

SCHEME OF TEACHING AND EXAMINATION (2022 Scheme)

V Semester

B.E. in Artificial Intelligence & Data Science

Batch:2022-2023

Sl. No.	Course and Course Code		Course Title	Teaching / Paper setting Dept.	Teaching hrs./week				Examination				Credits	
					Lecture	Tutorial	Practical/ Drawing	Self-Study Component	Duration in hrs.	CIE Marks	SEE Marks	Total Marks		
					L	T	P	S						
1.	HSMS	S5CCS01	Software Engineering and Project Management	Dept.	42	0	0	48	3	50	50	100	3	
2.	IPCC	S5CCSI01	Database Management System (I)	Dept.	42	0	28	50	3	50	50	100	4	
3.	IPCC	S5CCSI02	Machine Learning Techniques (I)	Dept.	42	0	28	50	3	50	50	100	4	
4.	PCCL	S5CCSL01	Data Mining and Visualization Lab	Dept.	0	0	28	2	3	50	50	100	1	
5.	PEC		Professional Elective Course-I	Dept.	42	0	0	48	3	50	50	100	3	
6.	PROJ	S5ADMP	Mini Project	Dept.	0	0	56	4	3	50	50	100	2	
7.	AEC	SHS04	Research Methodology and IPR (Board: IEM)	ME, IM, CH	42	0	0	48	3	50	50	100	3	
8.	HSMS	SHS05	Environmental Studies (Board: CV)	CV	28	0	0	32	3	50	50	100	2	
9.	AEC	ARAS	Aptitude Related Analytical Skill		0	0	28	2	1½	50	50	100	1	
10.	NCMC	SMCO1	National Service Scheme (NSS)	NSS CO	0	0	2			100	-	100	0	
		SMCO2	Physical Education (PE) (Sports and Athletics)	PED										
		SMCO3	Yoga	PED										
		SMCO4	National Cadet Corps (NCC)	NCC CO										
Total										550	350	900	23	
		AAP	AICTE Activity Points (Applicable for both Regular and Lateral Entry students)	40 hours community service to be documented and produced for the examination										

Note: **HSMS:** Humanity and Social Science and management Course **IPCC:** Integrated Professional Core Course, **PCCL:** Professional Core Course laboratory, **PEC:** Professional Elective Course; **PROJ:** Project/Mini Project; **AEC:** Ability Enhancement Course; **NCMC:** Non-Credit Mandatory Course, **L:** Lecture, **T:** Tutorial, **P:** Practical **S= SDA:** Skill Development Activity, **CIE:** Continuous Internal Evaluation, **SEE:** Semester End Evaluation.

Professional Elective Course (PEC) (Offered by the Department)

S5CCSPE01	Information retrieval	S5CCSPE04	Knowledge Representation and Reasoning
S5CCSPE03	Business Intelligence and Analytics	S5CCSPE05	Recommendation Systems

Professional Core Course (IPCC): Refers to Professional Core Course Theory Integrated with practical of the same course. Credit for IPCC can be 04 and its Teaching–Learning hours (L : T : P) can be considered as (3 : 0 : 2) or (2 : 2 : 2). The theory part of the IPCC shall be evaluated both by CIE and SEE. The practical part shall be evaluated by only CIE (no SEE). However, questions from the practical part of IPCC shall be included in the SEE question paper. For more details, the regulation governing the Degree of Bachelor

of Engineering (B.E.) 2022-23 may please be referred.

National Service Scheme /Physical Education/Yoga: All students have to register for any one of the courses namely National Service Scheme (NSS), Physical Education (PE)(Sports and Athletics), and Yoga(YOG) with the concerned coordinator of the course during the first Week of III semesters. Activities shall be carried out between III semester to the VI semester (for 4 semesters). Successful completion of the registered course and requisite CIE score is mandatory for the award of the Degree. The events shall be appropriately scheduled by the colleges and the same shall be reflected in the calendar prepared for the NSS, PE, and Yoga activities. These courses shall not be considered for vertical progression as well as for the calculation of SGPA and CGPA, but completion of the course is mandatory for the award of Degree.

Mini-project work: Mini Project is a laboratory-oriented/hands on course that will provide a platform to students to enhance their practical knowledge and skills by the development of small systems/applications etc. Based on the ability/abilities of the student/s and recommendations of the mentor, a single discipline or a multidisciplinary Mini-project can be assigned to an individual student or to a group having not more than 4 students.

CIE procedure for Mini-project:

- (i) **Single discipline:** The CIE marks shall be awarded by a committee consisting of the Head of the concerned Department and two faculty members of the Department, one of them being the Guide. The CIE marks awarded for the Mini-project work shall be based on the evaluation of the project report, project presentation skill, and question and answer session in the ratio of 50:25:25. The marks awarded for the project report shall be the same for all the batches mates.
- (ii) **Interdisciplinary:** Continuous Internal Evaluation shall be group-wise at the college level with the participation of all the guides of the project. The CIE marks awarded for the Mini-project, shall be based on the evaluation of the project report, project presentation skill, and question and answer session in the ratio 50:25:25. The marks awarded for the project report shall be the same for all the batch mates.

No SEE component for Mini-Project.

Professional Elective Courses (PEC): A professional elective (PEC) course is intended to enhance the depth and breadth of educational experience in the Engineering and Technology curriculum. Multidisciplinary courses that are added supplement the latest trend and advanced technology in the selected stream of Engineering. Each group will provide an option to select one course. The minimum number of students' strengths for offering a professional elective is 10. However, this conditional shall not be applicable to cases where the admission to the program is less than 10.

SCHEME OF TEACHING AND EXAMINATION (2022 Scheme)

VI Semester

B.E. in Artificial Intelligence & Data Science

Batch:2022-2023

Sl. No.	Course and Course Code		Course Title	Teaching / Paper setting Dept.	Teaching hrs.				Examination				Credits
					Lecture	Tutorial	Practical/ Drawing	Self-Study Component	Duration in hrs.	CIE Marks	SEE Marks	Total Marks	
					L	T	P	S					
1.	IPCC	S6ADI01	Real Time Big Data Analytics (I)		42	0	28	50	3	50	50	100	4
2.	PCC	S6CCS01	Computer Networks		42	28	0	50	3	50	50	100	4
3.	PEC	S6ADPE0X	Professional Elective Course-II		42	0	0	48	3	50	50	100	3
4.	OEC		Open Elective Course-I		42	0	0	48	3	50	50	100	3
5.	PROJ	S6ADMP-1	Major Project Phase I		0	0	60		3	100	-	100	2
6.	PCCL	S6CCSL01	Internet of Things Lab		0	0	28	2	3	50	50	100	1
7.	NCMC	SHS06	Soft Skills (Additional Course offered by SIT)	T&P	0	2	0		-	50	50	100	0
8.	NCMC	SMCO1	National Service Scheme (NSS)	NSS CO	0	0	2			100	-	100	0
		SMCO2	Physical Education (PE) (Sports and Athletics)	PED									
		SMCO3	Yoga	PED									
		SMCO4	National Cadet Corps (NCC)	NCC CO									
9.	HSS	SHS07	Indian Knowledge Systems		1	0	0	0					0
Total										500	300	800	17
		AAP	AICTE Activity Points (Applicable for both Regular and Lateral Entry students)	40 hours community service to be documented and produced for the examination									

Note: IPCC: Integrated Professional Core Course, PCC: Professional Core Course; PEC: Professional Elective Course; OEC: Open Elective Course; PROJ: Project Phase –I; PCCL: Professional Core Course laboratory; AEC: Ability Enhancement Course, SEC: Skill Enhancement Course; NCMC: Non Credit Mandatory Course; L: Lecture, T: Tutorial, P: Practical S= SDA: Skill Development Activity, CIE: Continuous Internal Evaluation, SEE: Semester End Evaluation.

Professional Elective Course (PEC) (Offered by the Department)

S6CCSPE01	Cloud Computing	S6ADPE02	Full Stack Development with Django
S6ADPE01	Distributed Data Storage Management	S6CCSPE04	AI –Driven Cyber Security

Professional Core Course (IPCC): Refers to Professional Core Course Theory Integrated with practical of the same course. Credit for IPCC can be 04 and its Teaching–Learning hours (L : T : P) can be considered as (3 : 0 : 2) or (2 : 2 : 2). The theory part of the IPCC shall be evaluated both by CIE and SEE. The practical part shall be evaluated by only CIE (no SEE). However, questions from the practical part of IPCC shall be included in the SEE question paper. For more details, the regulation governing the Degree of Bachelor of Engineering (B.E.) 2022-23 may please be referred.

National Service Scheme /Physical Education/Yoga: All students have to register for any one of the courses namely National Service Scheme (NSS), Physical Education (PE)(Sports and Athletics), and Yoga(YOG) with the concerned coordinator of the course during the first Week of III semesters. Activities shall be carried out between III semester to the VI semester (for 4 semesters). Successful completion of the registered course and requisite CIE score is mandatory for the award of the Degree. The events shall be appropriately scheduled by the colleges and the same shall be reflected in the calendar prepared for the NSS, PE, and Yoga activities. These courses shall not be considered for vertical progression as well as for the calculation of SGPA and CGPA, but completion of the course is mandatory for the award of Degree.

Professional Elective Courses (PEC): A professional elective (PEC) course is intended to enhance the depth and breadth of educational experience in the Engineering and Technology curriculum. Multidisciplinary courses that are added supplement the latest trend and advanced technology in the selected stream of Engineering. Each group will provide an option to select one course. The minimum number of students' strengths for offering a professional elective is 10. However, this conditional shall not be applicable to cases where the admission to the program is less than 10.

Open Elective Courses:

Students belonging to a particular stream of Engineering and Technology are not entitled to the open electives offered by their parent Department. However, they can opt for an elective offered by other Departments, provided they satisfy the prerequisite condition if any. Registration to open electives shall be documented under the guidance of the Program Coordinator/ Advisor/Mentor. The minimum numbers of students' strength for offering Open Elective Course is 10. However, this condition shall not be applicable to class where the admission to the program is less than 10.

Project Phase-I : Students have to discuss with the mentor /guide and with their help he/she has to complete the literature survey and prepare the report and finally define the problem statement for the project work.

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - V

SOFTWARE ENGINEERING AND PROJECT MANAGEMENT

Course Code	S5CCS01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	3
Lecture Hours	42Hrs	Practical Hours	-

Course objectives: This course will enable students to:

1.	Identify the process model for the development of software with its merits and demerits
2.	Identify the clear, correct and consistent requirements for the software.
3.	Design suitable data, architecture and user interface that copes with the software requirements.
4.	Understand the coding standard and apply clean code logics
5.	Formulate the cyclomatic complexity and design the corresponding test cases.
6.	Estimate project cost estimation and develop project planning.

UNIT I

(9 Hours)

SOFTWARE AND SOFTWARE ENGINEERING: The Nature of Software, Software Engineering, The Software Process, Software Engineering Practice.

THE SOFTWARE PROCESS: PROCESS MODELS:-A Generic Process Model, Prescriptive Process Models: The Waterfall Model, Incremental Process Models, Evolutionary Process Models, Concurrent Models, Final Word on Evolutionary Processes, Specialized Process Models, Component-Based Development, The Formal Methods Model, The Unified Process, Phases of the Unified Process, Personal and Team Process Models.

AGILE DEVELOPMENT: What Is Agility? Agility and the Cost of Change, What Is an Agile Process, Extreme Programming, Other Agile Process Models

Text Book 1: Chapter 1 – 1.1, 1.2, Chapter 3 -3.1, Chapter 4 – 4.1, 4.2, 4.3, 4.4, Chapter 5 – 5.1, 5.2, 5.3, 5.4, 5.5

UNIT II

(8 Hours)

UNDERSTANDING REQUIREMENTS: Requirements Engineering, Establishing the groundwork, Eliciting Requirements, Developing Use cases, Building the Analysis model, Negotiating Requirement, Validating Requirements, Avoiding Common mistakes.

REQUIREMENTS MODELING: Requirements Analysis, Scenario-Based Modeling, UML Models That Supplement the Use Case, Class-Based Modeling

Text Book 1: Chapter 8 – 8. To 8.9 Chapter 9, Chapter 10

Self Study Component : UML Design and Patterns to demonstrate Software Requirements

UNIT III

(8 Hours)

DESIGN CONCEPTS: Design within the Context of Software Engineering, The Design Process, Design Concepts, The Design Model

ARCHITECTURAL DESIGN: Software Architecture, Architectural Genres, Architectural Styles, Architectural Design

COMPONENT-LEVEL DESIGN: What Is a Component, Designing Class-Based Components, Conducting Component-Level Design, Designing Traditional Components and Component-Based Development.

USER INTERFACE DESIGN: The Golden Rules, User Interface Analysis and Design

Self Study Component: UML Design and patterns to demonstrate Architecture, Components, UI designs

Text Book 1: Chapter 12, 13, 14, 15

UNIT IV

(9 Hours)

WRITING THE PROGRAMS: Programming Standards and Procedures, Programming Guidelines, Documentation, Programming Process, Information Systems example, Real time example

Text Book 2 – Chapter 7

Clean Code: Meaningful names, Functions, Comments, Objects and Data Structures, Error Handling, Boundaries

Text Book 3 – Chapter 2, 3, 4, 6, 7, 8

TESTING CONVENTIONAL APPLICATIONS: Test Strategies for Conventional Software, Validation

Testing, System Testing, White-Box Testing, Basis Path Testing, Control Structure Testing and Black-Box

Testing and Problems on Cyclomatic complexity, Model based Testing, Testing for Real time systems

UNIT V (8 Hours)

PROJECT MANAGEMENT CONCEPTS: The management spectrum, People, Product, Process, W5HH principle.

ESTIMATION FOR SOFTWARE PROJECTS: Observations on estimation, project planning process, software scope and reliability, Resources, Software Project estimation, decomposition techniques and discussion on empirical estimation models, Software equation and problems

Project Scheduling: Scheduling, Earned Value Analysis

Text Book 1 – Chapter 31, 33 - , Chapter 34 - 34.5, 34.6

Course Outcomes: On Successful completion of this course, students will be able to

1.	Understand process models, agility.
2.	Develop software requirement specification by synthesizing requirement modeling concepts.
3.	Build high level architecture design and component level designs for software requirement.
4.	Use and practice coding standards while building codes.
5.	Design test cases and test procedures.
6.	Plan and schedule project cost , estimation and resources.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Software Engineering - A Practitioners approach	Roger G. Pressman	McGraw-Hill	8th Edition, 2019
2	Software Engineering – Theory and Practice	Shari Lawrence Pfleeger, Joanne M Atlee	Pearson	4 th Edition, 2013
3	Clean Code – A handbook of Agile Software Craftmanship	Robert C Martin	Pearson	1 st Edition,2011
Reference Books				
1	Integrated Approach to Software Engineering	Ian Sommer ville	Pearson Education Ltd	8th Edition -
2	Software Engineering	Shari Lawrence PfleegerJoanneM Atlee	Pearson	4 th Edition2013
3	Software Engineering – Programs, Documentation, Operating Procedure	K K Agarwal, Yogesh Singh	New Age International	3rd edition

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	2												2		
CO2	2								2	2			2		
CO3		2							2	2			2		
CO4		2							2	2			2		
CO5		2							2	2			2		
CO6										2	2		2		
Overall CO	2	2							2	2	2		2		

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	2	2							2	2	2		2		

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - V

DATABASE MANAGEMENT SYSTEM (I)

Course Code	S5CCSI01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:2)	SEE Marks	50
Credits	4	Exam Hours	3
Lecture Hours	42Hrs	Practical Hours	28Hrs

Course objectives: This course will enable students to:

1	To define a Database, characteristics and functions of Database Management System and distinguish between a Traditional File System and a Database System
2	To model the real world database systems using Entity Relationship Diagrams (ERD) from the requirements specification and transform it to a relational model.
3	To design SQL and NoSQL queries to perform CRUD (Create, Retrieve, Update and delete) operations on database.
4	To apply normalization techniques to normalize a Relational database
5	To illustrate how a DBMS handles transactions by enforcing recovery from failure and concurrency control

UNIT I (8 Hours)

DATABASES AND DATABASE USERS: Introduction; An example; characteristics of the database approach; when Not to use a DBMS. Chapter 1 : 1.1 – 1.8

DATABASE SYSTEM – CONCEPTS AND ARCHITECTURE: Data models, schemas, and instances; three schema architecture and data independence; database languages; centralized and client/server/architectures for DBMSs. Chapter 2 : 2.1 to 2.6

ENTITY-RELATIONSHIP MODEL: Using High-Level Conceptual Data Models for Database Design; An Example Database Application; Entity Types, Entity Sets, Attributes and Keys; Relationship types, Relationship Sets, Roles and Structural Constraints; Weak Entity Types; Refining the ER Design for the COMPANY Database; ER Diagrams, Naming Conventions and Design Issues. Relational Database Design using ER- to-Relational Mapping. Chapter 7 : 7.1 to 7.7, Chapter 9 : 9.1.

Self Study: Actors on the scene; Workers behind the scene; advantages of using the DBMS approach; DBMS interfaces; the database system environment; Classification of database management system.

UNIT II (8 Hours)

RELATIONAL MODEL

Relational Model Concepts (only definitions); Relational Model Constraints and Relational Database Schemas; Update Operations and Dealing with Constraint Violations: Chapter 3 : 3.1 to 3.3.

Relational Algebra:

Unary Relational Operations: SELECT and PROJECT, Relational Algebra Operations from Set Theory, Binary Relational Operations: JOIN and DIVISION, Additional Relational Operations - Aggregate Functions and Grouping, OUTER JOIN Operations : Chapter 6 : 6.1 to 6.5

Self Study: Examples of Queries in Relational Algebra

UNIT III (9 Hours)

SQL-THE RELATIONAL DATABASE STANDARD: SQL Data Definition and Data Types, Specifying Basic Constraints in SQL, Schema Change Statements in SQL; Basic Queries in SQL; More Complex SQL Queries; Insert, Delete and Update Statements in SQL; Specifying General Constraints as Assertion; Views (Virtual Tables) in SQL. Chapter 4 : 4.1 to 4.4, Chapter 5 : 5.1 to 5.3

MangoDB tutorial, MangoDB operators, DB commands, Database, collection, CURD

URL: www.javatpoint.com/nosql-databases

UNIT IV (8 Hours)

DATABASE DESIGN: Informal Design Guidelines for Relation Schemas; Functional Dependencies; Normal Forms Based on Primary Keys; General Definitions of Second and Third Normal Forms; Boyce-Codd Normal Form; Properties of Relational Decompositions. Chapter 15 : 15.1 to 15.5 , Chapter 16 : 16.2.

UNIT V	(9 Hours)
TRANSACTION PROCESSING CONCEPT: Introduction to transaction processing; transaction and system concepts; desirable properties of transactions, characterizing schedules based on recoverability and serializability; transaction support in SQL Chapter 21 : 21.1 to 21.6 CONCURRENCY CONTROL & DATABASE RECOVERY TECHNIQUES: Two phase locking techniques, Concurrency control based on Timestamp ordering; Recovery concepts; recovery based on deferred update and Immediate Update, Shadow Paging, ARIES Recovery Algorithm Chapter 22 : 22.1 – 22.2, Chapter 23: 23.1 to 23.5.	

Course Outcomes: On Successful completion of this course, students will be able to	
1	Describe the fundamentals of database technologies
2	Design an ER diagram and transform it to a relational model for a given database specification.
3	Design Relational algebra, SQL and NoSQL queries to perform CRUD (Create, Retrieve, Update and delete).
4	Apply Informal Design guidelines and normalization techniques to improve database design
5	Analyse Concurrency control and Database recovery techniques in transaction processing.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Fundamentals of Database Systems	Elmasri and Navathe	Pearson Education	6 th Edition, 2011
2				
Reference Books				
1	Data base System Concepts.	Silberschatz, Korth and Sudharshan.	McGraw-Hill	6 th Edition, 2010
2	Database Management Systems.	Raghu Ramakrishnan and Johannes Gehrke	McGraw-Hill.	3 th Edition, 2010

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
CO1	2		2									2		
CO2	2		2									2		
CO3	2		3									3		
CO4	2		2									2		
CO5	2		2									2		
Overall CO	2		3									3		

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
	2		3									3		

DATABASE MANAGEMENT SYSTEM LABORATORY

Sl. No	Experiments
1	<p>Suppose a movie_studio has several film crews. The crews might be designated by a given studio as crew1, crew 2, and so on. However, other studios might use the same designations for crews, so the attribute crew_number is not a key for crews. Movie_studio holds the information like name, branch and several locations. Each crew holds information like sector and strength.</p> <ol style="list-style-type: none"> i) Establish the database by normalizing up to 3NF and considering all schema level constraints ii) Write SQL insertion query to insert few tuples to all the relations iii) List all movie studios which are not used a single crews. iv) Retrieve the movie studio which uses highest strength crew. v) Write a before insert trigger to check maximum number of crews to any studio is limited to 5. vi) Write a procedure retrieve all crews used by specific studio. vii) Create “crew” collection and perform the following CRUD operations using MongoDB: <ul style="list-style-type: none"> • Create a document. • Create two or more documents at the same time. • Update a document with crew number 10 to 20. • Delete all the crews with strength 10. <p>Retrieve the crews with strength greater than or equal to 20.</p>
2	<p>The production company is organized into different studios. We store each studio’s name branch and location; every studio must own at least one movie. We store each movie’s title, sensor number and year of production. Star may act in any number of movies and we store each actors name and address.</p> <ol style="list-style-type: none"> i) Establish the database by normalizing up to 3NF and considering all schema level constraints ii) Write SQL insertion query to insert few tuples to all the relations iii) List all the studios of the movie “xyz”; iv) List all the actors , acted in a movie ‘xyz’ v) Write a procedure to list all movies produced during the specific year. vi) Write a deletion trigger, does not allow to deleting current year movies. vii) Create studio collection and perform the following CRUD operations using MongoDB: <ul style="list-style-type: none"> • Create a document. • Create two or more documents at the same time. • Update a document with studio name ‘std1’ to ‘std2’. • Delete all the studios with location ‘xyz’. <p>Retrieve the studio with location equal to ‘xyz’.</p>
3	<p>The production company is organized into different studios. We store each studio’s name branch and location; a studio own any number of Cartoon-serials. We store each Cartoon- Serial’s title, sensor number and year of production. Star may do voices in any number of Cartoon-Serials and we store each actors name and address.</p> <ol style="list-style-type: none"> i) Establish the database by normalizing up to 3NF and considering all schema level constraints ii) Write SQL insertion query to insert few tuples to all the relations iii) Find total no of actors, do voiced in a Cartoon-Serials ‘xyz’ iv) Retrieve name of studio, location and Cartoon-Serials title in which star “abc” is

	<p>voiced.</p> <p>v) Write a procedure to list all Cartoon-Serials produced during the specific year.</p> <p>vi) Write a deletion trigger, does not allow to deleting current year Cartoon-Serials.</p> <p>vii) Create “cartoon serial” collection and perform the following CRUD operations using MongoDB:</p> <ul style="list-style-type: none"> • Create a document. • Create two or more documents at the same time. • Update a document with cartoon serial title ‘std1’ to ‘std2’. • Delete all the cartoon serials with title ‘xyz’. <p>Retrieve the cartoon serials with sensor numbers lesser than 100.</p>
4	<p>Car marketing company wants keep track of marketed cars and their owner. Each car must be associated with a single owner and owner may have any number of cars. We store car’s registration number, model & color and owner’s name, address & SSN. We also store date of purchase of each car.</p> <p>i) Establish the database by normalizing up to 3NF and considering all schema level constraints</p> <p>ii) Write SQL insertion query to insert few tuples to all the relations</p> <p>iii) Find a person who owns highest number of cars</p> <p>iv) Retrieve persons and cars information purchased on the day 11-11-11</p> <p>v) Write a insertion trigger to check date of purchase must be less than current date (must use system date)</p> <p>vi) Write a procedure to list all cars and owner information purchased during the specific year.</p> <p>vii) Create “car” collection and perform the following CRUD operations using MongoDB:</p> <ul style="list-style-type: none"> • Create a document. • Create two or more documents at the same time. • Update a document with car reg.no 10 to 20. • Delete all the cars with model ‘xyz’. <p>Retrieve the cars with colour green.</p>
5	<p>Puppy pet shop wants to keep track of dogs and their owners. The person can buy maximum three pet dogs. We store person’s name, SSN and address and dog’s name, date of purchase and sex. The owner of the pet dogs will be identified by SSN since the dog’s names are not distinct.</p> <p>i) Establish the database by normalizing up to 3NF and considering all schema level constraints</p> <p>ii) Write SQL insertion query to insert few tuples to all the relations</p> <p>iii) List all pets owned by a person ‘Abhiman’.</p> <p>iv) List all persons who are not owned a single pet</p> <p>v) Write a trigger to check the constraint that the person can buy maximum three pet dogs</p> <p>vi) Write a procedure to list all dogs and owner details purchased on the specific date.</p> <p>vii) Create “dog” collection and perform the following CRUD operations using MongoDB:</p> <ul style="list-style-type: none"> • Create a document. • Create two or more documents at the same time. • Update a document with dog name ‘xyz’ to ‘abc’. • Delete all the dogs with gender male. • Retrieve the dogs with gender female.

Course Outcomes: On Successful completion of this course, students will be able to	
1	Apply the knowledge of database management system development process and conduct the experiments using SQL and NoSQL queries to find the solution for given database problem
2	Analyze and design solutions for database system components to meet the specified needs of online transaction processing and information systems like Banking systems, Ticket Reservation systems etc...
3	Develop code for stored programs and triggers.
4	Contribute to the team as a member, lead the team

Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
CO1	2		2		2							2		
CO2	2		2		2							2		
CO3			2		2							2		
CO4								2				2		
Overall CO	2		2		2			2				2		

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
	2		2		2			2				2		

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - V

MACHINE LEARNING TECHNIQUES (I)

Course Code	S5CCSI02	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:2)	SEE Marks	50
Credits	4	Exam Hours	3
Lecture Hours	42hrs	Practical Hours	28Hrs

Course objectives: This course will enable students to:

1	Define machine learning and understand the basic theory underlying machine learning.
2	Differentiate supervised, unsupervised and reinforcement learning.
3	Understand the basics of concept learning and decision trees.
4	Understand the working principles of neural networks models.
5	Understand the Bayesian learning, Instance based learning and genetic algorithms.
6	Understand the reinforcement learning.

UNIT I **9 Hours**

INTRODUCTION, CONCEPT LEARNING: Well Posed Learning problem, Perspectives and Issues in machine learning, A Concept Learning Task, Concepts Learning as Search, Fins-S: Finding a maximally specific Hypothesis, Version Spaces and Candidate Elimination Algorithm, Remarks on version space and Candidate Elimination.

DECISION TREE LEARNING: Decision Tree Representation, The Basic Decision Tree Algorithm
1.1,1.3,2.1-2.6, 3.1-3.4

UNIT II **9 Hours**

NEURAL NETWORKS: Introduction, Neural Network Representations, Appropriate problems for Neural Networks, Perceptrons: Representational Power of Perceptrons, Training Rule, Gradient Descent and Delta Rule, Multilayer Networks and Back Propagation Algorithms: A Differential Threshold Unit, The Back propagation Algorithm, Derivation of Back propagation Rule.

4.1-4.5

UNIT III **8 Hours**

BAYESIAN LEARNING: Introduction, Bayes Theorem, Bayes Theorem and Concept Learning, , Bayes Optimal Classifier, Gibbs Algorithm, Naive Bayes Classifier, Bayesian Belief networks, 6.1-6.3, 6.7-6.11 (6.11.1-6.11.4),

UNIT IV **8 Hours**

INSTANCE BASED LEARNING: Introduction K- Nearest Neighbor Learning, Locally Weighted Regression, Radial Basis Functions.

GENETIC ALGORITHMS: Motivation, Genetic Algorithms, Hypothesis Space Search, Genetic Programming.

8.1-8.4, 9.1-9.5,.

UNIT V **8 Hours**

REINFORCEMENT LEARNING: Introduction, Learning Task, Q-Learning, Nondeterministic Rewards and actions, Temporal Difference Learning, Generalizing from Examples, Relationship to Dynamic Programming.

13.1- 13.7

Course Outcomes: On Successful completion of this course, students will be able to

1	Apply concept learning and decision trees techniques to create models for classifying the given data.
2	Analyze the working principles of neural networks and effectively apply neural networks models to provide solution for the given applications.
3	Analyze the working of Bayesian models and apply the same to infer the observed data.
4	Apply instant based learning and genetic algorithms concepts to solve the given machine learning based applications.

5	Explain reinforcement learning concepts and apply concepts to solve the given applications.
---	---

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	2														2
CO2	2														2
CO3	2														2
CO4	2														2
CO5	2														2
Overall CO	2														2

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	2														2

MLT LAB Experiments: The following list of programs can be implemented using Python or Java programming language.

1. Implement and demonstrate the **FIND-S algorithm** for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file.
2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the **Candidate-Elimination algorithm** to output a description of the set of all hypotheses consistent with the training examples
3. Build an Artificial Neural Network by implementing the **Back propagation algorithm** and test the same using appropriate data sets.
4. Write a program to implement the **naïve Bayesian classifier** for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.
5. Write a program to construct a **Bayesian network** considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set.
6. Apply **EM algorithm** to cluster a set of data stored in a .CSV file. Use the same data set for clustering using **k-Means algorithm**. Compare the results of these two algorithms and comment on the quality of clustering.
7. Write a program to implement k-Nearest Neighbor algorithm to classify the iris data set. Print both correct and wrong predictions.
8. Implement the non-parametric **Locally Weighted Regression algorithm** in order to fit data points. Select appropriate data set for your experiment and draw graphs.

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - V

DATA MINING AND VISUALIZATION LABORATORY

Course Code	S5CCSL01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(0:0:2)	SEE Marks	50
Credits	1	Exam Hours	3
Lecture Hours	-	Practical Hours	28hrs

Course objectives: This course will enable students to:

1	To understand the basic concepts of data mining
2	To implement Data Mining techniques, their need, scenarios (situations) and scope of their applicability
3	Synthesize the solution through advanced data mining tool to solve data analytics problems

1. Experiment to be conducted using WEKA tool:

1	outlook	temperatu	humidity	windy	play
2					
3	sunny	85	85	FALSE	no
4	sunny	80	90	TRUE	no
5	overcast	83	86	FALSE	yes
6	rainy	70	96	FALSE	yes
7	rainy	68	80	FALSE	yes
8	rainy	65	70	TRUE	no
9	overcast	64	65	TRUE	yes
10	sunny	72	95	FALSE	no
11	sunny	69	70	FALSE	yes
12	rainy	75	80	FALSE	yes
13	sunny	75	70	TRUE	yes
14	overcast	72	90	TRUE	yes
15	overcast	81	75	FALSE	yes
16	rainy	71	91	TRUE	no

- i) *Preprocess(Data Cleaning, Data Integration, Data Transformation, Data Reduction)* and *Classify (Posteriori and Priori)* panels. Analyze Input and Output Attributes.
- ii) Calculate the information of the whole data set on the basis of whether play is held or not.
- iii) Draw the histogram to show how the values of the *play* class occurs for each value of the *outlook* attribute
- iv) Derive minimum and maximum values, mean, and standard deviation
- v) Perform operations such as filter, delete, invert, Pattern, Undo, Edit, search, Select, Conversions etc
- vi) Examine the Output , classification error and Kappa statistics
- vii) Visualize threshold curve
- viii) Apply Logistic Regression model to classify
- ix) Measure the log likelihood of the clusters of training data. (Consider large data set.)
- x) Derive Information gain
- xi) Build Decision Tree on Humidity attribute. Also demonstrate decision tree after analysis of
 - a) Sunny and Overcast dataset
 - b) Sunny, Overcast and Rainy Data set
- xii) Compute Gini Index representing with respect to Temperature, Humidity, and Windy attributes.
- xiii) Obtain the Prediction of Play 'Yes' as well as 'No' for an unknown instance
- xiv) Apply Naïve Bayes Classifier to the Weather play data set and derive the probability for play no given outlook rainy

	<p>xv) Apply classification on given data. Remove label and then apply clustering. Perform class to cluster evaluation. Apply classification on unlabelled dataset by removing play attribute. Prepare the analysis report.</p> <p>Interpret the results.</p>																																																																																				
2.	<p>Experiment using WEKA tool. Consider the following data set</p> <table border="1"> <thead> <tr> <th>No.</th> <th>eid Numeric</th> <th>ename Nominal</th> <th>salary Numeric</th> <th>exp Numeric</th> <th>address Nominal</th> </tr> </thead> <tbody> <tr><td>1</td><td>101.0</td><td>raj</td><td>10000.0</td><td>4.0</td><td>pdtr</td></tr> <tr><td>2</td><td>102.0</td><td>ramu</td><td>15000.0</td><td>5.0</td><td>pdtr</td></tr> <tr><td>3</td><td>103.0</td><td>anil</td><td>12000.0</td><td>3.0</td><td>kdp</td></tr> <tr><td>4</td><td>104.0</td><td>sunil</td><td>13000.0</td><td>3.0</td><td>kdp</td></tr> <tr><td>5</td><td>105.0</td><td>rajiv</td><td>16000.0</td><td>6.0</td><td>kdp</td></tr> <tr><td>6</td><td>106.0</td><td>sunitha</td><td>15000.0</td><td>5.0</td><td>nlr</td></tr> <tr><td>7</td><td>107.0</td><td>kavitha</td><td>12000.0</td><td>3.0</td><td>nlr</td></tr> <tr><td>8</td><td>108.0</td><td>suresh</td><td>11000.0</td><td>5.0</td><td>gtr</td></tr> <tr><td>9</td><td>109.0</td><td>ravi</td><td>12000.0</td><td>3.0</td><td>gtr</td></tr> <tr><td>10</td><td>110.0</td><td>ramana</td><td>11000.0</td><td>5.0</td><td>gtr</td></tr> <tr><td>11</td><td>111.0</td><td>ram</td><td>12000.0</td><td>3.0</td><td>kdp</td></tr> <tr><td>12</td><td>112.0</td><td>kavya</td><td>13000.0</td><td>4.0</td><td>kdp</td></tr> <tr><td>13</td><td>113.0</td><td>navya</td><td>14000.0</td><td>5.0</td><td>kdp</td></tr> </tbody> </table> <p>Use the data sources, like ARFF, XML ARFF files. Do the following</p> <ol style="list-style-type: none"> a. Classify , Invoke MultiLayerPerception b. Build neural network GUI as below ii) Beginning the process of editing the network to add a second hidden layer iii) The finished network with two hidden layers iv) Apply Lazy classifier, multi instance classifier v) Apply any MetaLearning Algorithm vi) Optimize base classifier's performance vii) Use clustering algorithm such as Cobweb, and Hierarchical Cluster viii) Select attribute by specifying an evaluator and a search method ix) Insert 30 more records in this file. Perform clustering on this new dataset. Identify optimum number of clusters. After identification of optimum number of clusters, prepare clustering on this number. x) Perform data analysis on the result obtained and prepare an analysis report for the same xi) Apply Apriori, Interpret the results. xii) Apply Association rules and interpret the results. xiii) Apply Association mining with the Apriori algorithm and find the best rules with threshold value of support of 50% and confidence of 70% xiv) Interpret the results obtained at Summary. Analyse Precision, Recall, F Score values. xv) Install R package at Weka environment and install rpart package. Implement decision tree. 	No.	eid Numeric	ename Nominal	salary Numeric	exp Numeric	address Nominal	1	101.0	raj	10000.0	4.0	pdtr	2	102.0	ramu	15000.0	5.0	pdtr	3	103.0	anil	12000.0	3.0	kdp	4	104.0	sunil	13000.0	3.0	kdp	5	105.0	rajiv	16000.0	6.0	kdp	6	106.0	sunitha	15000.0	5.0	nlr	7	107.0	kavitha	12000.0	3.0	nlr	8	108.0	suresh	11000.0	5.0	gtr	9	109.0	ravi	12000.0	3.0	gtr	10	110.0	ramana	11000.0	5.0	gtr	11	111.0	ram	12000.0	3.0	kdp	12	112.0	kavya	13000.0	4.0	kdp	13	113.0	navya	14000.0	5.0	kdp
No.	eid Numeric	ename Nominal	salary Numeric	exp Numeric	address Nominal																																																																																
1	101.0	raj	10000.0	4.0	pdtr																																																																																
2	102.0	ramu	15000.0	5.0	pdtr																																																																																
3	103.0	anil	12000.0	3.0	kdp																																																																																
4	104.0	sunil	13000.0	3.0	kdp																																																																																
5	105.0	rajiv	16000.0	6.0	kdp																																																																																
6	106.0	sunitha	15000.0	5.0	nlr																																																																																
7	107.0	kavitha	12000.0	3.0	nlr																																																																																
8	108.0	suresh	11000.0	5.0	gtr																																																																																
9	109.0	ravi	12000.0	3.0	gtr																																																																																
10	110.0	ramana	11000.0	5.0	gtr																																																																																
11	111.0	ram	12000.0	3.0	kdp																																																																																
12	112.0	kavya	13000.0	4.0	kdp																																																																																
13	113.0	navya	14000.0	5.0	kdp																																																																																
3.	Consider the data set given below																																																																																				

Relation: employee					
No.	age Nominal	income Nominal	stud Nominal	creditrate Nominal	buyscomp Nominal
1	L20	high	no	fair	yes
2	20-40	low	yes	fair	yes
3	G40	medium	yes	fair	yes
4	L20	low	no	fair	no
5	G40	high	no	excellent	yes
6	L20	low	yes	fair	yes
7	20-40	high	yes	excellent	no
8	G40	low	no	fair	yes
9	L20	high	yes	excellent	yes
10	G40	high	no	fair	yes
11	L20	low	yes	excellent	no
12	G40	high	yes	excellent	no
13	20-40	medium	yes	excellent	yes
14	L20	medium	yes	fair	yes
15	G40	high	yes	excellent	yes

- a. Load ARFF file and explore knowledge flow interface
- b. Configure the data source , check the status area after executing the configuration
- c. Perform operations such as Attribute Selection, Filter, Classify, Data Sink, Visualization and Evaluation
- d. Apply incremental learning and analyze the result
- e. Do clustering : use generator properties, two clustering schemes, and result panel
- f. Generate Confusion Matrix and Interpret the results
- g. Construct Decision tree
- h. Perform Linear Regresssion and Analyze , Validate and Visualize the data
Apply Association mining with the Apriori alforithm and find the best rules with threshold value of support of 50% and confidence of 70%
- i. Interpret the results obtained at Summary. Analyse Precision, Recall, F Score values.

I. Consider glass data set.

- i) How many attributes are there in the dataset? What are their names? What is the class attribute? Run the classification algorithm IBk (weka.classifiers.lazy.IBk). Use cross-validation to test its performance, leaving the number of folds at the default value of 10.
- ii) What is the accuracy of IBk (given in the Classifier Output box)? Run IBk again, but increase the number of neighboring instances to $k = 5$ by entering this value in the KNN field. Use cross-validation as the evaluation method.
- iii) What is the accuracy of IBk with five neighboring instances ($k = 5$)?
- iv) Obtain best accuracy higher than the accuracy obtained on the full dataset. Verify ,Is this best accuracy an unbiased estimate of accuracy on future data?
- v) Record the cross-validated accuracy estimate of IBk for 10 different percentages of class noise and neighborhood sizes
- vi) Analyze, What is the effect of increasing the amount of class noise?
- vii) Analyze, What is the effect of altering the value of k ?
- viii) Verify the amount of training data

Additional: For both problems defined under I. and II. Use R package installed at Weka environment, derive decision tree. Interpret the results.

4. Set up SQL database and insert sample data consisting of customer data. Integrate SQL database with Weka.
 - j. Perform data preprocessing, classification, clustering tasks on customer data.
 - k. Apply filters and interpret the results
 - l. Connect WEKA to a relational database using JDBC.

	<ul style="list-style-type: none"> m. Retrieve customer data directly using SQL queries. n. Load data into WEKA's Explorer interface. o. Apply classification algorithms (e.g., J48, Naive Bayes) to predict customer spending behavior. p. Use clustering (e.g., k-means) for market segmentation. q. Generate evaluation metrics (accuracy, precision, recall). r. Optionally automate data refresh using scripts or the WEKA KnowledgeFlow interface. Validate and analyse the result <p>Additional: Derive Fact Table, Star schema, Snowflake schema and do the comparison</p>
5.	<p>Using Tableau do the following. Consider Titanic Data Set. Perform the following using Tableau platform:</p> <ol style="list-style-type: none"> i. Perform Calculations: <ul style="list-style-type: none"> i) Calculate the survival rate. Hint: you can use the following formula $SUM(IIF([Survived] = 1, 1, 0)) / COUNT([PassengerId])$ ii) Calculate the average iii) Calculate the total fare 2. Perform Group Operations: <ul style="list-style-type: none"> i) Create a group to categorize passengers by age ii) Create a group to categorize passengers by fare iii) Create a set to include only passengers who survived 3. Perform Set Operations: <ul style="list-style-type: none"> i) Create a set to include only passengers who survived ii) Create a set to include only passengers who did not survive iii) Create a set to include only passengers who traveled in first class 4. Create Dashboard <ul style="list-style-type: none"> i). Survival Rate by Age Group: Create a bar chart to display the survival rate by age group. ii). Average Fare by Fare Group: Create a bar chart to display the average fare by fare group. iii). Survivors by Class: Create a pie chart to display the number of survivors by class 5. Add Additional Visualizations 6. Combine Dashboards
6.	<p>Using Tableau platform, do the following. Consider Brazilian E-Commerce Public Dataset by Olist from Kaggle. Perform the following using Tableau platform:</p> <ol style="list-style-type: none"> i) Create a dashboard that displays the total sales, average sales price, and sales quantity for each state in Brazil, filtered by city, with custom geocoding and map layers. Generate mobile responsive dashboard ii) Calculate Total Sales per State, Average Sales Price per State, Sales quantity per state, Forecast sales by region iii) Generate map view iv) Interpret the KPI card v) Demonstrate Advanced Mapping technique vi) Use custom TopoJSON map of Brazil with state and city boundaries. Add heatmap to show density of sales vii) Demonstrate slicers. Add dynamic slicers. Sync slicers across multiple pages <p>Additional: Enable data export by filtered view and configure email alerts or subscriptions when sales in any region drop below a threshold</p>
7.	<p>Consider Sales Data set. Perform the following using powerBI platform.</p> <ol style="list-style-type: none"> 1. Create tables with <ul style="list-style-type: none"> i) manufacturing, sum of sales, columns. ii) Product Category, sum of profits etc iii) Find Answers for the following: <ul style="list-style-type: none"> a) Which Manufacturer has the highest Sales? b) Which Product Category has the lowest Profit value? c) Which Channel has the highest Cost of Sales? d) Which Manufacturer has the highest Profit?

	<p>e) Which Promotion Name has the highest Sales? f) Which Product Sub Category has the highest Profit?</p> <p>iv) Perform cross filtering between tables v) Create matrix visualization for product category , Region, sum of sales vi) Create a card visualtion for total sales, Total Profits, Avg Profits, Highest Profits, Lower Profits etc. and apply formatting</p> <p>2. Apply appropriate visualization. Create calculations for Visualization i) Calculate the difference between sum of Sales and Sum of Profits ii) Calculalte Profit Ratio. (Hint: Sum of Profit/Sum of Sales). Also demonstrate using in built funciton, e.g., DIVIDE, AVERAGE etc iii) Calculalte the average sales per product for each of manufacturers. (Hint: Sum of sales/Count of Product Name) iv) Calculalte the Percentage of grand total using inbuilt function</p> <p>3. Apply Filters and Slicers</p> <p>4. Build Graphs. Draw Trend Analysis Graph. Show trends and forecasting. For example: sum of sales per year, month, quarter etc</p> <p>5. Create Interactive Dashboard i) Create an Interactive Report The Sales Manager would like you to please create the following interactive report in Power BI Desktop: a) Create a heading - Sales Report b) Create the following Card visualizations: i. Total Sales ii. Total Profit iii. Average Sales iv. Number of Products v. Create a Area graph displaying Sales by Year and Quarter vi. Create a Column graph displaying Profit by Product Category</p>
8.	<p>Consider Sales Data set. Integrate R environment to power BI platform and perform the following i) Create Smart Narrative ii) Using R packages, Build ML based Data Model iii) Perform Visulaization and build report</p>
9.	<p>Consider Sales Data set. Integrate R environment to power BI platform and perform the following i) Create Q&A Visuals ii) Using R packages, Build ML based classification Model iii) Perform Visulaization and build report</p>
10.	<p>Consider Sales Data set. Integrate R environment to power BI platform and perform the following i) Create decomposition tree ii) Using R packages, Build ML based Clustering Model iii) Perform Visulaization and build report</p>
11.	OPEN ENDED MICRO PROJECT

Course Outcomes: On Successful completion of this course, students will be able to	
1	To understand the need for Data Mining and advantages to the business world
2	To get a clear idea of various classes of Data Mining techniques, their need, scenarios (situations) and scope of their applicability
3	To learn the algorithms used for various type of Data Mining problems
4	To understand how to explore and communicate data using data visualization techniques using data analytics tools such as PowerBI and Tableau
5	To derive solutions where business intelligence analytics is applicable.

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
CO1			2											2
CO2			2											2
CO3			2											2
CO4			2											2
Overall CO			2											2

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
			2											2

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - V

RESEARCH METHODOLOGY AND IPR

Course Code	SHS04	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	3
Lecture Hours	42 hrs	Practical Hours	-

Course objectives: This course will enable students to:

1	Understand the main knowledge representation and their reasoning.
2	learn to solve different reasoning tasks being aware of their complexity
3	Identify the procedure to develop propositional logic and reasoning using horn clauses,
4	Learn to formulate rules in production system and inheritance networks.
5	Discuss object oriented representation and degree of belief to quantify uncertainty.
6	Analyze the procedure to planning in a situation calculus.

UNIT I (9 Hours)

RESEARCH METHODOLOGY: Objectives and motivation of research - Types of research - Research approaches - Significance of research - Research methods verses methodology - Research and scientific method - Importance of research methodology - Research process - Approaches of investigation of solutions for research problem, data collection, analysis, interpretation, necessary instrumentations- Criteria of good research. Defining the research problem: Definition of research problem - Problem formulation - Necessity of defining the problem - Technique involved in defining a problem.

UNIT II (9 Hours)

LITERATURE SURVEY AND DATA COLLECTION: Importance of literature survey - Sources of information - Assessment of quality of journals and articles - Information through internet. Effective literature studies approaches, analysis, plagiarism, and research ethics. Data - Preparing, Exploring, examining and displaying

UNIT III (8 Hours)

RESEARCH DESIGN AND ANALYSIS: Meaning of research design - Need of research design - Different research designs - Basic principles of experimental design - Developing a research plan - Design of experimental set-up - Use of standards and codes. Overview of Multivariate analysis, Hypotheses testing and Measures of Association. Presenting Insights and findings using written reports and oral presentation..

UNIT IV (8 Hours)

INTELLECTUAL PROPERTY RIGHTS (IPR): Nature of Intellectual Property: Patents, Designs, Trade and Copyright. Process of Patenting and Development: technological research, innovation, patenting, development. Role of WIPO and WTO in IPR establishments, Right of Property, Common rules of IPR practices, Types and Features of IPR Agreement, Trademark, Functions of UNESCO in IPR maintenance

UNIT V (8 Hours)

PATENT RIGHTS (PR): Patent Rights: Scope of Patent Rights. Licensing and transfer of technology. Patent information and databases. Geographical Indications. New Developments in IPR: Administration of Patent System, IPR of Biological Systems, Computer Software etc. Traditional knowledge Case Studies, IPR and IITs. Licenses, Licensing of related patents, patent agents, Registration of patent agents.

Course Outcomes: On Successful completion of this course, students will be able to

1	Describe the research process & formulate research problem
2	Perform literature review, manage data & practice research ethics
3	Practice basic principles of experimental design, use standard codes and carry out research analysis
4	Distinguish between types of innovation, describe patenting procedure, maintenance and role of IPR establishments
5	Identify the significance of patent rights, licensing, technology transfer & manage patenting system

Text Books:		
Sl. No.	Author/s	Title, Publisher, Edition, Year, ISBN
1.	Peter S. Menel Mark A. Lemley, Robert P. Merges	"Intellectual Property in the New Technological-Vol. I Perspectives, 2021.
2.	Laura R. Ford	"The Intellectual Property of Nations: Sociological and Historical Perspectives on a Modern Legal Institution Paperback -2021.
Reference Book:		
Sl. No.	Author/s	Title, Publisher, Edition, Year, ISBN
1.	R. Ganesan	"Research Methodology for Engineers", MJP Publishers, Chennai, 2011.
2.	Cooper Donald R, Schindler Pamela S and Sharma JK	"Business Research Methods", Tata McGraw Hill Education, 11 th Edition, 2012.
3.	Catherine J. Holland	"Intellectual property: Patents, Trademarks, Copyrights, Trade Secrets", Entrepreneur Press, 2007.
4.	David Hunt, Long Nguyen, Matthew Rodgers	"Patent searching: tools & techniques", Wiley, 2007.
5.	The Institute of Company Secretaries of India, Statutory body under an Act of parliament	"Professional Programme Intellectual Property Rights, Law and practice", September 2013.

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
 Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - V

INFORMATION RETRIEVAL

Course Code	S5CCSPE01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	3
Lecture Hours	42Hrs	Practical Hours	-

Course objectives: This course will enable students to:

1	Examine the Information Retrieval problems and describe the architecture of a search engine
2	Analyse Search structures of dictionaries, Wildcard queries and Index construction done information retrieval
3	Understand the scoring and ranking in Information retrieval system
4	Study the various Information Retrieval Evaluation Techniques and processes involved
5	Understand how web search, web crawling and link analysis is done for information retrieval on the web

UNIT I

(9 Hours)

Introduction

Search Engines and Information Retrieval : What Is Information Retrieval? , The Big Issues, Search Engines, Search Engineers

Architecture of a Search Engine: What Is an Architecture? ,Basic Building Blocks ,Breaking It Down, Text Acquisition Text Transformation , Index Creation, User Interaction, Ranking ,Evaluation, How Does It Really Work?

Boolean retrieval: An example information retrieval problem , A first take at building an inverted index, Processing Boolean queries, The extended Boolean model versus ranked retrieval

Vocabulary and postings lists: Document delineation and character sequence decoding: Obtaining the character sequence in a document, Choosing a document unit

Determining the vocabulary of terms: Tokenization, Dropping common terms: stop words, Normalization (equivalence classing of terms), Stemming and lemmatization, Faster postings list intersection via skip pointers

Positional postings and phrase queries: Biword indexes, Positional indexes, Combination schemes.

UNIT II

(9 Hours)

Dictionaries and tolerant retrieval

Search structures for dictionaries, Wildcard queries: General wildcard queries , k-gram indexes for wildcard queries , Spelling correction: Implementing spelling correction, Forms of spelling correction, Edit distance, k-gram indexes for spelling correction, Context sensitive spelling correction, Phonetic correction.

Index construction :

Hardware basics, Blocked sort-based indexing, Single-pass in-memory indexing, Distributed indexing

Index Compression:

Dictionary Compression, Dictionary as a string, Blocked storage., Postings file compression: Variable byte codes, γ codes.

UNIT III

(8 Hours)

Scoring, term weighting and the vector space model : Parametric and zone indexes : Weighted zone scoring , Learning weights, The optimal weight g, Term frequency and weighting: Inverse document frequency , Tf-idf weighting , The vector space model for scoring: Dot products, Queries as vectors, Computing vector scores, Variant tf-idf functions: Sublinear tf scaling, Maximumtf normalization, Document and query weighting schemes, Pivoted normalized document length

Computing scores in a complete search system

Efficient scoring and ranking: Inexact top K document retrieval, Index elimination, Champion lists, Static quality scores and ordering, Impact ordering, Cluster pruning.

Components of an information retrieval system: Tiered indexes, Query-term proximity, Designing parsing and scoring functions, Putting it all together , Vector space scoring and query operator interaction.

UNIT IV	(8 Hours)
<p>Evaluation in information retrieval: Information retrieval system evaluation, Standard test collections, Evaluation of unranked retrieval sets, Evaluation of ranked retrieval results, Assessing relevance : Critiques and justifications of the concept of relevance , A broader perspective-System quality and user utility: System issues, User utility, Refining a deployed system, Results snippets.</p> <p>XML retrieval: Basic XML concepts , Challenges in XML retrieval, A vector space model for XML retrieval, Evaluation of XML retrieval, Text-centric vs. data-centric XML retrieval.</p> <p>Language models for information retrieval : Language models: Finite automata and language models, Types of language models, Multinomial distributions over words , The query likelihood model: Using query likelihood language models in IR, Estimating the query generation probability, Ponte and Croft’s Experiments, Language modelling versus other approaches in IR, Extended language modelling approaches Understanding Large Language Models, Retrieval: The Role of Large, Language Models in a Post-Search Engine Era.</p>	

UNIT V	(8 Hours)
<p>Web search basics : Background and history Web characteristics: The web graph, Spam,Advertising as the economic model, The search user experience: User query needs, Index size and estimation, Near-duplicates and shingling https://www.youtube.com/watch?v=DkspjZRYD8s</p> <p>Web crawling and indexes: Overview: Features a crawler must provide, Features a crawler should provide, Crawling: Crawler architecture: DNS resolution, The URL frontier, Distributing indexes, Connectivity servers</p> <p>Link analysis: The Web as a graph: Anchor text and the web graph, PageRank: Markov chain, The PageRank computation, Topic-specific PageRank , Hubs and Authorities: Choosing the subset of the Web.</p>	

Course Outcomes: On Successful completion of this course, students will be able to	
1	Analyse the Information Retrieval problems and illustrate the architecture of a Search Engine
2	Apply Search structures of dictionaries, Wildcard queries and Index construction for information retrieval.
3	Perform scoring and ranking on a search system
4	Apply Information Retrieval Evaluation Techniques on Search engines.
5	Perform web search, web crawling and link analysis for information retrieval on the web

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Introduction to Information Retrieval	C. Manning, P. Raghavan, and H. Schutze, 2008.	Cambridge University Press	1st Edition, 2009
2	Search Engines: Information Retrieval in Practice	Bruce Croft, Donald Metzler and Trevor Strohman	Addison Wesley	2nd Edition, 2015
3	Build a Large Language Model (From Scratch) https://livebook.manning.com/book/build-a-large-language-model-from-scratch/chapter-1/v-2/ https://www.manning.com/books/build-a-large-language-model-from-scratch	Sebastian Raschka; C. Manning	Manning Books	MEAP August 2024

Additional Resource : https://medium.com/@daniele.nanni/revolutionizing-information-retrieval-the-role-of-large-language-models-in-a-post-search-engine-7dd370bdb62				
Reference Books				
1	Modern Information Retrieval: The Concepts and Technology behind Search	Ricardo Baeza -Yates and Berthier Ribeiro - Neto	ACM Press	2nd Edition, 2011
2	Information Retrieval Implementing and Evaluating Search Engines	Stefan Buttcher Charles L. A. Clarke Gordon V. Cormack	MIT Press	1st Edition, February 2016

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1		2											2		
CO2	2												2		
CO3			2										2		
CO4			2										2		
CO5		2											2		
Overall CO	2	2	2										2		

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	2	2	2										2		

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE Outcome Based Education (OBE) and Choice Based Credit System (CBCS) SEMESTER - V			
BUSINESS INTELLIGENCE AND ANALYTICS			
Course Code	S5CCSPE03	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	3
Lecture Hours	42Hrs	Practical Hours	-

Course objectives: This course will enable students to:	
1	Outline the Decision Support systems and Business Intelligence framework.
2	Synthesize Data warehousing, Data Mining and Web Mining with concepts, architecture and Extraction, Transformation, and Load (ETL) Processes.
3	Explore and Derive Solutions using Descriptive Analytics and Hypothesis Testing
4	Illustrate the significance of computerized Decision Support and understand the mathematical modelling behind decision support using Correlation Analysis, Prescriptive Analysis and Clustering algorithms.
5	Synthesize Business Process Management , its activities, approaches, and its implementation in Industrial application

UNIT I	(8 Hours)
An Overview of Business Intelligence, Analytics and Decision Support: Changing Business Environments and Computerized Decision Support, A Framework for Business Intelligence (BI), Transaction Processing Versus Analytic Processing Textbook 1 : Chapter 1: 1.2, 1.3, 1.5 Why Analytics? Business Analytics, Descriptive Analytics, Predictive Analytics, Prescriptive Analytics, Techniques, Big Data Analytics, Web and Social Media Analytics, Framework for Data Driven Decision Making, Analytics Capacity Building, Roadmap for Analytics Capability Building Textbook 2 : Chapter 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 1.10, 1.12, 1.13, 1.14	

UNIT II	(8 Hours)
Data Warehousing: Definitions and Concepts, Process Overview, Architectures, ETL processes Text book 1: Chapter 2.2, 2.3, 2.4, 2.5 Data Mining: Concepts and Applications, Process, Methods Text book 1: Chapter 4.2, 4.3, 4.4, 4.5 Text and Web Analytics: Text mining Process, Sentiment Analysis, Web mining Overview, Search Engines, Web Usage Mining, Social Analytics Text book 1: Chapter 5.5, 5.6, 5.7, 5.8, 5.9, 5.10	

UNIT III	(9 Hours)
Descriptive Analytics: Introduction, Data types and Scale of variable, Types of Variable Measurement scales, Population and Sample, Data Visualization Text book 2: 2.1, 2.2, 2.3, 2.4, 2.9 Hypothesis Testing: Introduction, Setting up a hypothesis testing, One-Tailed and Two-Tailed Tests, Type 1 Error, Type II Error and Power of Hypothesis Test Text Book 2: 6.1, 6.2, 6.3, 6.4 Analysis of Variance: Introduction to Anova, Multiple t-Tests for comparing Several Means, One-way Anova, Two- way Anova Text Book 2: 7.1, 7.2, 7.3, 7.4 Case Study: Testing Marketing Hypotheses at WSES (Text Book 2 : Chapter 6, pg. 160 to 164) Self Study Components: Measures of Central Tendency, Percentile, Decile and Quartile, Measures of Variation, Measures of Shape Text book 2: 2.5, 2.6, 2.7, 2.8	

UNIT IV	(9 Hours)
Correlation Analysis: Introduction, Pearson Correlation Coefficient Text Book 2: Chapter 8.1, 8.2	

Clustering: Clustering Algorithms, K-Means Clustering, Hierarchical Clustering, Case study: Mark down Optimization for an Indian Apparel Retailer
Text Book 2: Chapter 14.4, 14.5, 14.6

Prescriptive Analytics: Introduction, Linear programming, LP Model Building, Sensitivity Analysis in LPP, Primal-Dual Relationships, Multi Period Models, Multi Criteria Decision Making Problems, Case Study: Linen Management at Apollo Hospital
Text Book 2: Chapter 15.1, 15.2, 15.3, 15.6, 15.11, 15.12, 15.14)
Self Study Components: Distance and Similarity Measures, Linear Programming Problem Terminologies, Assumptions, Solving LP using Graphical Methods, ILP (Text book 2: Chapter 14.2, 15.4, 15.5, 15.7, 15.13)

UNIT V

8 Hours

Business Reporting, Visual Analytics, and Business Performance Management:

Opening Vignette – Self Service Reporting Environment Saves Millions For Corporate Customers, Business Reporting Definitions and Concepts, Data and Information Visualization, Emergence of Data Visualization and Visual Analytics, Performance Dashboards, Business Performance Management, Performance Management, Balanced Scorecards, Six Sigma as performance measurement System.

Text Book 1: Chapter 3: 3.1, 3.2, 3.3, 3.5, 3.6, 3.7, 3.8, 3.9, 3.10

Case Study: Era of Quality at the Akshaya Patra Foundation (Text book 2: Chapter 18)

Self Study Components:

Six Sigma: Six Sigma Measures, Defects per Million Opportunities, Yield, Sigma Score, DMAIC Methodology, Six Sigma Project Selection for DMAIC implementation, DMAIC Methodology – case of Armoured Vehicle, Tool box

Text book 2: Chapter 18

Course Outcomes: On Successful completion of this course, students will be able to

1	Apply the basics of data and business to understand Decision Support systems and Business Intelligence framework.
2	Describe the significance of computerised Decision Support, apply the basics of mathematics to understand the mathematical modelling behind decision support.
3	Explain Data warehousing, its architecture and Extraction, Transformation, and Load (ETL) Processes.
4	Analyse the importance of knowledge management and explain its activities, approaches and its implementation.
5	Describe the Expert systems and analyse its development, discuss areas suitable for application of experts' system.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Business Intelligence – A managerial Perspective on Analytics	Ramesh Sharda et.al	Pearson.	3rd Edition ,2019
2	Business Analytics – The Science of Data Driven Decision Making	U Dinesh Kumar	Wiley	2 th edition,2022
Reference Books				
1	Decision support systems and intelligent systems	Efraim Turban, Jay E. et a	PHI Editon	7th Edition, 2010
2	Data Science for Business	Foster Provost & Tom Fawcett	O'Reilly Media, Inc	2013

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	2	2													2
CO2		2													2
CO3		2													2
CO4		2													2
CO5		2													2
Overall CO	2	2													2

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	2	2													2

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE Outcome Based Education (OBE) and Choice Based Credit System (CBCS) SEMESTER - V			
KNOWLEDGE REPRESENTATION AND REASONING			
Course Code	S5CCSPE04	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	3
Lecture Hours	42Hrs	Practical Hours	-

Course objectives: This course will enable students to:	
1	Understand the main knowledge representation and their reasoning.
2	learn to solve different reasoning tasks being aware of their complexity
3	Identify the procedure to develop propositional logic and reasoning using horn clauses,
4	Learn to formulate rules in production system and inheritance networks.
5	Discuss object oriented representation and degree of belief to quantify uncertainty.
6	Analyze the procedure to planning in a situation calculus.

UNIT I	(9 Hours)
<p>Introduction: The Key concepts: Knowledge representation and Reasoning, why Knowledge representation and Reasoning?, the role of logic. The language of first order logic: Introduction, the syntax, the semantics, the pragmatics, the explicit and implicit beliefs.</p> <p>Expressing Knowledge: Knowledge engineering, vocabulary, basic facts, complex facts, terminology facts, entailments, abstract individuals, other sorts of facts.</p>	

UNIT II	(9 Hours)
<p>Resolution: The propositional case, handling variables and quantifiers, dealing with computational intractability. Reasoning with Horn clauses: Horn clauses, SLD resolution, computing SLD derivations.</p> <p>Procedural control of Reasoning: Facts and rules, rule formation and search strategy, algorithm design, specifying goal order, committing to proof methods, controlling backtracking, negation as failure, dynamic databases.</p>	

UNIT III	(8 Hours)
<p>Rules in production system: Production system basic operation, working memory, production rule, a first example, a second example, conflict resolution, making production system more efficient, application and advantages, some significant production rule systems.</p> <p>Inheritance: Inheritance networks, strategies for defeasible inheritance, a formal account of inheritance networks.</p>	

UNIT IV	(8 Hours)
<p>Object oriented representation: Objects and frames, a basic frame formalism, an example: using frames to plan a trip, beyond the basics.</p> <p>Vagueness, uncertainty, and degree of belief: Noncategorical reasoning, objective probability, subjective probability, vagueness.</p>	

UNIT V	(8 Hours)
<p>Planning: planning in a situation calculus, the STRIPS representation, planning as a reasoning task, beyond the basics.</p>	

Course Outcomes: On Successful completion of this course, students will be able to	
1	Define knowledge representation and reasoning the basic theory underlying knowledge engineering.
2	Illustrate propositional logic, horn clauses, and procedural control of reasoning.
3	Summarize the rules in production system and inheritance.
4	Analyse object oriented representation, vagueness, uncertainty and belief
5	Discuss planning in a situation calculus and STRIPS representation

TEXT BOOKS		
1	R. Brachman & H. Levesque,	Knowledge Representation and Reasoning," Morgan-Kaufmann, First edition, 2004.

REFERENCE BOOKS		
1	Dean Allemang, James Hendler	Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL, 2nd edition, 2011
2	Prof. Deepak Khemani	NOC: Artificial Intelligence: Knowledge Representation and Reasoning , IIT Madras Link: https://nptel.ac.in/courses/106106140

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	2		2										2		
CO2	2	2											2		
CO3	2		2										2		
CO4	2		2										2		
CO5	2	2											2		
Overall CO	2	2	2										2		

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	2	2	2										2		

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
 Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - V

RECOMMENDATION SYSTEM

Course Code	S5CCSPE05	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	03
Lecture Hours	42hrs	Practical Hours	-

Course objectives: This course will enable students to:

1	To understand basic techniques and problems in the field of recommender systems.
2	Evaluate Types of recommender systems.
3	Apply algorithms and techniques to develop Recommender Systems that are widely used.
4	To develop state-of-the-art recommender systems...

UNIT I (9 Hours)

Introduction: Introduction to basic concepts, Recent developments, **Collaborative recommendation:** User-based nearest neighbor recommendation, Item-based nearest neighbor recommendation, About ratings, Further model-based and preprocessing-based approaches, Recent practical approaches and systems. Attacks on collaborative recommender systems.

(Text Book-1: 1,2.1-2.5,9.1-9.6)

UNIT II (9 Hours)

Content-based recommendation: Content representation and content similarity, Similarity-based retrieval, Other text classification methods. **Knowledge-based recommendation:** Knowledge representation and reasoning, Interacting with constraint-based recommenders, Interacting with case-based recommenders, Example applications.

(Text Book-1: 3.1-3.3,4.1-4.5)

UNIT III (8 Hours)

Hybrid recommendation approaches: Opportunities for hybridization, Monolithic hybridization design, Parallelized hybridization design, Pipelined hybridization design. **Evaluating recommender systems:** Introduction, General properties of evaluation research, Popular evaluation designs, Evaluation on historical datasets, Alternate evaluation designs.

(Text Book-1: 5.1-5.4,7.1-7.5)

UNIT IV (8 Hours)

Structural Recommendations in Networks: Introduction, **Ranking Algorithms-** PageRank, Personalized PageRank, Applications to Neighborhood-Based Methods, Social Network Recommendations, Personalization in Heterogeneous Social Media, Traditional Collaborative Filtering, SimRank, The Relationship Between Search and Recommendation. **Recommendations by Collective Classification-** Iterative Classification Algorithm, Label Propagation with Random Walks, Applicability to Collaborative Filtering in Social Networks. **Recommending Friends: Link Prediction-** Neighborhood-Based Measures, Katz Measure Random Walk-Based Measures, Link Prediction as a Classification Problem, Matrix Factorization for Link Prediction, Symmetric Matrix Factorization, **Connections Between Link Prediction and Collaborative Filtering-** Using Link Prediction Algorithms for Collaborative Filtering, Using Collaborative Filtering Algorithms for Link Prediction.

(Text Book-1: 10)

UNIT V (8 Hours)

Advanced Topics in Recommender Systems: Introduction, Learning to rank, Multi armed Buidit Algorithms, Group Recommender Systems, Multi-Criteria Recommender Systems, Active learning in recommender systems, Privacy in recommender systems, Some interesting application domains.

(Text Book-2: 13)

Course Outcomes: On Successful completion of this course, students will be able to	
1	Describe the concept of collaborative recommendation system.
2	Describe the concept of content-based and knowledge-based recommendation system.
3	Describe the concept of hybrid recommendation and understand the evaluation methods for recommendation systems.
4	Describe the concept of recommendation for networks.
5	Describe some advanced topics of recommender systems like Group Recommender Systems.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Recommender Systems: An Introduction,	Jannach D., Zanker M. and FelFering A., Friedrich G.	Cambridge University Press.,	2011
2	Recommender Systems,	Charu C. Aggarwal	Springer International Publishing Switzerland,	2016.
Reference Books				
1	Recommender Systems Handbook	Ricci F., Rokach L., Shapira D., Kantor B.P.	Springer	2011.
2	Recommender Systems For Learning.	Manouselis N., Drachsler H., Verbert K., Duval E.	Springer	2013

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	2		2										2		
CO2	2		2										2		
CO3	2		2										2		
CO4	1		1										2		
CO5	2		2										2		
Overall CO	2		2										2		

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO 1-5	2		2										2		

VI SEM

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - VI

REAL TIME BIG DATA ANALYTICS(I)

Course Code	S6ADI01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:2)	SEE Marks	50
Credits	04	Exam Hours	03
Lecture Hours	42hrs	Practical Hours	28hrs

Course objectives: This course will enable students to:

1	Describe the basic paradigms, data model, evolution for Big Data (L2)
2	Explain the importance of a serialization framework and limitations of serialization frameworks for Big Data (L2)
3	Analyze how the data is stored on the batch layer
4	Design of the batch layer starting from ingesting new data to computing batch views
5	Illustrate how to build the serving layer for Big data
6	Describe the real time views of Cassandra's data model for Big data
7	Demonstrate how to implement the concepts of queuing and stream processing using real-world tools..

UNIT I

9 Hours

A new paradigm for Big Data: Scaling with a traditional database – No SQL is not a panacea - First principles - Desired properties of a Big Data system - The problems with fully incremental - Lambda Architecture - Recent trends in technology - Example application: SuperWebAnalytics.com.

Data model for Big Data: The properties of data - The fact-based model for representing data - Graph - A complete data model for SuperWebAnalytics.com.

UNIT II

9 Hours

Data storage on the batch layer: Storage requirements for the master dataset - Choosing a storage solution for the batch layer - How distributed file systems work - Storing a master dataset with a distributed file system - Vertical partitioning - Low-level nature of distributed file systems - Storing the SuperWebAnalytics.com master dataset on a distributed file system

Batch layer: Computing on the batch layer, Re-computation algorithms vs. incremental algorithms, Scalability in the batch layer, MapReduce: a paradigm for Big Data computing, Low-level nature of MapReduce, Pipe diagrams: a higher-level way of thinking about batch computation

UNIT III

8 Hours

Batch layer - Architecture and algorithms: Design of the SuperWebAnalytics.com batch layer - Workflow overview - Ingesting new data - URL normalization - User-identifier normalization - Deduplicate pageviews - Computing batch views

Batch layer: Implementation: Starting point - Preparing the workflow - Ingesting new data - URL normalization - User-identifier normalization - Deduplicate pageviews - Computing batch views

UNIT IV

8 Hours

Serving layer: Performance metrics for the serving layer - The serving layer solution to the normalization/denormalization problem - Requirements for a serving layer database - Designing a serving layer for SuperWebAnalytics.com - Contrasting with a fully incremental solution.

Realtime views : Computing realtime views - Storing realtime views - Challenges of incremental computation - Asynchronous versus synchronous updates - Expiring realtime views.

UNIT V

8 Hours

Queuing and stream processing: Queuing, Stream processing, Higher-level, one-at-a-time stream processing, SuperWebAnalytics.com speed layer

Queuing and stream processing: Illustration: Defining topologies with Apache Storm, Apache Storm clusters and deployment, Guaranteeing message processing

Course Outcomes: On Successful completion of this course, students will be able to

1	Apply the basic knowledge related to Big data to explain its elements, its analytics, its usage in business
---	---

	context..
2	Illustrate data storage on the batch layer using the Hadoop Distributed File System
3	Design and develop batch layer to the solution of various real world application problems in the context of master data
4	Design and develop serving layer to the solution of various real world application problems in the context of master data
5	Design and develop speed layer considering the concepts of queuing and stream processing using real-world tools.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Big Data - PRINCIPLES AND BEST PRACTICES OF SCALABLE REAL-TIME DATA SYSTEMS	NATHAN MARZ with JAMES WARREN	Manning Publications	2015 Edition
2	Spark in Action	Petar Zečević Marko Bonaći	Manning Publications	Nov 2016 Edition
Reference Books				
1	Hadoop: The Definitive Guide	Tom White	O'reilly Media	4 th Edition,2015
2	Big Data and Analytics	Seema Acharya, Subhashini Chellappan	Wiley India Publications,	May 2015
3	Big Data Black book	D T Editorial Services	Dream tech press	2016 Edition

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	3														2
CO2					2										2
CO3			3												3
CO4		2			2										2
CO5					2										2
Overall CO	3	2	3	-	2	-	-	-	-	-	-	-	-	-	3

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	3	2	3	-	2	-	-	-	-	-	-	-	-	-	3

Lab program using PySpark

<p>The fact-based model for representing data: Demonstrate the followings</p> <ul style="list-style-type: none"> Setting up the pyspark shell Reading and ingesting data into a data frame Exploring data in the DataFrame structure Moving from a sentence to a list of words Reshaping and Filtering rows
<p>Programs to work in batch mode: Demonstrate the following</p> <ul style="list-style-type: none"> Grouping records: Counting word frequencies Writing data from a data frame Putting it all together: counting launch the program in batch mode
<p>Analysing and processing tabular data : Demonstrate the following</p> <ul style="list-style-type: none"> The most common operations on a data frame (selection, deletion, and creation or Columns), Summarizing data frame Joining two data frames together

multi-dimensional data using compound types. Filling null values in your data frame
Use of SQL-like syntax within data frame methods: Creating a view in Spark SQL and in PySpark Unioning tables together in Spark SQL and in PySpark
Manipulating data the resilient distributed dataset (RDD) way: map, filter and reduce
Moving from large data sets in PySpark to small summaries in pandas for assessment: Exploring data and getting first feature columns Addressing data mishaps and building first feature set Getting data set ready for assembly: null imputation and casting

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE Outcome Based Education (OBE) and Choice Based Credit System (CBCS) SEMESTER - VI			
COMPUTER NETWORKS			
Course Code	S6CCS01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:2:0)	SEE Marks	50
Credits	04	Exam Hours	03
Lecture Hours	42hrs	Practical Hours	-

Course objectives: This course will enable students to:	
1	Understand the basic networking concepts and layers of TCP/IP model.
2	Interpret Line coding, error detection and correction techniques and access protocols.
3	Understand routing algorithms, congestion control and resource allocation.
4	Introduces internetworking and describes the key elements of the IP.
5	Analyse the transport-layer concepts: Transport-Layer services Reliable vs. un-reliable data transfer - TCP protocol -UDP protocol and QoS.

UNIT I	(9 Hours)
TCP/IP Protocol Suite , Layered Architecture, Layers in the TCP/IP Protocol Suite, Description of Each Layer, Encapsulation and Decapsulation, Addressing, Multiplexing and De-multiplexing. Data Rate Limits: Noiseless Channel: Nyquist Bit Rate, Noisy Channel: Shannon Capacity .Digital-To-Digital Conversion: Line Coding, Line Coding Schemes Analog-To-Digital Conversion: Pulse Code Modulation (PCM) Cyclic Codes: Cyclic Code Encoder Using binary and Polynomials Media Access Control (Mac):CSMA, CSMA/CD, CSMA/CA. Section: 2.2, 3.5, 4.1(4.1.1, 4.12), 4.2(4.2.1), 10.3 (10.3.1 to 10.3.3), 12.1 (12.1.2 to 12.1.4)	

UNIT II	(9 Hours)
Network Layer: Network-Layer Services: Packetizing, Routing and Forwarding Network-Layer Performance: Delay, Throughput, Packet Loss. Congestion Control. IPv4 Addresses: Address Space, Classful Addressing, Classless Addressing, Dynamic Host Configuration Protocol (DHCP), Network Address Resolution (NAT). Internet Protocol (IP): Datagram Format, Fragmentation, Security of IPv4 Datagrams. Section:18.1, 18.3, 18.4, 19.1	

UNIT III	(8 Hours)
Unicast Routing: Introduction: General Idea, Least-Cost Routing. Routing Algorithms: Distance-Vector Routing, Link-State Routing, Path-Vector Routing. Next Generation IP: IPv6 Addressing: Representation, Address Space, Address Space Allocation, Auto configuration. The IPv6 Protocol: Packet Format, Extension Header, Transition from IPv4 To IPv6: Strategies. Section:20.1, 20.2, 22.1, 22.2, 22.4	

UNIT IV	(8 Hours)
Transport Layer Protocols: Introduction: Services, Port Numbers. User Datagram Protocol: User Datagram, UDP Services, UDP Applications. TransmissionControl Protocol: TCP Services, TCP Features, Segment, A TCP Connection, State Transition Diagram, Windows in TCP, Flow Control, Error Control (except Sender and Receiver FSMs), TCP Congestion Control. Section:24.1, 24.2, 24.3 (24.3.1 to 24.3.9)	

UNIT V	(8 Hours)
<p>World Wide Web and HTTP: World Wide Web. Hypertext Transfer Protocol (HTTP), SMTP protocol, DNS: Name Space and Resolution, Telnet.</p> <p>Quality Of Service: Data-Flow Characteristics: Definitions, Sensitivity of Applications, Flow Classes. Flow Control To Improve QoS; Scheduling, Traffic Shaping or Policing, Resource Reservation, Admission Control. Integrated Services (Intserv): Flow Specification, Admission, Service Classes. Resource Reservation Protocol (RSVP), Problems with Integrated Services.</p> <p>Differentiated Services(DFSERV): DS Field, Per-Hop Behaviour, Traffic Conditioners.</p> <p>Section:26.1, 26.3, 26.4, 26.6, 30.1, 30.2, 30.3,30.4</p>	
LAB COMPONENT	
Study of basic Linux networking commands:	
1) ifconfig , 2) ip, 3) tracepath, 4) ping, 5) netstat, 6) ss, 7) dig, 8) nslookup, 9) route, 10) host, 11) arp, 12) hostname 13) wget, 14) curl	
Basic experiments in CISCO packet tracer	
1.Connecting Two PCs in Cisco Packet	
2.Connecting Two Different Networks using Router,	
3.Switch configuration	
3.DHCP Configuration	
Implement an Ethernet LAN using n nodes and set multiple traffic nodes and plot congestion window for different source / destination. packet analysis for the following network protocols:	
Hypertext Transfer Protocol, Domain Name Server, TCP, UDP, IP, ICMP and DHCP	
Using NS2/NS3 Simulator, implement the following	
a. Implement three nodes point – to – point network with duplex links between them. Set the queue size, vary the bandwidth and find the number of packets dropped.	
b. Implement transmission of ping messages/trace route over a network topology consisting of 6 nodes and find the number of packets dropped due to congestion.	
c. Implement an Ethernet LAN using n nodes and set multiple traffic nodes and plot congestion window for different source / destinationx) Simulate wc -l, cat f1 f2.	

Course Outcomes: On Successful completion of this course, students will be able to	
1	Apply the basics of computer networks technology and analyse the concepts of Digital transmission, error control protocols and random access protocols.
2	Apply the knowledge of Packet switching concepts in computer networking, Identify different categories of IP addresses and design subnets.
3	Analyse different Unicast routing mechanisms and protocols.
4	Analyse the transport-layer concepts and services -unreliable vs. reliable data transfer.
5	Examine various network protocols and Appraise existing QoS and application layer protocol/s.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Data Communications and Networking	Behrouz A. Forouzan	McGraw-Hill	5 th Edition, 2013
2	Computer Networks: A Systems Approach	Larry L Peterson and Bruce S Davie	Elsevier	5 th Edition, 2011
Reference Books				
1	Data and Computer Communications	William Stallings	Pearson Education	10 th Edition, 2013
2	Computer Networking: A Top-Down Approach	Kurose James F, Ross Keith W.	Pearson Education	6 th Edition, 2017
3	Computer Networks	Andrew S. Tanenbaum and David J. Wetherall	Pearson Educatin	5 th Edition,2011
4	Unix Network Programming, Interprocess Communications,	W Richard Stevens	Pearson Educatin	2 nd Edition

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	2														2
CO2	2		3												3
CO3		2													2
CO4		2													2
CO5	2	2													2
Overall CO	2	2	3							2	2				3

Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	2	2	3						2	2					3

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE Outcome Based Education (OBE) and Choice Based Credit System (CBCS) SEMESTER - VI			
INTERNET OF THINGS LABORATORY			
Course Code	S6CCSL01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(0:0:2)	SEE Marks	50
Credits	1	Exam Hours	03
Lecture Hours	-	Practical Hours	28hrs

Course objectives: This course will enable students to:	
1	Discover key IoT concepts including identification, sensors, localization, wireless protocols, data storage and security
2	Explore IoT technologies, architectures, standards, and regulation
3	Realize the value created by collecting, communicating, coordinating, and leveraging the data from connected devices
4	Synthesize the sensor data by deploying to the IoT cloud

Sl. No.	EXPERIMENTS
	Lab Set Questions
1	a) To interface LED/buzzer with Arduino and write a program to control lights (min 3 LED's) and turn The buzzer ON when all lights turned ON. Additional Program – Add switch interface to control b) Experiment to interface IR/LDR with Arduino and write a program to control IR sensor and turn LED on when the push button is pressed Additional Program – Add Buzzer and develop the logic to integrate buzzer
2	Experiment to interface temperature sensor DHT11 and write a program to print the temperature and humidity reading. Turn the LED and buzzer ON when the temperature threshold reaches beyond 35 degrees. Additional Program – add the push button switch interface and do the same using Raspberry Pi
3	Experiment to interface servo motor using Arduino. Control the positional reading through switch and add the indicators using buzzer. Additional Program – do the same using Raspberry Pi and ESP32
4	To interface Bluetooth with Arduino and write a program to send sensor data to smartphone using Bluetooth. Additional Program – do the same through Raspberry Pi
5	a) To interface GPS UNIT with Arduino and write a program to send location data to smartphone b) Send the smart phone's location sensor data to Arduinio through BT
6	Control LED light using Voice command . Use Arduinio Uno. Additional Program – do the same using Raspberry Pi/ ESP32
7	Set up a simple web server using ESP32 UNIT and monitor the live temperature in the web browser
8	Consider IR /LDR sensors. Set up the circuit using Raspberry Pi. On intrusion detection, send an email to the owner. Upload IR/LDR sensor data to Thing speak cloud.
9	Consider any sensor (DHT11/IR). Write a program on Raspberry Pi to publish/subscribe to MQTT broker for temperature data and print it on Raspberry Pi console. Upload temperature and humidity data to Thingspeak cloud.
10	Build Smart Environment Monitoring Using IoT and RESTful API. Design an IoT system that continuously monitors the surrounding environment using a DHT11/DHT22 sensor connected to a microcontroller (e.g., Raspberry Pi or ESP32). The collected sensor data should be: <ol style="list-style-type: none"> a) Read every 10 seconds. b) Sent to a backend server via a REST API using HTTP POST. c) Stored in a local or cloud database (e.g., SQLite, Firebase, or PostgreSQL). d) Accessible via HTTP GET to retrieve the last N readings or average of past hour. e) Visualized via a simple web interface or terminal-based output

Course Outcomes: On Successful completion of this course, students will be able to	
1	Understand key IoT core system architecture through the usage of sensors, actuators, embedded devices (Arduino/Raspberry Pi/ESP32), Physical link interfaces etc.
2	Explore IoT connection technologies through Bluetooth, wifi, UART(FTDI) interfaces
3	Realize the value created by collecting, communicating, coordinating, and leveraging the data from connected devices through communication protocol, MQTT broker.
4	Synthesize the sensor data by deploying to the IoT cloud

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1			2												2
CO2			2												2
CO3			2												2
CO4			2												2
CO5			2												2
Overall CO			2												2

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
			2												2

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - VI

CLOUD COMPUTING

Course Code	S6CCSPE01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	3
Lecture Hours	42hrs	Practical Hours	-

Course objectives: This course will enable students to:

1	Learning about cloud types, paradigm shift in cloud computing, attributes that make the cloud computing unique, SLA and licencing.
2	Understanding architecture and infrastructure of fog computing and cloud computing including SaaS, PaaS and IaaS.
3	Understanding various types of virtualization and learning about the capacity planning for the cloud.
4	Understanding how cloud data can be secured.
5	Understand implications of technologies such as Fog & Edge Computing ,IOT

UNIT I

(8 Hours)

Defining Cloud Computing: Cloud Types, The NIST model, The Cloud Cube Model, Deployment models, Service models, Examining the Characteristics of Cloud Computing, Paradigm shift, Benefits of cloud computing, Disadvantages of cloud computing; Assessing the value proposition: Early adopters and new applications, the laws of cloud economics, cloud computing obstacles, behavioural factors relating to cloud adoption, measuring cloud computing costs, specifying SLAs.

Textbook1: Chapter1,2

UNIT II

(8 Hours)

Cloud Infrastructure: Cloud Computing at Amazon, Cloud Computing: The Google Perspective, Microsoft Windows Azure and Online Services, Open-Source Software Platforms for Private Clouds, Cloud Storage Diversity and Vendor Lock-in, Cloud Computing Interoperability: The Intercloud, Energy Use and Ecological Impact of Large-Scale Data Centers, Service- and Compliance-Level Agreements, Responsibility Sharing Between User and Cloud Service Provider, User Experience, Software Licensing.

Textbook 2: Chapter3: (3.1 to 3.11)

UNIT III

(8 Hours)

Understanding Abstraction and Virtualization: Using Virtualization Technologies, Load balancing and Virtualization, Understanding Hypervisors; Capacity Planning: Defining Baseline and Metrics, Baseline measurements, System metrics, Load testing, Resource ceilings, Server and instance types, Network Capacity, Scaling

Textbook1: Chapter5,6

UNIT IV

(8 Hours)

Understanding Cloud Security: Securing the Cloud, The security boundary, Security service boundary, Security mapping, Securing Data, Brokered cloud storage access, Storage location and tenancy, Encryption, Auditing and compliance, Establishing Identity and Presence, Identity protocol standards, Windows Azure identity standards.

Textbook1: Chapter12

UNIT V

(8 Hours)

Fog Computing and its Applications: Introduction: Essential characteristics in fog computing, Fog nodes, Fog node deployment model. View of a Fog Computing Architecture: Node view, System view, Software view. **Fog Computing in IoT:** Importance of Fog Computing, Time sensitiveness in Fog Computing. Selected Applications of Fog Computing.

Textbook3: Chapter11

Edge Computing State-of-the-Art Interfaces and Devices: Middleware, Hydra, Aura, TinyDB, FiWare, Application Interfaces, Edge Computing Simulators: PureEdgeSim, IoTSim-Edge, iFogSim and Edge CloudSim.

Textbook4: Chapter 2(2.6 and 2.7)

Course Outcomes: On Successful completion of this course, students will be able to	
1	Articulate the key dimensions of Cloud Computing, characteristics, benefits and drawbacks of Cloud computing
2	List Services provided by various cloud vendors & analyse the importance of each service..
3	Analyse the impact of vendor lock –in ,SLA, Large scale data centres.
4	Analyse the importance virtualization in cloud for resource pooling.
5	Analyse the cloud security issues.
6	List the features of fog computing & edge computing & Analyse the relationship between fog computing, edge computing & IoT.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Cloud Computing Bible	Barrie Sosinsky	Wiley Publishing Inc.	2011
2	Cloud Computing Theory and Practice	Dan C. Marinescu	Morgan Kaufmann, Elsevier	2013
3	Introduction to IOT	SudipMisra, Anandarup Mukherjee, Arijit Roy	Cambridge University press	2020
4	Edge Computing Fundamentals Advances and Applications	K. Anitha Kumari, G.Sudha Sadasivam, D.Dharani and M. Niranjana Murthy	CRC Press	2022
Reference Books				
1	Cloud Computing Principles and Paradigms	RajkumarBuyya, James Broberg,AndrzejGoscinski	Wiley Publishing Inc.	2013
2	Cloud Computing and SOA Convergence in Your Enterprise:	David S. Linthicum	Addison-Wesley Professional	1 st Edition
3	Distributed and Cloud Computing	Kai Hwang, Geoffrey C. Fox, and Jack J. Dongarra	Morgan Kaufman Publishers	2012

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	2														2
CO2	2														2
CO3	2														2
CO4	2														2
CO5	2														2
Overall CO	2														2

Articulation Matrix:

Course	Program Outcomes	PSOs
--------	------------------	------

Outcomes	1	2	3	4	5	6	7	8	9	10	11	1	2	3
	2													2

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE Outcome Based Education (OBE) and Choice Based Credit System (CBCS) SEMESTER - VI			
DISTRIBUTED DATA STORAGE MANAGEMENT			
Course Code	S6ADPE01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	03
Lecture Hours	42hrs	Practical Hours	-

Course objectives: This course will enable students to:	
1	Understand Distributed Data Storage Management System and various Design Issues.
2	Explore several algorithms for processing queries in distributed environment.
3	Explore to various DDBMS such as parallel, distributed object, P2P, web and data streams.

UNIT I	9 Hours
Introduction: Distributed Data Processing, What is a Distributed Database System?, Promises of DDBSs, Design Issues, Distributed DBMS Architecture. Distributed Database Design: Top-Down Design Process, Fragmentation. (1.1,1.2,1.4,1.6,1.7,3.1,3.3)	

UNIT II	9 Hours
Database Integration: Bottom-Up Design Methodology, Schema Matching, Schema Integration, Schema Mapping. Overview of Query Processing: Query processing problem, Objectives of Query Processing, Complexity of Relational Algebra operations, Characterization of Query processors, Layers of Query Processing.	

(4.1-4.4,6.1-6.5)

UNIT III

8 Hours

Query Decomposition and Data Localization: Query Decomposition, Localization of Distributed Data.
Parallel Database Systems: Parallel Database System Architectures, Parallel Data Placement, Parallel Query Processing, Load Balancing, Database Clusters.
(7.1,7.2,14.1-14.5)

UNIT IV

8 Hours

Distributed Object Database Management: Fundamental Object concepts and Object models, Object distribution design, Architectural issues, Object management, Distributed object storage, Object query processing.
Peer-to-Peer Data Management: Infrastructure, Schema Mapping in P2P Systems.
(15.1-15.6.16.1,16.2)

UNIT V

8 Hours

Web Data Management: Web Graph Management, Web Search, Web Querying.
Data Stream Management.(17.1-17.3,18.1)

Course Outcomes: On Successful completion of this course, students will be able to

1	Describe the distributed database concepts, issues, and its structures.
2	Produce the transaction management and query processing techniques in DDBMS.
3	Analys the access control of data bases and design parallel database system.
4	Design and query the storage systems related to distributed object database design and management
5	Apply distributed data storage concepts to design and develop Web Data Management Systems..

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	“Principles of Distributed Database Systems”	M. Tamer Özsu Patrick Valduriez	Prentice Hall.	Third Edition,
Reference Books				
1	Distributed Databases - Principles and Systems	Stefano Ceri Guisepe Pelagatti	Tata McGraw Hill	1985
2	Fundamental of Database Systems	Elmasri Navathe	Pearson Education; Asia.	
	Database System Concepts	Korth Sudarshan	TMH	

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
CO1	2	2										2		
CO2		2	2									2		
CO3		2	2									2		
CO4		2	2									2		
CO5		2	2									2		
Overall CO	2	2	2									2		

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
	2	2	2									2		

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - VI

FULL STACK DEVELOPMENT WITH DJANGO

Course Code	S6ADPE02	CIE Marks	50
Teaching Hours/Week (L:T:P)	(03:0:0)	SEE Marks	50
Credits	03	Exam Hours	03
Lecture Hours	42hrs	Practical Hours	-

Course objectives: This course will enable students to:

1	Analyze the evolution of web technologies and compare modern technology stacks and methodologies for product design and deployment.
2	Apply essential front-end and back-end technologies to design, develop, and integrate user interfaces and data storage solutions.
3	Construct Django-based web applications by organizing projects, defining models, implementing views, and utilizing templates.
4	Manage and extend Django applications by configuring the admin interface, processing forms, and employing advanced data handling techniques.
5	Integrate advanced Django features and best practices to implement class-based views, user management, database operations, testing, and application security.

UNIT I

9 Hours

Foundations of Modern Web Development

- 1.1 State of the web (T1 Chapter 1): Rise of the web, mobile web, state of HTML, websites, and applications
- 1.2 Technology Stacks: Overview, MERN, MEAN, MEVN, MENG, LAMP, WAMP, XAMPP, Ruby on Rails stacks, Comparison of Stacks
- 1.3 Product Design (T1 Parts of Chapters 2 to 6, 8, 10-12): Requirements, Work Breakdown, User Experience Design, User Interaction Design, Systems Architecture, Component Interactions, Module Design, Design for Failures
- 1.4 Product Development and Deployment (T1 Chapter 13): Agile, Scrum, Twelve Factor Apps, Development Machines, Testing, Production Environments, Continuous Delivery and Continuous Deployment
- 1.5 Twelve Factor App Methodology (W01): Background, Twelve Factors

UNIT II

9 Hours

Essentials of Front-End and Back-End Technologies

- 2.1 HTML (R1): HTML5, Document Structure, Text formatting, Lists, Links & Anchors, Tables, Forms, Semantic & Structural Elements, UX Design best practices.
- 2.2 CSS (T1 Chapter 3, R1): Syntax, Location of Styles, Selectors, Cascade, Box Model, Text Styling, Flow, Positioning, CSS Layouts, Filters, Transitions, Animations
- 2.3 JavaScript (T1 Chapter 8, R1, R2): Language fundamentals, Variables, Data Types, Outputs, Conditions, Loops, Arrays, Objects, DOM & its manipulation, Events, Forms
- 2.4 JQuery (T1 Chapter 8, R1, R2): Syntax, Selectors, Events, Effects, Traversing, AJAX
- 2.5 Bootstrap (W02): Introduction to Bootstrap, Walk through Bootstrap Library
- 2.6 Introduction to SQLite3 (R3): Introduction, Data Types, Basic SQL Commands, Transactions
- 2.7 API Design (T1 Chapter 10 to 11, R4): API Responsibilities, Designing REST APIs, Securing APIs,

Event-Based APIs, Discovering APIs, and Using APIs.

UNIT III

8 Hours

Django Fundamentals and Core Components

- 3.1 Introduction to Django (T2 Chapter 3): Introduction, Feature Sets, Security, Scalability, Versatility, Applications, Project Structure
- 3.2 Django Models (T2 Chapter 4): Supported databases, Defining models, Basic Data Access, Creating relationships
- 3.3 Django Views (T2 Chapter 5): Introductions, Configuring URLs, Dynamic Content, Dynamic URLs
- 3.4 Django Templates (T2 Chapter 6): Template System Basics, Design Philosophy, Template inheritance, Displaying database data, Loading Static Content, Template Includes

UNIT IV

8 Hours

Django Administration, Forms, and Advanced Data Handling

- 4.1 The Django Admin (T2 Chapter 7): Accessing admin site, Registering Models, Editing tables, Managing Users
- 4.2 Django Forms (T2 Chapter 8): Basics, Model Forms
- 4.3 Advanced Models (T2 Chapter 9): Working with data, Methods that return QuerySets, Methods that do not return QuerySets, Field Lookups, Aggregate Functions, Model Methods, Model Inheritance
- 4.4 Advanced Views (T2 Chapter 10): Request and Response Objects, QueryDict objects, Template Response Objects, Middleware, Generating Non-HTML Content, Pagination
- 4.5 Advanced Templates (T2 Chapter 11): Tags, Filters, Handling invalid variables, Custom Tags, Custom Filters

UNIT V

8 Hours

Advanced Django Features and Best Practices

- 5.1 Class Based Views (T2 Chapter 13): View, RedirectView and TemplateView, Customizing Generic View Classes and Methods, Django's Class-based Generic Views
- 5.2 Advanced User Management (T2 Chapter 14): The User Model Class, Creating a Custom User Model, Custom Authentication
- 5.3 Advanced Forms (T2 Chapter 15): Customizing Forms, The Messages Framework, Django Formsets, Handling Multiple Forms
- 5.4 Working with Databases (T2 Chapter 16): Database Management Commands, Connecting to Database Engines (PostgreSQL and MySQL)
- 5.5 Debugging and Testing (T2 Chapter 17): Using Django's Error Page, Using the Messages Framework, Unit Testing
- 5.6 Other Topics (T2 Chapter 18): Django's Cache Framework, Sessions in Django, System Logging, Security in Django, Images in Django, Signals, Search

Course Outcomes: On Successful completion of this course, students will be able to

1	Explain the foundational concepts and methodologies of modern web development and product design.
2	Develop interactive web applications using front-end and back-end technologies, including HTML, CSS, JavaScript, and SQLite.
3	Build robust web applications using Django, incorporating models, views, templates, and forms.
4	Administer and enhance Django applications through advanced data handling, user management, and customization of the admin interface.
5	Evaluate and apply advanced Django features, including class-based views, testing, debugging, and security measures, to deliver production-ready web solutions.

EXERCISES

- 1. Biodata website using Django
- 2. Using Django implement a program to accept USN, Name, subject code, CIE marks and store information in a database and display students whose cie is less than 20
- 3. To accept student name, usn, semester, exam fee from web page and delete all the students who have not paid exam fee
- 4. Using Django create a HR database with the collection employees having fields like employee name, email, phone, hired date, job title, salary. Accept this fields and store it in database and display all employees details whose salary is greater than 50k

5. Create an exam management system using Django framework for creating student database and displaying students who have secured “O” grade.
6. To implement in final year’s database which accepts usn ,name ,and company name as fields and store it in database. Display the list of students who are placed for Amazon .
7. Program to accept the fields id, title, name and branch of a faculty and store it in database. Display all the faculty who belong to CSE branch and the title is professor.
8. Consider the student information like name ,usn, dept ,grade from a web page and store it in a database and update student grade with the name specified by the user and display the results
9. Using Django, perform CRUD operations for an employee details.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	The Full Stack Developer ISBN: 978-1-4842-4151-6	Chris Northwood	Apress Media LLC	2018
2	Mastering Django ISBN:978-0648884415	Nigel George	GNW Independent Publishing	2020
Reference Books				
1	HTML 5 Black Book	DT Editorial Services	Dreamtech Press	2016
2	JavaScript and JQuery: Interactive Front-End Web Development	Jon Duckett	Wiley	1 st Edition 2014
3	SQLite3	Vicente Hernando	McGraw Hill	1 st Edition 2017
W1	https://www.12factor.net/			
W2	https://getbootstrap.com/			

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1	2		2										2		
CO2	2		2										2		
CO3			2										2		
CO4			2										2		
CO5			3										2		
Overall CO	2		3										2		

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	2		2										2		

B.E. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND DATA SCIENCE)
 Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - VI

AI DRIVEN CYBER SECURITY

Course Code	S6CIPE01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	03
Lecture Hours	42hrs	Practical Hours	-

Course objectives: This course will enable students to:

1	Illustrate the understanding of Cyber Security Fundamentals.
2	Analyses the attacker motivation and the techniques used by them to break the security of the application.
3	Study the vulnerabilities in applications and networks. Analyses the possible attacks that can be built by the hackers.
4	Understand the Artificial Intelligence methods and principles can address cybersecurity challenges
5	Understand and Analyse AI methods and usecases suitable for solving to cybersecurity issues

UNIT I

(8 Hours)

Introduction and Overview of Cyber Crime, Nature and Scope of Cyber Crime, classification of Cyber Crime, Social Engineering, Categories of Cyber Crime, Property Cyber Crime. Cybercrime-Indian perspective/the Indian ITA 2000, Cyber Offenses: How criminals plan then.

UNIT II	(8 Hours)
Unauthorized Access to Computers, Computer Intrusions, White collar Crimes, Viruses and Malicious Code, Internet Hacking and Cracking, Virus Attacks, Pornography, Software Piracy, Intellectual Property, Mail Bombs, Exploitation, Stalking and Obscenity in Internet, Digital laws and legislation, Law Enforcement Roles and Responses, Cybercrime: Mobile and wireless Devices	

UNIT III	(9 Hours)
Tools and method used in cybercrime: Proxy servers and Anonymizers, Phishing: methods, techniques, spear phishing, types of phishing scams, toolkits and spy phishing, phishing countermeasures; Identity theft; Password cracking, keyloggers and spywares, Virus and worms, Trojan Horses and Backdoors, Steganography, DoS and DDoS Attacks, SQL Injection, Buffer Overflow, Attacks on wireless sensor networks, Cybercrime: Case Studies: Real-life examples and Online scams.	

UNIT IV	(9 Hours)
AI for Cybersecurity, The Use Cases Intend to Solve Various Cybersecurity Challenges through A Unified DL Pipeline, AI Conducts Two Reverse Engineering Tasks, Related Work, Model Architecture, Model Training Issues, Model Performance, Deployed Model, Source Code and Dataset, Remaining Issues.	

UNIT V	(8 Hours)
AI Detects DNS Cache Poisoning Attack, The Security Problem, Raw Data Generation and Collection, Labeling DNS Sessions, Feature Extraction and Data Sample Representation, Data Set Construction, Model Architecture, Parameter Tuning, Evaluation results, Model Deployment, Remaining Issues, Code and Data Resources.	

Course Outcomes: On Successful completion of this course, students will be able to	
1	Understand the basic concepts of crime, crime behavior, forensic science and its linkage to crime scenario.
2	Analyze the techniques used by hackers to create frauds
3	Determine and analyse software vulnerabilities and security solutions to reduce the risk of exploitation.
4	Apply the AI principles to solve cybersecurity challenges
5	Interpret and Analyse Deep learning methods for use cases intend to solve various cyber security challenges

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Cyber Security: Understanding Cyber Crimes, Computer Forensics and Legal Perspectives	Nina Godbole and Sunit Belapure	Wiley India Pvt Limited"	2011
2	Computer Forensics and Investigations, Cengage Learning	Nelson Phillips and EnfingerSteuart		New Delhi, 2009
3	An Incident-Based Approach to Forensic Investigations Malware forensics. In: Practical Cyber Forensics.	Niranjan Reddy	Apress, Berkeley,	CA (2019). https://doi.org/10.1007/978-1-4842-4460-9
4	AI for Cybersecurity A Handbook of Use Cases, Penn State Cyber Security Lab	Peng Liu, Tao Liu et al.,		
Reference Books				
1	Incident Response and Computer Forensics,	Kevin Mandia, Chris Prosize, Matt Pepe	Tata McGraw -Hill,	New Delhi, 2006
2	Software Forensics, Tata,	Robert M Slade	McGraw -Hill,.	New Delhi ,2005

3	“Understanding Forensics in IT ”,	Bernadette H Schell, Clemens Martin	Cybercrime, ABC – CLIO Inc, California, 2004.	NIIT Ltd,2005
---	-----------------------------------	--	---	---------------

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
CO1	2													2
CO2		2												2
CO3		2			2									2
CO4	2													2
CO5		2												2
Overall CO	2	2												2

Program Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs		
	1	2	3	4	5	6	7	8	9	10	11	1	2	3
	2	2			2									2

B.E. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE Outcome Based Education (OBE) and Choice Based Credit System (CBCS) SEMESTER - VI				
INDIAN KNOWLEDGE SYSTEMS				
Course Code	SHS07		CIE Marks	100
Teaching Hours/Week (L:T:P)	1: 0: 0		SEE Marks	-
Credits	-		Exam Hours	
Lecture Hours	15hrs		Practical Hours	-

Course objectives: This course will enable students to:	
1	Rejuvenate and mainstream Indian Knowledge Systems by integrating it into the fabric of education research ecosystem.
2	Bridge the gap between Traditional Indian Knowledge Systems and contemporary knowledge systems.
3	Make a tangible difference to the contemporary world by mainstreaming Indian Knowledge Systems.

UNIT I: Introduction to Indian Knowledge Systems (IKS)	(5 Hours)
Overview: Need of IKS, Importance of ancient knowledge. Defining IKS, History of IKS, Indian philosophical systems- Development and unique features, Yoga-darsana, Number system in India-Historical evidence, salient features of Indian number system, the concept of zero and its importance Reference book- 1 : 1.1; 1.2; 1.5; 3.1; 3.3.2; 6.1; 6.2; 6.2.1	

UNIT II: Fundamental concepts in Science in IKS	(5 Hours)
Knowledge: Framework and classification. The knowledge triangle. Prameya – A Vaisesikan approach to physical reality. Dravyas – the constituents of physical reality. Attributes – the properties of substances. Unique aspects of Indian Mathematics. Great Mathematicians and their contributions. Arithmetic. Square of a number. Unique aspects of Indian astronomy. Elements of Indian calendar. Pancanga – the Indian calendar system. Reference book- 1 : 7.1; 7.2; 7.2.1; 7.2.2; 8.1; 8.2; 8.3; 8.3.1; 9.1; 9.4; 9.6.	

UNIT III: Humanities and Social sciences in IKS (5 Hours)

Ayurveda-Definition of health, Tri- dosas - Relationship to health, yoga way of life –relevance to health and wellness, Indian approach to psychology, The tri-guna system, The body-mind-intellect – consciousness complex, consciousness- the true nature of an individual, Arthasastra- Governance and administration, the Kautilyan state
Reference book- 1: 13.1; 13.2;13.4;13.5; 13.6; 13.7; 13.8;14.1;14.2

Course Outcomes: On Successful completion of this course, students will be able to

1	Analyse the continuous unbroken knowledge traditions of Bhārata since time immemorial.
2	Recognize the relevance of traditional knowledge in science
3	Recognize the relevance of traditional knowledge in Humanities and Social sciences

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Reference Books				
1	Introduction to Indian Knowledge System- concepts and applications,	B Mahadevan, Vinayak Rajat Bhat, Nagendra Pavana R N	PHI Learning Private Ltd	2022
2	Traditional Knowledge System in India	Amit Jha	Atlantic Publishers and Distributors (P) Ltd	2009
3	Knowledge Traditions and Practices of India, Vol. 1	Kapil Kapoor, Avadesh Kumar Singh,	DK Print World (P) Ltd	2005
Suggested Web Links:				
1.	https://www.youtube.com/watch?v=LZP1StpYEPM			
2.	http://nptel.ac.in/courses/121106003/			
3.	http://www.iitkgp.ac.in/department/KS;jsessionid=C5042785F727F6EB46CBF432D7683B63 (Centre of Excellence for Indian Knowledge System, IIT Kharagpur)			
4.	https://www.wipo.int/pressroom/en/briefs/tk_ip.html			
5.	https://unctad.org/system/files/official-document/ditcted10_en.pdf			
6.	http://nbaindia.org/uploaded/docs/traditionalknowledge_190707.pdf			
7.	https://unfoundation.org/what-we-do/issues/sustainable-development-goals/?gclid=EAlaIqobChMImp-Jtb_p8gIVTeN3Ch27LAmPEAAAYASAAEgIm1vD_BwE			

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
CO1						1	1					1			
CO2	2					2	1					2			
CO3	1											2			
Overall CO	2					2	1					2			

Articulation Matrix:

Course Outcomes	Program Outcomes											PSOs			
	1	2	3	4	5	6	7	8	9	10	11	1	2	3	
	2					2	1					2			