SIDDAGANGA INSTITUTE OF TECHNOLOGY, TUMAKURU

(An autonomous institution affiliated to VTU, Belagavi, Approved by AICTE, New Delhi, Accredited by NAAC with 'A++' grade & ISO 9001:2015 Certified)

B.E. in Computer Science and Engineering (AI & ML)

SCHEME OF TEACHING AND EXAMINATION (2022 Scheme)

V Semester

B.E. in CSE(AI&ML) 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	S5CIMP	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	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.

SIDDAGANGA INSTITUTE OF TECHNOLOGY, TUMAKURU

(An autonomous institution affiliated to VTU, Belagavi, Approved by AICTE, New Delhi, Accredited by NAAC with 'A++' grade & ISO 9001:2015 Certified)

B.E. in Computer Science and Engineering (AI & ML)

SCHEME OF TEACHING AND EXAMINATION (2022 Scheme) (w.e.f. 2024-25)

VI Semester

B.E. in CSE(AI&ML)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	S6CII01	Natural Language Processing (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		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	S6CIMP-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									
		SMCO3	NCC	NCC CO									
9.	HSS	SHS07	Indian Knowledge Systems		1	0	0	0	-	100	-	100	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	**S6CCSPE04**	High Performance Computing
S6CIPE01	AI Driven Cyber Security	**S6CIPE02**	Real Time Big Data Analytics

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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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 – Chapter2, 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

Text Book 1 – Chapter 22- 22.3, 22.7, 22.8, 23.3. 23.4, 23.5, 23.6, 23.7, 23.9

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	4th Edition, 2013
3	Clean Code – A handbook of Agile Software Craftsmanship	Robert C Martin	Pearson	1stEdition,2011
Reference Books				
1	Integrated Approach to Software Engineering	Ian Sommerville	Pearson Education Ltd	8th Edition -
2	Software Engineering	Shari Lawrence PfleegerJoanneM Atlee	Pearson	4th 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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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 andtransform 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	6th Edition,2011
2				
Reference Books				
1	Data base System Concepts.	Silberschatz, Korth and Sudharshan.	McGraw-Hill	6th Edition, 2010
2	Database Management Systems.	Raghu Ramakrishnan and Johannes Gehrke	McGraw-Hill.	3th 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	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. **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: • 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. Retrieve the crews with strength greater than or equal to 20**.**
2	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. **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: • 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'. Retrieve the studio with location equal to 'xyz'.
3	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. **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 voiced. **v)** Write a procedure to list all Cartoon-Serials produced during the specific year. **vi)** Write a deletion trigger, does not allow to deleting current year Cartoon-Serials.

	vii) Create "cartoon serial" collection and perform the following CRUD operations using MongoDB: • 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'. Retrieve the cartoon serials with sensor numbers lesser than 100.
4	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. **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 a person who owns highest number of cars **iv)** Retrieve persons and cars information purchased on the day 11-11-11 **v)** Write a insertion trigger to check date of purchase must be less than current date (must use system date) **vi)** Write a procedure to list all cars and owner information purchased during the specific year. **vii)** Create "car" collection and perform the following CRUD operations using MongoDB: • 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'. Retrieve the cars with colour green.
5	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. **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 pets owned by a person 'Abhiman'. **iv)** List all persons who are not owned a single pet **v)** Write a trigger to check the constraint that the person can buy maximum three pet dogs **vi)** Write a procedure to list all dogs and owner details purchased on the specific date. **vii)** Create "dog" collection and perform the following CRUD operations using MongoDB: • 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**.**

B.E. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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:

	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
xv) Apply classification on given data. Remove label and then apply clustering. Perform class to cluster evaluation. Apply classification on un labelled dataset by removing play attribute.

	Prepare the analysis report. **Interpret the results.**
2.	Experiment using WEKA tool. Consider the following data set

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

Use the data sources, like ARFF, XML ARFF files. Do the following

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 alforithm 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.

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. <u>Consider glass data set.</u>** 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. 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 Additional: Derive Fact Table, Star schema, Snowflake schema and do the comparison
5.	Using Tableau do the following. Consider Titanic Data Set. Perform the following using Tableau platform: i. Perform Calculations: 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: 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: 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 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.	Using Tableau platform, do the following. Consider Brazilian E-Commerce Public Dataset by Olist from Kaggle. Perform the following using Tableau platform: 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 Additional: Enable **data export** by filtered view and configure **email alerts** or subscriptions when sales in any region drop below a threshold
7.	Consider Sales Data set. Perform the following using powerBI platform. 1. Create tables with i) manufacturing, sum of sales, columns. ii) Product Category, sum of profits etc iii) Find Answers for the following: 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? e) Which Promotion Name has the highest Sales? f) Which Product Sub Category has the highest Profit? iv) Perform cross flitering 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 2. Apply appropriate visualization. Create calculations for Visualization i) Calculate the difference between sum of Sales and Sum of Profits ii) Calcualte Profit Ratio. (Hint: Sum of Profit/Sum of Sales). Also demonstrate using in built funciton, e.g., DIVIDE, AVERAGE etc iii) Calcualte the average sales per product for each of manufacturers. (Hint: Sum of sales/Count of Product Name) iv) Calcualte the Percentage of grand total using inbuit function 3. Apply Filters and Slicers 4. Build Graphs. Draw Trend Analysis Graph. Show trends and forecasting. For example: sum of sales per year, month, quarter etc 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
8.	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
9.	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
10.	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
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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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 ni 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, 11th 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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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)
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. **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. **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.	

UNIT V	(8 Hours)
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 **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 **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.	

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	Ricardo Baeza -Yates	ACM Press	2nd Edition, 2011

	Retrieval: The Concepts and Technology behind Search	and Berthier Ribeiro - Neto		
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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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	2th 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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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)**

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.
Expressing Knowledge: Knowledge engineering, vocabulary, basic facts, complex facts, terminology facts, entailments, abstract individuals, other sorts of facts.

UNIT II **(9 Hours)**

Resolution: The propositional case, handling variables and quantifiers, dealing with computational intractability.
Reasoning with Horn clauses: Horn clauses, SLD resolution, computing SLD derivations.
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.

UNIT III **(8 Hours)**

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.

Inheritance: Inheritance networks, strategies for defeasible inheritance, a formal account of inheritance networks.

UNIT IV **(8 Hours)**

Object oriented representation: Objects and frames, a basic frame formalism, an example: using frames to plan a trip, beyond the basics.

Vagueness, uncertainty, and degree of belief: Noncategorical reasoning, objective probability, subjective probability, vagueness.

UNIT V **(8 Hours)**

Planning: planning in a situation calculus, the STRIPS representation, planning as a reasoning task, beyond the basics.

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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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 Bundit 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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - VI

NATURAL LANGUAGE PROCESSING (I)

Course Code	S6CII01	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	To Understand the NLP techniques like parsing, POS-tagging and Word sense disambiguation.
2	To explore language modeling techniques such as N-grams.
3	To explore the applications of NLP such as Machine translation, Information retrieval etc.
4	To understand the basic architecture of the NLG system and the role of NLP in a search engine.
5	Demonstrate the use of modern NLP techniques for processing of text like extracting the data. Text to Feature representation etc..

UNIT I **(9 Hours)**

Introduction to NLP: NLP in the Real-world, NLP Tasks, what is Language: Building Blocks of Language, Why NLP is Challenging, Machine Learning, Deep Learning, and NLP: An Overview, Approaches to NLP: Heuristic based NLP, Machine Learning for NLP, Deep Learning for NLP, Why Deep Learning is not Yet the Silver Bullet for NLP, An NLP Walkthrough: Conversational Agents.

NLP Pipeline: Data Acquisition, Text Extraction and Cleanup: HTML Parsing and Cleanup, Unicode Normalization, Spelling Correction, System-Specific Error Correction, Pre-Processing: Preliminaries, Frequent Steps, Other Pre-Processing Steps, Advanced Processing, Feature Engineering, Modeling, Evaluation, Post-Modeling Phases, Working with Other Languages, Case Study. **Extracting the Data:** Text data collection using APIs, Reading PDF file in Python, Reading word document, Reading JSON object, Regular expressions. **Exploring and Processing Text Data:** Lowercasing, Punctuation removal, Stop words removal, Text standardization.

(Text Book-1: 1,2)
(Text Book-2: Recipe 1.1-1.6, Recipe 2.1-2.4)
(Text Book-4: 3.4, 3.5, 3.7)

UNIT II **(9 Hours)**

Language Modeling: Introduction, Statistical Language Model- N-gram model, Add-one smoothing, Good-Turing smoothing. Generating N-grams. **Part-of-Speech Tagging**: Rule-based Tagger, Stochastic Tagger, Hybrid Tagger. Tagging Part of Speech. **Syntactic Analysis**: CFG, Parsing- Top-down parsing, Bottom-up parsing, Probabilistic Parsing- Estimating Rule probabilities. Recursive Descent Parsing, Shift-Reduce Parsing, The Left-Corner Parser, Dependencies and Dependency Grammar. **Tokenizing Text and WordNet Basics:** Introduction, Tokenizing text into sentences, Tokenizing sentences into words, Tokenizing sentences using regular expressions, training a sentence tokenizer, Filtering stop words in a tokenized sentence, Looking up Synsets for a word in WordNet, looking up lemmas and synonyms in WordNet, Calculating WordNet Synset similarity, Discovering word collocations.

(Text Book-3: 2.3.1, 2.3.2, 2.3.3, 3.7.1, 3.7.2, 3.7.3, 4.2, 4.4.1, 4.4.2)
(Text Book-2: Recipe 4-3. Recipe 3-3)
(Text Book-4: 8.4, 8.5)
(Text Book-5: 1)

UNIT III **(8 Hours)**

Information Retrieval: Introduction, Design features of Information Retrieval Systems, Information Retrieval models, Classical Information Retrieval models, Alternative models of IR, Evaluation of the IR system. **Cross-Lingual Information Retrieval. Converting Text to Features:** One Hot encoding, Count vectorizer, Co-occurrence matrix, Hash vectorizer, Word embedding, Implementing fastText. **Information retrieval using word embeddings.**

(Text Book-3: 9.1-9.4,9.6.1-9.6.2,9.7,10.6)
(Text Book-2: Recipe 3.1-3.2,3.4-3.8)

UNIT IV	**(8 Hours)**
Ambiguity, Word sense Disambiguation: Knowledge based approaches, Supervised Learning of WSD. **Disambiguating Text.** **Text Classification:** Applications, A Pipeline for Building Text Classification Systems, A Simple Classifier Without the Text Classification Pipeline, Using Existing Text Classification APIs, One Pipeline, Many Classifiers, Naive Bayes Classifier, Logistic Regression, Support Vector Machine, Using Neural Embeddings in Text Classification, Word Embeddings, Subword Embeddings and fastText, Document Embeddings, Deep Learning for Text Classification, CNNs for Text Classification, LSTMs for Text Classification, Text Classification with Large, Pre-Trained Language Models, Interpreting Text Classification Models, Explaining Classifier Predictions with Lime, Learning with No or Less Data and Adapting to New Domains, No Training Data, Less Training Data: Active Learning and Domain Adaptation, Case Study: Corporate Ticketing. **Machine Translation:** Introduction, Problems in Machine Translation, Machine translation approaches, Direct Machine translation, Rule-based machine translation, Corpus based MT. **(Text Book-3: 5.4.1,5.5.2,8.1-8.7)** **(Text Book-1: 4)** **(Text Book-2: Recipe 4.8)**	

UNIT V	**(8 Hours)**
Information Extraction: IE Applications, IE Tasks, The General Pipeline for IE, Keyphrase Extraction, Implementing KPE, Named Entity Recognition, Building an NER System, NER Using an Existing Library, NER Using Active Learning, Named Entity Disambiguation and Linking, NEL Using Azure API. **Topic Modeling, Text Summarization, Recommender Systems for Textual Data, Question-Answering System.** **Applied NLP: NLP in a Search Engine, Social Media:** Applications, Unique Challenges, NLP for Social Data, Word Cloud, Tokenizer for SMTD, Trending Topics, Understanding Twitter Sentiment, Pre-Processing SMTD, Text Representation for SMTD, Customer Support on Social Channels, Memes and Fake News. **(Text Book-3: 11.4)** **(Text Book-1: 5,8,7)** **(Text Book-2: Recipe 5.6)**	

Course Outcomes: On Successful completion of this course, students will be able to	
1	**Outline** the fundamental concepts of Natural Language Processing and **Apply** text data acquisition techniques and perform basic text preprocessing using Python.
2	**Outline** the basic NLP tasks like N-gram language modelling, POS tagging, Parsing, Tokenization and WordNet based synset similarity and explore the same using python libraries.
3	**Identify** the various text to feature representation methods and their usage using python libraries and **apply** the same to develop various information retrieval models.
4	**Outline** various applications of NLP like Machine Translation, information extraction, etc. explore the same using python libraries.

Sl. No	Title of the Book	Name of the Author/s	Name of the publisher	Edition & Year
Textbooks				
1	Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems	Sowmya Vajjala, Bodhisattwa Majumder, Anuj Gupta, and Harshit Surana	O'Reilly	June 2020: First Edition
2	Natural Language Processing Recipes Unlocking Text Data with Machine Learning and Deep Learning using Python	Akshay Kulkarni, Adarsha Shivananda.	Apress	2019
3	Natural Language Processing and Information Retrieval	Tanveer Siddiqui, U S Tiwary	Oxford University Press	2nd Edition, 2010.
4	Natural language processing with Python	Steven, Ewan Klein, and Edward Loper	O'Reilly Media	1st Edition, 2009.
5	Python 3 Text Processing with NLTK 3 Cookbook	Jacob Perkins	PACKT	Second edition: August 2014
Reference Books				
1	Speech and Language	Daniel Jurafsky and	Prentice Hall,,	Low Price Edition,

	Processing: An introduction to Natural Language Processing, Computational Linguistics and Speech Recognition,	James H Martin		2000.
2	Foundations of Statistical Natural Language Processing	Christopher D. Manning	MIT Press	1999.

NATURAL LANGUAGE PROCESSING LABORATORY	
Sl.no.	**Experiments**
	PART-A
1	Create a corpus of minimum five files with minimum of 5 sentences in each file, search for a given pattern using regular expression from the corpus and list all the sentences that have the searched pattern by highlighting the first occurrence of the pattern for each sentence and also print the name of the file each sentence belongs to.
2	Write a program that takes a DFA and a string as an input and checks for the validity of the string.
3	Write a program that takes an NFA and a string as an input and checks for the validity of the string using DFS/BFS strategy.
4	Explore NLTK/Spacy and any other equivalent tools on the following fundamentals: a) Perform sentence and word tokenization. b) Remove stopwords in a text. c) Remove punctuations. d) Tag the words in a given text using POS tagger. e) Stemming and Lemmatization.
5	Write a program for predicting next word in the sequence using n-grams.
6	Write a program to create and read an input file, perform basic cleanup operations on the text in the file like removing HTML tags, URLs, remove the duplicate texts, perform spelling correction, and remove the additional spaces. Finally write the cleaned text into an output file.
7	Write a program to read an input file, delete the odd numbers in texts and replace the even numbers with their equivalent words. Finally write the updated text into an output file.
	PART-B
8	Write a program that takes CFG for a language and a sentence belongs to a language as an input and generates parse tree for the same using various parsers available in NLTK and Spacy.
9	Write a program to Extract names, emails and phone numbers from a text.
10	Write a program to retrieve the information from a text file using verb/noun keywords as a search query.
11	Perform Information extraction over a given text that includes entity and relation extraction.
12	Classify a text as positive/negative sentiment.
13	Find Synonyms from NLTK WordNet.
14	Develop a gender classifier by using the existing classifiers.

Course outcomes:

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

1. Develop NLP programs in Python.
2. Demonstrate the use of modern NLP techniques for processing of text.
3. Explore tools like NLTK/Spacy in pre-processing and some advanced processing of texts.

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
	2		2										2	

B.E. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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 **(9L+6P)**

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 **(9L+6P)**

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 **(8L+6P)**

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 **(8L+5P)**

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 **(8L+5P)**

World Wide Web and HTTP: World Wide Web. Hypertext Transfer Protocol (HTTP), SMTP protocol, DNS: Name Space and Resolution, Telnet.
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.
Differentiated Services(DFFSERV): DS Field, Per-Hop Behaviour, Traffic Conditioners.
Section:26.1, 26.3, 26.4, 26.6, 30.1, 30.2, 30.3,30.4

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.Swtich 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	5thEdition, 2013
2	Computer Networks: A Systems Approach	Larry L Peterson and Bruce S Davie	Elsevier	5th Edition, 2011
Reference Books				
1	Data and Computer Communications	William Stallings	Pearson Education	10th Edition, 2013
2	Computer Networking: A Top-Down Approach	Kurose James F, Ross Keith W.	Pearson Education	6th Edition, 2017
3	Computer Networks	Andrew S. Tanenbaum and David J. Wetherall	Pearson Educatin	5th Edition,2011
4	Unix Network Programming, Interprocess Communications,	W Richard Stevens	Pearson Educatin	2nd 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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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 Ardunio through BT
6	Control LED light using Voice command . Use Ardunio 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 Thingspeak 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: 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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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 **(5 Hours)**

Introduction to Indian Knowledge Systems (IKS)

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 **(5 Hours)**

Fundamental concepts in Science in IKS

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 **(5 Hours)**

Humanities and Social sciences in IKS

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

Reference Books	
1	**Introduction to Indian Knowledge System- concepts and applications**, B Mahadevan, Vinayak Rajat Bhat, Nagendra Pavana R N, 2022, PHI Learning Private Ltd, ISBN-978-93-91818-21-0
2	**Traditional Knowledge System in India**, Amit Jha, 2009, Atlantic Publishers and Distributors (P) Ltd., ISBN-13: 978-8126912230,
3	**Knowledge Traditions and Practices of India**, Kapil Kapoor, Avadesh Kumar Singh, Vol. 1, 2005, DK Print World (P) Ltd., ISBN 81-246-0334,
	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=EAIaIQobChMInp-Jtb_p8gIVTeN3Ch27LAmPEAAYASAAEgIm1vD_BwE

Course Articulation Matrix (CO-PO and CO_PSO MAPPING)

CO/PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CO1	-	-	-	-	-	1	1	-	-	-	1
CO2	2	-	-	-	-	2	1	-	-	-	2
CO3	1	-			-	-	-	-	-	-	2

B.E. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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	1st 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 Outcomes	**Program Outcomes**											**PSOs**		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**1**	**2**	**3**
	2													2

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

HIGH PERFORMANCE COMPUTING

Course Code	**S6CSPE01**	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 Hour	-

Course objectives: The course will enable students to

1. To analyse typical parallel algorithm models and its application in scientific computing. (Analysis, Application)
2. To develop programs using message-passing paradigm. (Synthesis)
3. To learn how GPUs works using the CUDA architecture and its applications (Comprehension, Application)
4. To gain practical knowledge by giving hands on experience inGraphics Interoperability, CUDA C on multiple GPUs and CUDA toolkit (Synthesis)
5. To analyze the latest parallel computing techniques and research - prepare a technical document and make a presentation (Analysis, Syntheses and Evaluation)

To develop open ended solution for any of the identified high performance computing problems

UNIT 1 **8 Hours**

Principles of Parallel Algorithm Design: Preliminaries, Decomposition Techniques, Characteristics of Tasks and Interactions,Mapping Techniques for Load Balancing ,

Basic Communication Operations: One-to-All Broadcast and All-to-One Reduction , All-to-All Broadcast and Reduction
Analytical Modeling of Parallel Programs: Sources of Overhead in Parallel Programs(5.1) , the Effect of Granularity on Performance.

UNIT 2 **8 Hours**

Programming Using the Message-Passing Paradigm:
Principles of Message-Passing Programming , The Building Blocks: Sendand Receive Operations , MPI: the Message Passing Interface

UNIT 3 **8 Hours**

Programming Shared Address Space Platforms:
Thread Basics , Why Threads? , The POSIX Thread API,Thread Creation and Termination, Synchronization Primitives in Pthreads ,Controlling Thread and Synchronization Attributes , Thread Cancellation,Composite Synchronization Constructs, OpenMP: a Standard for Directive Based Parallel Programming

UNIT 4 **9 Hours**

Why CUDA? Why Now?: The Age of Parallel Processing, Central Processing UNITs,The Rise of GPU Computing, A brief history of GPUs, Early GPU computing, CUDA,What is CUDA architecture, using the CUDA architecture, Applications of CUDA,Medical Imaging, Computational Fluid Dynamics, Environmental Science, Introductionto CUDA C: A First Program, Hello world, A kernel call, Passing parameters, Queryingdevices, using device properties, Parallel Programming in CUDA C:CUDA parallel
programming, Summing vectors, A fun example.

UNIT 5 **9 Hours**

Graphics Interoperability: Graphics Interoperation, GPU Ripple with Graphics Interoperability - the GPUAnimBitmap structure; GPU Ripple Redux , Heat transfer with Graphics Interop, DirectX Interoperability
CUDA C ON multiple GPUS : Zero-Copy Host Memory -Zero-Copy Dot Product; Zero-Copy Performance, Using Multiple GPUs, Portable Pinned Memory
CUDA Tools: CUDA Toolkit- CUFFT, CUBLAS, NVIDIA GPU Computing SDK, Debugging CUDA C

Topics for Open Ended Activity :
Self Learning component

Parallel Computing models, parallel virtual machines and usage, Data-Parallel Algorithms, Graphics Interoperability, Parallel Computing Strategies, Linear Algebra Image/Video Processing computation, Data Compression, Physically-Based Simulation, Parallel pragma/directive based frameworks, Usage and Performance analysis and comparisons– FPGA, GPUs, GP-GPUS, CPU-GPUs, Analysis of HPC Benchmark Suite/Tools/Solutions/Standard etc, and so on.

Course outcomes:

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

1. Select and analyze the characteristics of various parallel computing platforms.

2. Choose a suitable platform for parallel computing.

3. Analyze simple parallel algorithm models.

4. Apply the principles of message-passing programming construct to solve engineering problems.

5. Design and develop parallel programs using CUDA and OpenMp programming interface

Sl. No.	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
Textbooks				
1	Introduction to parallel computing (chapters 01,2.1-2.5,3,4.1.1-4.1.3, 5.1, 5.2, 5.3, 6,7)	Ananth Grama,AnshulGupta, Vipinkumar,GeorgeKarypis	Pearson education publishers	second edition, 2015
2	CUDA by example (Chapters 1 ,3, 4, 5, 8, 11, 12)	Jason Sanders Edward Kandrot	NVIDIA Corporation	2015
Reference Books				
1	Parallel Programming for Multicore and cluster systems	Thomas Rauber and GudulaRunger	Springer	International Edition,2009
2	Introduction to Parallel Computing	Niranjan N. Chiplunkar, Raju K	Wiley	2020

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**							**2**		
CO5			**2**		**2**							**2**		
Overall CO	2	**2**	**2**		**2**							**2**		

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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
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 Perpectives	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. 100 7/978-1-4842-4460-9
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 Prosise, 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. COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - VI

REAL TIME BIG DATA ANALYTICS

Course Code	S6CIPE02	CIE Marks	50
Teaching Hours/Week (L:T:P)	(03:0:0)	SEE Marks	50
Credits	03	Exam Hours	03
Lecture Hours	42hrs	CIE Marks	-

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 serializationframeworks 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 Bigdata .
6	**Describe** the real time views of Cassandra's data model for Bigdata
7	**Demonstrate** how to implement the concepts of queuing and stream processing using real-world tools

UNIT I **8 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 **8 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 **7 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	4th Edition,2015
2	Big Data and Analytics	SeemaAcharya, Subhashini Chellappan	Wiley India Publications,	May 2015

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