anonymity, differential privacy, and data masking, and apply them to a real-world dataset.
6. Privacy Policy Analysis: Students can analyze the privacy policies of different companies and identify gaps or areas for improvement.
7. Privacy-Enhancing Technologies: Students can explore privacy-enhancing technologies (PETs), such as virtual private networks (VPNs), Tor, and secure messaging apps, and evaluate their effectiveness in protecting privacy.
8. Privacy Breach Response Plan: Students can develop a privacy breach response plan for a company or organization, including steps to take in the event of a data breach and strategies for communicating with affected parties.
9. Ethical Considerations: Students can explore ethical considerations in data privacy, such as the balance between privacy and security, the impact of data collection and analysis on marginalized communities, and the role of data ethics in technology development.
10. Case Studies: Students can analyze case studies of privacy breaches or successful privacy protection strategies, and identify key lessons and takeaways.

DISCIPLINE SPECIFIC ELECTIVE COURSE: Unix Network Programming

Credit distribution, Eligibility and Pre-requisites of the Course

Course title & Code	**Credits**	**Credit distribution of the course**			**Eligibility criteria**	
		Lecture	**Tutorial**	**Practical/ Practice**		
Unix Network Programming	**4**	**3**	**0**	**1**	Pass in Class XII	DSC 04 Object Oriented Programming with C++/ GE 1a Programming using C++ / GE1b Programming with Python/ DSC 01 Programming using Python/ GE 3b: Java Programming