Program Elective II

COURSE TITLE: Recommender Systems

Course Code:				**CSE-241051EDS**		**Examination Scheme**			
Total number of Lecture Hours: 56						**External**	**80**		
						Internal	**20**		
Lecture (L):	**4**	**Practicals (P):**	**4**	**Tutorial (T):**	**-**	**Total Credits**	**6**		

Course Objectives

- Students will be able to explain key concepts in information retrieval, including retrieval models, search techniques, and the role of relevance feedback and user profiles in enhancing search effectiveness.
- Students will develop skills in designing and implementing content-based filtering systems, including feature extraction, item profiling, and user profile learning methods.
- Students will evaluate different collaborative filtering approaches, including user-based and item-based methods, as well as hybrid models, while understanding the challenges and potential vulnerabilities of these systems.
- Students will learn to assess the performance of various recommender systems using established evaluation metrics, and classify systems into categories such as personalized web search, knowledge-based, and social tagging recommender systems.

Course Content	TEACHING HOURS
UNIT 1: Foundations of Information Retrieval and Recommender Systems	**14 Hrs.**
Introduction: Overview of Information Retrieval, Retrieval Models, Search and Filtering Techniques: Relevance Feedback, User Profiles, Recommender system functions, Matrix operations, covariance matrices, Understanding ratings, Applications of recommendation systems, Issues with recommender system.	
UNIT 2: Content-Based Filtering: Techniques and Applications	**14 Hrs.**
Content-based Filtering: High level architecture of content-based systems, Advantages and drawbacks of content-based filtering, Item profiles, Discovering features of documents, pre-processing and feature extraction, Obtaining item features from tags, Methods for learning user profiles, Similarity based retrieval, Classification algorithms.	
UNIT 3: Collaborative Filtering: Strategies and Hybrid Approaches	**14 Hrs.**
Collaborative Filtering: User-based recommendation, Item-based recommendation, Model based approaches, Matrix factorization, Attacks on collaborative recommender systems. Hybrid approaches: Opportunities for hybridization, Monolithic hybridization design: Feature combination, Feature augmentation, Parallelized hybridization design: Weighted, Switching, Mixed, Pipelined hybridization design: Cascade Meta-level, Limitations of hybridization strategies	

UNIT 4: Evaluating and Classifying Recommender Systems	**14 Hrs.**
Evaluating Recommender System: Introduction, General properties of evaluation research, Evaluation designs: Accuracy, Coverage, confidence, novelty, diversity, scalability, serendipity, Evaluation on historical datasets, Offline evaluations. Types of Recommender Systems: Recommender systems in personalized web search, knowledge-based recommender system, Social tagging recommender systems, Trust-centric recommendations, Group recommender systems.	

Textbooks

Jannach D., Zanker M. and FelFering A., Recommender Systems: An Introduction, Cambridge University Press (2011), 1st ed.

Reference Books

1. Charu C. Aggarwal, Recommender Systems: The Textbook, Springer (2016), 1st ed.
2. Ricci F., Rokach L., Shapira D., Kantor B.P., Recommender Systems Handbook, Springer(2011), 1st ed.
3. Manouselis N., Drachsler H., Verbert K., Duval E., Recommender Systems For Learning, Springer (2013), 1st ed.

COURSE OUTCOMES (CO):

CO1: Critical Analysis: Students will be able to critically analyze different information retrieval and recommender system models, identifying their strengths and weaknesses in various application contexts.

CO2: Technical Proficiency: Students will demonstrate proficiency in implementing content-based and collaborative filtering techniques, including feature extraction and similarity-based retrieval, using relevant programming tools and frameworks.

CO3: Hybrid System Design: Students will design and evaluate hybrid recommender systems, applying appropriate methodologies for combining different filtering techniques to enhance recommendation quality.

CO4: Evaluation Competence: Students will proficiently evaluate recommender systems using established metrics, interpreting the results to make informed decisions about system improvements and identifying suitable applications for different types of recommender systems.

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	1	2	3	4	5	6	7	8	9	10	11	12
1	3	3	2	2	2	2	1	1	1	1	1	1
2	2	2	3	2	3	1	1	1	2	2	1	1
3	2	2	3	2	2	3	3	2	2	2	1	1
4	1	2	2	3	2	1	1	1	1	3	1	2

Recommender Systems (LAB)

The lab should cover the following topics:

1 Introduction to Information Retrieval
2 Retrieval Models
3 Search Techniques and Relevance Feedback
4 User Profiles and Recommender System Functions
5 Content-Based Filtering Techniques
6 Similarity-Based Retrieval
7 Collaborative Filtering: User-Based and Item-Based
8 Matrix Factorization Techniques
9 Hybrid Recommender Systems
10 Evaluating Recommender Systems
11 Advanced Evaluation Techniques
12 Types of Recommender Systems

Note: Additional Resources:

- **Suggested software tools: Python, Jupyter Notebooks, scikit-learn, Surprise library for recommender systems, and evaluation libraries.**
- **Recommended datasets: MovieLens, Goodreads, or any domain-specific datasets relevant to the course.**

This structure provides a thorough hands-on experience across all aspects of the syllabus, ensuring that students develop practical skills alongside theoretical knowledge.

COURSE TITLE: Machine Learning								
Course Code:				**CSE-241052EDS**		**Examination Scheme**		
Total number of Lecture Hours: 56						**External**	**80**	
						Internal	**20**	
Lecture (L):	**4**	**Practicals(P):**	4	**Tutorial (T):**	-	**Total Credits**	6	

Course Objectives

- To learn the concept of how to learn patterns and concepts from data without being explicitly programmed in various nodes.
- To design and analyse various machine learning algorithms and techniques with a modern outlook focusing on recent advances.
- Explore supervised and unsupervised learning paradigms of machine learning.
- To explore Deep learning technique and various feature extraction strategies.

Course Content	TEACHING HOURS
UNIT 1: Supervised Learning (Regression/Classification)	**14-Hrs**
• K Nearest-Neighbor Classifier • Decision Trees (ID3, SAFARI). • Linear Regression, Logistic Regression • Support Vector Machines, Nonlinearity and Kernel Methods • Beyond Binary Classification: Multi-class Outputs.	
UNIT 2: Unsupervised Learning	**14-Hrs**
• Distance-based methods • Clustering: K-means • Dimensionality Reduction: PCA • Generative Models	
UNIT 3:	**14-Hrs**
• Ensemble Methods • Boosting • Bagging • Random Forests	
UNIT 4:	**14-Hrs**
• Semi-supervised Learning, • Active Learning, • Reinforcement Learning, • Introduction to Bayesian Learning	

Textbooks

1. Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007.

Reference Books

1. Kevin Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012
2. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Springer 2009 (freely available online)
3. Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007

COURSE OUTCOMES (CO):

After completion of course, students would be able to:

CO1: Extract features that can be used for a particular machine learning approach in various applications.

CO2: To compare and contrast pros and cons of various machine learning techniques and to get an insight of when to apply a particular machine learning approach.

CO3: To mathematically analyze various machine learning approaches and paradigms.

CO4: To discover Deep learning method and different feature extraction approaches.

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	3	2	3	3	2	1	1	-	1	2	1	2
2	2	-	3	-	2	2	1	-	1	2	2	3
3	3	3	2	3	2	1	1	-	1	3	2	3
4	2	2	3	3	2	1	1	3	3	2	1	2

Lab Manual-Machine learning

Students are encouraged to perform hands-on lab exercises on the following topics:

- Implementation of Distance-Based Methods (e.g., K-Nearest Neighbors).
- Building Decision Trees for Classification and Regression (ID3).
- Linear Regression and Logistic Regression using Python/Scikit-learn.
- Training Support Vector Machines with Linear and Nonlinear Kernels.
- Implementing Multi-class Classification Techniques.
- K-means Clustering on real-world datasets (e.g., customer segmentation).
- Dimensionality Reduction using PCA and visualizing results in 2D/3D.
- Implementing Generative Models for unsupervised learning.
- Implementing Bagging and Random Forests for robust classification.
- Experimenting with Boosting Techniques (e.g., AdaBoost, Gradient Boosting).
- Hyperparameter tuning and model evaluation using Cross-Validation.

- Exploring Semi-supervised Learning techniques.
- Implementing Active Learning strategies.
- Reinforcement Learning: Implementing Q-Learning for decision-making.
- Introduction to Bayesian Learning and probabilistic models.

COURSE TITLE: Data Storage Technologies & Networks							
Course Code:			**CSE-241053EDS**			**Examination Scheme**	
Total number of Lecture Hours: 56						**External**	**80**
						Internal	**20**
Lecture (L):	**4**	**Practicals(P):**	**4**	**Tutorial (T):**	**-**	**Total Credits**	**6**

Course Objectives

- To understand the fundamental principles of various storage media and technologies, including magnetic, optical, and semiconductor-based systems, and their read/write techniques and limitations.
- To explore the design considerations for storage usage and access, emphasizing memory hierarchy positioning, hardware/software integration, and performance optimization.
- To analyze the architecture and scalability challenges of large storage systems, including hard disks, network-attached storage, and storage partitioning techniques.
- To investigate the components, architectures, and quality of service (QoS) considerations of storage area networks (SANs), including reliability, performance, security, and recent trends like copy data management.

Course Content	TEACHING HOURS
UNIT 1: Storage Media and Technologies	**-Hrs**
Storage Media and Technologies – Magnetic, Optical and Semiconductor Media, Techniques for read/write Operations, Issues and Limitations.	14
UNIT 2: Usage and Access	**- Hrs**
Usage and Access – Positioning in the Memory Hierarchy, Hardware and Software Design for Access, Performance issues.	14
UNIT 3: Large Storages & Storage Architecture	**-Hrs**
Large Storages – Hard Disks, Networked Attached Storage, Scalability issues, Networking issues. Storage Architecture - Storage Partitioning, Storage System Design, Caching, Legacy Systems.	14
UNIT 4: Storage Area Networks	**-Hrs**
Storage Area Networks – Hardware and Software Components, Storage Clusters/Grids. Storage QoS–Performance, Reliability, and Security issues. Recent Trends related to Copy data management	14

Textbooks

1. Data Storage Networking: Real World Skills for the CompTIA Storage by Nigel Poulton

Reference Books

1. The Complete Guide to Data Storage Technologies for Network-centric Computing Paperback– Import, Mar 1998 by Computer Technology Research Corporation

COURSE OUTCOMES (CO):

CO1: Demonstrate knowledge of various storage media technologies, their operation techniques, and limitations, and evaluate their applications in modern systems.

CO2: Apply principles of memory hierarchy and storage access to design efficient hardware and software solutions for optimized performance in storage systems.

CO3: Analyze and design scalable storage architectures by addressing partitioning, caching, and networking issues in large storage systems and legacy setups.

CO4: Evaluate the hardware and software components of SANs, ensuring performance, reliability, and security, while exploring innovative approaches like storage clusters and copy data management.

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	1	2	3	4	5	6	7	8	9	10	11	12
1	3	3	2	1	3	1	2	1	-	1	-	2
2	3	3	3	2	3	1	2	1	1	2	1	2
3	3	3	3	3	3	2	3	3	1	2	2	3
4	3	3	3	3	3	3	2	2	1	2	2	3

Lab Manual

1. **Comparison of Storage Media**
2. **File Read/Write Operations**
3. **Error Handling in Storage Media**
4. **Storage Limitations Analysis**
5. **Memory Hierarchy Simulation**
6. **Performance Benchmarking of Storage Access**
7. **File System Structure Analysis**
8. **Caching Mechanism Implementation**
9. **HDD vs. NAS Performance Testing**
10. **RAID Implementation and Analysis**
11. **Storage Partitioning Experiment**
12. **Storage Caching Performance Test**
13. **Basic SAN Configuration**
14. **Storage QoS Analysis**
15. **Copy Data Management (CDM) Experiment**
16. **Storage Cluster Performance Analysis**

Program Elective III

COURSE TITLE: Data Visualization

Course Code:			**CSE-242041EDS**		**Examination Scheme**		
Total number of Lecture Hours: 56					**External**	**80**	
					Internal	**20**	
Lecture (L):	**4**	**Practical (P):**	**-**	**Tutorial (T):**	**-**	**Total Credits**	**4**

Course Objectives

- Introduce students to key principles of visual perception, Gestalt principles, and how they influence the design and interpretation of visualizations.
- Equip students with the skills to design, develop, and evaluate various types of data visualizations, focusing on effective visual mapping, interaction techniques, and addressing common issues like information overload and misleading representations.
- Provide students with the tools and knowledge necessary to visualize complex data types such as volumetric data, vector fields, and geographic information.
- Expose students to recent trends in data perception and visualization, emphasizing new techniques and technologies.

Course Content	TEACHING HOURS
UNIT 1: Introduction to Visual Perception and Data Representation	**14Hrs**
Introduction of visual perception, visual representation of data, Gestalt principles, information overloads. Creating visual representations, visualization reference model, visual mapping, visual analytics, Design of visualization applications.	
UNIT 2: Classification of Visualization Systems and Interaction Techniques	**14 Hrs**
Classification of visualization systems, Interaction and visualization techniques misleading, Visualization of one, two and multi-dimensional data, text and text documents. Visualization of groups, trees, graphs, clusters, networks, software, Metaphorical visualization	
UNIT 3: Advanced Visualization Techniques for Specialized Data	**14 Hrs**
Visualization of volumetric data, vector fields, processes and simulations, Visualization of maps, geographic information, GIS systems, collaborative visualizations, Evaluating visualizations	
UNIT 4 : Emerging Trends and Data Structures in Visualization	**14 Hrs**
Recent trends in various perception techniques, various visualization techniques, data structures used in data visualization.	

Textbooks

1. WARD, GRINSTEIN, KEIM, .Interactive Data Visualization: Foundations, Techniques, and Applications. Natick : A K Peters, Ltd.

Reference Books

2. E. Tufte, The Visual Display of Quantitative Information, Graphics Press.

COURSE OUTCOMES (CO):

CO1: Apply Gestalt Principles and Visual Mapping to Create Effective Visual Representations: Students will be able to apply visual perception theories (such as Gestalt principles) and visual mapping techniques to create clear and intuitive visualizations for complex datasets.

CO2: Design, Develop, and Evaluate Visualization Applications: Students will develop interactive visualizations using appropriate techniques and tools, and critically evaluate the effectiveness of these visualizations in terms of clarity, accuracy, and user engagement.

CO3: Visualize Complex and Multi-dimensional Data Using Advanced Techniques: Students will be able to visualize multidimensional and specialized data types (such as 3D, geographic, and text data) using advanced techniques and tools like GIS, vector fields, and simulation data.

CO4: Identify and Apply Emerging Trends in Data Visualization and Evaluation: Students will recognize and apply emerging trends in data visualization, including new perception techniques and cutting-edge visualization methods. They will also be able to assess the usability and effectiveness of visualizations through evaluation metrics and real-world application.

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	1	2	3	4	5	6	7	8	9	10	11	12
1	2	2	3	1	3	2	-	2	3	3	2	2
2	3	3	3	-	3	3	1	2	3	3	2	1
3	3	2	3	3	3	2	1	2	3	2	-	1
4	2	2	2	3	2	-	1	2	2	3	3	3

COURSE TITLE: Big Data Analytics

Course Code:	CSE-242042EDS	Examination Scheme	T	P
Total number of Lecture Hours:56 Total number of Practical Hours:-		External	80	
		Internal	20	

Lecture (L):	4	Practicals(P):	-	Tutorial (T):	-	Total Credits	4

Course Objectives:

- Understand the significance and applications of big data across various industries, including marketing, healthcare, and finance.
- Explore the fundamentals of NoSQL databases, including different data models and advanced concepts such as sharding and replication.
- Gain proficiency in data management and analysis using Hadoop, including its architecture, HDFS, and data processing techniques.
- Learn MapReduce workflows and integrate NoSQL databases, such as HBase and Cassandra, with Hadoop for effective data analytics

Course Content	TEACHING HOURS
UNIT 1: Introduction to Big Data and Its Applications Across Industries	**-14 Hrs**
What is big data, why big data, convergence of key trends, unstructured data, industry examples of big data, web analytics, big data and marketing, fraud and big data, risk and big data, credit risk management, big data and algorithmic trading, big data and healthcare, big data in medicine, advertising and big data, big data technologies, introduction to Hadoop, open source technologies, cloud and big data, mobile business intelligence, Crowd sourcing analytics, inter and trans firewall analytics.	
UNIT 2: NoSQL Databases and Data Processing Techniques	**- 14 Hrs**
Introduction to NoSQL, aggregate data models, aggregates, key-value and document data models, relationships, graph databases, schemaless databases, materialized views, distribution models, sharding, master-slave replication, peer peer replication, sharding and replication, consistency, relaxing consistency, version stamps, map-reduce, partitioning and combining, composing map-reduce calculations.	
UNIT 3: Data Management and Analysis with Hadoop	**-14 Hrs**
Data format, analyzing data with Hadoop, scaling out, Hadoop streaming, Hadoop pipes, design of Hadoop distributed file system (HDFS), HDFS concepts, Java interface, data flow, Hadoop I/O, data integrity, compression, serialization, Avro, file-based data structures	
UNIT 4: MapReduce Workflows and NoSQL Integration	**-14 Hrs**
MapReduce workflows, unit tests with MRUnit, test data and local tests, anatomy of MapReduce job run, classic Map-reduce, YARN, failures in classic Map-reduce and YARN, job scheduling, shuffle and sort, task execution, MapReduce types, input formats, output formats. Hbase, data model and implementations, Hbase clients, Hbase examples, praxis.Cassandra, Cassandra data model, Cassandra examples, Cassandra clients, Hadoop integration.	

Lab Manual

Lab should be covered on the following topics:

- Analyze datasets of varying size and structure (CSV, JSON, unstructured text). Identify volume, velocity, and variety characteristics.
- Use web analytics tools (e.g., Google Analytics) on sample data to perform customer segmentation and understand behavior patterns.
- Apply data mining techniques on a financial dataset to detect anomalies indicative of fraud (e.g., clustering for anomaly detection).
- Explore a healthcare dataset (e.g., patient records) to identify trends, assess risk factors, and understand applications in predictive analytics.
- Set up a Hadoop cluster on a local machine or cloud environment. Store and retrieve data on HDFS, exploring storage distribution.
- Use MongoDB or Redis to create and manipulate key-value pairs and document structures. Experiment with CRUD operations and indexing.
- Build a basic graph database with Neo4j. Load and query data with relationships (e.g., social network connections) to see graph traversal.
- Implement sharding and replication in MongoDB or Cassandra. Evaluate how sharding improves performance and replication enhances reliability.
- Run MapReduce tasks on NoSQL data, such as counting document types or aggregating values in a collection.
- Upload and retrieve files in HDFS. Experiment with replication factors, block sizes, and permissions to see their effect on storage and performance.
- Develop MapReduce workflows for a specific task (e.g., log parsing), set up unit tests with MRUnit, and perform local tests.
- Set up Hbase and Cassandra clients with Hadoop. Create data models and run queries, observing data distribution and access.

Textbooks

- **"Big Data: Principles and Best Practices of Scalable Real-Time Data Systems"**, Nathan Marz and James Warren, 1st Edition, 2015
- **Hadoop: The Definitive Guide"**, Tom White, 4th Edition, 2015, O'Reilly Media

Reference Books

- **"NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence"**, Pramod J. Sadalage and Martin Fowler, 1st Edition, 2012, Addison-Wesley
- **Data-Intensive Text Processing with MapReduce"**, Jimmy Lin and Chris Dyer, 1st Edition, 2010, Morgan & Claypool Publishers.
- **"HBase: The Definitive Guide"**, Lars George, 1st Edition, 2011, O'Reilly Media.
- **"Cassandra: The Definitive Guide"** Eben Hewitt, 2nd Edition, 2016, O'Reilly Media

COURSE OUTCOMES (CO):

CO1: Describe big data and use cases from selected business domains
CO2: Explain NoSQL big data management
CO3: Install, configure, and run Hadoop and HDFS
CO4: Perform map-reduce analytics using Hadoop, Use Hadoop related tools such as HBase, Cassandra

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	1	2	3	4	5	6	7	8	9	10	11	12
1	3	2	-	-	2	1	-	1	-	2	-	-
2	3	3	2	-	3	-	-	-	-	-	-	1
3	3	3	3	3	3	1	-	-	1	-	2	-
4	3	3	3	3	3	-	-	-	1	-	2	1

COURSE TITLE: Data Warehouse and Data Mining							
Course Code:		**CSE-242043EDS**			**Examination Scheme**		
Total number of Lecture Hours: 56					**External**	**80**	
					Internal	**20**	
Lecture(L):	**4**	**Practicals(P):**	**0**	**Tutorial(T):**	**0**	**Total Credits**	**4**

Course Objectives

- Comprehend the architecture, components, and processes involved in building and managing data warehouses.
- Learn to design a data warehouse schema and use ETL (Extract, Transform, Load) processes to populate it.
- Gain knowledge in various data mining techniques and algorithms for pattern discovery and predictive modeling.
- Develop hands-on experience with popular data mining tools and techniques for real-world applications.
- Use data mining and data warehouse knowledge to create insights and make data-driven decisions.
- Implement data mining algorithms to solve business problems.

Course Content	TEACHING HOURS
UNIT 1:	**- Hrs**
Overview of Data Warehousing: Definition, purpose, and benefits. Data Warehouse Architecture: Layers of data warehouse, components (data sources, staging area, data warehouse, and data marts). OLAP vs OLTP: Differences and use cases. Data Warehouse Models: Star Schema, Snowflake Schema, and Fact Constellation Schema. ETL Process: Extraction, Transformation, and Loading of data.	**14**
UNIT 2:	**- Hrs**
Data Modeling for Data Warehouses: Fact tables, dimension tables, and normalization/denormalization. Schema Design and Indexing: Star schema design, dimensional modeling, surrogate keys. Data Integration and ETL Tools: Techniques for cleaning, transforming, and loading data into data warehouses. OLAP Operations: Roll-up, drill-down, slice-and-dice, pivot.	**14**
UNIT 3:	**- Hrs**
Overview of Data Mining: Types of data mining, classification, regression, clustering, association rules. Data Mining Process: Problem definition, data preparation, model building, evaluation, deployment. Data Mining Techniques: Decision trees, k-means clustering, k-nearest neighbour (KNN), and association rule mining.	**14**
UNIT 4:	**- Hrs**
Mining Time series Data, Periodicity Analysis for time related sequence data, Trend analysis, Similarity search in Time-series analysis; Mining Data Streams, Methodologies for stream data processing and stream data systems, Frequent pattern mining in stream data, Sequential Pattern Mining in Data Streams, Classification of dynamic data streams, Class Imbalance Problem; Graph Mining; Social Network Analysis.	**14**

Textbooks

1. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. 3rd ed., Wiley, 2013.
2. Data Mining: Concepts and Techniques. 3rd ed., Morgan Kaufmann (Elsevier), 2012.
3. Data Warehousing: Concepts, Techniques, Products, and Applications. 2nd ed., Wiley, 2010.

Reference Books

1. Fundamentals of Data Warehouses. 1st ed., Springer, 2019.
2. Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner. 2nd ed., Wiley, 2017.
3. Big Data: Principles and Paradigms. 1st ed., Wiley, 2016.
4. Data Mining: Practical Machine Learning Tools and Techniques. 4th ed., Morgan Kaufmann (Elsevier), 2016.

COURSE OUTCOMES(CO):

CO1: Demonstrate an understanding of the concepts, techniques, and tools used in data warehousing and data mining.

CO2: Design, implement, and query a data warehouse, including OLAP (Online Analytical Processing) operations.

CO3: Develop and apply data mining algorithms to extract patterns, trends, and models from large datasets.

CO4: Evaluate the performance and accuracy of data mining models using appropriate metrics.

CO5: Use advanced data mining techniques such as classification, clustering, and association rule mining.

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	1	2	3	4	5	6	7	8	9	10	11	12
1	3	3	2	2	2	2	1	1	2	2	1	2
2	3	3	2	3	3	2	1	1	2	2	2	3
3	3	3	3	3	3	2	1	1	2	3	2	3
4	2	2	2	2	3	1	1	1	3	3	2	2
5	3	3	3	3	2	3	2	2	3	3	2	3

Program Elective IV

COURSE TITLE: Data Security and Access Control							
Course Code:			**CSE-242051EDS**			**Examination Scheme**	
Total number of Lecture Hours: 56						**External**	**80**
						Internal	**20**
Lecture (L):	**4**	**Practicals(P):**	**4**	**Tutorial (T):**	**-**	**Total Credits**	**6**

Course Objectives

- The objective of the course is to provide fundamentals of database security. Various access control techniques mechanisms were introduced along with application areas of access control techniques.
- Analyze the structure and application of RBAC models and compare them with DAC and MAC access control policies.
- Examine integrity and security models while integrating RBAC into enterprise IT infrastructures.
- Explore smart card technology, its security mechanisms, and emerging trends in data security management.

Course Content	TEACHING HOURS
UNIT 1: Introduction to Access Control	**14 Hrs**
Introduction to Access Control, Purpose and fundamentals of access control, brief history, Policies of Access Control, Models of Access Control, and Mechanisms, Recent trends in access control mechanisms, Discretionary Access Control (DAC), Non- Discretionary Access Control, Mandatory Access Control (MAC). Capabilities and Limitations of Access Control Mechanisms: Access Control List (ACL) and Limitations, Capability List and Limitations.	
UNIT 2: Role-Based Access Control	**14 Hrs**
Role-Based Access Control (RBAC) and Limitations, Core RBAC, Hierarchical RBAC, Statically Constrained RBAC, Dynamically Constrained RBAC, Limitations of RBAC. Comparing RBAC to DAC and MAC Access control policy. Case study of Role-Based Access Control (RBAC) systems.	
UNIT 3: Security Models and Role-Based Access Control in Enterprise Systems	**14 Hrs**
Biba's integrity model, Clark-Wilson model, Domain type enforcement model, mapping the enterprise view to the system view, Role hierarchies-inheritance schemes, hierarchy structures and inheritance forms, using SoD in real system Temporal Constraints in RBAC, MAC AND DAC. Integrating RBAC with enterprise IT infrastructures: RBAC for WFMSs, RBAC for UNIX and JAVA environments Case study: Multi-line Insurance Company.	
UNIT 4: Smart Card Technology and Emerging Trends in Data Security Management	**14 Hrs**
Smart Card based Information Security, Smart card operating system fundamentals, design and implantation principles, memory organization, smart card files, file management, atomic operation, smart card data transmission ATR, PPS Security techniques- user identification, smart card security, quality assurance and testing, smart card life cycle-5 phases, smart	

card terminals. Recent Trends related to data security management, vulnerabilities in different DBMS.	

Textbooks
Computer Security: Principles and Practice" (4th Edition) by William Stallings and Lawrie Brown
Reference Books
Role Based Access Control: David F. Ferraiolo, D. Richard Kuhn, Ramaswamy Chandramouli. Second Edition
COURSE OUTCOMES (CO): **CO1:** In this course, the students will be enabled to understand and implement classical models and algorithms **CO2:** They will learn how to analyse the data, identify the problems, and choose the relevant models and algorithms to apply. **CO3:** They will learn how to analyse the data, identify the problems, and choose the relevant models and algorithms to apply.
LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	1	-	2	1	1	-	-	3	1	2	2	3
2	2	1	1	-	-	-	2	1	3	2	1	3
3	-	2	2	1	-	1	1	-	-	1	1	3

Lab Data Security and Access Control

Lab Topics

1) Simulate Access Control Lists (ACLs) using Python dictionaries, and compare Discretionary Access Control (DAC) with Mandatory Access Control (MAC).
2) Implement simple access control policies using Python scripts, and test them with simulated user requests.
3) Simulate Attribute-Based Access Control (ABAC) using user attributes in Python, and compare it with traditional ACL methods.
4) Analyze a real-world access control implementation (e.g., in healthcare or government) and prepare a brief report on its effectiveness.

5) Develop a basic Role-Based Access Control (RBAC) system by mapping users, roles, and permissions using Python.
6) Extend the RBAC system to include role hierarchies and permission inheritance for more efficient access management.
7) Compare the behavior of RBAC, DAC, and MAC models through simple Python simulations and discuss their respective strengths and weaknesses.
8) Review a case study of a real-world RBAC implementation (e.g., in a banking environment) and document its key features in a report.
9) Simulate Biba's integrity model in Python.
10) Simulate the Clark-Wilson security model using Python
11) Integrate the RBAC model with a simulated enterprise workflow management system using Python.
12) Automate routine access control tasks using Python scripting to improve efficiency.
13) Implement basic encryption mechanisms to secure sensitive data.
14) Create a secure login system incorporating multi-factor authentication using Python.
15) Analyze security logs and perform basic security audits to evaluate the effectiveness of implemented access control mechanisms.

Computing Resources

- **Operating Systems:** Windows, Linux (e.g., Ubuntu)
- **Programming Language:** Python
- **Virtualization Tools (Optional):** VirtualBox or Docker

COURSE TITLE: Web Analytics and Development									
Course Code:				**CSE-242052EDS**		**Examination Scheme**			
Total number of Lecture Hours: 56 + 14 Lab Hours						**External**	**80**		
						Internal	**20**		
Lecture (L):	**4**	**Practicals(P):**	**4**	**Tutorial (T):**		**Total Credits**	**6**		

Course Objectives

- Introduce fundamental concepts of social networks, web data structures, and analytical methods to measure individuals and networks.
- Equip students with knowledge of web analytics tools for analyzing user behavior, conducting A/B testing, and gathering insights through online surveys.
- Develop a comprehensive understanding of web search and retrieval processes, including SEO, web crawling, indexing, and ranking algorithms.
- Explore the principles of link analysis, network evolution, social connections, and the diffusion of innovation to understand the dynamics of online communities.

Course Content	TEACHING HOURS
UNIT 1:	**- Hrs**
Introduction – Social network and Web data and methods, Graph and Matrices, Basic measures for individuals and networks, Information Visualization	14
UNIT 2:	**- Hrs**
Web Analytics tools: Click Stream Analysis, A/B testing, Online Surveys	14
UNIT 3:	**-Hrs**
Web Search and Retrieval: Search Engine Optimization, Web Crawling and indexing, Ranking Algorithms, Web traffic models	14
UNIT 4:	**-Hrs**
Making Connection: Link Analysis, Random Graphs and Network evolution, Social Connects: Affiliation and identity, Connection Search, Collapse, Robustness Social involvements and diffusion of innovation	14

Textbooks

1. Hansen, Derek, Ben Sheiderman, Marc Smith. 2011. Analyzing Social Media Networks with NodeXL: Insights from a Connected World. Morgan Kaufmann, 304

Reference Books

1. Avinash Kaushik. 2009. Web Analytics 2.0: The Art of Online Accountability.
2. Easley, D. & Kleinberg, J. (2010). Networks, Crowds, and Markets: Reasoning About a Highly Connected World. New York: Cambridge University Press. http://www.cs.cornell.edu/home/kleinber/networks-book/
3. Wasserman, S. & Faust, K. (1994). Social network analysis: Methods and applications. New York: Cambridge University Press. Monge, P. R. & Contractor, N. S. (2003). Theories of communication networks. New York: Oxford University Press.

COURSE OUTCOMES (CO):

CO1: Demonstrate understanding of social network structures, metrics, and visualization techniques to interpret data from networks and web sources.

CO2: Apply web analytics tools to assess user interaction, perform A/B testing, and utilize online survey results for decision-making.

CO3: Implement web search and retrieval strategies, including SEO, crawling, and ranking algorithms, to improve content discoverability and traffic management.

CO4: Analyze social connections and network evolution, using link analysis and random graphs to study social dynamics, robustness, and the spread of innovations in networks.

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	1	2	3	4	5	6	7	8	9	10	11	12
1	3	2	2	2	2	1	2	1	2	2	1	2
2	3	3	2	2	3	-	-	-	2	1	-	2
3	3	2	3	2	3	-	-	-	2	1	2	2
4	3	3	2	3	2	1	2	2	2	2	-	2

Lab Manual

The lab should cover the following topics:

1. **Graph Representation of a Social Network**
2. **Degree Centrality Calculation**
3. **Network Visualization**
4. **Community Detection**
5. **Clickstream Analysis**
6. **A/B Testing Simulation**

To be effective from year-2024

7. **Online Survey Data Analysis**
8. **Basic Search Engine**
9. **Web Crawler Implementation**
10. **PageRank Algorithm Simulation**
11. **Web Traffic Analysis**
12. **Link Prediction in Social Networks**
13. **Random Graph Model Simulation**
14. **Diffusion of Innovation Simulation**

COURSE TITLE: Knowledge Discovery

Course Code:				**CSE-242053EDS**		**Examination Scheme**	
Total number of Lecture Hours: 56 + 14 Lab hours						**External**	**80**
						Internal	**20**
Lecture (L):	**4**	**Practicals(P):**	4	**Tutorial (T):**	**-**	**Total Credits**	**6**

Course Objectives

- Understand the fundamentals and ethical considerations of data mining and knowledge discovery.
- Explore various methods of knowledge representation for effective data analysis.
- Develop skills in building and evaluating decision trees for predictive modeling.
- Learn techniques for constructing and evaluating classification and association rules.

Course Content	TEACHING HOURS
UNIT 1: Introduction KDD and Data Mining	**14 Hrs**
Data Mining and Machine Learning, Machine Learning and Statistics, Generalization as Search,Data Mining and Ethics 7	
UNIT 2: Knowledge Representation	**14 Hrs**
Decision Tables, Decision Trees, Classification Rules, Association Rules, Rules involving Relations, Trees for Numeric Predictions, Neural Networks, Clusters 10	
UNIT 3: Decision Trees	**14 Hrs**
Divide and Conquer, Calculating Information, Entropy, Pruning, Estimating Error Rates, The C4.5 Algorithm Evaluation of Learned Results- Training and Testing, Predicting Performance, Cross-Validation 9	
UNIT 4: Classification Rules	**14 Hrs**
Inferring Rudimentary Rules, Covering Algorithms for Rule Construction, Probability Measure for Rule Evaluation, Association Rules, Item Sets, Rule Efficiency	

Textbooks

- Data mining and knowledge discovery handbook by Maimon, oded(et al.)

Reference Books

- Data Cleansing : A Prelude to knowledge Discovery

To be effective from year-2024

COURSE OUTCOMES (CO):

CO1: Explain the basics of data mining, machine learning, and their ethical implications.

CO2: Represent data knowledge using decision tables, trees, and neural networks.

CO3: Construct and evaluate decision trees using algorithms like C4.5 and cross-validation.

CO4: Develop and assess classification and association rules for effective data insights.

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	1	2	3	4	5	6	7	8	9	10	11	12
1	2	2	2	3	3	3	-	1	2	1	1	2
2	2	2	2	3	3	3	-	2	2	3	-	1
3	3	3	3	3	2	2	2	-	2	2	2	2
4	2	2	2	2	2	3	3	3	1	1	2	2

Lab Manual

Lab should cover the following topics:

- Data Mining and Machine Learning
- Machine Learning and Statistics
- Generalization as Search
- Data Mining and Ethics
- Decision Tables
- Decision Trees
- Classification Rules
- Association Rules
- Rules Involving Relations
- Trees for Numeric Predictions
- Neural Networks
- Clusters

Program Elective V

COURSE TITLE: GPU Computing

Course Code:			**CSE-243011EDS**			**Examination Scheme**	
Total number of Lecture Hours: 56						**External**	**80**
						Internal	**20**
Lecture(L):	**4**	**Practicals(P):**	**0**	**Tutorial(T):**	**0**	**Total Credits**	**4**

Course Objectives

1. Explore the fundamental architecture and design principles of Graphics Processing Units (GPUs) and their role in high-performance computing (HPC).
2. Gain hands-on experience with parallel programming using GPU-specific languages and libraries such as CUDA, OpenCL, and others.
3. Learn how to optimize algorithms for GPUs to achieve significant speedups in a wide range of applications (scientific computing, machine learning, data processing, etc.).
4. Understand and optimize memory management techniques on GPUs, including shared memory, global memory, and registers.
5. Get introduced to advanced GPU features like multi-threading, kernel optimization, and the use of modern GPU libraries.
6. Implement GPU-based solutions for real-world computational problems, including image processing, simulations, and machine learning.

Course Content	TEACHING HOURS
UNIT 1:	**- Hrs**
Understanding the need for parallel computing and the limitations of CPU-based approaches. GPU Architecture: Key differences between CPUs and GPUs; introduction to SIMD (Single Instruction, Multiple Data) architecture, CUDA cores, streaming multiprocessors, memory hierarchy. GPU Programming Models: Overview of parallel computing models and programming paradigms. CUDA and OpenCL: Introduction to CUDA programming model and OpenCL framework for GPU computing.	**14**
UNIT 2:	**- Hrs**
Introduction to CUDA: Setting up a CUDA development environment, compiling and running CUDA programs. Basic CUDA Programming: Writing simple CUDA kernels, understanding thread and block structure. Memory Models: Understanding global, shared, constant, and local memory in CUDA. Basic Performance Optimization: Optimizing simple CUDA programs for memory bandwidth and thread synchronization.	**14**
UNIT 3:	**- Hrs**
Thread Organization: Managing grids, blocks, and threads; understanding the importance of the grid/block/thread model. Defining and launching CUDA kernels; multi-threading and execution configuration. Synchronization and Deadlocks: Synchronizing threads, managing shared resources, and avoiding deadlocks in parallel execution. Introduction to OpenCL architecture and programming model; comparison with CUDA.	**14**

UNIT 4:	**- Hrs**
Using GPUs for large-scale scientific simulations, solving differential equations, Monte Carlo simulations. GPU acceleration in machine learning algorithms, using frameworks like TensorFlow and cuDNN for deep learning. Implementing GPU-accelerated algorithms for image filtering, FFT, and other signal processing tasks. Implementing a real-world GPU application (e.g., a simulation or deep learning model).	**14**

Textbooks

5. Programming Massively Parallel Processors: A Hands-on Approach. Kirk, David B., and Wen-mei W. Hwu. 2nd ed., Morgan Kaufmann, 2016.
6. CUDA Programming: A Developer's Guide to Parallel Computing with GPUs; Shane Cook; Morgan Kaufman; 2012.
7. CUDA by Example: An Introduction to General-Purpose GPU Programming. Sanders, Jason, and Edward Kandrot. Addison-Wesley, 2010.

Reference Books

1. Hands-On GPU-Accelerated Machine Learning with Python. Miller, Michael B. Packt Publishing, 2020.
2. Patterson. Computer Architecture: A Quantitative Approach. Hennessy, John L., and David A. 6th ed., Elsevier, 2017.
3. OpenCL Programming Guide. Munshi, Aaftab, et al. 1st ed., Addison-Wesley, 2011.

COURSE OUTCOMES(CO):

CO1: Demonstrate knowledge of GPU hardware, architecture, and programming models.

CO2: Write efficient programs for GPUs using CUDA or OpenCL to leverage parallelism and optimize performance.

CO3: Adapt and optimize algorithms to run efficiently on GPUs, including managing data transfer between CPU and GPU.

CO4: Effectively use various GPU memory types (e.g., global, shared, constant memory) and optimize memory access patterns to improve performance.

CO5: Implement practical applications that benefit from GPU acceleration, including tasks such as image processing, scientific simulations, and deep learning.

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	3	3	2	2	2	2	1	1	2	2	1	2
2	3	3	2	3	3	2	1	1	2	2	2	3
3	3	3	3	3	3	2	1	1	2	3	2	3
4	2	2	2	2	3	1	1	1	3	3	2	2
5	3	3	3	3	2	3	2	2	3	3	2	3

COURSE TITLE: Cloud Computing							
Course Code:				**CSE-243012EDS**		**Examination Scheme**	
Total number of Lecture Hours: 56						**External**	**80**
						Internal	**20**
Lecture (L):	**4**	**Practicals(P):**	**0**	**Tutorial (T):**	**0**	**Total Credits**	**4**

Course Objectives

- Understand core cloud computing concepts and service models.
- Gain practical skills in deploying and managing cloud applications.
- Understand how to manage cloud service performance, reliability, and security.
- Analyze the cost and benefits of different cloud platforms.

Course Content	TEACHING HOURS
UNIT 1: CLOUD COMPUTING FUNDAMENTALS	**14Hrs**
Introduction to Cloud Computing, Online Social Networks and Applications, Cloud introduction and overview, Different clouds, Risks, Novel applications of cloud computing Cloud Computing definition;, private, public and hybrid cloud. Cloud types; IaaS, PaaS, SaaS. Benefits and challenges of cloud computing, public vs private clouds, Business Agility: Benefits and challenges to Cloud architecture. Application availability, performance, security and disaster recovery; next generation Cloud Applications.	
UNIT 2: VIRTUALIZATION AND CLOUD APPLICATIONS	**14 Hrs**
VIRTUALIZATION: Role of virtualization in enabling the cloud : Types of Virtual Machines, Advantages of Virtualization, Components of Virtualization, cloud applications : Technologies and the processes required when deploying web services; Deploying a web service from inside and outside a cloud architecture, advantages and disadvantages	
UNIT 3: MANAGEMENT OF CLOUD SERVICES	**14Hrs**
Reliability, availability and security of services deployed from the cloud. Performance and scalability of services, tools and technologies used to manage cloud services deployment; Cloud Economics: Cloud Computing infrastructures available for implementing cloud based services. Economics of choosing a Cloud platform for an organization, based on application requirements, economic constraints and business needs. Privacy Issues, Data Life Cycle, Key Privacy Concerns in the Cloud, Protecting Privacy	
UNIT 4: APPLICATION DEVELOPMENT	**14Hrs**
Audit and Compliance Internal Policy Compliance, Governance, Risk, and Compliance (GRC), Regulatory/External Compliance, Cloud Security Alliance, Auditing the Cloud for Compliance, Security-as-a-Cloud Application Development: Design and implementation in cloud environments. Development Platforms: AWS, Azure, Google App Engine. Deployment and management strategies for cloud applications.	

Textbooks

- Gautam Shroff, "Enterprise Cloud Computing: Technology, Architecture, Applications", Cambridge University Press; 2nd Edition [ISBN: 9780521137355], 2023.
- Toby Velte, Anthony Velte, Robert Elsenpeter, "Cloud Computing: A Practical Approach" McGraw-Hill Education; 2nd Edition [ISBN: 9780071826400], 2022.
- Dimitris N. Chorafas, "Cloud Computing Strategies" CRC Press; 2nd Edition [ISBN: 9780367338611], 2021.

Reference Books

- Thomas Erl, "Cloud Computing: Concepts, Technology & Architecture" Prentice Hall; 3rd Edition [ISBN: 9780133994164], 2024.
- Rajkumar Buyya, Christian Vecchiola, and Selvi, S. Thamarai, "Mastering Cloud Computing: Foundations and Applications Programming" Morgan Kaufmann; 3rd Edition [ISBN: 9780128180747], 2022.

COURSE OUTCOMES (CO):

CO1: Explain cloud computing principles and service models.
CO2: Successfully deploy and manage cloud-based applications.
CO3: Apply best practices for cloud service management.
CO4: Assess the economic aspects of cloud computing platforms

LEVEL OF CO-PO MAPPING TABLE

	POs											
COs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	3	2	1	2	3	1	2	-	1	2	-	3
2	2	3	2	3	3	2	1	1	2	2	1	2
3	1	2	3	2	2	3	3	1	3	1	2	3
4	2	1	2	3	1	2	1	3	1	3	3	2