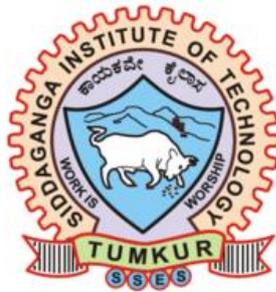


Scheme of Teaching, Examination and Syllabus
B.E. COMPUTER SCIENCE AND ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

Batch: 2022-23

Third Year
(V and VI SEMESTER)
(Effective from the academic year 2024-2025)



SreeSiddaganga Education Society®

Siddaganga Institute of Technology

(An Autonomous institute affiliated to Visvesvaraya Technological University, Belagavi)

(Approved by AICTE, New Delhi, Accredited by NAAC and ISO 9001-2015 certified)

B.H. Road, Tumakuru-572 103, Karnataka, India

Phone: Direct +91-816-2282696, Fax: +91-816-2282994

E-mail: principal@sit.ac.in web: www.sit.ac.in

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

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		Software Engineering and Project Management	Dept.	3	0	0	3.5(48 hrs)	3	50	50	100	3
2.	IPCC	S5CCSI01	Database Management System (I)	Dept.	3	0	2	3.5(50 hrs)	3	50	50	100	4
3.	IPCC	S5CCSI02	Artificial Intelligence and Machine Learning (I)	Dept.	3	0	2	3.5(50 hrs)	3	50	50	100	4
4.	PCCL	S5CCSL01	Data Mining and Visualization Lab	Dept.	0	0	2		3	50	50	100	1
5.	PEC		Professional Elective Course-I	Dept.	3	0	0	3.5(48 hrs)	3	50	50	100	3
6.	PROJ		Mini Project / Extension Survey Project	Dept.	0	0	4		3	100	-	100	2
7.	AEC		Research Methodology and IPR (Board: IEM)	ME, IM, CH	2	2	0	2.0(34hrs)	3	50	50	100	3
8.	HSMS	SHS06	Environmental Studies (Board: CV)	CV	2	0	0	2.0(32hrs)	3	50	50	100	2
9.	NCMC		Soft Skills (Additional Course offered by SIT)	T&P	0	2	0		-	100	-	100	0
10.	NCMC	NS	National Service Scheme (NSS)	NSS CO	0	0	2			100	-	100	0
		PE	Physical Education (PE) (Sports and Athletics)	PED									
		YO	Yoga	PED									
			Total						550	350	900	22	
	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	S5CCSPE03	Business Intelligence and Analysis
S5CCSPE02	Social Network Analysis	S5CCSPE04	Knowledge Representation and Reasoning

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

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

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

CIE procedure for Mini-project:

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

No SEE component for Mini-Project.

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

SCHEME OF TEACHING AND EXAMINATION (2022 Scheme) (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./week				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)		3	0	2	3.5(50 hrs)	3	50	50	100	4
2.	PCC	S6CCS01	Computer Networks		3	2	0	3.5(50 hrs)	3	50	50	100	4
3.	PEC		Professional Elective Course-II		3	0	0	3.5(48 hrs)	3	50	50	100	3
4.	OEC		Open Elective Course-I		3	0	0	3.5(48 hrs)	3	50	50	100	3
5.	PROJ		Major Project Phase I		0	0	4		3	100	-	100	2
6.	PCCL	S6CCSL01	Internet of Things Lab		0	0	2		3	50	50	100	1
7.	AEC		Aptitude Related Analytical Skill		0	0	2		1½	50	50	100	1
8.	NCMC	NS	National Service Scheme (NSS)	NSS CO	0	0	2			100	-	100	0
		PE	Physical Education (PE) (Sports and Athletics)	PED									
		YO	Yoga	PED									
			Total							500	300	800	18
		AAP	AICTE Activity Points (Applicable for both Regular and Lateral Entry students)		40 hours community service to be documented and produced for the examination								
<p>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.</p>													
Professional Elective Course (PEC) (Offered by the Department)													
S6CCSPE01	Cloud Computing			S6CIPE02	Real Time Big Data Analytics								
S6CIPE01	AI Driven Cyber Security			S6CCSPE02	Recommender System								
<p>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.</p>													
<p>National Service Scheme /Physical Education/Yoga: All students have to register for any one of the courses namely National Service Scheme (NSS), Physical Education</p>													

(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.

Scheme of Teaching, Examination and Syllabus
B.E. COMPUTER SCIENCE AND ENGINEERING
(ARTIFICIAL INTELLIGENCE AND DATA SCIENCE)

Batch: 2021-22

V SEMESTER
(Effective from the academic year 2023-2024)

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) Outcome Based Education (OBE) And Choice Based Credit System (CBCS) SEMESTER-V			
SOFTWARE ENGINEERING AND PROJECT MANAGEMENT			
Course Code		CIE Marks	50
Teaching Hours/Week(L:T:P)	(3:0:0)	SEE Marks	50
Credits	03	Exam Hours	03
Lecture Hours	40Hrs	Practical Hours	26Hrs
Course objectives: The course will enable students to <ol style="list-style-type: none"> To define software engineering and explain importance of project management. To explain process of gathering software requirements and modeling complex software systems. To explain process of designing complex software products, implementing the software design and testing the developed product. To explain the importance of project management To learn about software engineering and project management in an industrial context. 			
UNIT-1		(8L+0THrs)	
1.1 Software Engineering: Engineering, Creativity and Craft (R1-2); Professional SoftwareDevelopment (T1-1.1); Software Engineering (R1-1, R2-1.1); Birth of Software Engineering(R1-1); Foundations of Software Engineering Discipline (R1-3); Experts and Learning (R1-3);ExpertsatManagingComplexity(R1-3);SoftwareEngineeringBodyofKnowledge(R2-14.4);			
1.2 Software Processes: Software Process Models (T1-2.1, R2-2.2); Process Activities (T1-2.2);Copingwith Change(T1-2.3);			
1.3 Agile Software Development: Agile Methods (T1-3.1); Agile Development Techniques (T1-3.2);Agile Project Management (T1-3.3);			
1.4 Project Management: The Project Management Body of Knowledge (W2); ProjectManagementKnowledgeAreas(W3)			
UNIT-2		(8L+0THrs)	
2.1 Requirements Engineering: Requirements Engineering (T1-4); Functional and Non-Functional Requirements (T1-4.1); Requirements Engineering Process (T1-4.2); RequirementsElicitation (T1-4.3); Requirements Specification (T1-4.4); Requirements Validation (T1-4.5,R2-4);Requirements Change (T1-4.6);			
2.2 System Modelling: Context Models (T1-5.1); Interaction Models (T1-5.2); Structural Models(T1-5.3);Behavioural Models (T1-5.4);Model Driven Engineering (T1-5.5);			
2.3 Architectural Design: Architectural Design Decisions (T1-6.1); Architectural Views (T1-6.2);ArchitecturalPatterns (T1-6.3);Application Architectures (T1-6.4,R2-5.48.1)			
UNIT-3		(8L+0THrs)	
3.1 Design and Implementation: Incrementalism in Software Development (R1-6); ObjectOriented Design using UML (T1-7.1); Design Patterns (T1-7.2); Achieving Quality Attributes(R2-5.5);Writing Programs (R2-7);			
3.2 Software Testing: Development testing (T1-8.1, R2-8); Test driven development (T1-8.2);Releasetesting (T1-8.3); User testing (T1-8.4);			

3.3	Software Evolution: Software Evolution (T1-9.1); Legacy Systems (T1-9.2); SoftwareMaintenance (T1-9.3)
UNIT-4 (8L+0THrs)	
4.1	Project Management: Overview (W3); Risk Management (T1-22.1); Boehm's Top Ten RiskItems (R2-3.4); Members of the Development Team (R2-1.7); Teamwork (T1-22.3); ManagingPeople(T1-22.2);
4.2	Project Planning: Working Iteratively (R1-4); Plan Driven Development (T1-23.2); TheProject Plan (R2-3.5); Agile Planning (T1-23.4); Estimation Techniques (T1-23.5, R2-3.3);ProjectScheduling (T1-23.3); COCOMO CostModeling (T1-23.6);
4.3	Complexity Management: Software Complexity (R1-III); Methods of Managing Complexity(R1-III,R2-6.2);
4.4	Quality Management: What is Good Software? (R2-1.3); Feedback in Software Development(R1-5); Software Quality (T1-24.1); McCall's Quality Model (R2-1.3); Software Standards (T1-24.2); Reviews and Inspections (T1-24.3); Quality Management in Agile Development (T1-24.4);SoftwareMeasurements (T1-24.5);
4.5	Configuration Management: Version Management (T1-25.1); System Building (T1-25.2);ChangeManagement (T1-25.3); ReleaseManagement (T1-25.4)
UNIT-5 (7L+0THrs)	
5.1	MLOps: Need and benefits of MLOps, vs DevOps, MLOps Phases, MLOps architecture and components
5.2	Industrial Case Study: Defining, architecting, designing, developing, testing, releasing, maintaining a complex software product and managing the associated project.

Courseoutcomes:

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

1. Analyze fundamental processes of software engineering and project management.
2. Analyze the functional and non-functional requirements.
3. Practice the software estimation, architecture, and design principles.
4. Implement software design and test the quality of software products.
5. Identify and practice specific techniques of project management.

Sl. No.	TitleoftheBook	NameoftheAuthor/s	NameofthePublisher	Edition andYear
Textbooks				
T1	SoftwareEngineering ISBN:978-93-325-8269-9	IanSommerville	Pearson EducationLimited	10 th Edition2017
Reference Books				

R1	Modern Software Engineering: Doing What Works to Build Better Software Faster ISBN:978-0-13-731491-1	David Farley	Addison-Wesley	2022
R2	Software Engineering: Theory and Practice ISBN:978-81-317-6062-8	Shari Lawrence Pfleeger Joanne M Atlee	Pearson	4 th Edition 2013
Web Resources				
W1	Supporting material for T1 https://software-engineering-book.com/			
W2	PMBOK 7th Edition Summary https://www.projecttimes.com/articles/the-pmbok-guide-seventh-edition-summary/			
W3	PMBOK Knowledge Areas https://www.projectengineer.net/the-10-pmbok-knowledge-areas/			

Course Articulation matrix (CO-PO Mapping)

Course Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
CO1	2													2	
CO2		2												2	
CO3			3											2	
CO4											3			2	
CO5			3								3			2	
Overall CO	2	2	3								3			2	

Degree of compliance 1: Low 2: Medium 3: High

Program Articulation Matrix: (PO-PSO Mapping)

Course	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
SE & PM	2	2	3								3			2	

Degree of compliance 1: Low 2: Medium 3: High

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) Outcome Based Education (OBE) And Choice Based Credit System (CBCS) SEMESTER – V			
DATABASE MANAGEMENT SYSTEM			
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	40Hrs	Practical Hours	26Hrs
<p>Course objectives: The course will enable students to</p> <ol style="list-style-type: none"> 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-1		(6hrs)	
<p>Databases and Database Users: Introduction; An example; characteristics of the database approach; actors on the scene; workers behind the scene; advantages of using the DBMS approach; A brief history of database Applications; when to use a DBMS.</p> <p>Text 1 : Chapter 1 : 1.1 – 1.8</p> <p>Database System – Concepts And Architecture: Data models, schemas, and instances; three schema architecture and data independence; database languages and interfaces; the database system environment; centralized and client/server/architectures for DBMSs. Classification of database management system.</p> <p>Text 1 : Chapter 2 : 2.1 to 2.6</p>			
UNIT-2		(8hrs)	
<p>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.</p> <p>Text 1 : Chapter 7 : 7.1 to 7.7</p> <p>Relational Model: Relational Model Concepts; Relational Model Constraints and Relational Database Schemas; Update Operations and Dealing with Constraint Violations; Relational Database Design using ER- to-Relational Mapping.</p> <p>Text 1 : Chapter 3 : 3.1 to 3.3, Chapter 9 : 9.1</p>			
UNIT-3		(9hrs)	
<p>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; Additional Features of 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</p> <p>MangoDB tutorial, MangoDB operators, DB commands, Database, collection, CURD</p> <p>URL: www.javatpoint.com/nosql-databases</p>			
UNIT-4		(8 hrs)	

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.
Text 1 : Chapter 15 : 15.1 to 15.5 , Chapter 16 : 16.2

UNIT-5 **(9 hrs)**

Transaction Processing Concept: Introduction to transaction processing; transaction and system concepts; desirable properties of transactions, characterizing schedules based on recoverability and serializability; transaction support inSQLText1 : 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
Text1 : Chapter 22 : 22.1 – 221.2, Chapter 23: 23.1 to 23.5

Course outcomes:

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

CO1: Describe the fundamentals of database technologies.

CO2: Design an ER diagram and transform it to a relational model for a given database specification.

CO3: Design SQL and NoSQL queries to perform CRUD (Create, Retrieve, Update and delete) operations on database.

CO4: Apply Informal Design guidelines and normalization techniques to improve database design

CO5: 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 and Year
Textbooks				
1	Fundamentals of Database Systems	Elmasri and Navathe	Pearson Education	6 th Edition, 2011
Reference Books				
1	Data base System Concepts.	Silberschatz, Korth and Sudharshan.	McGraw-Hill	6 th Edition, 2010
2	Database Management Systems.	Raghu Ramakrishnan and Johannes Gehrke	McGraw-Hill.	3 th Edition, 2010

URL: www.javatpoint.com/nosql-databases

Course Articulation matrix(CO-PO and CO-PSO mapping)

Course Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
CO1	2	2	2										2		
CO2	2		2										2		
CO3	2		3										3		
CO4	2	2											2		
CO5	2		2										2		
Overall CO	2	2	3										3		

Program articulation matrix:

Course Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
CO 1-5	2	2	3										3		

Degree of compliance 1: Low 2: Medium 3: High

DATA BASE MANAGEMENT SYSTEMS LABORATORY

Course objectives: After the completion of the course, the student will be able to:

1. **Apply** the knowledge of database management system development process and conduct the experiments using SQL and NoSQL queries to find the solution for a givendatabase problem.
2. **Analyze** and design solutions for database system components to meet the specified needs of online transaction processing and information systems like Banking systems, Ticket Reservation systemsetc..
3. **Develop** code for stored programs, triggers assertions and to generatereport
4. **Contribute** to the team as a member, lead theteam.

Sl. no.	Experiments
1	<p>Suppose a movie_studio has several film crews. The crews might be designated by a given studio as crew1, crew 2, and so on. However, other studios might use the same designations for crews, so the attribute crew_number is not a key for crews. Movie_studio holds the information like name, branch and several locations. Each crew holds information like sector andstrength.</p> <ol style="list-style-type: none"> i) Establish the database by normalizing up to 3NF and considering all schema level constraints ii) Write SQL insertion query to insert few tuples to all therelations iii) List all movie studios which are not used a singlecrews. iv) Retrieve the movie studio which uses highest strengthcrew. 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 specificstudio.
2	<p>The production company is organized into different studios. We store each studio's name branch and location; every studio must own at least one movie. We store each movie's title, sensor number and year of production. Star may act in any number of movies and we store each actors name and address.</p> <ol style="list-style-type: none"> i) Establish the database by normalizing up to 3NF and considering all schema level constraints ii) Write SQL insertion query to insert few tuples to all therelations iii) List all the studios of the movie“xyz”;

	<p>iv) List all the actors , acted in a movie ‘xyz’</p> <p>v) Write a procedure to list all movies produced during the specific year.</p> <p>vi) Write a deletion trigger, does not allow to deleting current year movies.</p>
3	<p>The production company is organized into different studios. We store each studio’s name branch and location; a studio own any number of Cartoon-serials. We store each Cartoon-Serial’s title, serial number and year of production. Star may do voices in any number of Cartoon-Serials and we store each actors name and address.</p> <p>i) Establish the database by normalizing up to 3NF and considering all schema level constraints</p> <p>ii) Write SQL insertion query to insert few tuples to all the relations</p> <p>iii) Find total no of actors, do voiced in a Cartoon-Serials ‘xyz’</p> <p>iv) Retrieve name of studio, location and Cartoon-Serials title in which star “abc” is voiced.</p> <p>v) vii. Write a procedure to list all Cartoon-Serials produced during the specific year.</p> <p>vi) v. Write a deletion trigger, does not allow to deleting current year Cartoon-Serials.</p>
4	<p>Car marketing company wants keep track of marketed cars and their owner. Each car must be associated with a single owner and owner may have any number of cars. We store car’s registration number, model & color and owner’s name, address & SSN. We also store date of purchase of each car.</p> <p>i) Establish the database by normalizing up to 3NF and considering all schema level constraints</p> <p>ii) Write SQL insertion query to insert few tuples to all the relations</p> <p>iii) Find a person who owns highest number of cars</p> <p>iv) Retrieve persons and cars information purchased on the day 11-11-11</p> <p>v) Write a insertion trigger to check date of purchase must be less than current date (must use systemdate)</p> <p>vi) Write a procedure to list all cars and owner information purchased during the specific year.</p>
5	<p>Puppy pet shop wants to keep track of dogs and their owners. The person can buy maximum three pet dogs. We store person’s name, SSN and address and dog’s name, date of purchase and sex. The owner of the pet dogs will be identified by SSN since the dog’s names are not distinct.</p> <p>i) Establish the database by normalizing up to 3NF and considering all schema level constraints</p> <p>ii) Write SQL insertion query to insert few tuples to all the relations</p> <p>iii) List all pets owned by a person ‘Abhiman’.</p>

	<p>iv) List all persons who are not owned a singlepet</p> <p>v) Write a trigger to check the constraint that the person can buy maximum three pet dogs</p> <p>vi) Write a procedure to list all dogs and owner details purchased on the specificdate.</p>
6	<p>No SQL:</p> <p>Lab 1. Installation and set up of MongoDB client and server</p> <p>Lab 2. Create a database collection using MongoDB environment. For example a documentcollection meant for analyzing Restaurant records can have fields like restaurant_id, restaurant_name, customer_name, locality, date, cuisine, grade, comments. etc.</p> <p>Lab 3. Create database using INSERT, UPDATE, UPSERTS, DELETE and INDEX.</p> <p>Lab 4. Practice writing simple MongoDB queries such as displaying all the records, display selected records with conditions</p> <p>Lab 5. Experiment with MongoDB comparison and logical query operators - \$gt, \$gte, \$lt, \$lte, \$in, #nin, \$ne, \$and, \$or, \$not</p> <p>Lab 6. Practice exercise on element, array based and evaluation query operators - \$exists, \$type, \$mod, \$regex</p>
	<p>Course outcomes:</p> <p>After the completion of the course, the student will be able to:</p> <ol style="list-style-type: none"> 1. Apply the knowledge of database management system development process and conduct the experiments using SQL and NoSQL queries to find the solution for givendatabase problem. 2. Design ER Model & its mapping to relational for a given problem. 3. Develop code for stored programs& triggers
	<p>Conduct of Practical Examination:</p> <ol style="list-style-type: none"> 1. All laboratory experiments are to be included for practical examination. 2. Breakup of marks and the instructions printed on the cover page of answer script to be strictly adhered bythe examiners. 3. Students can pick one experiment from the questions lot prepared by the examiners. 4. Change of experiment is allowed only once and 20% Marks is to be deducted.

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) Outcome Based Education (OBE) And Choice Based Credit System (CBCS) SEMESTER – V			
ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (I)			
Course Code	S5CCSI02	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:2)	SEE Marks	50
Credits	4.0	Exam Hours	3
Lecture Hours	40 hrs	Practical Hours	26hrs
Course objectives: The course will enable students to <ol style="list-style-type: none"> 1. Explain the basics of Artificial Intelligence and Machine Learning algorithms 2. Identify the problems where Artificial Intelligence and Machine Learning techniques are applicable. 3. Discuss knowledge representation issues and different kind of learning algorithms. 4. Compare learning strategy adopted by various kinds of machine learning algorithms. 			
UNIT-1			(8hrs)
What is artificial intelligence? What is AI? Acting humanly and thinking humanly, thinking rationally and acting rationally, Intelligent Agents: Agents and Environments, Good Behavior: The concept of Rationality: Rationality, Omniscience, Learning and autonomy, The nature of Environments: specifying the task environment, properties of task environments, The structure of Agents: Agent Programs, simple reflex agents, Model-based reflex agents, Goal-based agents, Utility-based agents, Learning agents, How the components of agents programme work, Solving problems by Searching: problem-solving agents, well-defined problems and solutions, Example problems. Searching for Solutions: infrastructure for search algorithms, measuring problem-solving performance. Textbook-1: Chapter 1: 1.1 to 1.5, Chapter 2: 2.1 to 2.5			
UNIT-2			(8hrs)
Search strategies: Uninformed Search strategies: BFS, uniform-cost search, DFS, depth-limited search, iterative deepening depth-first search, bidirectional search, comparing uniformed search strategies, Informed search strategies: Greedy best-first search, A* search, Memory-bounded heuristic search, learning to search better, Constraint satisfaction problems: ; Example problem: Map coloring, Example problem: Job-shop scheduling, Variations on the CSP formalism, constraint propagation: Inference in cps: Node consistency, Arc consistency, Path consistency, K-consistency, Global constraints, Backtracking search for CSPs; Variable and value ordering, Interleaving search and inference, Intelligent backtracking: Looking backward Textbook-1: Chapter 3: 3.1 to 3.5, Chapter 6: 6.1 to 6.5			
UNIT-3			(8hrs)
Concept Learning Introduction, Concept Learning: Well Posed Learning problem, Designing Learning systems, Perspectives and Issues in machine learning, Concept Learning: Introduction, A Concept Learning Task, Concepts Learning Search, Version Spaces and Candidate Elimination Algorithm, Remarks on version space and Candidate Elimination. Textbook-2: Chapter 2: 2.1 to 2.7			
UNIT-4			(8 hrs)
Bayesian Learning: Introduction, Bayes Theorem, Bayes Theorem and Concept Learning, Maximum Likelihood and least squared error hypotheses, Minimum Description Length Principle, Bayes Optimal Classifier, and Naive Bayes Classifier, An Example: Learning to Classify Text, Bayesian Belief network, EM Algorithm- General Statements of EM Algorithm. Textbook-2: Chapter 6: 6.1 to 6.12			

UNIT-5**(8 hrs)**

Neural Networks: Introduction, Neural Network Representations, Appropriate problems for Neural Networks, Perceptron's, Multilayer Networks and Back Propagation Algorithms.

Reinforcement Learning: Introduction, The learning task, Q-Learning, Nondeterministic rewards and actions, and Temporal difference learning.

Textbook-2: Chapter 4: 4.1 to 4.6, Chapter 13: 13.1 to 13.5

Course outcomes:

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

1. Analyse and develop Artificial Intelligent agents for simple applications.
2. Apply searching algorithms to develop Artificial Intelligent agents.
3. Analyse and apply concept learning tasks to solve applications of ML.
4. Apply Bayesian learning for classification problems.
5. Apply neural networks and reinforcement learning concepts to demonstrate applications in ML

Sl. No.	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
Textbooks				
1	Artificial Intelligence: A Modern Approach	Stuart RusselPeterNorvig	Pearson Education	3 rd Edition, 2013/ 4 th edition 2020
2	Machine Learning	Tom M Mitchell	McGraw Hill Education	1 st Edition, 2017
Reference Books				
1	Machine Learning	S. Sridhar, M.vijayalakshmi	Oxford University Press	1th Edition, 2021
2	Hands-On Machine Learning with Scikit-Learn and Tensor Flow: Concepts, Tools, and Techniques to Build Intelligent Systems	AurelienGeron	Shroff/O'Reilly Media	3 rd Edition, 2022.
3	Introduction to Machine Learning	EthemAlpaydin	PHI Learning Pvt. Ltd	2 nd Edition, 2014.
4	StructuresandStrategiesforComplex ProblemSolving	GeorgeFLuger	Pearson Education	5 th Edition, 2011

Lab Syllabus:

Implementation of programs on the following Artificial Intelligence concepts:

1. Agent Programs (simple reflex agents, Goal-based agents)
2. Breadth First Search
3. Depth First Search
4. Best First Search
5. A* Search
6. Constraint Satisfaction Problems (CSPs)
7. Candidate Elimination Algorithm
8. Bayes Theorem
9. Bayesian Belief Network
10. EM Algorithm
11. Back Propagation Algorithm

12. Q-learning
13. Temporal Difference Learning

Course outcomes for lab:

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

1. Apply AI techniques to solve search problems.
2. Design and implement searching and CSP problem using C/Python/PROLOG.
3. Implement concept learning tasks to solve applications of ML.
4. Implement Bayesian learning for classification problems.
5. Implement neural networks and reinforcement learning concepts to demonstrate applications in ML

COURSE ARTICULATION MATRIX															
Course Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
CO1	1		2											2	
CO2	1		2											2	
CO3		1	2											2	
CO4	1	1												1	
CO5	1		2											2	
Overall CO	1	1	2											2	
PROGRAM ARTICULATION MATRIX															
Course Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
Overall CO	1	1	2											2	

B.E COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE & MACHINE LEARNING)
Outcome Based Education (OBE) And Choice Based Credit System (CBCS)
SEMESTER – V

DATA MINING & VISUALIZATION

Course Code	S5CCSL01	CIE Marks	50
TeachingHours/Week (L:T:P)	(0:0:2)	SEE Marks	50
Credits	1	Exam Hours	3
Lecture Hours	-	Practical Hours	26hrs

Course objectives: The course will enable students to

1. Learn to setup Android application development environment and AI technologies.
2. Develop native mobile apps to extend databases and use them with respect to AI context.
3. Learn to develop user interfaces for interacting with apps and triggering actions.
4. Interpret tasks used in handling multiple activities.
5. Identify options to save persistent application data.

Sl. no.	Experiments																																																																																																						
1	<p>1. Experiment to be conducted using WEKA tool:</p> <table border="1" style="margin-left: 40px;"> <thead> <tr> <th>1</th> <th>outlook</th> <th>temperatu</th> <th>humidity</th> <th>windy</th> <th>play</th> </tr> </thead> <tbody> <tr><td>2</td><td></td><td></td><td></td><td></td><td></td></tr> <tr><td>3</td><td>sunny</td><td>85</td><td>85</td><td>FALSE</td><td>no</td></tr> <tr><td>4</td><td>sunny</td><td>80</td><td>90</td><td>TRUE</td><td>no</td></tr> <tr><td>5</td><td>overcast</td><td>83</td><td>86</td><td>FALSE</td><td>yes</td></tr> <tr><td>6</td><td>rainy</td><td>70</td><td>96</td><td>FALSE</td><td>yes</td></tr> <tr><td>7</td><td>rainy</td><td>68</td><td>80</td><td>FALSE</td><td>yes</td></tr> <tr><td>8</td><td>rainy</td><td>65</td><td>70</td><td>TRUE</td><td>no</td></tr> <tr><td>9</td><td>overcast</td><td>64</td><td>65</td><td>TRUE</td><td>yes</td></tr> <tr><td>10</td><td>sunny</td><td>72</td><td>95</td><td>FALSE</td><td>no</td></tr> <tr><td>11</td><td>sunny</td><td>69</td><td>70</td><td>FALSE</td><td>yes</td></tr> <tr><td>12</td><td>rainy</td><td>75</td><td>80</td><td>FALSE</td><td>yes</td></tr> <tr><td>13</td><td>sunny</td><td>75</td><td>70</td><td>TRUE</td><td>yes</td></tr> <tr><td>14</td><td>overcast</td><td>72</td><td>90</td><td>TRUE</td><td>yes</td></tr> <tr><td>15</td><td>overcast</td><td>81</td><td>75</td><td>FALSE</td><td>yes</td></tr> <tr><td>16</td><td>rainy</td><td>71</td><td>91</td><td>TRUE</td><td>no</td></tr> <tr><td>17</td><td></td><td></td><td></td><td></td><td></td></tr> </tbody> </table> <ol style="list-style-type: none"> 1. Preprocess and Classify panels 2. Draw the histogram to show how the values of the play class occurs for each value of the outlook attribute 3. Derive minimum and maximum values, mean, and standard deviation 4. Perform operations such as filter, delete, invert, Pattern, Undo, Edit, search, Select, Conversions etc 5. Build the decision tree and analyze the weather data. 6. Examine the Output , classification error and Kappa statistics 7. Visualize threshold curve 8. Apply Logistic Regression model to classify 	1	outlook	temperatu	humidity	windy	play	2						3	sunny	85	85	FALSE	no	4	sunny	80	90	TRUE	no	5	overcast	83	86	FALSE	yes	6	rainy	70	96	FALSE	yes	7	rainy	68	80	FALSE	yes	8	rainy	65	70	TRUE	no	9	overcast	64	65	TRUE	yes	10	sunny	72	95	FALSE	no	11	sunny	69	70	FALSE	yes	12	rainy	75	80	FALSE	yes	13	sunny	75	70	TRUE	yes	14	overcast	72	90	TRUE	yes	15	overcast	81	75	FALSE	yes	16	rainy	71	91	TRUE	no	17					
1	outlook	temperatu	humidity	windy	play																																																																																																		
2																																																																																																							
3	sunny	85	85	FALSE	no																																																																																																		
4	sunny	80	90	TRUE	no																																																																																																		
5	overcast	83	86	FALSE	yes																																																																																																		
6	rainy	70	96	FALSE	yes																																																																																																		
7	rainy	68	80	FALSE	yes																																																																																																		
8	rainy	65	70	TRUE	no																																																																																																		
9	overcast	64	65	TRUE	yes																																																																																																		
10	sunny	72	95	FALSE	no																																																																																																		
11	sunny	69	70	FALSE	yes																																																																																																		
12	rainy	75	80	FALSE	yes																																																																																																		
13	sunny	75	70	TRUE	yes																																																																																																		
14	overcast	72	90	TRUE	yes																																																																																																		
15	overcast	81	75	FALSE	yes																																																																																																		
16	rainy	71	91	TRUE	no																																																																																																		
17																																																																																																							

9. Measure the log likelihood of the clusters of training data. (Consider large data set.)

2

Consider the following data set

Relation: employee					
No.	age Nominal	income Nominal	stud Nominal	credtrate 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

- i) Load ARFF file and explore knowledge flow interface
- ii) configure the data source , check the status area after executing the configuration
- iii) Perform operations such as Attribute Selection, Filter, Classify, Data Sink, Visualization and Evaluation
- iv) Apply incremental learning and analyze the result
- v) do clustering : use generator properties, two clustering schemes, and result panel
- vi) Generate classification Matrix and Construct Decision tree
- vii) Perform Linear Regresssion and Analyze , Validate and Visualize the data

3

Consider the data set

No.	eid Numeric	ename Nominal	salary Numeric	exp Numeric	address Nominal
1	101.0	raj	10000.0	4.0	pdr
2	102.0	ramu	15000.0	5.0	pdr
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

- i) Classify , Invoke MultiLayerPerception
- ii) Build neural network GUI as below
 - a) Beginning the process of editing the network to add a second hidden layer
 - b) The finished network with two hidden layers
- iii) Apply Lazy classifier, multi instance classifier
- iv) Apply any MetaLearning Algorithm
- v) Optimize base classifier's performance
- vi) Use clustering algorithm such as Cobweb, and Hierarchical Cluster
- vii) Select attribute by specifying an evaluator and a search method

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

Lab Cycle 2:

- 5 To do Data Visualization using Tableau. Perform the following:
- i) Apply the concept of Group and Set
 - ii) Advanced Table Calculations
 - iii) Advanced data preparation and analytics
 - iv) Animations

6	To do Data Visualization using Tableau. Perform the following: i) Detailed Calculations ii) Advanced Mapping Techniques
Lab Cycle 3:	
7	To do the visualization using Power BI. Perform the following i) Explore the data through Business Intelligence and create semantic model ii) Model data for ML
8	To do the visualization using Power BI. Perform the following i) Build and train a binary prediction model ii) Build and train general classification ML model
Course outcomes: On successful completion of this course, students will be able to: <ol style="list-style-type: none"> 1. Apply and Synthesize the knowledge of data mining using WEKA tool 2. Analyze and visualize the data using Tableau 3. Analyze and visualize the data using powerBI 4. CO4. Perform the data mining and visualization for large data set as Open Ended Project(any tools). 	
Conduct of Practical Examination: <ol style="list-style-type: none"> 1. All laboratory experiments are to be included for practical examination. 2. Breakup of marks and the instructions printed on the cover page of answer script to be strictly adhered bythe examiners. 3. Students can pick one experiment from the questions lot prepared by the examiners. 4. Change of experiment is allowed only once and 20% Marks is to be deducted. 	

Mapping of Course outcomes (COs) to Program outcomes (Pos):

Program Articulation Matrix															
Course Outcomes	Program Outcomes												Program Specific Outcome		
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1			2												2
CO2			2												2
CO3			2												2
CO4			2												2
over all			2												2

Degree of compliance 1: Slight (Low) 2: Moderate (Medium) 3: Substantial (High)

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & 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	03	Exam Hours	03
Lecture Hours	40hrs	Practical Hours	-
Course objectives: The course will enable students to			
<ul style="list-style-type: none"> • Identify the Information Retrieval problems and describe the architecture of a search engine • Analyse Search structures of dictionaries, Wildcard queries and Index construction done information retrieval • Analyse the scoring and ranking mechanisms used in Information retrieval systems • Study the various Information Retrieval Evaluation Techniques and processes involved. • Study how web search, web crawling and link analysis is done for information retrieval on the web 			
UNIT-1		7 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-2		7 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-3		7 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, Maximum tf 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-4

9 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-5

9 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:

At the end of the course the student will be able to:

- Analyse the Information Retrieval problems and describe the architecture of a Search Engine
- Apply Search structures of dictionaries, Wildcard queries and Index construction for information retrieval.
- Apply scoring and ranking mechanisms to design an efficient Search Engine
- Apply suitable evaluation techniques and language models in the design of Search Engine
- Analyse web search, web crawling and link analysis mechanisms for information retrieval on the web

Sl No	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
Textbook				
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)	Sebastian Raschka; C. Manning	Manning Books	MEAP August 2024
Additional Resource : https://medium.com/@daniele.nanni/revolutionizing-information-retrieval-the-role-of-large-language-models-in-a-post-search-engine-7dd370bdb62				
Reference Books				
1	Modern Information Retrieval: The Concepts and Technology	Ricardo Baeza - Yates and Berthier Ribeiro - Neto	ACM Press	2nd Edition, 2011
2	Information Retrieval Implementing and Evaluating Search Engines	Stefan Buttcher Charles L. A. Clarke Gordon V. Cormack	MIT Press	1st Edition, February 2016

Course Articulation matrix(CO-PO and CO-PSO mapping)

Course Out	PROGRAM OUTCOMES												PSO			
	1	2	3	4	5	6	7	8	9	10	11	12	PSC	PSC	PSC	
CO1		2												2		
CO2	2													2		
CO3			2											2		
CO4			2											2		
CO5		2												2		
Overall CO	2	2	2											2		

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) Outcome Based Education (OBE) And Choice Based Credit System (CBCS) SEMESTER – V			
SOCIAL NETWORK ANALYSIS			
Course Code	S5CCSPE02	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	03	Exam Hours	03
Lecture Hours	40hrs	Practical Hours	-
Course objectives: The course will enable students to <ol style="list-style-type: none"> 1. To understand the components of the social network. 2. To model and visualize the social network. 3. To mine the users in the social network. 4. To understand the evolution of the social network. 5. To know the applications in real time systems. 			
UNIT-1		(8 hrs)	
Introduction: Introduction to Web - Limitations of current Web – Development of Semantic Web – Emergence of the Social Web – Statistical Properties of Social Networks -Network analysis - Development of Social Network Analysis - Key concepts and measures in network analysis - Discussion networks - Blogs and online communities - Web-based networks.			
UNIT-2		(7 hrs)	
Modelling and Visualization: Visualizing Online Social Networks - A Taxonomy of Visualizations - Graph Representation - Centrality- Clustering - Node-Edge Diagrams - Visualizing Social Networks with Matrix- Based Representations- Node-Link Diagrams - Hybrid Representations - Modelling and aggregating social network data – Random Walks and their Applications –Use of Hadoop and Map Reduce - Ontological representation of social individuals and relationships.			
UNIT-3		(9 hrs)	
Mining Communities: Aggregating and reasoning with social network data, Advanced Representations – Extracting evolution of Web Community from a Series of Web Archive - Detecting Communities in Social Networks - Evaluating Communities – Core Methods for Community Detection & Mining - Applications of Community Mining Algorithms - Node Classification in Social Networks.			
UNIT-4		(8 hrs)	
Evolution: Evolution in Social Networks – Framework - Tracing Smoothly Evolving Communities - Models and Algorithms for Social Influence Analysis - Influence Related Statistics - Social Similarity and Influence - Influence Maximization in Viral Marketing - Algorithms and Systems for Expert Location in Social Networks - Expert Location without Graph Constraints - with Score Propagation – Expert Team Formation - Link Prediction in Social Networks - Feature based Link Prediction – Bayesian Probabilistic Models - Probabilistic Relational Models.			
UNIT-5		(7 hrs)	

Applications:

A Learning Based Approach for Real Time Emotion Classification of Tweets, A New Linguistic Approach to Assess the Opinion of Users in Social Network Environments, Explaining Scientific and Technical Emergence Forecasting, Social Network Analysis for Biometric Template Protection.

Course outcomes:

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

1. Work on the internal components of the social network.
2. Model and visualize the social network.
3. Analyse the behaviour of the users in the social network.
4. Predict the possible next outcome of the social network.
5. Apply social network in real time applications.

Sl. no.	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
Textbooks				
1	Social Network Analysis	Tanmoy Chakraborty	Wiley	2021
2	Computational Social Network Analysis: Trends, Tools and Research Advances	Ajith Abraham Aboul Ella Hassanien, Václav Snášel	Springer	2012
3	Handbook of Social Network Technologies and Applications	Borko Furht	Springer	1 st Edition 2011
4	Social Network Data Analytics	Charu C. Aggarwal	Springer	2014
Reference Books				
1	Advances in Social Network Mining and Analysis	Giles Mark Smith John Yen	Springer	2010
2	Web Mining and Social Networking – Techniques and applications	Guandong Xu Yanchun Zhang	Springer	1 st Edition 2012
3	Social Networks and the Semantic Web	Peter Mik	Springer	1 st Edition 2007
4	Applications of Social Media and Social Network Analysis	Przemyslaw Kazienko, Nitesh Chawla	Springer	2015

Course Articulation matrix(CO-PO and CO-PSO mapping)

Course Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
CO1	2	2													2
CO2		2	2												2
CO3		2	2												2
CO4		2													2
CO5		2	2												2
Overall CO	2	2	2												2

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) Outcome Based Education (OBE) And Choice Based Credit System (CBCS) SEMESTER – V			
BUSINESS INTELLIGENCE AND ANALYSIS			
Course Code	S5CCSPE03	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	03	Exam Hours	03
Lecture Hours	40hrs	Practical Hours	-
Course objectives: The course will enable students to <ol style="list-style-type: none"> 1. Explain the Decision Support systems and Business Intelligence framework. 2. Illustrate the significance of computerised Decision Support and understand the mathematical modelling behind decision support. 3. Explain Data warehousing, its architecture and Extraction, Transformation, and Load (ETL) Processes. 4. Explore knowledge management, explain its activities, approaches, and its implementation. 5. Describe the Expert systems, areas suitable for application of experts' system. 			
UNIT-1		(8L+0Thrs)	
Decision Support and Business Intelligence: Opening Vignette , Changing Business Environments and Computerized Decision Support, Managerial Decision Making, Computerized Support for Decision Making, An Early Framework for Computerized Decision Support, The Concept of Decision Support Systems (DSS), A Framework for Business Intelligence (BI), A Work System View of Decision Support Textbook 1 : Chapter 1: 1.1,1.2,1.3,1.4,1.5,1.6			
UNIT-2		(8L+0T hrs)	
Decision Making Systems, Modelling and Support: Decision Making, Models, Phases of the Decision-Making Process, The Intelligence Phase, The Design Phase, The Choice Phase, The Implementation Phase, How Decisions Are Supported, personality types, The decision makers. Decision support system development: Introduction to DSS development, The traditional system development life cycle, Alternative development life cycle, Prototyping: The DSS development methodologies. Textbook 2 : Chapter 6:6.1,6.2,6.3,6.4,6.5			
UNIT-3		(8L+0T hrs)	
Business intelligence: Data Warehousing, Data Acquisition, Business Analytics & Visualization: The Nature and Sources of Data, Data Collection, Problems and Quality, The Web/Internet and Commercial Database Services, Database Management System in Business Intelligence, Data Warehousing, Data Marts, Business Intelligence, Online Analytical Processing, Data Mining, Data Visualization, Multidimensionality and Real Time Analytics, Business Intelligence, and the Web Textbook 2 : Chapter 5: 5.1,5.2,5.3,5.4,5.5,5.6,5.7,5.8,5.9,5.10,5.11,5.12,5.14			
UNIT-4		(8L+0T hrs)	
Knowledge Management: Introduction to Knowledge Management, Organizational learning and transformation, Knowledge management initiatives, Approaches to knowledge management, Information technology in knowledge management, Knowledge management system implementation, roles of people in knowledge management, ensuring success of knowledge management. Textbook 2 : Chapter 9 :9.2,9.3,9.4,9.5,9.6,9.7,9.8,9.9			
UNIT-5		(8L+0T hrs)	
Expert system: Basics concepts of expert system, Applications of expert system, Structure of expert systems, How expert system works, Problems areas suitable for expert systems, Benefits and			

capabilities of expert systems, Problems and limitations of expert system, Expert system success factors, Types of expert systems, Expert systems on the web

Textbook 2 : Chapter 10:10.5,10.6,10.7,10.8,10.9,10.10,10.11,10.12,10.13,10.14

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 and Year
Textbooks				
1	Business Intelligence, A Managerial Perspective on Analytics	Sharda, R, Delen D, Turban E.	Pearson.	2014
2	Decision support systems and intelligent systems	Efraim Turban , Jay E. Ting-Peng Liang	PHI	7 th edition,2010
Reference Books				
1	Business Intelligence, Analytics, and Data Science,	Ramesh Sharda,DursunDele	Pearson Education	2019
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	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
CO1	2	2				2									2
CO2		2													2
CO3		2													2
CO4		2				2									2
CO5		2													2
Overall CO	2	2													2

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & 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	03	Exam Hours	03
Lecture Hours	40hrs	Practical Hours	-
Course objectives: This course will enable students to: <ol style="list-style-type: none"> 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-1		(08 hrs)	
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-2		(08 hrs)	
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-3		(08 hrs)	
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-4		(08 hrs)	
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-5		(08 hrs)	
Planning: planning in a situation calculus, the STRIPS representation, planning as a reasoning task, beyond the basics			

Course outcomes:

Upon completion of this course the student 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.

Sl. No.	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
Textbooks				
1	Knowledge Representation and Reasoning	R. Brachman & H. Levesque,	Morgan-Kaufmann,	First edition, 2004
Reference Books				
1	Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL	Dean Allemang, James Hendler		2nd edition, 2011
2	NOC: Artificial Intelligence: Knowledge Representation and Reasoning	Prof. Deepak Khemani	IIT Madras	Link: https://nptel.ac.in/courses/106106140,

Course articulation matrix(CO-PO and CO-PSO mapping)

Course Outcome	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3
CO1	2		2											2	
CO2	2	2												2	
CO3	2		2											2	
CO4	2		2											2	
CO5	2	2												2	

Program articulation matrix:

Course Outcomes	PROGRAM OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
	2	2	2											2	

Degree of compliance 1: Low 2: Medium 3: High

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) Outcome Based Education (OBE) And Choice Based Credit System (CBCS) SEMESTER – V			
Research Methodology and IPR			
Course Code	S5IMA01	CIE Marks	50
Teaching Hours/Week (L:T:P)	(2:2:0)	SEE Marks	50
Credits	3	Exam Hours	3
Lecture Hours	40 hrs	Practical Hour	-
UNIT 1			6 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 2			5 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 3			5 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 4			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 5			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

CO – PO Mapping:

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3	PSO4
CO1		3	2									2				
CO2		3	2					3				2				
CO3		3	3									2				
CO4		3	2									2				
CO5		3	2									2				

Text Books:

Sl. No.	Author/s	Title, Publisher, Edition, Year, ISBN
1.	Peter S. Menel Mark A. Lemley, Robert P. Merges	"Intellectual Property in the New Technological-Vol. I Perspectives, 2021.
2.	Laura R. Ford	"The Intellectual Property of Nations: Sociological and Historical Perspectives on a Modern Legal Institution Paperback -2021.

Reference Book:

Sl. No.	Author/s	Title, Publisher, Edition, Year, ISBN
1.	R. Ganesan	"Research Methodology for Engineers", MJP Publishers, Chennai, 2011.
2.	Cooper Donald R, Schindler Pamela S and Sharma JK	"Business Research Methods", Tata McGraw Hill Education, 11 th Edition, 2012.
3.	Catherine J. Holland	"Intellectual property: Patents, Trademarks, Copyrights, Trade Secrets", Entrepreneur Press, 2007.
4.	David Hunt, Long Nguyen, Matthew Rodgers	"Patent searching: tools & techniques", Wiley, 2007.
5.	The Institute of Company Secretaries of India, Statutory body under an Act of parliament	"Professional Programme Intellectual Property Rights, Law and practice", September 2013.

Scheme of Teaching, Examination and Syllabus

B.E. COMPUTER SCIENCE AND ENGINEERING

(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

Batch: 2022-23

VI SEMESTER

(Effective from the academic year 2024-2025)



SreeSiddaganga Education Society®

Siddaganga Institute of Technology

(An Autonomous institute affiliated to Visvesvaraya Technological University, Belagavi)

(Approved by AICTE, New Delhi, Accredited by NAAC and ISO 9001-2015 certified)

B.H. Road, Tumakuru-572 103, Karnataka, India

Phone: Direct +91-816-2282696, Fax: +91-816-2282994

E-mail: principal@sit.ac.in web: www.sit.ac.in

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & 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	40hrs	Practical Hours	26hrs
Course objectives: The course will enable students to <ol style="list-style-type: none"> To Understand the NLP techniques like parsing, POS-tagging and Word sense disambiguation. To explore language modeling techniques such as N-grams. To explore the applications of NLP such as Machine translation, Information retrieval etc. To understand the basic architecture of the NLG system and the role of NLP in a search engine. Demonstrate the use of modern NLP techniques for processing of text like extracting the data. Text to Feature representation etc.. 			
UNIT-1		(08 hrs)	
Introduction: What is NLP? Origins of NLP, Language and Knowledge, The challenges of NLP, NLP Applications, Some successful Earley NLP systems. Word level Analysis: Introduction, Regular Expressions, Finite-state Automata, Morphological Parsing. Extracting the Data: Text data collection using APIs, Reading PDF file in Python, Reading word document, Reading JSON object, Reading HTML page and HTML parsing, Regular expressions, String handling, Web scraping. Exploring and Processing Text Data: Lowercasing, Punctuation removal, Stop words removal, Text standardization, Spelling correction, Tokenization, Stemming, Lemmatization, Exploratory data analysis, End-to-end processing pipeline. (Text Book-1: 1.1-1.5,1.7,1.8,3.1-3.4) (Text Book-2: 1,2)			
UNIT-2		(08 hrs)	
Language Modeling: Introduction, Statistical Language Model- N-gram model, Add-one smoothing, Good-Turing smoothing. Part-of-Speech Tagging: Rule-based Tagger, Stochastic Tagger, Hybrid Tagger. Syntactic Analysis: CFG, Parsing- Top-down parsing, Bottom-up parsing, The Earley Algorithm, Probabilistic Parsing- Estimating Rule probabilities. (Text Book-1: 2.1,2.3,3.7,4.2,4.4.1-4.4.4,4.5.1)			
UNIT-3		(08 hrs)	
Information Retrieval-1: Introduction, Design features of Information Retrieval Systems, Information Retrieval models, Classical Information Retrieval models, Non-classical models of IR, Alternative models of IR, Evaluation of the IR system. Information Retrieval-2: Natural Language Processing in IR, 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-1: 9.1-9.7,10.2,10.6) (Text Book-2: 3)			
UNIT-4		(08 hrs)	

Ambiguity, Word sense Disambiguation: Context-based WSD Approaches, Knowledge based approaches, Supervised Learning of WSD, Bayesian Classification, Testing, K-Nearest Neighbour or Memory-based Learning, Bootstrapping, Bilingual Corpora, Unsupervised methods of WSD.

Machine Translation: Introduction, Problems in Machine Translation, Characteristics of Indian Languages, Machine translation approaches, Direct Machine translation, Rule-based machine translation, Corpus based MT, Semantic or Knowledge-based MT systems, Translation involving Indian Languages.

(Text Book-1: 5.4,5.5.2,8.1-8.9)

UNIT-5

(08 hrs)

Natural Language Generation: Introduction, Architectures of NLG systems, Generation tasks and representations (Except “Approach based on functional unification grammar”).

Other Applications of NLP: Introduction, Information Extraction, Automatic Text Summarization, Question-Answering System.

Lexical Resources: WordNet, FrameNet, Stemmers, POS taggers, Research Corpora.

Disambiguating word sense using Wordnet, NLP in a Search Engine.

(Text Book-1: 7.1-7.3, 11.1-11.4, 12.2-12.6)

(Text Book-2: 4.8,5.6)

Course outcomes:

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

1. Design finite state automata and context free grammars for word level and syntax level analysis respectively.
2. Describe and Apply N-grams Language model to predict the next word in the text sequence.
3. Outline Natural Language Generation techniques and various lexical resources.
4. Describe basics of NLP and identify various applications of NLP like Machine Translation, information Retrieval, etc.
5. Describe the use of various NLP techniques like POS tagging, WSD etc. for text processing and develop python code for the same.

Sl. No.	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
Textbooks				
1	Natural Language Processing and Information Retrieval	Tanveer Siddiqui, U S Tiwary	Oxford University Press	2 nd Edition, 2010.
2	Natural Language Processing Recipes Unlocking Text Data with Machine Learning and Deep Learning using Python	Akshay Kulkarni, Adarsha Shivananda.	Apress	2019
Reference Books				
1	Speech and Language Processing: An introduction to Natural Language Processing, Computational Linguistics and Speech Recognition,	Daniel Jurafsky and James H Martin	Prentice Hall,,	Low Price Edition, 2000.

2	Foundations of Statistical Natural Language Processing	Christopher D. Manning	MIT Press	1999.
3	Natural language processing with Python	Steven, Ewan Klein, and Edward Loper	O'Reilly Media	1st Edition, 2009.

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:	
<ol style="list-style-type: none"> 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 Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
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	

B.E COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE & 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	40hrs	Tutorial Hours	26hrs

Course objectives: The 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-1 (08L+ 4P hrs)

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, 4.1.2), 4.2(4.2.1), 10.3 (10.3.1 to 10.3.3), 12.1 (12.1.2 to 12.1.4)

UNIT-2 (08L+ 4P hrs)

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-3 (08L+ 6P hrs)

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-4 (08L+ 6P hrs)

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-5		(08L+ 6P hrs)		
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 Digitaltransmission, error control protocols and random access protocols. 2. Apply the knowledge of Packet switching concepts in computer networking, Identify differentcategories 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 and Year
Textbooks				
1	Data Communications and Networking	Behrouz A. Forouzan	McGraw-Hill	5 th Edition, 2013
2	Computer Networks: A Systems Approach	Larry L Peterson and	Elsevier	5 th Edition, 2011
Reference Books				

1	Data and Computer Communications	William Stallings	Pearson Education	10 th Edition, 2013
2	Computer Networking: A Top-Down Approach	Kurose James F, Ross Keith W.	Pearson Education	6 th Edition, 2017
3	Computer Networks	Andrew S. Tanenbaum and David J.	Pearson Education	5 th Edition, 2011
4	Unix Network Programming, Interprocess Communications,	WRichard Stevens	Pearson Education	2nd Edition

COURSE ARTICULATION MATRIX (CO-PO AND CO-PSO MAPPING):

COURSE ARTICULATION MATRIX															
Course Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
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

B.E COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE & 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	30hrs	Practical Hours	-

Course objectives: The 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.

UNIT-1 **(8 hrs)**

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-2 **(9 hrs)**

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-3 **(8 hrs)**

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-4 **(8 hrs)**

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-5 **(7 hrs)**

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

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 & Analyse the relationship between fog computing & IoT.

Sl. No	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and
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
Reference Books				
1	Cloud Computing Principles and Paradigms	RajkumarBuyya, James Broberg, Andrzej Goscinski	Wiley Publishing Inc.	2013
2	Cloud Computing and SOA Convergence in Your Enterprise:	David S. Linthicum	Addison-Wesley Professional	1 st Edition
3	Distributed and Cloud Computing	Kai Hwang, Geoffrey C. Fox, and Jack J. Dongarra	Morgan Kaufman Publishers	2012
4	Enterprise Cloud Computing Technology Architecture Applications	GautamShroff	Cambridge University Press	1 st Edition
5	Cloud Computing, A Practical Approach	Toby Velte, Anthony Velte, Robert Elsenpeter	McGraw-Hill Education	1 st Edition

Mapping of Course outcomes (COs) to Program outcomes (Pos):

Program Articulation Matrix															
Course Outcomes	Program Outcomes												Program Specific Outcome		
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	2														2
CO2	2														2
CO3	2														2
CO4	2														2
over all	2														2

Degree of compliance 1: Slight (Low) 2: Moderate (Medium) 3: Substantial (High)

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & 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	40hrs	Practical Hours	-
Course objectives:			
This course will enable students to:			
<ol style="list-style-type: none"> 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-1		(08 hrs)	
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-2		(08 hrs)	
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-3		(08 hrs)	
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-4		(08 hrs)	
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-5		(08 hrs)	
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:

Upon completion of this course the student 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 and Year
Textbooks				
1	Cyber Security: Understanding Cyber Crimes, Computer Forensics and Legal Perspectives	Nina Godbole and Sunit Belapure	Wiley India Pvt Limited"	2011
2	Computer Forensics and Investigations, Cengage Learning	Nelson Phillips and EnfingerSteuart		New Delhi,2009
3	An Incident-Based Approach to Forensic Investigations Malware forensics. In: Practical Cyber Forensics.	Niranjan Reddy	Apress, Berkeley,	CA (2019). https://doi.org/10.1007/978-1-4842-4460-9
4	AI for Cybersecurity A Handbook of Use Cases, Penn State Cyber Security Lab	Peng Liu, Tao Liu et al.,		
Reference Books				
1	Incident Response and Computer Forensics,	Kevin Mandia, Chris Prosize, Matt Pepe	Tata McGraw - Hill,	New Delhi, 2006
2	Software Forensics, Tata,	Robert M Slade	McGraw -Hill,.	New Delhi ,2005
3	"Understanding Forensics in IT ",	Bernadette H Schell, Clemens Martin	Cybercrime, ABC – CLIO Inc, California, 2004.	NIIT Ltd,2005

Course Outcomes	PROGRAM OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
CO1	2														2
CO2		2													2
CO3		2			2										2
CO4	2														2
CO5		2													2
Overall CO	2	2													2

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & 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
<p>Course objectives: The course will enable students to</p> <ol style="list-style-type: none"> Describe the basic paradigms, data model, evolution for Big Data (L2). Explain the importance of a serialization framework and limitations of serialization frameworks for Big Data (L2) Analyze how the data is stored on the batch layer Design of the batch layer starting from ingesting new data to computing batch views. Illustrate how to build the serving layer for Bigdata. Describe the real time views of Cassandra's data model for Bigdata Demonstrate how to implement the concepts of queuing and stream processing using real-world tools 			
UNIT-1		(08 hrs)	
<p>A new paradigm for Big Data: Scaling with a traditional database - NoSQL 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.</p> <p>Data model for Big Data: The properties of data - The fact-based model for representing data - Graph - A complete data model for SuperWebAnalytics.com.</p> <p>Data model for Big Data - illustration: Why a serialization framework? - Apache Thrift - Limitations of serialization frameworks.</p>			
UNIT-2		(08 hrs)	
<p>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</p> <p>Data storage on the batch layer – Illustration: Using the Hadoop Distributed File System - Data storage in the batch layer with Pail - Basic Pail operations - Serializing objects into pails - Batch operations using Pail - Vertical partitioning with Pail - Pail file formats and compression – Summarizing the benefits of Pail - Storing the master dataset for SuperWebAnalytics.com</p> <p>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</p>			
UNIT-3		(08 hrs)	
<p>Batch layer: Illustration: An illustrative example - Common pitfalls of data-processing tools - An introduction to JCasalog – Composition</p> <p>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</p> <p>Batch layer: Implementation: Starting point - Preparing the workflow - Ingesting new data - URL normalization - User-identifier normalization - Deduplicate pageviews - Computing batch views.</p>			
UNIT-4		(08 hrs)	
<p>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.</p> <p>Serving layer: Illustration: Basics of ElephantDB - Building the serving layer for</p>			

SuperWebAnalytics.com.

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

UNIT-5

(07 hrs)

Realtime view – Illustration:Cassandra’s data model 220 - Using Cassandra.

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 , its elements , its analytics , computing in Big data etc. to the solutions of complex real world engineering problems.
2. Select and apply appropriate modern tools of Hadoop ecosystem to the solution of various problems in storage, processing, accessing, managing and analysing the Big data.
3. Design and Develop Map Reduce programs to the solution of various real world application problems.
4. Identify the importance of Big data Stack architecture and Analyse the merits of u modern data warehouses against the limitations of Traditional Databases.
5. Design and Develop Spark programs to the solution of various problems.

Sl. No.	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
Textbooks				
1	Big Data - PRINCIPLES AND BEST PRACTICES OF SCALABLE REAL-TIME DATA SYSTEMS	NATHAN MARZ with JAMES WARREN	Manning Publications	2015 Edition
2	Spark in Action	Petar Zečević Marko Bonaći	Manning Publications	Nov 2016 Edition
Reference Books				
1	Hadoop: The Definitive Guide	Tom White	O’reilly Media	4 th Edition,2015
2	Big Data and Analytics	SeemaAcharya, Subhashini Chellappan	Wiley India Publications,	May 2015

Course Articulation matrix(CO-PO and CO-PSO mapping)

Course Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
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	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
BIG DATA(RCSE32)	3	2	3	-	2	-	-	-	-	-	-	-	-	-	3

Degree of compliance 1: Low 2: Medium 3: High

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) Outcome Based Education (OBE) And Choice Based Credit System (CBCS) SEMESTER – VI			
RECOMMENDER SYSTEM			
Course Code	S6CCSPE02	CIE Marks	50
Teaching Hours/Week (L:T:P)	(3:0:0)	SEE Marks	50
Credits	3	Exam Hours	03
Lecture Hours	40hrs	Practical Hours	-
Course objectives: This course will enable students to: <ol style="list-style-type: none"> 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-1		(08 hrs)	
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-2		(08 hrs)	
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-3		(08 hrs)	
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-4		(08 hrs)	
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-5 (08 hrs)
Advanced Topics in Recommender Systems: Introduction, Learning to rank, Multi armed Bandit 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: Upon completion of this course the student will be able to: <ol style="list-style-type: none"> 1. Describe the concept of collaborative recommendation system. 1. Describe the concept of content-based and knowledge-based recommendation system. 2. Describe the concept of hybrid recommendation and understand the evaluation methods for recommendation systems. 3. Understand the concept of recommendation for networks. 4. Understand 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 and 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 Outcomes	PROGRAMME OUTCOMES												PSO		
	1	2	3	4	5	6	7	8	9	10	11	12	PSO1	PSO2	PSO3
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	

B.E COMPUTER SCIENCE & ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) Outcome Based Education (OBE) And Choice Based Credit System (CBCS) SEMESTER – VI			
INTERNET OF THINGS LABORATORY			
CourseCode	S6CCSL01	CIEMarks	50
TeachingHours/Week(L:T:P)	(0:0:2)	SEE Marks	50
Credits	1	Exam Hours	03
Lecture Hour	-	Practical Hour	26
Course objectives: The course will enable students to			
<ul style="list-style-type: none"> • SynthesizeIoTworkinglogicsusingIoTcomponents. • ToexploreIoTtechnologies,architectures • Tomanageandprocesscomplexrawdata 			
Sl. No.	Experiments		
LAB SET QUESTIONS			
Day 1	Introduction to IoT toolkit – Familiarization with Arduino board and perform necessary SW installations, Sensors, Actuators etc. Build simple IoT project using Tinker CAD		
1.	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.		
2.	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		
3.	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.		
4.	Experiment to interface servomotor using Arduino. Control the positional reading through switch and add the indicators using buzzer.		
5.	To interface Bluetooth with Arduino and write a program to send sensor data to smartphone using Bluetooth.		
6.	To interface GPS UNIT with Arduino and write a program to send location data to smartphone		
7.	Set up a simple web server using ESP32 UNIT and monitor the live temperature in the web browser		
8.	Send SMS/Email using ESP32		
9.	Control light through Voice command using Arduino		
10.	Publish DHT11/IR Sensor Reading to ThingSpeak cloud using ESP32		
Revised Bloom's Taxonomy Level	L ₂ -Understanding, L ₃ - Applying		
Course outcomes:			
At the end of the course the student will be able to:			
<ul style="list-style-type: none"> • Discover key IoT concepts including identification, sensors, localization, wireless protocols, data storage and security • Explore IoT technologies, architectures, standards, and regulation • Realize the value created by collecting, communicating, coordinating, and leveraging the data from connected devices • Understand how to develop and implement IoT technologies, solutions, and applications 			

Conduct of Practical Examination:

1. All laboratory experiments are to be included for practical examination.
2. Breakup of marks and the instruction printed on the coverpage of answer script to be strictly adhered by the examiners.
3. Student scanpick on the experiment from the question slot prepared by the examiners.
4. Change of experiments allowed only once and 20% of the Marks to be deducted for the same.

Program Articulation Matrix

Course Outcomes	Program Outcomes												Program Specific Outcome		
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1		2													2
CO2		2													2
CO3		2													2
CO4		2													2
over all		2													2