MAR ATHANASIUS COLLEGE OF ENGINEERING

(Government Aided and Autonomous)

Kothamangalam-686666

Affiliated to APJ Abdul Kalam Technological University

Thiruvananthapuram

Master of Technology (M. Tech)

Curriculum - 2024

COLLEGE VISION AND MISSION

VISION

Excellence in education through resource integration.

MISSION

The institution is committed to transform itself into a centre of excellence in Technical Education upholding the motto ”Knowledge is Power.” This is to be achieved by imparting quality education to mould technically competent professionals with moral integrity, ethical values and social commitment, and by promoting innovative activities in the thrust areas emerging from time to time.

MAR ATHANASIUS COLLEGE OF ENGINEERING
(GOVT. AIDED & AUTONOMOUS)
M.TECH CURRICULUM AND SCHEME-2024
Department of Computer Science and Engineering
(Computer Science and Engineering)

PROGRAM OUTCOMES – PO

Outcomes are the attributes that are to be demonstrated by a graduate after completing the programme

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

The departments conducting the M.Tech programme shall define their own PSOs, if required, and evaluation shall also be done for the same.

SEMESTER I

Slot	COURSE CODE	Courses	Marks		L-T-P-S	Hours	Credit
			CIE	ESE			
A	M24CS1T101	COMPUTATIONAL INTELLIGENCE	40	60	4-0-0-4	4	4
B	M24CS1T102	ADVANCED DATABASE MANAGEMENT	40	60	4-0-0-4	4	4
C	M24CS1T103	ADVANCED MACHINE LEARNING	40	60	4-0-0-4	4	4
D	M24CS1E104A	PROGRAM ELECTIVE 1	40	60	3-0-0-3	3	3
E	M24CS1E105A	PROGRAM ELECTIVE 2	40	60	3-0-0-3	3	3
J	M24CS1R106	RESEARCH METHODOLOGY AND IPR	40	60	2-0-0-2	2	2
G	M24CS1L107	ADVANCED MACHINE LEARNING LAB	60	40	0-0-3-3	3	2
Total			300	400		23	22

Teaching Assistance: 7 hours
Self-study- 23 Hrs

PROGRAM ELECTIVE I						
SLOT	SL NO	COURSE CODE	COURSE NAME	L-T-P-S	HOURS	CREDIT
D	1	M24CS1E104A	OBJECT ORIENTED SOFTWARE ENGINEERING	3-0-0-3	3	3
	2	M24CS1E104B	FUNDAMENTALS OF CRYPTOGRAPHY	3-0-0-3	3	3
	3	M24CS1E104C	ADVANCED DATA MINING	3-0-0-3	3	3

	4	M24CS1E104D	SOCIAL NETWORK ANALYSIS	3-0-0-3	3	3
	5	M24CS1E104E	COMPUTATIONAL BIOLOGY	3-0-0-3	3	3

PROGRAM ELECTIVE II						
SLOT	SL NO	COURSE CODE	COURSE NAME	L-T-P-S	HOURS	CREDIT
E	1	M24CS1E105A	FORMAL METHODS IN SOFTWARE ENGINEERING	3-0-0-3	3	3
	2	M24CS1E105B	BLOCKCHAIN AND CRYPTOGRAPHY	3-0-0-3	3	3
	3	M24CS1E105C	PATTERN RECOGNITION	3-0-0-3	3	3
	4	M24CS1E105D	ADVANCED COMPUTER NETWORKS	3-0-0-3	3	3
	5	M24CS1E105E	WIRELESS SENSOR NETWORKS	3-0-0-3	3	3

SEMESTER II

SLOT	COURSE CODE	COURSE NAME	MARKS		L-T-P-S	HOURS	CREDIT
			CIE	ESE			
A	M24CS1T201	ADVANCED DATA STRUCTURES AND ALGORITHMS	40	60	4-0-0-4	4	4
B	M24CS1T202	ADVANCED OPERATING SYSTEMS	40	60	4-0-0-4	4	4

C	M24CS1E203A	PROGRAM ELECTIVE 3	40	60	3-0-0-3	3	3
D	M24CS1E204A	PROGRAM ELECTIVE 4	40	60	3-0-0-3	3	3
E	M24CS1S205	INDUSTRY INTEGRATED COURSE -	40	60	3-0-0-3	3	3
G	M24CS1P206	MINI PROJECT	60	40	0-0-3-3	3	2
H	M24CS1L207	MODERN DATABASE LAB	60	40	0-0-3-3	3	2
Total			320	380		23	21

Teaching Assistance: 7 hours
Self-study- 23 Hrs

PROGRAM ELECTIVE III						
SLOT	SL NO	COURSE CODE	COURSE NAME	L-T-P-S	HOURS	CREDIT
C	1	M24CS1E203A	SOFTWARE TESTING	3-0-0-3	3	3
	2	M24CS1E203B	CYBER FORENSICS AND INFORMATION SECURITY	3-0-0-3	3	3
	3	M24CS1E203C	DEEP LEARNING	3-0-0-3	3	3
	4	M24CS1E203D	INFORMATION RETRIVAL	3-0-0-3	3	3
	5	M24CS1E203E	MACHINE LEARNING IN COMPUTATIONAL BIOLOGY	3-0-0-3	3	3

PROGRAM ELECTIVE IV						
SLOT	SL NO	COURSE CODE	COURSE NAME	L-T-P-S	HOURS	CREDIT
D	1	M24CS1E204A	SOFTWARE ARCHITECTURE AND DESIGN PATTERNS	3-0-0-3	3	3
	2	M24CS1E204B	DATA PRIVACY AND SECURITY	3-0-0-3	3	3
	3	M24CS1E204C	WEB MINING	3-0-0-3	3	3
	4	M24CS1E204D	SEMANTIC WEB ARCHITECTURE	3-0-0-3	3	3
	5	M24CS1E204E	REINFORCEMENT LEARING	3-0-0-3	3	3

INDUSTRY INTEGRATED COURSE						
SLOT	SL NO	COURSE CODE	COURSE NAME	L-T-P-S	HOURS	CREDIT
E	1	M24CS1S205A	CLOUD PRACTITIONER ESSENTIALS	3-0-0-3	3	3
	2	M24CS1S205A	CLOUD INFRASTRUCTURE ENGINEERING	3-0-0-3	3	3

Teaching Assistance: 6 hours

*MOOC Course of minimum 8 weeks duration to be successfully completed before the end of fourth semester (starting from semester 1).

**Internship- mandatory internship of 6 to 8 weeks

Internship - mandatory internship of more than 16 weeks

###Dissertation Phase 1 – Should be done in Industry

TRACK 1 / TRACK 2

In second year, the students can choose either of the two tracks: TRACK 1 or TRACK 2. Track 1 is conventional M.Tech programme in which the dissertation Phase 1 is conducted in college. Track 2 is M.Tech programme designed for students who undergone long term internship (not less than 16 weeks) in industry. An aspirant in track 2 needs to do the dissertation in the industry. The candidates should also be good with performing in-depth research and colluding the conclusions of research led by them. Such students are expected to have the following skills: Technical Skills, Research Skills, Communication Skills, Critical Thinking Skills, and Problem Solving Skills.

The eligibility for Track 2:

- Shall have qualified in the GATE or have a SGPA above 8.0 during the first semester, and
- Qualify an interview during the end of second semester by an expert committee constituted by the College.

SEMESTER III

SLOT	COURSE CODE	COURSE NAME	MARKS		L-T-P-S	HOURS	CREDIT
			CIE	ESE			
TRACK 1							
A	M24CS1M301	*MOOC	To be completed successfully		--	--	2
B	M24CS1E302A	PROGRAM ELECTIVE 5	40	60	3-0-0-3	3	3
K	M24CS1I303	**INTERNSHIP	50	50	--	--	3
P	M24CS1P304	DISSERTATION PHASE 1	100	--	0-0-17-0	17	11
Total			190	110		20	19
TRACK 2							

A	M24CS1M305	*MOOC 1	To be completed successfully		--	--	2
B	M24CS1M306	*MOOC 2	To be completed successfully		--	--	2
K	M24CS1I307	##INTERNSHIP	50	50	---	--	4
P	M24CS1P308	###DISSERTION PHASE 1	100	--	---	--	11
Total			150	50			19

PROGRAM ELECTIVE V						
SLOT	SL NO	COURSE CODE	COURSE NAME	L-T-P-S	HOURS	CREDIT
B	1	M24CS1E302A	BIG DATA ANALYTICS	3-0-0-3	3	3
	2	M24CS1E302B	DIGITAL IMAGE PROCESSING	3-0-0-3	3	3
	3	M24CS1E302C	COMPUTER VISION	3-0-0-3	3	3
	4	M24CS1E302D	DISTRIBUTED ALGORITHMS	3-0-0-3	3	3
	5	M24CS1E302E	NATURAL LANGUAGE PROCESSING	3-0-0-3	3	3

Teaching Assistance : 6 Hours

*MOOC Course of minimum 8 weeks duration to be completed before the end of fourth semester (starting from semester 1).

**internship - mandatory internship of 6 to 8 weeks

##internship - mandatory internship more than 16 weeks

###Dissertation Phase 1 - Should be done in industry

TRACK 1 / TRACK 2

In second year, the students can choose either of the two tracks: TRACK 1 or TRACK 2. Track 1 is conventional M.Tech programme in which the dissertation Phase 1 is conducted in college. Track 2 is M.Tech programme designed for students who undergone long term internship (not less than 16 weeks) in industry. An aspirant in track 2 needs to do the dissertation in the industry. The candidates should also be good with performing in-depth research and colluding the conclusions of research led by them. Such students are expected to have the following skills: Technical Skills, Research Skills, Communication Skills, Critical Thinking Skills, and Problem Solving Skills.

The eligibility for Track 2:

- ➢ Shall have qualified in the GATE or have a SGPA above 8.0 during the first semester, and
- ➢ Qualify an interview during the end of second semester by an expert committee constituted by thE College.

SEMESTER IV

SLOT	COURSE CODE	COURSE NAME	MARKS		L-T-P-S	HOURS	CREDIT
			CIE	ESE			
TRACK 1							
P	M24CS1P401	DISSERTION PHASE II	100	100	0-0-27-24	27	18
TOTAL			100	100		27	18
TRACK 2							
P	M24CS1P402	##DISSERTION PHASE II	100	100	--	--	18
TOTAL			100	100			18
Total credit in all four semesters							80

##Dissertation Phase II- Should be done in Industry

COURSE NUMBERING SCHEME

The course number consists of digits/alphabets. The pattern to be followed is

- **For General Courses - MYYBBXCSNN**
- **For Elective Courses - MYYBBXCSNNA**

- M: MASTERS
- YY: Last two digits of year of regulation
- BB: DEPARTMENT

Sl.No	Department	Course Prefix
01	Civil Engg	CE
02	Computer Science	CS
03	Electrical & Electronics	EE
04	Electronics & Communication	EC
05	Mechanical Engg	ME
06	Any	GE
07	External (Industry/NPTEL etc)	EX

- X :Specialization number
- C: Course Type

  T- Core Course
  E- Elective Course
  R- Research Methodology & IPR
  L- Laboratory Course
  S- Industry Integrated Course
  I- Internship
  M- MOOC
  P- Project/Dissertation

- S : Semester of Study
  1- Semesters 1
  2- Semester 2

3- Semester 3
4- Semester 4

- NN: Course sequence number
- A: Elective sequence number - A/B/C/D/E

It is illustrated below: Examples:

M24CE1T202 is a second core course of first specialization offered by the Civil Department in semester 2

M24EC1R106 is Research Methodology & IPR offered in semester 1

M24EC1E104A is the first subject of Elective 1 of first specialization offered by the EC Department in semester 1

EVALUATION PATTERN

(i) CORE COURSES

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

Micro project/Course based project	:10 marks
Course based task/Seminar/Quiz	:10 marks
Test paper 1 (Module 1 and Module 2)	:10 marks
Test paper 2 (Module 3 and Module 4)	:10 marks

The project shall be done individually (Preferable).

End Semester Examination: 60marks

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

(ii) ELECTIVE COURSES

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Continuous Internal Evaluation: 40 marks

Seminar*	:	10 marks
Course based task/Micro Project//Data collection and interpretation/Case study	:	10 marks
Test paper 1 (Module 1 and Module 2)	:	10 marks
Test paper 2 (Module 3 and Module 4)	:	10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course.

End Semester Examination: 60 marks

The end semester examination will be conducted by the College. There will be two parts; Part A and Part B. Part A will contain 5 numerical/short answer questions with 1 question from each module, having 4 marks for each question (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

(iii) RESEARCH METHODOLOGY & IPR

Continuous Internal Evaluation : 40 marks

Preparing a review article based on peer reviewed Original publications in the relevant discipline (minimum 10 publications shall be referred) : 10 marks

Course based task/Seminar/Quiz : 10 marks

Test paper 1 (Module 1 and Module 2) : 10 marks

Test paper 2 (Module 3 and Module 4) : 10 marks

End Semester Examination : 60 marks

The end semester examination should be conducted by the college. The time duration will be for 3 Hrs and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

(iv) INTERNSHIP

Internships are educational and career development opportunities, providing practical experience in a field or discipline. They are structured, short-term, supervised placements often focused around particular tasks or projects with defined timescales. An internship may be compensated or non-compensated by the organization providing the internship. The internship has to be meaningful and mutually beneficial to the intern and the organization. It is important that the objectives and the activities of the internship program are clearly defined and understood. The internship offers the students an opportunity to gain hands-on industrial or organizational exposure; to integrate the knowledge and skills acquired through the coursework; interact with professionals and other interns; and to improve their presentation, writing, and communication skills. Internship often acts as a gateway for final placement for many students.

A student shall opt for carrying out the Internship at an Industry/Research Organization or at another institute of higher learning and repute (Academia). The organization for Internship shall be selected/decided by the students on their own with prior approval from the faculty advisor/respective PG Programme Coordinator/Guide/Supervisor. Every student shall be assigned an internship Supervisor/Guide at the beginning of the Internship. The training shall be related to their specialization after the second semester for a minimum duration of six to eight weeks. On completion of the course, the student is expected to be able to develop skills in facing and solving the problems experiencing in the related field.

Objectives

- Exposure to the industrial environment, which cannot be simulated in the classroom and hence creating competent professionals for the industry.

- Provide possible opportunities to learn understand and sharpen the real time technical / managerial skills required at the job.
- Exposure to the current technological developments relevant to the subject area of training.
- Create conducive conditions with quest for knowledge and its applicability on the job.
- Understand the social, environmental, economic and administrative considerations that influence the working environment.
- Expose students to the engineer's responsibilities and ethics.

Benefits of Internship

Benefits to Students

- An opportunity to get hired by the Industry/ organization.
- Practical experience in an organizational setting & Industry environment.
- Excellent opportunity to see how the theoretical aspects learned in classes are integrated into the practical world. On-floor experience provides much more professional experience which is often worth more than classroom teaching.
- Helps them decide if the industry and the profession is the best career option to pursue.
- Opportunity to learn new skills and supplement knowledge.
- Opportunity to practice communication and teamwork skills.
- Opportunity to learn strategies like time management, multi-tasking etc in an industrial setup.
- Makes a valuable addition to their resume.
- Enhances their candidacy for higher education/placement.
- Creating network and social circle and developing relationships with industry people.
- Provides opportunity to evaluate the organization before committing to a full time position.

Benefits to the Institute

- Build industry academia relations.
- Makes the placement process easier.
- Improve institutional credibility & branding.
- Helps in retention of the students.
- Curriculum revision can be made based on feedback from Industry/ students.
- Improvement in teaching learning process.

Benefits to the Industry

- Availability of ready to contribute candidates for employment.
- Year round source of highly motivated pre-professionals.
- Students bring new perspectives to problem solving.
- Visibility of the organization is increased on campus.
- Quality candidate's availability for temporary or seasonal positions and projects.
- Freedom for industrial staff to pursue more creative projects.
- Availability of flexible, cost-effective workforce not requiring a long-term employer commitment.
- Proven, cost-effective way to recruit and evaluate potential employees.
- Enhancement of employer's image in the community by contributing to the educational enterprise.

Types of Internships

- Industry Internship with/without Stipend
- Govt / PSU Internship (BARC/Railway/ISRO etc)
- Internship with prominent education/research Institutes
- Internship with Incubation centres /Start-ups

Guidelines

- All the students need to go for internship for minimum duration of 6 to 8 weeks.
- Students can take mini projects, assignments, case studies by discussing it with concerned authority from industry and can work on it during internship.
- All students should compulsorily follow the rules and regulations as laid by industry.
- Every student should take prior permissions from concerned industrial authority if they want to use any drawings, photographs or any other document from industry.
- Student should follow all ethical practices and SOP of industry.
- Students have to take necessary health and safety precautions as laid by the industry.
- Student should contact his /her Guide/Supervisor from college on weekly basis to communicate the progress.
- Each student has to maintain a diary/log book
- After completion of internship, students are required to submit
    - Report of work done
    - Internship certificate copy
    - Feedback from employer / internship mentor
    - Stipend proof (in case of paid internship).

Total Marks 100: The marks awarded for the Internship will be on the basis of (i) Evaluation done by the Industry (ii) Students diary (iii) Internship Report and (iv) Comprehensive Viva Voce.

Continuous Internal Evaluation: 50 marks

Student's diary - 25 Marks
Evaluation done by the Industry - 25 Marks

Student's Diary/ Daily Log: The main purpose of writing daily diary is to cultivate the habit of documenting and to encourage the students to search for details. It develops the students' thought process and reasoning abilities. The students should record in the daily training diary the day to day account of the observations, impressions, information gathered and suggestions given, if any. It should contain the sketches & drawings related to the observations made by the students. The daily training diary should be signed after every day by the supervisor/ in charge of the section where the student has been working. The diary should also be shown to the Faculty Mentor visiting the industry from time to time and got ratified on the day of his visit. Student's diary will be evaluated on the basis of the following criteria:

- Regularity in maintenance of the diary
- Adequacy & quality of information recorded
- Drawings, design, sketches and data recorded
- Thought process and recording techniques used
- Organization of the information.

The format of student's diary

Name of the Organization/Section:

Name and Address of the Section Head:

Name and Address of the Supervisor:

Name and address of the student:

Internship Duration: From .... To ....

Brief description about the nature of internship:

Day	Brief write up about the Activities carried out: Such as design, sketches, result observed, issues identified, data recorded, etc.
1	
2	
3	

Signature of Industry Supervisor *Signature of Section Head/HR Manager*

Office Seal

Attendance Sheet

Name of the Organization/Section:

Name and Address of the Section Head:

Name and Address of the Supervisor:

Name and address of the student:

Internship Duration: From .... To ....

Month & Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	...
Month & Year																					
Month & Year																					

Signature of Industry Supervisor *Signature of Section Head/HR Manager*

Office Seal

Note:

- Student's Diary shall be submitted by the students along with attendance record and an evaluation sheet duly signed and stamped by the industry to the Institute immediately after the completion of the training.
- Attendance Sheet should remain affixed in daily training diary. Do not remove or tear it off.
- Student shall sign in the attendance column. Do not mark 'P'.
- Holidays should be marked in red ink in the attendance column. Absent should be marked as 'A' in red ink.

Evaluation done by the Industry (Marks 25)

Format for Supervisor Evaluation of Intern

Student Name :______________________Date:___________________

Supervisor Name :___________________Designation:_____________

Company/Organization :____
Internship Address: ____

Dates of Internship: From______________To ______________________

Please evaluate intern by indicating the frequency with which you observed the following parameters:

Parameters Marks	Needs improvement (0 – 0.25 mark)	Satisfactory (0.25 – 0.50 mark)	Good (0.75 mark)	Excellent (1 mark)
Behavior				
Performs in a dependable Manner				
Cooperates with coworkers and supervisor				
Shows interest in work				
Learns quickly				
Shows initiative				
Produces high quality work				
Accepts responsibility				
Accepts criticism				
Demonstrates organizational skills				
Uses technical knowledge and expertise				
Shows good judgment				
Demonstrates creativity/originality				
Analyzes problems effectively				
Is self-reliant				
Communicates well				
Writes effectively				
Has a professional attitude				
Gives a professional appearance				
Is punctual				
Uses time effectively				

Overall performance of student

Intern (Tick one) : Needs improvement (0 - 0.50 mark) / Satisfactory (0.50 – 1.0 mark) / Good (1.5 mark) / Excellent (2.0 mark)

Additional comments, if any (2 marks) :

Signature of Industry Supervisor *Signature of Section Head/HR Manager*

Office Seal

End Semester Evaluation (External Evaluation): 50 Marks

Internship Report	-	25 Marks
Viva Voce	-	25 Marks

Internship Report: After completion of the internship, the student should prepare a comprehensive report to indicate what he has observed and learnt in the training period and should be submitted to the faculty Supervisor. The student may contact Industrial Supervisor/ Faculty Mentor for assigning special topics and problems and should prepare the final report on the assigned topics. Daily diary will also help to a great extent in writing the industrial report since much of the information has already been incorporated by the student into the daily diary. The training report should be signed by the Internship Supervisor, Programme Coordinator and Faculty Mentor.

The Internship report (25 Marks) will be evaluated on the basis of following criteria:

- Originality
- Adequacy and purposeful write-up
- Organization, format, drawings, sketches, style, language etc.
- Variety and relevance of learning experience
- Practical applications, relationships with basic theory and concepts taught in the course

Viva Voce (25 Marks) will be done by a committee comprising Faculty Supervisor, PG Programme Coordinator and an external expert (from Industry or research/academic Institute). This committee will be evaluating the internship report also.

(v) LABORATORY COURSES

Lab work and Viva-voce : 60 marks

Final evaluation Test and Viva voce : 40 marks

The laboratory courses will be having only Continuous Internal Evaluation and carries 100 marks. Final evaluation shall be done by two examiners; one examiner will be a senior faculty from the same department.

(vi) INDUSTRY INTEGRATED COURSE

Engineering students frequently aspire to work in areas and domains that are key topics in the industry. There are concerns by recruiters that skill sets of engineering students did not match with the Industry requirements, especially in the field of latest topics. In response to their desires, the College has incorporated Industry integrated course in the curriculum.

The evaluation pattern for Industry based electives is as follows:

Continuous Internal Evaluation : 40 marks

Seminar : 10 marks

Course based task/Seminar/Data collection

and interpretation/Case study : 10marks

Test paper 1 (Module 1 and Module 2) : 10 marks

Test paper 2 (Module 3 and Module 4) : 10 marks

End Semester Examination : 60 marks

The examination will be conducted by the College with the question paper provided by the Industry. The examination will be for 3 Hrs and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks. The valuation of the answer scripts shall be done by the expert in the Industry handling the course.

(vii) MOOC COURSES

The MOOC course shall be considered only if it is conducted by the agencies namely AICTE/NPTEL/SWAYAM or NITTTR. The MOOC course should have a minimum duration of 8 weeks and the content of the syllabus shall be enough for at least 40 hours of teaching. The course should have a proctored/offline end semester examination. The students can do the MOOC according to their convenience, but shall complete it before the end of fourth semester. The list of MOOC courses will be provided by the concerned BoS if at least 70% of the course content match with the area/stream of study. The course shall not be considered if its content has more than 50% of overlap with a core/elective course in the concerned discipline or with an open elective.

MOOC Course to be successfully completed before the end of fourth semester (starting from semester 1). A credit of 2 will be awarded to all students whoever successfully completes the MOOC course as per the evaluation pattern of the respective agency conducting the MOOC.

(viii) MINIPROJECT

Total marks: 100, only CIA

Mini project can help to strengthen the understanding of student's fundamentals through application of theoretical concepts and to boost their skills and widen the horizon of their thinking. The ultimate aim of an engineering student is to resolve a problem by applying theoretical knowledge. Doing more projects increases problem- solving skills. The introduction of mini projects ensures preparedness of students to undertake dissertation. Students should identify a topic of interest in consultation with PG Programme Coordinator. Demonstrate the novelty of the project through the results and outputs. The progress of the mini project is evaluated based on three reviews, two interim reviews and a final review. A report is required at the end of the semester.

Interim evaluation: 40 (20 marks for each review), final evaluation by a Committee (will be evaluating the level of completion and demonstration of functionality/specifications, clarity of presentation, oral examination, work knowledge and involvement): 35, Report (the committee will be evaluating for the technical content, adequacy of references, templates followed and permitted plagiarism level is not more than 25%): 15, Supervisor/Guide: 10

(ix) DISSERTATION

Dissertation: All Students should carry out the dissertation in the college or can work either in any CSIR/Industrial R&D organization/any other reputed Institute which have facilities for dissertation work in the area proposed.

Dissertation outside the Institute: For doing dissertation outside the Institution, the following conditions are to be met:

- They have completed successfully the course work prescribed in the approved curriculum up to the second semester.
- They should choose Track 2 in semester 3 and 4
- The student has to get prior approval from the DLAC and CLAC.
- Facilities required for doing the dissertation shall be available in the Organization/Industry (A certificate stating the facilities available in the proposed organization and the time period for which the facilities shall be made available to the student, issued by a competent authority from the Organization/Industry shall be submitted by the student along with the application).
- They should have an external as well as an internal supervisor. The internal supervisor should belong to the parent institution and the external supervisor should be Scientists or Engineers from the Institution/Industry/ R&D organization with which the student is associated for doing the dissertation work. The external supervisor shall be with a minimum post graduate degree in the related area.
- The student has to furnish his /her monthly progress as well as attendance report signed by the external guide and submit the same to the concerned Internal guide.
- The external guide is to be preferably present during all the stages of evaluation of the dissertation.

Note1- Students availing this facility should continue as regular students of the College itself.

Note 2-The course work in the 3rd semester is to be completed as per the curriculum requirements (i) MOOC can be completed as per the norms mentioned earlier

Internship leading to Dissertation: The M. Tech students who after completion of 6 to 8 weeks internship at some reputed organization are allowed to continue their work as dissertation for the third and fourth semester after getting approval from the DLAC. Such students shall make a brief presentation regarding the work they propose to carry out before the DLAC for a detailed scrutiny and to resolve its suitability for accepting it as an M.Tech dissertation. These students will be continuing as regular students of the Institute in third semester for carrying out all academic requirements as per the curriculum/regulation. However, they will be permitted to complete their dissertation in the Industry/Organization (where they have successfully completed their internship) during fourth semester.

Dissertation as part of Employment: Students may be permitted to discontinue the programme and take up a job provided they have completed all the courses till second semester (FE status students are not permitted) prescribed in the approved curriculum. The dissertation work can be done during a later period either in the organization where they work if it has R & D facility, or in the Institute. Such students should submit application with details (copy of employment offer, plan of completion of their project etc.) to the Dean (PG) through HoD. The application shall be vetted by CLAC before granting the approval. When the students are planning to do the dissertation work in the organization with R & D facility where they are employed, they shall submit a separate application having following details:

- Name of R&D Organization/Industry
- Name and designation of an external supervisor from the proposed Organization/Industry (Scientists or Engineers with a minimum post graduate degree in the related area) and his/her profile with consent
- Name and designation of a faculty member of the Institute as internal supervisor with his/her consent
- Letter from the competent authority from the Organization/Industry granting permission to do the dissertation
- Details of the proposed work
- Work plan of completion of project

DLAC will scrutinize the proposal and forward to CLAC for approval.

When students are doing dissertation work along with the job in the organization (with R & D facility) where they are employed, the dissertation work shall be completed in four semesters normally (two semesters of dissertation work along with the job may be considered as equivalent to one semester of dissertation work at the Institute). Extensions may be granted based on requests from the student and recommendation of the supervisors such that he/she will complete the M. Tech programme within four years from the date of admission as per the regulation. Method of evaluation and grading of the dissertation will be the same as in the case of regular students. The course work in the 3rd semester for such students are to be completed as per the curriculum requirements (i) MOOC can be completed as per the norms mentioned earlier. However, for self learning students, all evaluations shall be carried out in their parent Institution as in the case of regular students.

Mark Distribution:

Phase 1: Total marks: 100, only CIA

Phase 2: Total marks: 200, CIA = 100 and ESE = 100 marks

(x) TEACHING ASSISTANCESHIP (TA)

All M.Tech students irrespective of their category of admission, shall undertake TA duties for a minimum duration as per the curriculum. Being a TA, the student will get an excellent opportunity to improve their expertise in the technical content of the course, enhance communication skills, obtain a hands-on experience in handling the experiments in the laboratory and improve peer interactions.

The possible TA responsibilities include the following: facilitate a discussion section or tutorial for a theory/ course, facilitate to assist the students for a laboratory course, serve as a mentor for students, and act as the course web-master. TAs may be required to attend the instructor's lecture regularly. A TA shall not be employed as a substitute instructor, where the effect is to relieve the instructor of his or her teaching responsibilities.

For the tutorial session:

(i) Meet the teacher and understand your responsibilities well in advance, attend the lectures of the course for which you are a tutor, work out the solutions for all the tutorial problems yourself, approach the teacher if you find any discrepancy or if you need help in solving the tutorial problems, use reference text books, be innovative and express everything in English only.

(ii) Try to lead the students to the correct solutions by providing appropriate hints rather than solving the entire problem yourself, encourage questions from the students, lead the group to a discussion based on their questions, plan to ask them some questions be friendly and open with the students, simultaneously being firm with them.

(iii) Keep track of the progress of each student in your group, give a periodic feedback to the student about his/her progress, issue warnings if the student is consistently under-performing, report to the faculty if you find that a particular student is consistently underperforming, pay special attention to slow-learners and be open to the feedback and comments from the students and faculty.

(iv) After the tutorial session you may be required to grade the tutorials/assignments/tests. Make sure that you work out the solutions to the questions yourself, and compare it with the answer key, think and work out possible alternate solutions to the same question, understand the marking scheme from the teacher. Consult the teacher and make sure that you are not partial to some student/students while grading. Follow basic ethics.

Handling a laboratory Session:

(i) Meet the faculty – in- charge a few days in advance of the actual lab class and get the details of the experiment, get clarifications from him/her regarding all aspects of the experiment and the expectations, prepare by reading about the theoretical background of the experiment, know the physical concepts involved in the experiment, go to the laboratory and check out the condition of the equipment/instrumentation, perform the laboratory experiment at least once one or two days before the actual laboratory class, familiarize with safety/ security aspects of the experiment / equipment/laboratory, prepare an instruction sheet for the experiment in consultation with the faculty, and keep sufficient copies ready for distribution to students for their reference.

(ii) Verify condition of the equipment/set up about 30 minutes before the students arrive in the class and be ready with the hand outs, make brief introductory remarks about the experiment, its importance, its relevance to the theory they have studied in the class, ask the students suitable questions to know there level of preparation for the experiment, discuss how to interpret results, ask them comment on the results.

(iii) Correct/evaluate/grade the submitted reports after receiving suitable instructions from the faculty in charge, continue to interact with students if they have any clarifications regarding any aspect of the laboratory session, including of course grading, Carefully observe instrument and human safety in laboratory class, Preparing simple questions for short oral quizzing during explanation of experiments enables active participation of students, facilitate attention, provides feedback and formative evaluation.

POINTS TO REMEMBER

1. Arrange an awareness programme to all M.Tech students on day 1 regarding the curriculum and the regulation.
2. The departments should prepare the list of MOOC courses suitable to their programmes and encourage the students to complete at the earliest.
3. Make a tie up with industries by the middle of semester for Industry Integrated Course. While choosing the course, it should be ensured that the programme is relevant and updated in that discipline. The Industry expert handling the course shall be a postgraduate degree holder. The evaluation procedure shall also be clearly explained to them.
4. Each department offering M.Tech programme should be careful in selecting the mini project in semester 2.
5. The departments should invite the Industries/research organizations during first semester and inform them about the mandatory 6-8 weeks internship that the students should undergo after their second semester. The possibility of doing their dissertation at the Industry shall also be explored. They should also be made aware about the evaluation procedure of the Internships. They may also be informed that it is possible to continue internship provided if it leads to their dissertation. Proposals may be collected from them for allotting to students according to their fields of interest.
6. Make sure that all internal evaluations and the end semester examinations to be conducted by the college are carried out as per the evaluation procedure listed in the curriculum. Any dilution from the prescribed procedure shall be viewed seriously.

7. Teaching assistance shall be assigned to all students as per the curriculum. However, a TA shall not be employed as a substitute instructor, where the effect is to relieve the instructor of his or her teaching responsibilities.
8. The possible TA responsibilities include the following: facilitate a discussion section or tutorial for a theory/ course, facilitate to assist the students for a laboratory course, serve as a mentor for students, and act as the course web-master.

****

SEMESTER I

M.TECH

COMPUTER SCIENCE AND ENGINEERING

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1T101	**COMPUTATIONAL INTELLIGENCE**	CORE	4	0	0	4	4

Preamble : The aim of this course is to provide the students with the knowledge and skills required to design and implement effective and efficient Computational Intelligence solutions to problems for which a direct solution is impractical or unknown. This course covers concepts of fuzzy logic, genetic algorithms, and swarm optimization techniques. The learners will be able to provide Fuzzy and AI –based solutions to real world problems.

Prerequisites : Nil

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Apply fuzzy logic to handle uncertainty and solve engineering problems. (Cognitive Knowledge Level : Apply)
CO 2	Apply Fuzzy Logic Inference methods in building intelligent machines. (Cognitive Knowledge Level : Apply)
CO 3	Design genetic algorithms for optimized solutions in engineering problems. (Cognitive Knowledge Level : Analyze)
CO 4	Analyze the problem scenarios and apply Ant colony system to solve real optimization problems. (Cognitive Knowledge Level : Analyze)
CO 5	Apply PSO algorithm to solve real world problems. (Cognitive Knowledge Level : Apply)

Mapping of Course Outcomes With Program Outcomes

	PO 1	**PO 2**	**PO 3**	**PO 4**	**PO 5**	**PO 6**
CO 1						✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	**Continuous Internal Evaluation Tests**		**End Semester Examination (% Marks)**
	Test 1 (%Marks)	**Test 2 (%Marks)**	
Remember			
Understand			
Apply	60-80%	60-80%	60-80%
Analyse	20-40%	20-40%	20-40%
Evaluate			
Create			

Mark Distribution

Continuous Internal Evaluation Pattern : Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Micro project/Course based project	:	10 marks
Course based task/Seminar/Quiz	:	10 marks
Test paper 1	:	10 marks
Test paper 2	:	10 marks

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (10 hours)

Crisp sets, Fuzzy sets, Logic Operations and application of Fuzzy sets, Membership functions-basic operations, linguistic variables and hedges , Fuzzy logic laws. Arithmetic operation on fuzzy, Complement forms on fuzzy set, Extension principle, Operations on fuzzy relations, Fuzzy composition- Max- min, Max – product, Max-star, Alpha-cut representation.

MODULE 2 (8 hours)

Fuzzy interpretations-Mamdani, Zadeh, Fuzzy Reasoning – GMP and GMT, Fuzzy Inference System -Mamdani FIS, Larsen Model, Defuzzification methods - Centroid method, Centre of Sums, Mean of Maxima.

MODULE 3 (8 hours)

Introduction to Genetic Algorithms - Theoretical foundation, Gene representation and fitness function, Selection, Recombination(crossover), Mutation, Elitism, Applications of GA, Multi-objective Genetic Algorithm – Pareto–optimal Ranking,Goldberg's Ranking, Fonseca Flemming Method, Lu and Yen method

MODULE 4 (9 hours)

Swarm intelligent systems –Fundamentals, Introduction to Ant colony systems,Development of the ant colony system-Ant colony optimization -Pheromone updation,Pheromone Evaporation,Pheromone intensification , Types of Ant Colony Models, ACO algorithm application - TSP

MODULE 5 (10 hours)

Basics of Particle swarm optimization (PSO), Global Best PSO, Local Best PSO- Comparison of 'gbest' to 'lbest', PSO Algorithm Parameters-Swarm size, Velocity components, Acceleration coefficients, Problem Formulation of PSO algorithm-Velocity Updation. Velocity clamping,Inertia weight, Application of PSO.

Text Books

1. Samir Roy, Udit Chakraborty, Introduction to Soft Computing Neuro- Fuzzy Genetic Algorithms, Pearson, 2013
2. N.P. Padhy, Artificial Intelligence and Intelligent systems, Oxford Press, New Delhi, 2005.

Reference Books

1. Xin-She Yang School of Science and Technology, Middlesex University London, NatureInspired Optimization Algorithms, Elsevier, First edition, 2014
2. Satyobroto Talukder, Blekinge Institute of Technology, Mathematical Modelling and Applications of Particle Swarm Optimization, February 2011
3. Mitchell Melanie, An Introduction to Genetic Algorithm, Prentice Hall,1998
4. Andries Engelbrecht, Computational Intelligence: An Introduction, Wiley, 2007

5. Marco Dorigo and Thomas Stutzle, “Ant Colony optimization”, Prentice Hall of India, New Delhi 2005

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	10 hours
1.1	Crisp sets	1 hour
1.2	Fuzzy sets, Logic, Operations and application of Fuzzy sets	1 hour
1.3	Membership functions- basic operations	1 hour
1.4	Linguistic variables and hedges	1 hour
1.5	Fuzzy logic laws	1 hour
1.6	Arithmetic operation on fuzzy, Complement-forms on fuzzy set	1 hour
1.7	Extension principle	1 hour
1.8	Operations on fuzzy relations	1 hour
1.9	Fuzzy composition- Max- min, Max – product, Max-stare	1 hour
1.10	Alpha-cut representation	1 hour
2	**Module 2**	8 hours
2.1	Fuzzy interpretations Mamdani, Zadeh	2 hours
2.2	Fuzzy Reasoning – GMP and GMT	2 hours
2.3	Fuzzy Inference System -Mamdani FIS, Larsen Model	2 hours
2.4	Defuzzification methods - Centroid method , Centre of Sums ,Mean of Maxima.	2 hours
3	**Module 3**	8 hours
3.1	Introduction to Genetic Algorithms- Theoretical foundation	1 hour

3.2	Gene representation and Fitness function	1 hour
3.3	Selection	1 hour
3.4	Recombination(crossover)	1 hour
3.5	Mutation	1 hour
3.6	Elitism	1 hour
3.7	Applications of GA	1 hour
3.8	Multi-objective Genetic Algorithm – Pareto–optimal Ranking,Goldberg's Ranking, Fonseca Flemming Method, Lu and Yen method	1 hour
4	**Module 4**	9 hours
4.1	Swarm intelligent systems –Fundamentals	1 hour
4.2	Introduction to Ant colony systems	1 hour
4.3	Development of the ant colony system	1 hours
4.4	Ant colony optimization -Pheromone updation,Pheromone Evaporation,Pheromone intensification	2 hours
4.5	Types of Ant Colony Models	2 hours
4.6	ACO algorithm application	2 hours
5	**Module 5**	10 hours
5.1	Basics of Particle swarm optimization (PSO)	1 hour
5.2	Global Best PSO	1 hour
5.3	Local Best PSO	1 hour
5.4	Comparison of 'gbest' to 'lbest'	1 hour
5.5	PSO Algorithm Parameters	1 hour
5.6	Velocity Updating	1 hour
5.7	Velocity clamping	1 hour
5.8	Inertia weight	1 hour
5.9	Application of PSO	1 hour

5.10	Problem Formulation of PSO algorithm -Velocity Updation. Veloc- ity clamping,Inertia weight, Application of PSO.	1 hour
	Total	45 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 4**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1T101

Course Name: :COMPUTATIONAL INTELLIGENCE

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Consider the set of students S = John, Ramesh, Alice, Diya, Job.

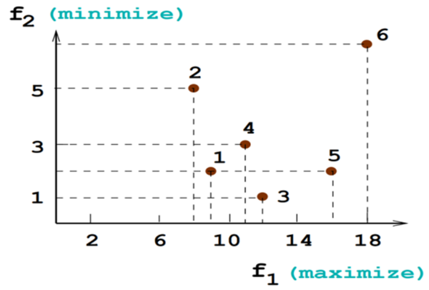
$$\text{Let Hardworking Students } H = \frac{0.6}{\text{John}} + \frac{0.8}{\text{Ramesh}} + \frac{0.9}{\text{Alice}} + \frac{0.8}{\text{Divya}} + \frac{0.2}{\text{Job}}$$

$$\text{Intelligent Students } I = \frac{0.4}{\text{John}} + \frac{0.9}{\text{Ramesh}} + \frac{0.8}{\text{Alice}} + \frac{0.4}{\text{Divya}} + \frac{0.9}{\text{Job}}$$

   Use appropriate operations to find students who are Hard working or Intelligent, Hardworking but not Intelligent and Highly Intelligent.

2. Develop a membership function for “Tall”. Based on that devise membership function for “Very Tall”. Explain how it is done.

3. Consider two objective functions f_1 to be minimized and f_2 to be maximized . Determine the non-dominating solutions among 1,2,3,4,5 and 6. Graphically represent the Pareto-Optimal Front and Pareto-Optimal Ranking of these two objective functions.



4. Describe how pheromone is updated. What is elitist / elastic ants ? Are they useful in this scenario?

5. What is velocity clamping in Particle Swarm Optimization? How it helps in exploration of the search space? Which one is more suitable for optimization problems with multiple optima, global best PSO or local best PSO ?

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) The membership function of a fuzzy set HIG, where HIG stands for high income group, is defined as follows: (4 marks)

   (b) Find the alpha cut representation for

$$uHIG(i) = \begin{cases} 0 & \text{if } i \leq 3 \\ \frac{i-3}{3} & \text{if } 3 < i \leq 6 \\ 1 & \text{if } i > 6 \end{cases}$$

(4 marks)

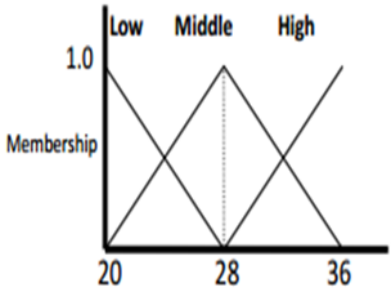
7. Apply the fuzzy GMP rule to deduce Rotation is quite slow Given that: If temperature is High then rotation is Slow. The temperature is Very High Given X = 30, 40, 50, 60, 70, 80, 90, 100 be the set of temperatures. Y = 10, 20, 30, 40, 50, 60 be the set of rotations per minute. The fuzzy set High(H), Very High (VH), Slow(S) and Quite Slow

(QS) are given below. H = (70, 1),(80, 1),(90, 0.3) VH = (90, 0.9),(100, 1) S = (30, 0.8),(40, 1.0),(50, 0.6) QS = (10, 1),(20, 0.8).Determine R using Zadeh's interpretation

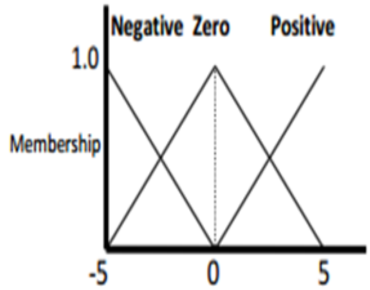
(8 marks)

8. A fuzzy system is used to control the temperature in a chamber. The inputs to the controller are the 'Temperature' and 'Temperature Change'. The output is the 'Fan Voltage'. The membership functions and the rule table are given below. If the temperature is 22 degree Celsius and the change of temperature is 0, calculate the fan voltage using Mamdani inference and Mean of maxima defuzzification method.

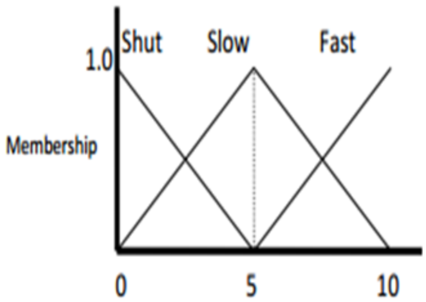
(8 marks)



Temp. (C)



Temp Change (C/min)



Fan Voltage (V)

Table Rules		Temperature		
		Low	Middle	High
Temp. Change	Neg	Fan Shut	Fan Shut	Fan Slow
	Zero	Fan Shut	Fan Slow	Fan Fast
	Pos	Fan Slow	Fan Fast	Fan Fast

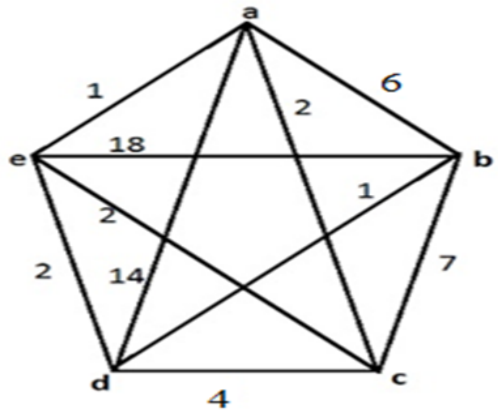
9. (a) Mention the importance of objective (fitness) function in genetic algorithm

(4 marks)

(b) Describe two methods used to select individuals from a population for the mating pool in Genetic Algorithms

(4 marks)

10. (a) Consider the TSP with the following edge costs. Given the evaporation factor ρ =0.02 and initial pheromone at all edges Tij=100 (4 marks)



(b) Using the equation

$$T_{ij}(t+1) = (1-\rho)T_{ij}(t) + \Delta T_{ij}(t+1)$$

compute the Tijof the edge when 10 ants uses the edges , using the following models:
i.Ant Density Model (Constant Q=10)
ii. Ant Quantity Model(Constant Q=100) where Q is the constant related to the pheromone updation

(4 marks)

11. What are the different ways in which pheromone trail updating process can be carried out in ant systems? Explain pheromone updation in AS ranking model. (8 marks)

12. Apply PSO to minimize

$$2x^2 - 2xy + y^2 \quad \text{for} \quad -5 < x < 1 \quad \text{and} \quad -1 \leq y \leq 1$$

Assume that c0 ,c1 and c2 are 1. Maximum velocity in any direction is 20. After nth iteration, assume a population of p1(-40,40), p2(-20, 80) and p3(10,50) and best solutions for p1, p2 and p3 are p1'(-20,40), p2'(-40,90) and p3'(10,0),current velocity of all particles(-10,+10).Determine the best solution after 1 iteration. Show intermediate steps. (8 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1T102	**ADVANCED DATABASE MANAGEMENT**	CORE	4	0	0	4	4

Preamble : This course provides an exposure to the concepts and techniques in advanced database management. Various strategies regarding query processing and optimization are discussed in this curriculum. An optimum insight of database security is provided. Different layouts of database system architecture and distributed system architecture, along with semistructured data is included for better understanding of advanced data management. This course helps the learners to develop applications that manage data efficiently with the help of suitable data models and techniques.

Prerequisites : M24CS1T102 DATABASE MANAGEMENT SYSTEMS.
Knowledge of basic database management techniques and SQL queries.

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Identify various measures of query processing and optimization. (Cognitive Knowledge Level: Apply)
CO 2	Analyze and implement security mechanisms to secure a database system. (Cognitive Knowledge Level: Analyze)
CO 3	Apply knowledge and awareness of the different database architectures in different scenarios. (Cognitive Knowledge Level: Apply)
CO 4	Analyze implementation aspects of distributed system on database architecture. (Cognitive Knowledge Level: Analyze)
CO 5	Make use of semi structured data, XML and XML queries for data management. (Cognitive Knowledge Level: Apply)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓		✓
CO 2	✓		✓	✓		✓
CO 3	✓		✓	✓		✓
CO 4	✓		✓	✓		✓
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	20	20	10
Apply	50	50	50
Analyse	30	30	30
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Micro project/Course based project	: 10 marks
Course based task/Seminar/Quiz	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (8 hours)

Query Processing and Optimization: Review of indexing and Hashing - Overview- Measures of query cost- Algorithms for Selection and Join with cost analysis- Evaluation of expressions- Optimization of RA expressions

MODULE 2 (7 hours)

Database Security: Threats to databases, control measures, database security and DBA, Discretionary access control, Mandatory access control (role-based only), SQL injection.

MODULE 3 (9 hours)

Database System Architectures: Centralized and Client-Server Architectures Centralized server systems - Server System Architectures - Parallel Systems- Parallel storage - Data partitioning, replication and indexing in Parallel Databases- Parallel query processing.

MODULE 4 (10 hours) **Distributed System Architecture**: Distributed System architecture- Distributed storage - Distributed file systems – Distributed RDB design- Transparency– Distributed Transactions - Commit Protocols – Concurrency Control - Distributed Query Processing Advanced Indexing Techniques: Bloom filter - Bitmap indices - Indexing spatial data - Hash indices.

MODULE 5 (6 hours)

Semi-structured Data: Semi-structured Data and XML Databases: XML Data Model – XSD – XPath and XQuery – Example Queries. Native XML databases, Object Relational Systems

Text Books

1. R. Elmasri, S.B. Navathe, "Fundamentals of Database Systems", 7/e, PearsonEducation/Addison Wesley, 2016

Reference Books

1. Thomas Cannolly and Carolyn Begg, "Database Systems, A Practical Approach to Design, Implementation and Management", 3/e, Pearson Education, 2010.
2. Henry F Korth, Abraham Silberschatz, S. Sudharshan, "Database System Concepts", 7/e,Tata McGraw Hill, 2019.
3. Joe Fawcett, Danny Ayers, Liam R. E. Quin, Beginning XML, 5/e, John Wiley & Sons, 2012
4. Grigoris Antoniou. Frank van Harmelen, "A Semantic Web Primer", The MIT Press, Cambridge, Massachusetts, 2003

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	8 hours
1.1	Introduction to Query Processing and Optimization	1 hour
1.2	Review of indexing	1 hour
1.3	Hashing – Overview	1 hour
1.4	Measures of query cost	1 hour
1.5	Algorithms for Selection with cost analysis	1 hour
1.6	Algorithms for Join with cost analysis	1 hour
1.7	Evaluation of expressions	1 hour
1.8	Optimization of RA expressions	1 hour
2	**Module 2**	7 hours
2.1	Introduction to Database Security	1 hour
2.2	Threats to databases	1 hour
2.3	Control measures	1 hour
2.4	Database security and DBA	1 hour
2.5	Discretionary access control	1 hour
2.6	Mandatory access control (role-based only)	1 hour

2.7	SQL injection	1 hour
3	**Module 3**	9 hours
3.1	Introduction to Database System Architectures	1 hour
3.2	Overview of Centralized and Client-Server Architectures	1 hour
3.3	Centralized server systems	1 hour
3.4	Server System Architectures	1 hour
3.5	Parallel Systems	1 hour
3.6	Parallel storage	1 hour
3.7	Data partitioning, replication in Parallel Databases	1 hour
3.8	Indexing in Parallel Databases	1 hour
3.9	Parallel query processing	1 hour
4	**Module 4**	10 hours
4.1	Introduction to Distributed System architecture	1 hour
4.2	Distributed storage & Distributed file systems	1 hour
4.3	Distributed RDB design & its Transparency	1 hour
4.4	Distributed Transactions	1 hour
4.5	Commit Protocols & Concurrency Control	1 hour
4.6	Distributed Query Processing	1 hour
4.7	Advanced indexing Techniques: Bloom filter	1 hour

4.8	Bitmap indices	1 hour
4.9	Indexing spatial data	1 hour
4.10	Hash indices	1 hour
5	**Module 5**	6 hours
5.1	Introduction to Semi-structured Data and XML Databases	1 hour
5.2	XML Data Model – XSD	1 hour
5.3	XPath and XQuery	1 hour
5.4	Example Queries	1 hour
5.5	Native XML databases	1 hours
5.6	Object Relational Systems	1 hour
	Total	40 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1T102

Course Name: *ADVANCED DATABASE MANAGEMENT*

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Why is it not desirable to force users to make an explicit choice of a query processing strategy? Are there cases in which it is desirable for users to be aware of the costs of competing query-processing strategies? Explain your answer.

2. What are the relative merits of using Discretionary Access Control or Mandatory Access Control? What is role-based access control? In what ways is it superior to DAC and MAC?

3. Suppose relation r is stored partitioned and indexed on A, and s is stored partitioned and indexed on B. Consider the query: r.Cγcount(s.D)((σA>5(r)) $\bowtie$r.B=s.B s)
   i Give a parallel query plan using the exchange operator, for computing the subtree of the query involving only the select and join operators.
   ii Now extend the above to compute the aggregate. Make sure to use preaggregation to minimize the data transfer.

4. Insert and query on a Bloom filter of size m = 10 and number of hash functions k= 3. Let H(x) denote the result of the three hash functions which will write as a set of three values h 1 (x), h 2 (x), h 3 (x) Has functions used: A= x mod 10, B= x mod 7, C= (sum of digits) mod 9.

5. Design an XML document for storing hostel mess food details (meals taken such as breakfast, lunch, dinner) with their charges for the month of June 2022. Charges may vary depending on the food taken. Students can opt not to take any meals on certain days
   i Write a sample XML for 2 students for 2 days.
   ii Write a XQuery to return the lunch details of all

Arial*PART B*

Arial***Answer any five question from each module. Each question carries 8 marks.***

6. (a) Consider the issue of interesting orders in optimization. Suppose you are given a query that computes the natural join of a set of relations S. Given a subset S1 of S, what are the interesting orders of S1? (4 marks)

   (b) Suppose you want to get answers to r s sorted on an attribute of r, and want only the top K answers for some relatively small K. Give a good way of evaluating the query:
   i When the join is on a foreign key of r referencing s, where the foreign key attribute is declared to be not null.
   ii When the join is not on a foreign key (4 marks)

7. A database relation may have the values of certain attributes encrypted for security. Why do database systems not support indexing on encrypted attributes? Using your answer to this question, explain why database systems do not allow encryption of primary-key attributes. (8 marks)

8. If a parallel data-store is used to store two relations r and s and we need to join r and s, it may be useful to maintain the join as a materialized view. What are the benefits and overheads in terms of overall throughput, use of space, and response time to user queries? Explain in detail. (8 marks)

9. Consider the bitmap representation of the free-space map, where for each block in the file, two bits are maintained in the bitmap. If the block is between 0 and 30 percent full the bits are 00, between 30 and 60 percent the bits are 01, between 60 and 90 percent the bits are 10, and above 90 percent the bits are 11. Such bitmaps can be kept in memory even for quite large files.
   i Outline two benefits and one drawback to using two bits for a block, instead of one byte as described earlier in this chapter.
   ii Describe how to keep the bitmap up to date on record insertions and deletions.
   iii Outline the benefit of the bitmap technique over free lists in searching for free space and in updating free space information.

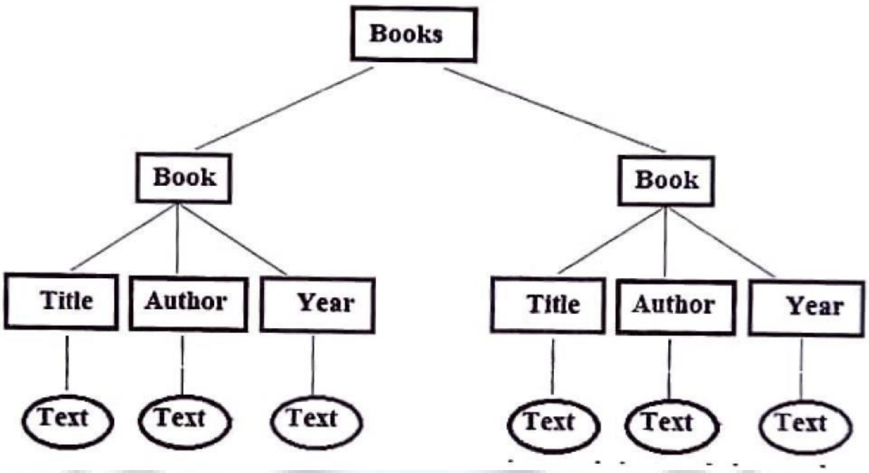
   (8 marks)

10. Assume a relationship RAB at site 1 and relationship SCD at site 2 as follows: R= (1,2), (3,4), (5, 6), (7, 8), (9, 10) S= (1, 0), (8,1), (9, 2), (10, 3), (11, 4) Compute R S using bloom join with A=C and explain the intermediate steps. Show the tuples transferred with the hash function mod 4. (8 marks)

11. Consider the following XML Tree:

    Write an XML schema for the above, and also provide an XQuery expression to get the books published in the year 1992. (8 marks)

    (4 marks)



12. (a) Consider the country data. Write XPath for the following:
i Return the area of India.
ii Return the names of all countries with population greater than 100 million.
iii Return the names of all countries whose population is less than one thousandth that of some city (in any country). iv Return the names of all cities that have the same name as the country in which they are located.

(4 marks)

(b) Consider the country data. Write XQuery for the following:
i Return the name of the country with the highest population.
ii Return the name of the country that has the city with the highest population.
iii Return the average population of Russian-speaking countries.

(4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1T103	**ADVANCED MACHINE LEARNING**	CORE	4	0	0	4	4

Preamble : This course introduces machine learning concepts and popular machine learning algorithms. It will cover the standard and most popular supervised learning algorithms including linear regression, logistic regression, decision trees, k-nearest neighbour, an introduction to Bayesian learning and the naive Bayes algorithm, support vector machines and kernels and basic clustering algorithms. Dimensionality reduction methods and some applications to real world problems will also be discussed. It helps the learners to develop application machine learning based solutions for real world applications.

Prerequisites : Basic understanding of probability theory, linear algebra, multivariate calculus and multivariate probability theory.

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Analyse the Machine Learning concepts, basic parameter estimation methods, Optimization and regularization techniques. (Cognitive Knowledge Level: Analyse)
CO 2	Illustrate the concepts of regression and classification techniques (Cognitive Knowledge Level: Apply)
CO 3	Describe unsupervised learning concepts and dimensionality reduction techniques. (Cognitive Knowledge Level: Apply)
CO 4	Explain Support Vector Machine concepts and markov models. (Cognitive Knowledge Level: Apply)
CO 5	Choose suitable model parameters for different machine learning techniques and to evaluate a model performance. (Cognitive Knowledge Level: Apply)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	**Continuous Internal Evaluation Tests**		**End Semester Examination (% Marks)**
	Test 1 (%Marks)	**Test 2 (%Marks)**	
Remember			
Understand			
Apply	60-80%	60-80%	60-80%
Analyse	20-40%	60-80%	60-80%
Evaluate			
Create			

Mark Distribution

Total Marks	**CIE Marks**	**ESE Marks**	**ESE Duration**
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Micro project/Course based project	: 10 marks
Course based task/Seminar/Quiz	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (8 hours)

Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning. Basics of parameter estimation: Maximum Likelihood Estimation (MLE), Maximum a Posteriori Estimation (MAP). Gradient Descent Algorithm, normal equations and closed form solution. Practical aspects in machine learning: data preprocessing, overfitting, Underfitting, parameter and model selection, Bias-Variance tradeoff, Regularization techniques - LASSO and RIDGE.

MODULE 2 (11 hours)

Regression algorithms: Simple linear regression, Multiple linear regression, Polynomial regression Classification algorithms: Logistic regression, Naive Bayes, Decision trees, K-Nearest Neighbour(KNN). Neural networks : Concept of Artificial neuron, , Perceptrons , Linear Separability , Multi-layer Perceptron , Back propagation algorithm.

MODULE 3 (7 hours)

Unsupervised learning: clustering, k-means, Hierarchical clustering, Density-based spatial clustering of applications with noise (DBSCAN). Gaussian mixture models: Expectation Maximization (EM) algorithm for Gaussian mixture model. Dimensionality Reduction – Linear Discriminant Analysis, Principal Component Analysis, Independent Component Analysis.

MODULE 4 (9 hours)

Support vector machines and kernels: Support Vector Machine, Optimal Separating hyper plane, Softmargin hyperplane, Nonlinear SVM and the kernel trick, nonlinear decision boundaries, Kernel functions. Discrete Markov Processes, Hidden Markov models

MODULE 5 (10 hours)

Classification Performance Evaluation Metrics: Accuracy, Precision, Precision, Recall, Specificity, False Positive Rate (FPR), F1 Score, Receiver Operator Characteristic (ROC) Curve, AUC. Regression Performance Evaluation Metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R Squared/Coefficient of Determination. Clustering Performance Evaluation Metrics: Purity, Jaccard index, Normalized Mutual Information, Clustering Accuracy, Silhouette Coefficient, Dunn's Index. Resampling methods: cross-validation, bootstrap. Ensemble methods: bagging, boosting, random forest, Stacking. Boosting: AdaBoost, gradient boosting machines, XGBoost

Reference Books

1. Christopher Bishop. Neural Networks for Pattern Recognition, Oxford University Press, 1995
2. Kevin P. Murphy. Machine Learning: A Probabilistic Perspective, MIT Press 2012.
3. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Second edition Springer 2007.
4. Ethem Alpaydin, Introduction to Machine Learning, 2nd edition, MIT Press 2010.

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	8 hours
1.1	Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning.	1 hour
1.2	Basics of parameter estimation: Maximum Likelihood Estimation (MLE).	1 hour
1.3	Basics of parameter estimation: Maximum a Posteriori Estimation (MAP).	1 hour
1.4	Gradient Descent Algorithm, normal equations and closed form solution.	1 hour
1.5	Generalisation and overfitting	1 hour
1.6	Bias-Variance tradeoff, parameter and model selection.	1 hour
1.7	Resampling methods: cross-validation, bootstrap.	1 hour
1.8	Regularization techniques - LASSO and RIDGE.	1 hour
2	**Module 2**	11 hours
2.1	Regression algorithms: Simple linear regression	1 hour
2.2	Multiple linear regression	1 hour
2.3	Polynomial regression	1 hour
2.4	Classification algorithms: Logistic regression	1 hour

2.5	Naive Bayes	1 hour
2.6	Decision trees	1 hour
2.7	K-Nearest Neighbour(KNN)	1 hour
2.8	Neural networks : Concept of Artificial neuron	1 hour
2.9	Perceptrons , Linear Separability	1 hour
2.10	Multi-layer Perceptron	1 hour
2.11	Back propagation algorithm	1 hour
3	**Module 3**	7 hours
3.1	Unsupervised learning: clustering, k-means	1 hour
3.2	Hierarchical clustering,	1 hour
3.3	Density-based spatial clustering of applications with noise (DBSCAN).	1 hour
3.4	Gaussian mixture models: Expectation Maximization (EM) algorithm for Gaussian mixture model	1 hour
3.5	Dimensionality Reduction – Linear Discriminant Analysis	1 hour
3.6	Principal Component Analysis	1 hour
3.7	Independent Component Analysis	1 hour
4	**Module 4**	9 hours
4.1	Support vector machines and kernels: Support Vector Machine,	1 hour
4.2	Optimal Separating hyper plane, Softmargin hyperplane,	1 hour

4.3	Nonlinear SVM and the kernel trick,	1 hour
4.4	nonlinear decision boundaries	1 hour
4.5	, Kernel functions	1 hour
4.6	Discrete Markov Processes,	1 hour
4.7	Hidden Markov models	1 hour
4.8	Hidden Markov models	1 hour
4.9	Hidden Markov models	1 hour
5	**Module 5**	10 hours
5.1	Classification Performance Evaluation Metrics: Accuracy, Precision, Precision, Recall, Specificity, False Positive Rate (FPR), F1 Score, Receiver Operator Characteristic (ROC) Curve, AUC.	1 hour
5.2	Classification Performance Evaluation Metrics - Examples	1 hour
5.3	Regression Performance Evaluation Metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R Squared/Coefficient of Determination.	1 hour
5.4	Regression Performance Evaluation Metrics - Examples	1 hour
5.5	Clustering Performance Evaluation Metrics: Purity, Jaccard index, Normalized Mutual Information, Clustering Accuracy, Silhouette Coefficient, Dunn's Index	1 hour
5.6	Clustering Performance Evaluation Metrics - Examples	1 hour
5.7	Ensemble methods: bagging, boosting	1 hour
5.8	random forest, Stacking	1 hour

5.9	Boosting: AdaBoost	1 hour
5.9	Gradient boosting machines, XGBoost	1 hour
	Total	45 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1T103

Course Name: :ADVANCED MACHINE LEARNING

Max. Marks: 60 Duration: 3 hours

Arial***PART A***

Arial***Answer all questions. Each question carries 4 marks.***

1. Suppose that X is a discrete random variable with the following probability mass function: where $0 \leq \theta \leq 1$ is a parameter. The following 10 independent observations were taken from such a distribution: (3, 0, 2, 1, 3, 2, 1, 0, 2, 1). What is the maximum likelihood estimate of θ.

X	0	1	2	3
P(X)	$\frac{2\theta}{3}$	$\frac{\theta}{3}$	$\frac{2(1-\theta)}{3}$	$\frac{(1-\theta)}{3}$

2. In a two-class logistic regression model, the weight vector w = [4, 3, 2, 1, 0]. We apply it to some object that we would like to classify; the vectorized feature representation of this object is x = [2, 0,3, 0.5, 3]. What is the probability, according to the model, that this instance belongs to the positive class?

3. Expectation maximization (EM) is designed to find a maximum likelihood setting of the parameters of model when some of the data is missing. Does the algorithm converge? If so, do you obtain a locally or globally optimal set of parameters?

4. Suppose that you have a linear support vector machine(SVM) binary classifier. Consider a point that is currently classified correctly, and is far away from the decision boundary. If you remove the point from the training set, and re-train the classifier, will the decision boundary change or stay the same? Justify your answer.

5. Suppose there are three classifiers A,B and C. The (FPR, TPR) measures of the three classifiers are as follows – A (0, 1), B (1, 1) , C (1,0.5). Which can be considered as a perfect classifier? Justify your answer

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Explain the principle of the gradient descent algorithm (4 marks)

   (b) Suppose that the lifetime of an electrical device is modelled by an exponential distribution with (unknown) parameter λ. We test 5 devices and find they have lifetimes of 10, 8, 7, 9, and 4 years, respectively. What is the MLE for λ? The probability mass function of exponential distribution $f(x) = \lambda e^{-\lambda x}$. (4 marks)

7. Implement AND function with bipolar input and target using perceptron (8 marks)

8. Apply the DBSCAN algorithm on where minimum point = 3 and epsilon = 1.5. (8 marks)

9. Describe Principal Component Analysis. What criterion does the method minimize? What is the objective of the method? Give a way to compute the solution from a matrix X encoding the features (8 marks)

10. What is ensemble learning? Can ensemble learning using linear classifiers learn classification of linearly non-separable sets? (8 marks)

11. How does random forest classifier work? Why is a random forest better than a decision tree? (8 marks)

Point	X	Y
P1	3	7
P2	4	6
P3	5	5
P4	6	4
P5	7	3
P6	6	2

12. Consider a two-class classification problem of predicting whether a photograph contains a man or a woman. Suppose we have a test dataset of 10 records with expected outcomes and a set of predictions from our classification algorithm. Compute the confusion matrix, accuracy, precision, recall, sensitivity and specificity on the following data

SI NO	Actual	Predicted
1	man	Woman
2	man	man
3	woman	woman
4	man	man
5	man	woman
6	woman	woman
7	woman	man
8	man	man
9	man	woman
10	woman	woman

(8 marks)

SEMESTER I

PROGRAM ELECTIVE I

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E104A	**OBJECT ORIENTED SOFTWARE ENGINEERING**	ELECTIVE	3	0	0	3	3

Preamble : This course will emphasize on the systematic application of scientific and technological knowledge, methods, and experience to the design, implementation testing, and documentation of software.

Prerequisites :

Fundamental programming and software engineering concepts.

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Make use of project organization and management concepts and analyse the various tasks carried out. **(Cognitive Level: Apply)**
CO 2	Identify and select suitable process model for a given problem.**(Cognitive Level: Apply)**
CO 3	Analyse the requirements of a given software project and examine the various designing principles and patterns of a software product. **(Cognitive Level: Analyse)**
CO 4	Build the mapping of product design to code, its testing and maintenance. **(Cognitive Level: Apply)**
CO 5	Design, analyse object models and dynamic models for a given problem statement. **(Cognitive Level: Create)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓		✓
CO 2	✓		✓	✓	✓	
CO 3	✓		✓	✓	✓	
CO 4	✓		✓	✓	✓	
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand			
Apply	40	40	25
Analyse	40	40	25
Evaluate			
Create	20	20	10

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Classical Paradigm: System Design Concepts – Project Organization Concepts: Project Organizations, Roles, Tasks and Work Products, Schedule – Project Communication concepts: Planned Communication, Unplanned Communication, Communication Mechanism.

MODULE 2 (8 hours)

Process Models: Life cycle models: Sequential Activity Centered Models, Iterative Activity Centered models, Entity Centered models – Unified Process – Iterative and Incremental – Workflow – Agile Processes.

MODULE 3 (7 hours)

Analysis: Requirements Elicitation Concepts – An Overview of Unified Modeling Language – Analysis Concepts: Analysis Object Model and Analysis Dynamic Models – Non-functional requirements – Analysis Patterns – Executable specification.

MODULE 4 (7 hours)

Design: System Design, Architecture – Design Principles – Design Patterns – Dynamic Object Modeling Static Object Modeling – Model based approach vs Document based approach – Interface Specification – Object Constraint Language.

MODULE 5 (7 hours)

Implementation, Deployment And Maintenance: Mapping Design (Models) to Code – Testing – Usability – Deployment – Configuration Management – Maintenance.

Text Books

1. Bernd Bruegge, Alan H Dutoit, Object-Oriented Software Engineering, 2nd edition, Pearson Education, 2004.

Reference Books

1. Craig Larman, Applying UML and Patterns 3rd edition, Pearson Education, 2005
2. Stephen Schach, Software Engineering 7th ed, McGraw-Hill, 2007.
3. Ivar Jacobson, Grady Booch, James Rumbaugh, The Unified Software Development Process, Pearson Education, 1999.
4. Alistair Cockburn, Agile Software Development 2nd ed, Pearson Education.

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1: Classical Paradigm**	**7 hours**
1.1	System Design Concepts	2
1.2	Project Organization Concepts- Project Organizations	1
1.3	Roles, Tasks, Work Products and Schedule	1
1.4	Project Communication concepts	1
1.5	Planned Communication , Unplanned Communication	1
1.6	Communication Mechanism	1
1.7	Communication Mechanism	1
2	**Module 2: Process Models**	**8 hours**
2.1	Life cycle models	1
2.2	Sequential Activity Centered Models	1
2.3	Iterative Activity Centered models	1
2.4	Entity Centered models	1
2.5	Unified Process	1
2.6	Iterative and Incremental	1
2.7	Workflow	1

2.8	Agile Processes	1
3	**Module 3: Analysis**	**7 hours**
3.1	Requirements Elicitation Concepts	1
3.2	An Overview of Unified Modeling Language	1
3.3	Analysis Concepts	1
3.4	Analysis Object Model and Analysis Dynamic Models	1
3.5	Non-functional requirements	1
3.6	Analysis Patterns	1
3.7	Executable specification	1
4	**Module 4: Design**	**7 hours**
4.1	System Design, Architecture	1
4.2	Design Principles	1
4.3	Design Patterns	1
4.4	Dynamic Object Modeling	1
4.5	Static Object Modeling	1
4.6	Model Based approach vs Document based approach	1
4.7	Interface Specification	1
5	**Module 5: Implementation, Deployment and Maintenance**	**7 hours**
5.1	Mapping Design (Models) to Code	1

5.2	Mapping Design (Models) to Code(Continued)	1
5.3	Testing	1
5.4	Usability	1
5.5	Deployment	1
5.6	Configuration Management	1
5.7	Maintenance	1
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: M24CS1E104A

Course Name: Object Oriented Software Engineering

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Being the member of the design team to develop an interface for an online registration portal, you are not sure about the mandatory fields. People in what role can help you out. Whether planned or unplanned communication will be more beneficial in this situation. Why?

2. For what type of project, the spiral model suit's best. Why?

3. Describe the different Requirement Elicitation techniques

4. Discuss the design Principles of System Design

5. Demonstrate the different steps of software deployment

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Explain Work Breakdown Structure. (4 marks)

   (b) What are the types of project organizations? (4 marks)

7. (a) Explain Iterative Activity Centered Models. (4 marks)

   (b) Explain Agile process? (4 marks)

8. (a) Problem statement: Student Attendance Monitoring System. Student Attendance Monitoring System is used to track the attendance of students in an Institute. Faculty advisor can add students into the system, which is verified and approved by HOD. Once the students list is approved, the teachers can mark attendance on the system. Students can apply for duty leaves to the faculty advisor. Faculty advisor forwards the application to the HOD for approval. Students, teachers, faculty advisor and HOD can view the attendance reports of every student. Draw the sequence diagram for this problem statement. (8 marks)

9. (a) Identify any four functional and non-functional requirements of KTU website. (4 marks)

   (b) Compare the Dynamic Object Modelling with the Static Object Modelling. (4 marks)

10. (a) Design patterns speed up the development process quiet a lot. Illustrate with example. (4 marks)

    (b) Discuss about the Object Constraint Language? (4 marks)

11. (a) Discuss System Documentation. (4 marks)

    (b) What is skill matrix? Briefly describe the project management activities? (4 marks)

12. (a) Consider a method that will return the fare of a transport bus, given the source, destination and number of passengers. The source and destination are specified as integers. 1 represents station A, 2 represents station B etc. The total number of seats is 30. Generate test cases for Unit testing the system. (4 marks)

    (b) List out the benefits of model transformation. (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E104B	**FUNDAMENTALS OF CRYPTOGRAPHY**	ELECTIVE	3	0	0	3	3

Preamble : The course on Fundamentals of Cryptography aims at exploring various algorithms deployed in offering confidentiality, integrity, authentication and non-repudiation services. This course covers classical encryption techniques, symmetric and public key crypto-system, key exchange and management, and authentication functions. The concepts covered in this course enable the learners in effective use of cryptographic algorithms for real life applications.

Prerequisites : A sound background in Number Theory.

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Identify the security services provided for different types of security attacks. **(Cognitive Knowledge Level : Understand)**
CO 2	Summarize the classical encryption techniques for information hiding. **(Cognitive Knowledge Level: Apply)**
CO 3	Illustrate symmetric / asymmetric key cryptographic algorithms for secure communication.**(Cognitive Knowledge Level: Apply)**
CO 4	Interpret key management techniques for secure communication.**(Cognitive Knowledge Level: Understand)**
CO 5	Summarize message authentication functions in a secure communication scenario.**(Cognitive Knowledge Level: Understand)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓	✓				✓
CO 2	✓	✓	✓	✓	✓	✓
CO 3	✓	✓	✓	✓	✓	✓
CO 4	✓	✓				✓
CO 5	✓	✓				✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember	30	30	30
Understand	30	30	30
Apply	40	40	40
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Introduction to the Concepts of Security: Need for security, Security approaches, Principles of security, Types of attacks, OSI Security Architecture, Classical encryption techniques - Substitution techniques, Transposition techniques. Stream cipher, Block cipher, Public key cryptosystems vs. Symmetric key cryptosystems, Encrypting communication channels.

MODULE 2 (9 hours)

Symmetric Key Cryptosystems: Overview of symmetric key cryptography, Block cipher principles, Data Encryption Standard (DES), Differential and Linear cryptanalysis, Double DES, Triple DES, International Data Encryption Algorithm (IDEA), Advanced Encryption Algorithm (AES), Block cipher modes of operation, Stream cipher, RC4.

MODULE 3 (6 hours)

Public Key Cryptosystems: Principles of public key cryptosystems, RSA algorithm, RSA illustration, Attacks, ElGamal cryptographic system, Knapsack algorithm, Diffie-Hellman key exchange algorithm, Elliptical curve cryptosystems.

MODULE 4 (7 hours)

Key Management: Symmetric key distribution using symmetric encryption, Symmetric key distribution using asymmetric encryption, Distribution of public keys, Generating keys, transferring keys, Verifying keys, Updating keys, Storing keys, Backup keys, Compromised keys, Public key infrastructure.

MODULE 5 (7 hours)

Authentication: Authentication requirements, Authentication functions, Message authentication codes (MAC), Hash functions, Security of Hash functions and MAC, Message Digest 5 (MD5), Secure Hash Algorithm (SHA)-512, Hash-based Message Authentication Code (HMAC), Cipher-based Message Authentication Code (CMAC), X.509 Authentication services.

Text Books

1. William Stallings, Cryptography and Network Security Principles and Practice, Pearson Edu, 6e.

2. Bruice Schneier, Applied Cryptography Protocols, Algorithms and source code in C, Wiley,2e.

Reference Books

1. Behrouz A. Forouzan, Cryptography and Network Security, McGraw Hill, 2e.
2. Johannes A. Buchmann, Introduction to Cryptography, Springer, 2e.
3. DouglasR. Stinson, Cryptography Theory and Practice, 3e,Chapman & Hall/CRC, 2006.
4. Bernard Menezes, Network Security and Cryptography, Cengage Learning, 2011.

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	7 hours
1.1	Need for security, Security approaches	1 hour
1.2	Principles of security, Types of attacks	1 hour
1.3	OSI Security Architecture	1 hour
1.4	Classical encryption techniques: Substitution techniques	1 hour
1.5	Transposition techniques	1 hour
1.6	Stream cipher, Block cipher	1 hour
1.7	Public- key cryptosystems vs. Symmetric key cryptosystems, Encrypting communication channels	1 hour
2	**Module 2**	9 hours

2.1	Overview of symmetric key cryptography	1 hour
2.2	Block cipher principles	1 hour
2.3	Data Encryption Standard (DES)	1 hour
2.4	Differential and Linear cryptanalysis	1 hour
2.5	Double DES, Triple DES	1 hour
2.6	IDEA	1 hour
2.7	Advanced Encryption Algorithm (AES structure)	1 hour
2.8	Block cipher modes of operation	1 hour
2.9	Stream cipher, RC4	1 hour
3	**Module 3**	6 hours
3.1	Principles of public key cryptosystems	1 hour
3.2	RSA algorithm, RSA illustration, Attacks	1 hour
3.3	ElGamal cryptographic system	1 hour
3.4	Knapsack algorithm	1 hour
3.5	Diffie-Hellman key exchange algorithm	1 hour
3.6	Elliptical curve cryptosystems(Elliptical curve arithmetic)	1 hour
4	**Module 4**	7 hours
4.1	Symmetric key distribution using symmetric encryption	1 hour
4.2	Symmetric key distribution using asymmetric encryption	1 hour

4.3	Distribution of public keys	1 hour
4.4	Generating keys, Transferring keys	1 hour
4.5	Verifying keys, Updating keys	1 hour
4.6	Storing keys, Backup keys, Compromised keys	1 hour
4.7	Public key infrastructure	1 hour
5	**Module 5**	7 hours
5.1	Authentication requirements, Authentication functions	1 hour
5.2	Message Authentication Codes (MAC)	1 hour
5.3	Hash functions Security of Hash functions and MAC	1 hour
5.4	MD5	1 hours
5.5	SHA-512	1 hour
5.6	HMAC, CMAC	1 hour
5.7	X.509 Authentication services	1 hour
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E104B

Course Name: FUNDAMENTALS OF CRYPTOGRAPHY

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Define Substitution Cipher. Encrypt using one time pad M = HONORS and K = CIPHER.
2. Differentiate between diffusion and confusion.
3. Perform encryption using RSA Algorithm for the following p=7; q=11; e=13; M=5
4. Define a certificate authority and its relation to public key cryptography.
5. Distinguish between integrity and message authentication.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) With a neat sketch, Explain OSI Security architecture model. (4 marks)

   (b) How does link encryption differ from end-to-end encryption? Explain. (4 marks)

7. (a) Explain the block cipher modes i) Cipher feedback mode ii) Output feedback mode. (4 marks)

   (b) Describe the four types of transformations in AES. (4 marks)

8. (a) Define a session key and show how a KDC can create can create a session key between Alice and Bob. (4 marks)

   (b) What are the requirements for the use of a public-key certificate scheme? (4 marks)

9. (a) Specify the format for X.509 certificate. Explain the steps required to obtain user's certificate.

   (4 marks)

   (b) With suitable block diagrams, explain the types of functions that may be used to produce an authenticator. (4 marks)

10. (a) Describe how SHA-512 logic produce message digest (4 marks)

    (b) Distinguish between HMAC and CMAC. (4 marks)

11. (a) What are the core components of a PKI? Briefly describe each component.

    (4 marks)

    (b) Describe the following (i) Updating keys (ii) Compromised Keys

    (4 marks)

12. (a) User A and B use the Diffie-Hellman key exchange technique with a common prime q=71 and primitive root alpha=7.
    If user A has private key XA =3, What is A's public key YA?

    (4 marks)

    (b) If user B has private key XB =6, What is A's public key YB? (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E104C	**ADVANCED DATA MINING**	ELECTIVE	3	0	0	3	3

Preamble : This course provides exposure to the concepts, principles and techniques of data mining. This course will enable the learners to identify the key process of Data mining and Warehousing, apply appropriate techniques to convert raw data into suitable format for practical data mining tasks, apply various data mining algorithms in appropriate domain, analyze the performance using performance metrics and extend data mining methods to the new domains of data.

Prerequisites :
Basic knowledge in Data Mining

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Summarize basic concepts of Data mining and Illustrate feature vector representation for a given data collection **(Cognitive Knowledge Level: Understand)**
CO 2	Design Data Warehouse for problems in various domains. **(Cognitive Knowledge Level: Apply)**
CO 3	Implement Association Rules for analyzing Transactional databases **(Cognitive Knowledge Level: Apply)**
CO 4	Implement major Classification And Clustering Algorithms to a given problem. **(Cognitive Knowledge Level: Analyze)**
CO 5	To develop Data Mining system and analyze the performance **(Cognitive Knowledge Level: Create)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓			✓
CO 2	✓		✓			✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand			
Apply	60-80%	60-80%	60-80%
Analyse	20-40%	20-40%	20-40%
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (6 hours)

Data Mining and Knowledge Discovery: Desirable Properties of Discovered Knowledge – Knowledge representation, Data Mining Functionalities, Motivation and Importance of Data Mining, Classification of Data Mining Systems, Integration of a Data Mining System with a Database or Data Warehouse System, Classification, Clustering, Regression, Data Pre-processing: Data Cleaning, Data Integration and Transformation, normalization, standardization, Data Reduction, Feature vector representation. importance of feature engineering in machine learning; forward selection and backward selection for feature selection; curse of dimensionality; data imputation techniques; No Free Lunch theorem in the context of machine learning, Data Discretization and Concept Hierarchy Generation

MODULE 2 (7 hours)

Data Warehouse and OLAP Technology for Data Mining : Data warehouses and its Characteristics - Data warehouse Architecture and its Components, Data Warehouse Design Process, Data Warehouse and DBMS, Data marts, Metadata, Data Cube and OLAP, Extraction - Transformation – Loading - Schemas for Multidimensional Database: Stars, Snowflakes and Fact constellations, OLAP Cube - OLAP Operations - OLAP Server Architecture - Data Warehouse Implementation - From Data Warehousing to Data Mining, Trends in data warehousing

MODULE 3 (7 hours)

Association Pattern Mining : Mining Frequent Patterns, Associations and Correlations –Mining Methods – Mining Various Kinds of Association Rules – Correlation Analysis – Constraint Based Association Mining, Single Dimensional Boolean Association Rules From Transaction Databases, Multilevel Association Rules from transaction databases – Multi dimension Association Rules from Relational Database and Data Warehouses, Frequent Item Set Generation, Apriori Algorithm, Improved Apriori Algorithm for Association Rules Mining, Methods to improve Apriori, FP Growth Algorithm - Generating association rules from frequent itemset, Compact Representation of Frequent Item set - Maximal Frequent Item Set - Closed Frequent Item Sets. Pattern Evaluation Methods- Relationship Between FP-Growth and Enumeration-Tree Methods From Association Analysis to Correlation

MODULE 4 (8 hours)

Classification and Prediction : Classification Techniques, Decision Tree - Decision tree Construction, Measures for Selecting the Best Split - Algorithm for Decision tree Induction - CART, Bayesian Belief Networks, Instance-Based Learning, K-Nearest neighbor classification, Accuracy and Error measures, Multiclass Classification, Semi-Supervised Classification, Multi class Learning, Rare class learning, Active Learning, Transfer Learning, Fuzzy Set Approaches for Classification, Rough Set Approaches, Techniques to improve classification accuracy-Ensemble methods, Bias-Variance Trade-off, Improving classification accuracy of class imbalanced data

MODULE 5 (8 hours)

Cluster Analysis : Desired features of cluster Analysis, Types of data in cluster analysis, Categorization of Major Clustering Methods, Density-Based Methods, Clustering High Dimensional Data, Constraint Based Cluster Analysis, GA based clustering, Dealing with Large Databases, Probabilistic Model Based Clustering, Clustering with Constraints, Semi-supervised clustering, Cluster Ensembles, Quality and validity of cluster analysis methods, Outlier Analysis-Statistical Approaches, Proximity Based Approaches Advanced Mining: Multimedia Data Mining - Text Mining, Graph Mining and Social Network Analytics - Geospatial Data Mining, Temporal Mining, Data Mining Applications - Social Impacts of Data Mining.

Text Books

1. Kevin Murphy, Machine Learning: A Probabilistic Perspective (MLAPP), MIT Press, 2012
2. Jiawei Han, Micheline Kamber, Jian Pei, “Data Mining: Concepts and Techniques”, Morgan Kaufmann, 2nd Ed., 2005

Reference Books

1. Christopher Bishop, Pattern Recognition and Machine Learning (PRML), Springer, 2007.
2. Charu C. Aggarwal, Data Mining, Springer, ISBN 978-3-319-14141-1,2015
3. Data Mining Techniques, Arun K Puari, Universities Press,2001
4. Margaret H. Dunham, “Data Mining: Introductory and Advanced Topics”, Prentice Hall, 1st Ed., 2002.
5. David G. Stork, Peter E. Hart, and Richard O. Duda. Pattern Classification (PC), Wiley-Blackwell, 2000
6. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning (ESL), Springer, 2009
7. G. K. Gupta “Introduction to Data Mining with Case Studies”, Eastern Economy Edition, Prentice Hall of India, 2006.
8. Soumen Chakrabarti, "Mining the Web: Discovering Knowledge from Hypertext Data", Morghan Kaufmann, 1st Ed., 2005.

9. Da Ruan, Guoqing Chen, Etienne E. Kerre, Geert Wets, "Intelligent Data Mining: Techniques and Applications (Studies in Computational Intelligence)", Springer, 1st Ed., 2010.

10. Masoud Mohammadian, "Intelligent Agents for Data Mining and Information Retrieval", Idea Group Publishing, 2004.

11. I. H. Witten and E. Frank. Data Mining: Practical Machine Learning Tools and Techniques Morgan Kaufmann, 2000.

12. D. Hand, H. Mannila and P. Smyth. Principles of Data Mining. Prentice-Hall. 2001

13. Valliappa Lakshmanan, Jordan Tigani, Google Big Query: The Definitive Guide: Data Warehousing, Analytics, and Machine Learning, O'Reilly Media, Inc.", 2019

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	6 hours
1.1	Data Mining Functionalities, Motivation and Importance of Data Mining	1 hour
1.2	Integration of a Data Mining System with a Database or Data Warehouse System, Major Issues in Data Mining. Classification, Clustering, Regression	1 hour
1.3	Data Pre-processing: Data Cleaning, Data Integration and Transformation, normalization	1 hour
1.4	Data Reduction, Different techniques	1 hour

1.5	Feature vector representation. importance of feature engineering in machine learning;	1 hour
1.6	Forward selection and backward selection for feature selection;	1 hour
2	**Module 2**	7 hours
2.1	Data warehouses and its Characteristics - Data warehouse Architecture and its Components	1 hour
2.2	Data Warehouse and DBMS, Data marts, Metadata Extraction - Transformation – Loading in DW,	1 hour
2.3	Multidimensional model	1 hour
2.4	Schemas for Multidimensional Database: Stars, Snowflakes Fact constellations	1 hour
2.5	Design Data Warehouse for problems in different domains	1 hour
2.6	OLAP Cube - OLAP Operations	1 hour
2.7	OLAP Server Architecture - Data Warehouse Implementation	1 hour
3	**Module 3**	7 hours
3.1	Mining Frequent Patterns, Associations and Correlations	1 hour
3.2	Mining Various Kinds of Association Rules – Correlation Analysis – Constraint Based Association Mining	1 hour
3.3	Multilevel Association Rules from transaction databases – Multi dimension Association Rules from Relational Database and Data Warehouses	1 hour

3.4	Frequent Item Set Generation, Apriori Algorithm, Apriori Algorithm-illustration with example	1 hour
3.5	Methods to improve Apriori, FP Growth Algorithm	1 hour
3.6	FP Growth Algorithm- illustration with example, Compact Representation of Frequent Item set	1 hour
3.7	Pattern Evaluation Methods, Association Analysis to Correlation Analysis, Lift	1 hour
4	**Module 4**	8 hours
4.1	Classification Techniques, Decision Tree - Decision tree Construction Measures for Selecting the Best Split	1 hour
4.2	Decision tree Induction - illustration with example Algorithm for Decision tree Induction - CART	1 hour
4.3	Bayesian Belief Networks - Training	1 hour
4.4	K-Nearest neighbor classification, Accuracy and Error measures	1 hour
4.5	Multiclass Classification, Semi-Supervised Classification	1 hour
4.6	Active Learning, Transfer Learning	1 hour
4.7	Fuzzy Set Approaches for Classification, Rough Set Approaches	1 hour
4.8	Ensemble methods. Improving classification accuracy of class imbalanced data	1 hour
5	**Module 5**	8 hours
5.1	Desired features of cluster Analysis, Types of data in cluster analysis,	1 hour

5.2	Categorization of Major Clustering Methods, Density-Based Methods,	1 hour
5.3	Semi supervised clustering, Clustering High Dimensional Data, Constraint Based Cluster Analysis,	1 hour
5.4	GA based clustering, Probabilistic Model Based Clustering	1 hour
5.5	Quality and validity of cluster analysis methods, Outlier Analysis-Statistical Approaches, Proximity Based Approaches	1 hour
5.6	Multimedia Data Mining, Text Mining	1 hours
5.7	Graph Mining and Social Network Analytics	1 hour
5.8	Geospatial Data Mining, Temporal Mining	1 hour
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E104C

Course Name: ADVANCED DATA MINING

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Differentiate between classification and regression with example
2. Explain concept hierarchy generation. With a suitable example show how is it done for categorical data.
3. How can you generate association rules from frequent item sets?
4. Why are nearest neighbor algorithms called lazy learners? What are the disadvantages of a lazy learner?
5. How do we relate text mining and web mining? Differentiate between spatial and non spatial data with example

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Why feature engineering is important? What is the output of feature engineering in machine learning? (4 marks)

   (b) Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70.
   i) Use smoothing by bin means to smooth the above data, using a bin depth of 3. Illustrate your steps.
   ii) How might you determine outliers in the data (4 marks)

7. (a) How do data warehousing relate to data mining? Discuss. (4 marks)

   (b) Suppose that a data warehouse consists of the three dimensions time, doctor, and patient, and the two measures count and charge, where charge is the fee that a doctor charges a patient for a visit.
   i) List three classes of schemas that are popularly used for modeling data warehouses.
   ii) Draw a schema diagram for the above data warehouse using one of the schema classes listed in(i). Starting with the base cuboid [day, doctor, patient], what a specific OLAP operations should be performed in order to list the total fee collected by each doctor in 2022? (4 marks)

8. (a) Why is the FP growth algorithm so efficient? (4 marks)

   (b) Discuss FP growth algorithm. Using Apriori and FP growth algorithm find the frequent itemsets from the following transactional database? (min_sup$\bar{2}$, confidence 70%). Compare the two processesing order to list the total fee collected by each doctor in 2022?

TID	List of item-IDs
T100	I1, I2, I3
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5

(4 marks)

9. (a) What is the metric for classification tasks in CART? How to use the CART algorithm for classification?

   (4 marks)

   (b) Differentiate between different types of ensemble methods for classification with example. (4 marks)

10. (a) How is the parameter "Distance-function" estimated in the DBSCAN Algorithm? What are the advantages and disadvantages of DBSCAN algorithm?

    (4 marks)

    (b) What is the purpose of cluster ensemble? How do you create a cluster ensemble? Discuss with example. (4 marks)

11. (a) Apply the Apriori algorithm for discovering frequent itemsets from the following data set minimum support of 50% and minimum confidence of 75%.

    (8 marks)

Transaction ID	Items
100	Bread, Cheese
200	Bread, Cheese, Juice
300	Bread, Milk
400	Cheese, Juice, Milk

12. (a) Compare the R-tree to the R*-tree.

(4 marks)

(b) Discuss different spatial datamining primitives with example. (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E104D	**SOCIAL NETWORK ANALYSIS**	ELECTIVE	3	0	0	3	3

Preamble : This course provides an exposure to the concepts and techniques in Social Network Analysis. This course covers various types of modeling, visualization and mining techniques used in social networks. This course helps the learners to analyze social media data using appropriate data/web mining techniques.

Prerequisites :
Basic knowledge in Graph Theory.

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Understand the concepts and properties of social networks. **(Cognitive Knowledge Level: Understand)**
CO 2	Analyze the concepts of evolution and privacy in social networks. **(Cognitive Knowledge Level: Analyze)**
CO 3	Model and visualize social networks. **(Cognitive Knowledge Level: Apply)**
CO 4	Mine the behaviour of users in the Social Networks. **(Cognitive Knowledge Level: Analyze)**
CO 5	Use Multimedia Information Networks in Social Media. **(Cognitive Knowledge Level: Analyze)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓					✓
CO 2	✓		✓		✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	30	30	30
Apply	40	40	40
Analyse	30	30	30
Evaluate			
Create			

Mark Distribution

Continuous Internal Evaluation Pattern

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Seminar* : 10 marks

Course based task/Micro Project

Data collection and interpretation/Case study : 10 marks

Test paper 1 (Module 1 and Module 2) : 10 marks

Test paper 2 (Module 3 and Module 4) : 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (8 hours)

Introduction: Introduction to Social Network Data Analytics: Introduction, Online Social Networks: Research Issues, Research Topics in Social Networks. Statistical Properties of Social Networks: Preliminaries, Static Properties, Dynamic Properties. Random Walks in Social Networks and their Applications: Random Walks on Graphs: Background, Application in Computer Vision, Text Analysis, Collaborative Filtering.

MODULE 2 (8 hours)

Evolution in Social Networks: Evolution in Social Networks: Framework, Challenges of Social Network Streams, Incremental Mining for Community Tracing, Tracing Smoothly Evolving Communities. Models and Algorithms for Social Influence Analysis: Influence Related Statistics, Social Similarity and Influence. Privacy in Social Networks: Privacy breaches in social networks, Privacy-preserving mechanisms.

MODULE 3 (8 hours)

Visualizing Social Networks: Visualizing Social Networks: A Taxonomy of Visualizing Social Networks: A Taxonomy of Visualizations. Data Mining in Social Media: Methods for Social Media, Ethnography and Netnography, Event Maps. Text Mining in Social Networks: Keyword Search, Classification and Clustering Algorithms, Transfer Learning in Heterogeneous Networks.

MODULE 4 (8 hours)

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.

MODULE 5 (8 hours)

Multimedia Information Networks in Social Media: Multimedia Information Networks in Social Media: Links from Semantics, Links from Community Media. Network of Personal Photo Albums, Geographical Information, Inference Methods. Social Tagging and Applications: Tags: Why What, Tagging System Design, Tag analysis, Visualization of Tags, Applications of Tags.

Text Books

1. Charu C. Aggarwal, "Social Network Data Analytics", Springer.

Reference Books

1. Peter Mika, "Social Networks and the Semantic Web", Springer, 1St edition 2007.

2. BorkoFurht, "Handbook of Social Network Technologies and Applications", Springer, 1st edition, 2010.

3. GuandongXu, Yanchun Zhang and Lin Li, "Web Mining and Social Networking Techniques and applications", Springer, 1St edition, 2011.

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	
1.1	Introduction to Social Network Data Analytics	1
1.2	Online Social Networks: Research Issues	1
1.3	Research Topics in Social Networks	1
1.4	Statistical Properties of Social Networks: Preliminaries, Static Properties, Dynamic Properties	1
1.5	Random Walks in Social Networks and their Applications	1
1.6	Random Walks on Graphs: Background	1
1.7	Application in Computer Vision	1
1.8	Text Analysis, Collaborative Filtering	1
2	**Module 2**	
2.1	Framework, Challenges of Social Network Streams	1
2.2	Incremental Mining for Community Tracing	1
2.3	Tracing Smoothly Evolving Communities	1

2.4	Models and Algorithms for Social Influence Analysis: Influence Related Statistics	1
2.5	Social Similarity and Influence	1
2.6	Privacy in Social Networks: Privacy breaches in social networks	1
2.7	Privacy-preserving mechanisms	1
3	**Module 3**	
3.1	A Taxonomy of Visualizing Social Networks	1
3.2	A Taxonomy of Visualizations	1
3.3	Data Mining in Social Media: Methods for Social Media	1
3.4	Ethnography and Netnography, Event Maps	1
3.5	Text Mining in Social Networks: Keyword Search	1
3.6	Classification and Clustering Algorithms	1
3.7	Transfer Learning in Heterogeneous Networks	1
4	**Module 4**	
4.1	Aggregating and reasoning with social network data	1
4.2	Advanced Representations - Extracting evolution of Web Community from a Series of Web Archive	1
4.3	Detecting Communities in Social Networks	1
4.4	Evaluating Communities	1
4.5	Core Methods for Community Detection & Mining	1

4.6	Core Methods for Community Detection & Mining	1
4.7	Applications of Community Mining Algorithms	1
4.8	Node Classification in Social Networks	1
5	**Module 5**	
5.1	Links from Semantics, Links from Community Media	1
5.2	Network of Personal Photo Albums	1
5.3	Geographical Information, Inference Methods	1
5.4	Social Tagging and Applications: Tags: Why What	1
5.5	Tagging System Design, Tag analysis	1
5.6	Visualization of Tags, Applications of Tags	1
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: M24CS1E104D

Course Name: SOCIAL NETWORK ANALYSIS

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. What are the static properties of unweighted graphs? (4 marks)
2. Briefly describe various Node-Edge Diagrams. (4 marks)
3. Explain the techniques for keyword search over graph data. (4 marks)
4. What are the applications of community mining algorithms? (4 marks)
5. Differentiate between Categorizers and Describers. (4 marks)

PART B

Answer any five questions. Each question carries 8 marks.

6. (a) Explain the properties of dynamic weighted and unweighted graphs. (4 marks)
   (b) Describe privacy breaches and privacy preserving mechanisms in social networks. (4 marks)
7. Describe privacy breaches and privacy preserving mechanisms in social networks. (8 marks)
8. Classify the visualization of social networks. (8 marks)
9. (a) Given a social network with labels on some nodes, how to provide a high-quality labelling for every node? (4 marks)
   (b) What are the types of changes in the evolution of the web community? (4 marks)
10. (a) Describe the different kinds of tags. (4 marks)
    (b) Explain the applications of tags. (4 marks)
11. Describe Ethnography, Netnography and event maps. (8 marks)
12. (a) Multimedia information networks can be viewed as a marriage of multimedia content and social networks. Justify. (4 marks)
    (b) Explain ontology based learning. (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E104E	**COMPUTATIONAL BIOLOGY**	ELECTIVE	3	0	0	3	3

Preamble : It covers Bio-molecules- DNA, RNA, Protein and amino acids, RNA processing , Central Dogma Bioinformatics databases, Secondary nucleotide sequence databases. Nature and scope of Computational ,Basic algorithms in Computational Biology has covered in this course. computational biology helps to sequence the human genome, it also helps to create accurate models of the human brain, map the 3D structure of genomes, and model biological systems.

Prerequisites :
Bioinformatics, Computational Intelligence

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Identify Chromosome-Genome-Genes-Databases and RNA processing **(Cognitive Knowledge Level : Understand)**
CO 2	Summarize the Nature and scope of Computational ,Basic algorithms in Computational Biology, and sequence alignment . **(Cognitive Knowledge Level: Apply)**
CO 3	Illustrate Sequence Representation& Analysis.**(Cognitive Knowledge Level: Apply)**
CO 4	Interpret Sequence alignment: Pair-wise sequence alignment.**(Cognitive Knowledge Level: Apply)**
CO 5	Summarize message authentication functions in a secure communication scenario.**(Cognitive Knowledge Level: Understand)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓	✓	✓			
CO 2	✓	✓	✓		✓	✓
CO 3	✓	✓	✓	✓	✓	✓
CO 4	✓	✓	✓	✓	✓	✓
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember	30	30	30
Understand	30	30	30
Apply	40	40	40
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Chromosome-Genome-Genes-Databases: Bio-molecules- DNA, RNA, Protein and amino acids, Chargaff's Rules, Codon bias, GC content. Central Dogma: Replication, Transcription, Translation, Post transcriptional & post translational modifications, RNA processing, RNA splicing and RNA editing. Sense/coding and anti-sense/template strands, Genetic code, wobble hypothesis. Introduction to DNA and Protein sequencing, Human Genome Project

MODULE 2 (7 hours)

Introduction to Computational Biology : Nature and scope of Computational Biology ,Basic algorithms in Computational Biology, Introduction to sequence alignment - Local and global, pair wise and multiple, BLAST. Basic file formats: FASTA, GenBank, EMBL, GCG, PIR, Phylip, Nexus file formats etc. Sequence Data Bases, detailed study of GenBank of NCBI- typical Gen Bank (DDBJ+EMBL) for DNA and RNA. , Bioinformatics databases, Type of databases, Nucleotide sequence databases, Primary nucleotide sequence databases,Secondary nucleotide sequence databases.

MODULE 3 (6 hours)

Sequence Representation & Analysis : Basic gene statistics–base counts, word, frequencies, , vector contamination analysis, experiments using Perl scripts, gene finding, splice site recognition, transcription factor binding site identification, SNPs, microsatellite, minisatellite, sequence profiles, sequence logos, sequence chromatograms.

MODULE 4 (8 hours)

Sequence alignment: Pair-wise sequence alignment, Need of Scoring schemes- Penalizing gaps – Linear and Affine gap penalty; Effect of scoring schemes, Scoring matrices for amino acid sequence alignment, PAM Probability matrix- Log odds matrix; BLOSUM; Dot-plot visualization; Smith –Waterman algorithm for local alignment, Needleman-Wunsch algorithm, Statistics of Sequence alignment score: E- values, bit scores and sensitivity, specificity; BLAST and FASTA.

MODULE 5 (7 hours)

Multiple sequence alignment : Need for MSA, SP measure- n dimensional dynamic programming- Heuristics algorithm for multiple sequence alignment - Progressive alignment, Iterative alignment - Tools for local, global and MSA: Muscle, T-Coffee, and ClustalW. Transcriptomics: Concept of Transcriptome, transcriptome analysis and Gene Expression-An Overview-introduction to microarrays; Types of non-coding RNA's- lncRNAs, miRNAs, piRNAs, siRNAs ceRNAs .

Text Books

1. Bioinformatics: Sequence and Genome Analysis by Mount D., Cold Spring Harbor

Laboratory Press, New York. 2004

2. Bioinformatics- a Practical Guide to the Analysis of Genes and Proteins by Baxevanis, A.D. and Francis Ouellellette, B.F., Wiley India Pvt Ltd. 2009
3. Introduction to Bioinformatics by Teresa K. Attwood, David J. Parry-Smith. Pearson

Reference Books

1. **Bioinformatics: Sequence and Genome Analysis** by David W. Mount
2. **Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids** by Richard Durbin, Sean R. Eddy, Anders Krogh, and Graeme Mitchison
3. **Computational Biology: An Introduction** by R. D. M. Page and E. C. Holmes

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	8 hours
1.1	Bio-molecules- DNA, RNA, Protein and amino acids	1 hour
1.2	Chargaff's Rules, Codon bias, GC content.	1 hour
1.3	Central Dogma: Replication, Transcrip- tion, Translation R	1 hour
1.4	Post transcriptional & post translational modifications, R	1 hour
1.5	RNA processing, RNA splicing and RNA editing.	1 hour

1.6	Sense/coding and anti-sense/template strands,	1 hour
1.7	Genetic code, wobble hypothesis	1 hour
1.8	Introduction to DNA and Protein sequencing, Human Genome Project	1 hour
2	**Module 2**	7 hours
2.1	Nature and scope of Computational Biology ,Basic algorithms in Computational Biology	1 hour
2.2	Introduction to sequence alignment - Local and global, pair wise and multiple, BLAST	1 hour
2.3	Basic file formats: FASTA, GenBank	1 hour
2.4	EMBL, GCG, PIR, Phylip, Nexus file formats etc.	1 hour
2.5	Sequence Data Bases, detailed study of Gen- Bank of NCBI-typical Gen Bank (DDBJ+EMBL) for DNA and RNA.	1 hour
2.6	ioinformatics databases, Type of databases, Nucleotide sequence databases	1 hour
2.7	Primary nucleotide sequence databases,Secondary nucleotide sequence databases.	1 hour
3	**Module 3**	7 hours
3.1	Basic gene statistics–base counts, word, fre- quencies,	1 hour
3.2	ector contamination analysis, experiments using Perl scripts,	1 hour
3.3	gene finding, splice site recognition	1 hour
3.4	transcription factor binding site identification, SNPs	1 hour

3.5	microsatellite, min- isatellite	1 hour
3.6	sequence profiles, sequence logos, sequence chromatograms.	1 hour
4	**Module 4**	8 hours
4.1	Pair-wise sequence alignment, Need of Scoring schemes	1 hour
4.2	Penalizing gaps – Linear and Affine gap penalty; Effect of scoring schemes	1 hour
4.3	Scoring matrices for amino acid sequence alignment	1 hour
4.4	PAM Probability matrix- Log odds matrix	1 hour
4.5	BLOSUM; Dot-plot vi- sualization	1 hour
4.6	Smith –Waterman algorithm for local alignment, Needleman-Wunsch algorithm	1 hour
4.7	Statistics of Sequence alignment score: E- values, bit scores and sensitivity, specificity; BLAST and FASTA.	1 hour
5	**Module 5**	8 hours
5.1	Need for MSA, SP measure- n dimensional dynamic programming	1 hour
5.2	Heuristics algorithm for multiple sequence alignment - Progressive alignment,	1 hour
5.3	Progressive alignment, Iterative alignmen	1 hour
5.4	Tools for local, global and MSA	1 hour
5.5	Muscle, T-Coffee, and ClustalW.	1 hour

5.6	Transcriptomics: Concept of Transcriptome, transcriptome analysis and Gene Expression-	1 hours
5.7	An Overview-introduction to microarrays;	1 hour
5.8	Types of non-coding RNA's- lncRNAs, miRNAs, piRNAs, siRNAs ceRNAs .	1 hour
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E104E

Course Name: COMPUTATIONAL BIOLOGY

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Explain how the wobble hypothesis contributes to the redundancy of the genetic code and its evolutionary significance.
2. Explain the Needleman-Wunsch algorithm and its application in bioinformatics.

3. Describe the process of splice site recognition and its importance in gene annotation.

4. Compare BLAST and FASTA in terms of their algorithms and applications in sequence searching.

5. What is the Sum-of-Pairs (SP) measure in MSA, and how is it used to evaluate alignment quality?

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Outline the steps of the central dogma of molecular biology, including replication, transcription, and translation. How do these processes ensure the flow of genetic information from DNA to protein? (4 marks)

   (b) Explain the genetic code and its characteristics. How does the wobble hypothesis contribute to the redundancy of the genetic code? (4 marks)

7. (a) Describe the role of basic algorithms in computational biology. Provide an example of a commonly used algorithm and its application. (4 marks)

   (b) Provide a detailed overview of the GenBank sequence database. How does it relate to the DDBJ and EMBL databases, and what are the typical contents of a GenBank record? (4 marks)

8. (a) Explain the role of sequence chromatograms in DNA sequencing. What challenges are involved in interpreting chromatograms, and how are these addressed computationally? (4 marks)

   (b) Compare and contrast microsatellites and minisatellites in terms of structure, function, and their use in genetic studies. How are these elements detected and analyzed computationally? (4 marks)

9. (a) Discuss the computational methods used for splice site recognition. What are the key features of splice sites, and how do algorithms identify these regions? (4 marks)

   (b) How are single nucleotide polymorphisms (SNPs) detected and analyzed in genomic studies? Discuss the impact of SNPs on gene function and their relevance in personalized medicine. (4 marks)

10. (a) Explain the key differences between global and local pair-wise sequence alignment. In what scenarios would you use the Needleman-Wunsch algorithm versus the Smith-Waterman algorithm?

    (4 marks)

    (b) What are linear and affine gap penalties, and how do they affect the outcome of sequence alignment? Provide an example of when each type of penalty might be preferable. (4 marks)

11. (a) Discuss the different approaches used by the MSA tools Muscle, T-Coffee, and ClustalW. How do these tools address the challenges of aligning multiple sequences?

    (4 marks)

    (b) Define the transcriptome and describe the importance of transcriptome analysis in understanding gene expression. How does transcriptome analysis differ from other genomic analyses? (4 marks)

12. (a) What is the BLOSUM scoring matrix, and how does it differ from the PAM matrix? Provide an example of when you would use BLOSUM over PAM.

    (4 marks)

    (b) Explain how dot-plot visualization is used in sequence alignment. What are its strengths and limitations? (4 marks)

SEMESTER I

PROGRAM ELECTIVE II

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E105A	**FORMAL METHODS IN SOFTWARE ENGINEERING**	ELECTIVE	3	0	0	3	3

Preamble : The purpose of this course is to introduce the necessary background, the basics of Formal Methods in Software Engineering. This course helps the learner to familiarize various aspects of Software Engineering

Prerequisites : Nil

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Explain software requirements using logical formalisms with an emphasis on Alloy. (Cognitive Knowledge Level: Understand)
CO 2	Apply algorithmic verification techniques using Spin to check the correctness of software designs and models against specified properties.(Cognitive Knowledge Level: Apply)
CO 3	Implement functional correctness of software components using Abstract Data Types (ADTs) (Cognitive Knowledge Level: Apply)
CO 4	Execute white-box testing applications using JPF focusing on internal structure and logic.(Cognitive Knowledge Level: Apply)
CO 5	Illustrate grey-box testing using AFL combining both white-box and black-box testing strategies.(Cognitive Knowledge Level: Understand)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1			✓	✓	✓	
CO 2			✓	✓	✓	
CO 3			✓	✓	✓	
CO 4			✓		✓	
CO 5			✓		✓	

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	30	30	30
Apply	30	30	30
Analyse	40	40	40
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Introduction to Formal Methods in Software Engineering, Importance of Formal Specification and Verification, Overview of Conceptual Modeling Techniques, Tool: Introduction to Alloy for formal specification,Logic-Based Modeling, Conceptual modeling of entities and relationships using Alloy,Constraints and Invariants in Alloy

MODULE 2 (7 hours)

Introduction to Algorithmic Verification, Fundamentals of Model-Checking,Temporal Logic and its Application in Model Checking, Tool: Introduction to Model-Checking using Spin, The Promela Modeling Language,Specifying Properties for Verification,Optimizing Model Checking with Spin

MODULE 3 (8 hours)

Introduction to Abstract Data Types (ADTs),Importance of ADTs in Software Design, Refinement of ADTs and Program Correctness, Tool: Rodin for ADT Specification and Verification, Introduction to Hoare Logic and its Application, Tool: VCC for Hoare Logic Assertions and Program Refinement,Combining ADTs with Formal Methods,Case Studies and Practical Applications

MODULE 4 (7 hours)

Introduction to White-Box Testing, Principles of White-Box Testing, Test Coverage Criteria: Path, Branch, and Condition Coverage, Tool: JPF for Automated White-Box Testing,Overview of Java Path Finder (JPF),Model Checking and Its Role in White-Box Testing,JPF Architecture and Components

MODULE 5 (7 hours)

Introduction to Grey-Box Testing, Applications of Grey-Box Testing, Combining White-Box and Black-Box Testing Approaches, Tool: AFL (American Fuzzy Lop) for Grey-Box Testing and Fuzzing, Security Testing and Vulnerability Identification, Advanced Topics in Grey-Box Testing and Fuzzing, Challenges and Limitations

Reference Books

1. Daniel Jackson, "Software Abstractions: Logic, Language, and Analysis", The MIT Press

2. Edmund M. Clarke, Orna Grumberg, and Doron A. Peled," Model Checking" The MIT Press

3. Gerard J. Holzmann," The Spin Model Checker: Primer and Reference Manual", The MIT Press

4. Frank M. Carrano and Henry Korth," Data Abstraction and Problem Solving with C++: Walls and Mirrors" Pearson

5. Chandrasekhar Kamkha, Daniel B. R. de Oliveira, and Sandeep Joshi ,"Fuzzing for Software Security Testing and Quality Assurance", CRC Press

6. Paul C. Jorgensen, " Software Testing: A Craftsman's Approach", CRC Press

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
	Module 1	
1.1	Introduction to Formal Methods in Software Engineering	1
1.2	Importance of Formal Specification and Verification	1
1.3	Overview of Conceptual Modeling Techniques	1
1.4	Tool: Introduction to Alloy for formal specification	1
1.5	Logic-Based Modeling	1
1.6	Conceptual modeling of entities and relationships using Alloy	1
1.7	Constraints and Invariants in Alloy	1
	Module 2	
2.1	Introduction to Algorithmic Verification	1
2.2	Fundamentals of Model-Checking.	1

2.3	Temporal Logic and its Application in Model Checking	1
2.4	Tool: Introduction to Model-Checking using Spin	1
2.5	The Promela Modeling Language	1
2.6	Specifying Properties for Verification	1
2.7	Optimizing Model Checking with Spin	1
	Module 3	
3.1	Introduction to Abstract Data Types (ADTs)	1
3.2	Importance of ADTs in Software Design	1
3.3	Refinement of ADTs and Program Correctness	1
3.4	Tool: Rodin for ADT Specification and Verification	1
3.5	Introduction to Hoare Logic and its Application	1
3.6	Tool: VCC for Hoare Logic Assertions and Program Refinement	1
3.7	Combining ADTs with Formal Methods	1
3.8	Case Studies and Practical Applications	1
	Module 4	
4.1	Introduction to White-Box Testing	1
4.2	Principles of White-Box Testing,	1
4.3	Test Coverage Criteria: Path, Branch, and Condition Coverage	1

4.4	Tool: JPF for Automated White-Box Testing	1
4.5	Overview of Java Path Finder (JPF)	1
4.6	Model Checking and Its Role in White-Box Testing	1
4.7	JPF Architecture and Components	1
	Module 5	
5.1	Introduction to Grey-Box Testing	1
5.2	Applications of Grey-Box Testing	1
5.3	Combining White-Box and Black-Box Testing Approaches	1
5.4	Tool: AFL (American Fuzzy Lop) for Grey-Box Testing and Fuzzing	1
5.5	Security Testing and Vulnerability Identification,	1
5.6	Advanced Topics in Grey-Box Testing and Fuzzing,	1
5.7	Challenges and Limitations	1

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: M24CS1E105A

Course Name: Formal Methods in Software Engineering

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Briefly describe Alloy and its purpose in formal specification.
2. What is algorithmic verification, and why is it important in software engineering?
3. Define an Abstract Data Type (ADT) and provide an example.
4. Define white-box testing and explain its key characteristics.
5. Define grey-box testing and explain how it differs from white-box and black-box testing.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Discuss the importance of formal specification and verification in the development of critical software systems (4 marks)

   (b) Compare and contrast various conceptual modeling techniques used in software engineering (4 marks)

7. (a) Explain how Alloy can be used for conceptual modeling of entities and relationships in a software system. Provide an example (4 marks)

   (b) Evaluate the role of formal methods in improving software quality. Include a discussion on the challenges associated with their adoption in the industry. (4 marks)

8. (a) Discuss the process of specifying properties for verification in Spin using temporal logic. Provide an example. (4 marks)

   (b) Explain the techniques used in Spin to handle the state space explosion problem. (4 marks)

9. Discuss the process of refining an Abstract Data Type (ADT) from its abstract specification to a concrete implementation. Include an example of how this refinement is achieved. (8 marks)

10. (a) Describe the key features of the Rodin tool and how it supports formal verification of ADTs (4 marks)

    (b) Discuss the VCC tool's approach to program verification using Hoare Logic assertions. Include an example of how VCC verifies a C program. (4 marks)

11. (a) Explain the process of setting up and configuring JPF for white-box testing. Include steps for integrating JPF with a Java development environment

    (4 marks)

    (b) Discuss how JPF handles the state space explosion problem in model checking. Include a description of techniques used to manage large state spaces. (4 marks)

12. (a) . Discuss the benefits of combining white-box and black-box testing approaches in a software testing strategy. Provide an example of how this combination can be applied. (4 marks)

    (b) Explain the process of setting up and using AFL for fuzzing a software application. Include key steps and considerations

    (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E105B	**BLOCKCHAIN AND CRYPTOGRAPHY**	ELECTIVE	3	0	0	3	3

Preamble : The purpose of this course is to create awareness and understanding among students on the foundation of blockchain technology. The course introduces the cryptographic principles behind blockchain and helps the students understand concepts like consensus, crypto-currency, smart contracts, use cases etc. The course enables students to develop simple decentralized applications using blockchain networks such as Ethereum.

Prerequisites : Basic knowledge in data structures and operating systems

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Illustrate the cryptographic building blocks of blockchain technology. **(Cognitive Knowledge Level : Understand)**
CO 2	Explain the fundamental concepts of blockchain technology. **(Cognitive Knowledge Level: Understand)**
CO 3	Summarize the classification of consensus algorithms.**(Cognitive Knowledge Level: Understand)**
CO 4	Explain the concepts of first decentralized cryptocurrency bitcoin.**(Cognitive Knowledge Level: Understand)**
CO 5	Explain the use of smart contracts and its use cases.**(Cognitive Knowledge Level: Understand)**
CO 6	Develop simple applications using Solidity language on Ethereum platform.**(Cognitive Knowledge Level: Apply)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓	✓	✓			
CO 2	✓	✓	✓		✓	✓
CO 3	✓	✓	✓	✓	✓	✓
CO 4	✓	✓	✓	✓	✓	✓
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember	30	30	30
Understand	30	30	30
Apply	40	40	40
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Fundamentals of Cryptography: Introduction to Cryptography, Symmetric cryptography – AES. Asymmetric cryptography – RSA. Elliptic curve cryptography, Digital signatures – RSA digital signature algorithms. Secure Hash Algorithms – SHA-256. Applications of cryptographic hash functions – Merkle trees, Distributed hash tables.

MODULE 2 (6 hours)

Fundamentals of Blockchain Technology: HBlockchain – Definition, architecture, elements of blockchain, benefits and limitations, types of blockchain. Consensus – definition, types, consensus in blockchain. Decentralization – Decentralization using blockchain, Methods of decentralization, Routes to decentralization, Blockchain and full ecosystem decentralization.

MODULE 3 (7 hours)

Consensus Algorithms and Bitcoin: Consensus Algorithms, Crash fault-tolerance (CFT) algorithms – Paxos, Raft. Byzantine fault- tolerance (BFT) algorithms – Practical Byzantine Fault Tolerance (PBFT), Proof of work (PoW), Proof of stake (PoS), Types of PoS. Bitcoin – Definition, Cryptographic keys – Private keys, public keys, addresses. Transactions – Lifecycle, coinbase transactions, transaction validation. Blockchain – The genesis block. Mining – Tasks of miners, mining algorithm, hash rate. Wallets – Types of wallets.

MODULE 4 (7 hours)

Smart Contracts and Use cases Smart Contracts – Definition, Smart contract templates, Oracles, Types of oracles, Deploying smart contracts. Decentralization terminology – Decentralized applications, Decentralized Autonomous Organizations. Use cases of Blockchain technology – Government, Health care, Finance, Supply chain management. Blockchain and allied technologies – Blockchain and Cloud Computing, Blockchain and Artificial Intelligence.

MODULE 5 (9 hours)

Ethereum and Solidity: Ethereum – The Ethereum network. Components of the Ethereum ecosystem – Keys and addresses, Accounts, Transactions and messages. The Ethereum Virtual Machine, Blocks and blockchain. The Solidity language – The layout of a Solidity source code, Structure of a smart contract, variables, data types, control structures, events, inheritance, libraries, functions, error handling. Smart contracts Case study: Voting, Auction.

Text Books

1. 1.Imran Bashir, Mastering Blockchain: A deep dive into distributed ledgers, consensus

protocols, smart contracts, DApps, cryptocurrencies, Ethereum, and more, Packt Publishing, Third edition, 2020.

Reference Books

1. Ritesh Modi, Solidity Programming Essentials: A beginner's guide to build smart contracts for Ethereum and blockchain, Packt Publishing, First edition, 2018.
2. Kumar Saurabh, Ashutosh Saxena, Blockchain Technology: Concepts and Applications, First Edition, Wiley Publications, First edition, 2020.
3. Chandramouli Subramanian, Asha A George, et al, Blockchain Technology, Universities Press (India) Pvt. Ltd, First edition, August 2020.
4. 5.Lorne Lantz, Daniel Cawrey, Mastering Blockchain: Unlocking the Power of Cryptocurrencies, Smart Contracts, and Decentralized Applications, O'Reilly Media, First edition, 2020.
5. Andreas M. Antonopoulos, Gavin Wood, Mastering Ethereum: Building Smart Contracts and DApps, O'Reilly Media, First edition, 2018.

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	7 hours
1.1	Introduction to cryptography	1 hour
1.2	Symmetric cryptography, AES	1 hour
1.3	Asymmetric cryptography, RSA	1 hour

1.4	Elliptic curve cryptography	1 hour
1.5	Digital signatures – RSA digital signature algorithm	1 hour
1.6	Secure Hash Algorithms – SHA-256	1 hour
1.7	Applications of cryptographic hash functions – Merkle trees, Distributed hash tables	1 hour
2	**Module 2**	6 hours
2.1	Blockchain – definition and architecture	1 hour
2.2	Elements of blockchain.	1 hour
2.3	Blockchain – benefits and limitations, types.	1 hour
2.4	Consensus – definition, types, consensus in blockchain	1 hour
2.5	Decentralization using blockchain, Methods of decentralization	1 hour
2.6	Routes to decentralization, Blockchain and full ecosystem decentralization	1 hour
3	**Module 3**	7 hours
3.1	Consensus Algorithms – Crash fault-tolerance (CFT) algorithms – Paxos, Raft (working is expected).	1 hour
3.2	Byzantine fault-tolerance (BFT) algorithms – Practical Byzantine Fault Tolerance (PBFT) (working is expected).	1 hour
3.3	Proof of work (PoW), Proof of stake (PoS), Types of PoS	1 hour
3.4	Bitcoin – Definition, Cryptographic keys – Private keys, public keys, addresses.	1 hour

3.5	Transactions – Lifecycle, coinbase transactions, transaction validation	1 hour
3.6	Blockchain – The genesis block. Mining – Tasks of miners	1 hour
3.7	Mining – mining algorithm, hash rate. Wallets – Types of wallets.	1 hour
3.8	Elliptical curve cryptosystems(Elliptical curve arithmetic)	1 hour
4	**Module 4**	7 hours
4.1	Smart Contracts – Definition, Smart contract templates	1 hour
4.2	Oracles, Types of oracles, Deploying smart contracts.	1 hour
4.3	Decentralization terminology –Decentralized applications, Decentralized Autonomous Organizations.	1 hour
4.4	Use cases of Blockchain technology – Government, Health care.	1 hour
4.5	Use cases of Blockchain technology – Finance, Supply chain management.	1 hour
4.6	Blockchain and Allied Technologies – Blockchain and Cloud Computing, Blockchain and Artificial Intelligence.	2 hour
5	**Module 5**	9 hours
5.1	Ethereum - The Ethereum network, Components of the Ethereum ecosystem – Keys and addresses, Accounts	1 hour
5.2	Components of the Ethereum ecosystem – Transactions and messages	1 hour
5.3	The Ethereum Virtual Machine	1 hour

5.4	Ethereum Blocks and blockchain	1 hour
5.5	The Solidity language – The layout of a Solidity source code, Structure of a smart contract, variables, data types	1 hour
5.6	The Solidity language – control structures, events, inheritance, libraries	1 hours
5.7	The Solidity language – functions, error handling	1 hour
5.8	Smart contracts Case study: Voting.	1 hour
5.9	Smart contracts Case study: Auction.	1 hour
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E105B

Course Name: BLOCKCHAIN AND CRYPTOGRAPHY

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Discuss the role of secure hash functions in blockchain.
2. Illustrate the blockchain based decentralized system.
3. Explain how Proof of Stake can achieve consensus among peers.
4. Explain the concept of blockchain-based digital identity cards.
5. Explain error handling in Solidity language.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Illustrate and explain how blockchain works using a neat diagram. (4 marks)

   (b) Explain the benefits, features and limitations of blockchain. (4 marks)

7. Explain consensus mechanisms used in blockchain. List out any six consensus algorithms used in the context of blockchain. (8 marks)

8. Define blockchain. Explain how decentralization of computing or processing power is achieved by a blockchain. (8 marks)

9. (a) Describe the various fields that make up a transaction in Bitcoin. (4 marks)

   (b) What is the role of a Bitcoin miner? Explain the mining algorithm used in Bitcoin with the help of a flowchart. (4 marks)

10. Explain the use of blockchain technology in supply chain management. (8 marks)

11. (a) Using Solidity l anguage, create a simple bank contract that allows a user to deposit, withdraw and view balance. (4 marks)

    (b) Using Solidity language, create a simple voting smart contract where a chairperson will give the right to vote to each address individually. (4 marks)

12. (a) Explain the concept of Gas in Ethereum. Explain how transaction cost can be calculated in an Ethereum blockchain network. (4 marks)

    (b) Explain how hash functions are used to build Merkle trees in blockchain. (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E105B	**PATTERN RECOGNITION**	ELECTIVE	3	0	0	3	3

Preamble : Study of this course enables the learner to get an overview of the concepts and algorithms involved in deep learning. The course covers the basic concepts in deep learning, optimization techniques, regularization techniques, convolutional neural networks, recurrent neural networks, graphical models, and deep generative models. This course helps the students to implement deep learning algorithms to solve real-world problems.

Prerequisites : NIL

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Apply probability and numerical methods in statistical pattern recognition. **(Cognitive Knowledge Level: Apply)**
CO 2	CO2 Apply statistical methods in feature selection. **(Cognitive Knowledge Level: Apply)**
CO 3	Apply linear algebra and statistical methods in parameter and non-parameter estimation. **(Cognitive Knowledge Level: Apply)**
CO 4	Apply linear algebra and statistical methods in parameter and non-parameter estimation. **(Cognitive Knowledge Level: Apply)**
CO 5	CO5 Design, Develop, Implement and Present innovative ideas in problem solving with various pattern recognition techniques. **(Cognitive Knowledge Level: Create)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand			
Apply	70	70	70
Analyse	30	30	30
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Introduction to Pattern Recognition: Basics of pattern recognition systems, various applications, Machine Perception, classification of pattern recognition systems. Design of Pattern recognition system, Pattern recognition Life Cycle. Statistical Pattern Recognition: Review of probability theory, Gaussian distribution. Normal density and discriminant functions.

MODULE 2 (8 hours)

Feature Selection: Outlier removal – Data normalization – Missing data, The Peaking phenomenon, Feature selection using statistical hypothesis testing- Hypothesis testing basics – Application of t-Test in feature selection. Class separability measures-Divergence-Chernoff bound and Bhattacharya distance- Feature subset selection –Scalar feature selection, Feature vector selection.

MODULE 3 (7 hours)

Clustering algorithms: Unsupervised learning and clustering - Criterion functions for clustering. Cluster validation. Fuzzy clustering algorithms- Point representatives- quadratic surfaces and representatives – hyper plane representatives. Binary morphology clustering algorithms (BMCAs) – Discretization – Morphological operations - Determination of clusters in a discrete binary set.

MODULE 4 (6 hours)

Artificial neural networks and Pattern Classification : Auto encoders, Variational Auto-Encoder , Deep generative models - Boltzmann machines, Restricted Boltzmann Machines, Deep Belief Networks, Deep Boltzmann Machines, Generative Adversarial Networks, Auto-Regressive Networks. Transfer Learning and Domain Adaptation.

MODULE 5 (8 hours)

Artificial neural networks and Pattern Classification : Auto encoders, Variational Auto-Encoder , Deep generative models - Boltzmann machines, Restricted Boltzmann Machines, Deep Belief Networks, Deep Boltzmann Machines, Generative Adversarial Networks, Auto-Regressive Networks. Transfer Learning and Domain Adaptation.

Text Books

1. S.Theodoridis and K.Koutroumbas, "Pattern Recognition", 4th Ed., Academic Press, 2009
2. C.M.Bishop, "Pattern Recognition and Machine Learning", Springer, 2006
3. R.O.Duda, P.E.Hart and D.G.Stork, "Pattern Classification", John Wiley, 2001

Reference Books

1. Hastie, T., Tibshirani, R. and Friedman, J. “The Elements of Statistical Learning”. Springer. 2001

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	7 hours
1.1	Basics of pattern recognition systems, applications	1 hours
1.2	Machine Perception, Classification of pattern recognition systems	1 hours
1.3	Design of Pattern recognition systems	1 hours
1.4	Pattern recognition Life Cycle	1 hours
1.5	Statistical Pattern Recognition	1 hours
1.6	Normal density and discriminant functions (Lecture 1)	1 hours
1.7	Normal density and discriminant functions (Lecture 2)	1 hours
2	**Module 2**	8 hours
2.1	Feature selection – Outlier removal	1 hours
2.2	Data normalization – Missing data	1 hours
2.3	The peaking phenomenon	1 hours
2.4	Feature selection using statistical hypothesis testing	1 hours
2.5	Hypothesis testing basics – Application of tTest in feature selection	1 hours

2.6	Class separability measures-Divergence	1 hours
2.7	Chernoff bound and Bhattacharya distance	1 hours
2.8	Feature subset selection –Scalar feature selection, Feature vector selection	1 hours
3	**Module 3**	7 hours
3.1	Unsupervised learning and clustering	1 hours
3.2	Criterion functions for clustering. Cluster validation	1 hours
3.3	Fuzzy clustering algorithms- Point representatives	1 hours
3.4	Quadratic surfaces and representatives – hyper plane representatives.	1 hours
3.5	Binary morphology clustering algorithms (BMCAs)	1 hours
3.6	Discretization	1 hours
3.7	Morphological operations - Determination of clusters in a discrete binary set	1 hours
4	**Module 4**	6 hours
4.1	Principal component analysis - its relationship to Eigen analysis	1 hours
4.2	Fisher discriminant analysis	1 hours
4.3	Generalised Eigen analysis	1 hours
4.4	Eigen vectors/Singular vectors as dictionaries	1 hours
4.5	Total variability space - a dictionary learning method	1 hours
4.6	Non negative matrix factorization - a dictionary learning method	1 hours
5	**Module 5**	8 hours

5.1	Artificial neural networks: A brief introduction to deep neural networks	1 hours
5.2	Convolutional neural networks	1 hours
5.3	Recurrent neural networks	1 hours
5.4	Pattern classification basics	1 hours
5.5	Non-metric methods for pattern classification: Non-numeric data or nominal data (lecture 1)	1 hours
5.6	Non-metric methods for pattern classification: Non-numeric data or nominal data (lecture 2)	1 hours
5.7	Decision trees: Classification and Regression Trees (CART)	1 hours
5.8	Decision trees: Classification and Regression Trees (CART)	1 hours

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E105C

Course Name: :Pattern Recognition

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. How does morphological operations play a role in pattern recognition?
2. In a town it was estimated that 3% of people have a particular disease. A diagnosis test was conducted for all the people, which yielded 8% false positive and 92% true positive results. A person is found as positive after the test. What is the probability that this person is truly having the disease?
3. How does a decision tree handle continuous attributes?
4. Define the terms: weights, bias, activations with respect to neural networks.

PART B

Arial***Answer any five question from each module. Each question carries 8 marks.***

5. Illustrate the design principles of pattern recognition system with an example. (8 marks)
6. Derive the fuzzy C spherical shells (FCSS) algorithm for the case that spherical clusters are to be identified. (8 marks)
7. Show that in the case of Gaussian distributions the Chernoff bound becomes

$$\epsilon_{CB} = \exp(-b(s))$$

where

$$b(s) = \frac{s(1-s)}{2}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \left[s\boldsymbol{\Sigma}_j + (1-s)\boldsymbol{\Sigma}_i\right]^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \frac{1}{2}\ln\frac{|s\boldsymbol{\Sigma}_j + (1-s)\boldsymbol{\Sigma}_i|}{|\boldsymbol{\Sigma}_j|^s|\boldsymbol{\Sigma}_i|^{1-s}}$$

Then take the derivative with respect to s and show that for equal covariance matrices the optimum is achieved for $s = 1/2$. Thus, in this case $b(s)$ equals the Bhattacharyya distance.

8. Discuss how the morphological open and morphological closure work using a suitable example. (8 marks)
9. Let N_1, N_2 be the available values of a feature in two classes, respectively. The feature is assumed to follow a Gaussian distribution with the same variance in each class. Define the test statistic

$$q = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s_z\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

where

$$s_z^2 = \frac{1}{N_1 + N_2 - 2}\left(\sum_{i=1}^{N_1}(x_i - \bar{x})^2 + \sum_{i=1}^{N_2}(y_i - \bar{y})^2\right)$$

and μ_1, μ_2 are the respective true mean values. Show that q follows the t-distribution with $N_1 + N_2 - 2$ degrees of freedom.

(8 marks)

10. Discuss the significance of pre-processing in feature selection. Illustrate any two methods used for pre-processing. (8 marks)

11. How can artificial neural networks be applied in Pattern recognition? Also illustrate the features of recurrent neural networks. (8 marks)

12. Construct a decision tree using the following data. (8 marks)

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Cloudy	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cold	Normal	False	Yes
Sunny	Cold	Normal	True	No
Cloudy	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Cool	Normal	True	Yes
Cloudy	Mild	High	True	Yes
Cloudy	Mild	Normal	False	Yes
Sunny	Hot	High	True	No

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E105A	**ADVANCED COMPUTER NETWORKS**	ELECTIVE	3	0	0	3	3

Preamble : This course enables the learners to get a good grasp of emerging technologies in the field of computer networks. The syllabus dwells at length on wireless networking, as well as solutions for problems faced while efficiently routing data. Newer networking applications and protocols particularly in multimedia are introduced. The learners are given a glimpse of recent trends in networking like software defined networking. The course enables the learners to analyze network protocols and develop network based applications.

Prerequisite : Computer Networks

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Examine the problem of scalability for routing and also identify the challenges in mobile and multicast routing. **(Cognitive Knowledge Level: Analyze)**
CO 2	Choose the technique that provides the Quality of Service needs of a particular application. **(Cognitive Knowledge Level: Apply)**
CO 3	Survey various wired and wireless networking technologies including wireless cellular technologies. **(Cognitive Knowledge Level: Analyze)**
CO 4	Classify the multimedia applications in the Internet and compile the various protocols handling these applications. **(Cognitive Knowledge Level: Analyze)**
CO 5	Describe examples of current networking trends and identify the technological gaps. **(Cognitive Knowledge Level: Evaluation)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand			
Apply	30%	40%	70%-80%
Analyse	70%	60%	30%-40%
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

The Global Internet, Routing Areas, Interdomain Routing -BGP, IP Version 6,Multicast, Multicast Addresses, Multicast Routing -DVMRP-PIM-MSDP, Routing to a mobile node, Mobile IP, TCP and Mobility

MODULE 2 (7 hours)

QoS Architectural Framework - Integrated Services Architecture – RSVP - Differentiated

Services, Multiprotocol Label Switching- Destination-Based Forwarding - Explicit Routing ,Virtual Private Networks and Tunnels, Performance issues in networks

MODULE 3 (8 hours)

Wired: DSL, Cable Networks, SONET Wireless: Satellite Networks, WiMAX. Cellular Networks: Introduction-Wireless links and Network characteristics -CDMA, Cellular Internet access -An overview of cellular network architecture, 3G cellular data networks, 4G LTE Cellular networks - LTE Protocol Stacks -LTE Radio Access Network -Additional LTE functions, 5G Cellular networks, Managing mobility in cellular networks, Wireless and Mobility-Impact on higher level protocols, Personal Area Networks: Bluetooth, Zigbee

MODULE 4 (7 hours)

Multimedia in the Internet: Streaming stored audio/video, Streaming live audio/video, Real time interactive audio/video, Real time Interactive Protocols: RTP- RTCP-SIP-H.323, SCTP Compression: Audio Compression, Image compression- JPEG, Video Compression- MPEG.

MODULE 5 (7 hours)

Overlay Networks: Routing overlays -Resilient overlay networks, Peer-Peer Networks – Bit Torrent-Distributed Hash Tables, Content Distribution networks, Software Defined Networks: Architecture – Control and Data Planes – Open Flow – SDN Controllers, Network Function Virtualization.

Reference Books

1. Larry Peterson and Bruce Davie, Computer Networks - A Systems Approach, Morgan Kaufmann, 6th edition, 2022

2. James F. Kurose and Keith W. Ross, Computer Networking A Top-Down Approach, Pearson, 8th edition, 2022

3. Jochen Schiller, Mobile Communications, Addison-Wesley, 2nd edition, 2003

4. William Stallings, Data and Computer Communications, Pearson, 5th edition, 2017

5. Andrew Tanenbaum and David Wetherall, Computer Networks, Pearson, 5th edition, 2010

6. Behrouz A Forouzan, Data Communications and Networking, McGraw Hill, 5th edition, 2017

7. Thomas D. Nadeau and Ken Gray, SDN – Software Defined Networks, O'Reilly, 2013

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1: Advanced Internetworking**	7 hours
1.1	The Global Internet, Routing Areas	1 hours
1.2	Interdomain Routing -BGP	1 hours
1.3	IP Version 6	1 hours
1.4	Multicast, Multicast Addresses	1 hours
1.5	Multicast Routing – DVMRP	1 hours
1.6	PIM, MSDP	1 hours
1.7	Routing to a mobile node, Mobile IP , TCP and Mobility	1 hours
2	**Module 2 : Internetwork Quality of Service**	7 hours
2.1	QoS Architectural Framework	1 hours
2.2	Integrated Services Architecture	1 hours
2.3	RSVP - Differentiated Services	1 hours
2.4	Multiprotocol Label Switching	1 hours

2.5	Virtual Private Networks and Tunnels	1 hours
2.6	Destination-Based Forwarding - Explicit Routing	1 hours
2.7	Performance issues in networks	1 hours
3	**Module 3 :Networking Technologies**	8 hours
3.1	Wired: DSL, Cable Networks, SONET	1 hours
3.2	Wireless: Satellite Networks, WiMAX	1 hours
3.3	Cellular Networks: Introduction-Wireless links and Network haracteristics -CDMA	1 hours
3.4	Cellular Internet access-An overview of cellular network architecture, 3G cellular data networks	1 hours
3.5	4G LTE Cellular networks - LTE Protocol Stacks -LTE Radio Access Network -Additional LTE functions	1 hours
3.6	5G Cellular networks	1 hours
3.7	Managing mobility in cellular networks, Wireless and Mobility-Impact on higher level protocols	1 hours
3.8	Personal Area Networks: Bluetooth, Zigbee	1 hours
4	**Module 4:Networking Applications**	7 hours
4.1	Multimedia in the Internet: Streaming stored audio/video, Streaming live audio/video,	1 hours
4.2	Real time interactive audio/video	1 hours
4.3	Real time Interactive Protocols: RTP- RTCP	1 hours
4.4	H-323	1 hours

4.5	SIP, SCTP	1 hours
4.6	Compression: Audio Compression, Image compression-JPEG	1 hours
4.7	Video Compression- MPEG	1 hours
5	**Module 5 Current Topics in Networking**	7 hours
5.1	Overlay Networks: Routing overlays	1 hours
5.2	Resilient overlay networks	1 hours
5.3	Peer-Peer Networks – Bit Torrent – Distributed Hash Tables	1 hours
5.4	Content Distribution networks	1 hours
5.5	Software Defined Networks: Architecture – Control and Data Planes	1 hours
5.6	Open Flow, SDN Controllers	1 hours
5.7	Network Function Virtualization	1 hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code : M24CS1E105D

Course Name: Advanced Computer Networks

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Illustrate with an example how standard TCP can be enhanced to support mobile users.
2. Explain the architectural framework for supporting Quality of Service in packet networks.
3. Examine the role of core network in 3G cellular data network.
4. There is one sender and eight receivers in a real time multimedia communication system. If the sender is sending multimedia data at 2 Mbps, how many RTCP packets can be sent by the sender and each receiver in a second? The system allocates 75 percent of the RTCP bandwidth to the receivers and 25 percent to the sender. The average size of each RTCP packet is 125 bytes.
5. Define OpenFlow specification used in SDN.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) X, Y, Z are three ASs. X and Z are connected through Y. X has a peering agreement with Y and Y with Z. Z moves all traffic from Y but does not forward traffic from X. Can Z use BGP to implement this policy? (5 marks)

   (b) How does PIM solve the scalability problem of existing multicast protocols (3 marks)

7. (a) Derive the hexadecimal form of representation of the following link local multicast address:

   i. a permanently-assigned multicast group address of 66

   ii. a transient multicast group address of 316

   (4 marks)

(b) A foreign network has a foreign agent. Explain if it is possible for two mobile nodes in the foreign network to use the same care-of address in mobile IP. (4 marks)

8. (a) Justify the need for Resource Reservation in multicast transmission. (4 marks)

   (b) How is VPN implemented using MPLS? (4 marks)

9. (a) Elaborate on the various elements of 4G LTE network and the interaction between them. (5 marks)

   (b) Calculate the minimum time required to download 2*106 bytes using ADSL modem with minimum rate (3 marks)

10. (a) Sketch the superframe format of Zigbee 802.15.4 standard. (3 marks)

    (b) Name some applications which use Zigbee standard and justify its use (4 marks)

11. (a) Describe H323 architectural model for Internet Telephony. (8 marks)

12. (a) Comment on the statement "Distributed Hash Tables are said to build structured P2P networks" (4 marks)

    (b) Explain Data Center Networking (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E105E	**INFORMATION RETRIEVAL**	ELECTIVE	3	0	0	3	3

Preamble : The course focuses on the representation, storage, association and access to data items utilizing different IR calculations and systems. It emphasizes the working of data recovery frameworks for reports in order to recover important or valuable data from them. It help the students to develop applications and evaluate recommender systems.

Prerequisites : - Machine Learning / Machine Learning and Artificial Intelligence here

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Utilize the information retrieval models. (Cognitive Knowledge Level: Apply)
CO 2	Acquire knowledge on pre-processing of web page. (Cognitive Knowledge Level: Understand)
CO 3	Organize data from semantic web. (Cognitive Knowledge Level: Apply)
CO 4	Operate on various search engine systems. (Cognitive Knowledge Level: Analyse)
CO 5	Implement different techniques of recommender system.(Cognitive Knowledge Level: Analyse)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓	✓	✓			✓
CO 2	✓	✓	✓	✓	✓	✓
CO 3	✓	✓	✓			✓
CO 4	✓	✓	✓	✓		
CO 5	✓	✓	✓	✓		✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	30	30	30
Apply	30	30	30
Analyse	40	40	40
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (9 hours)

Information Vs Data Retrieval – Information Retrieval Models: Boolean Model – Vector Space Model – Language Model – Probabilistic Model – Evaluation Measures – Precision – Recall curve – LEMUR.

MODULE 2 (9 hours)

Pre-processing Technique – Inverted index and its comparison – Latent Semantic Indexing – Web Search – Web Spamming – Sentiment Analysis – Privacy Issues - NLTK (Natural Language Toolkit).

MODULE 3 (9 hours)

Semantic Web Technologies – Structured Web Documents – Describing Web Resources – Applications and Challenges – Rule Interchange Format(RIF) – SPARQL -Web Ontology Language – OWL.

MODULE 4 (x hours)

Search Engine Architectures – Cluster based Architecture – Distributed Architectures – Search Engine Ranking – Link based Ranking – Simple Ranking Functions – Learning to Rank – Evaluations — Search Engine Ranking – Search Engine User Interaction – Browsing – Applications of a Web Crawler - Evaluation. INDRI SEARCH ENGINE.

MODULE 5 (x hours)

Recommender Systems Functions – Data and Knowledge Sources – Recommendation Techniques – Basics of Content-based Recommender Systems – High Level Architecture – Advantages and Drawbacks of Content-based Filtering – Collaborative Filtering – Matrix factorization models – Neighborhood models - Scout Portal Toolkit(SPT).

TEXT BOOKS

1. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, —Modern Information Retrieval: The Concepts and Technology behind Search, Second Edition, ACM Press Books, 2011.

2. Ricci, F, Rokach, L. Shapira, B.Kantor, —Recommender Systems Handbook, First Edition, 2011.

Reference Books

1. C. Manning, P. Raghavan, and H. Schütze, —Introduction to Information Retrieval, Cambridge University Press, 2012.

2. Stefan Buettcher, Charles L. A. Clarke and Gordon V. Cormack, —Information Retrieval: Implementing and Evaluating Search Engines, The MIT Press, 2010.

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	
1.1	Information Vs Data Retrieval	1 hours
1.2	Information Retrieval Models	1 hours
1.3	Boolean Model	1 hours
1.4	Vector Space Model	1 hours
1.5	Language Model	1 hours
1.6	Probabilistic Model	1 hours
1.7	Evaluation Measures	1 hours
1.8	Precision	1 hours
1.9	Recall curve – LEMUR.	1 hours
2	**Module 2**	
2.1	Pre-processing Technique	1 hours
2.2	Inverted index and its comparison	1 hours
2.3	Latent Semantic Indexing	1 hours
2.4	Web Search	1 hours

2.5	Web Spamming	1 hours
2.6	Sentiment Analysis	1 hours
2.7	Privacy Issues	1 hours
2.8	NLTK (Natural Language Toolkit).	1 hours
2.9	NLTK (Natural Language Toolkit).	1 hours
3	**Module 3**	
3.1	Semantic Web Technologies	1 hours
3.2	Structured Web Documents	1 hours
3.3	Describing Web Resources	1 hours
3.4	Applications and Challenges	1 hours
3.5	Rule Interchange Format(RIF)	1 hours
3.6	SPARQL	1 hours
3.7	Web Ontology Language	1 hours
3.8	Web Ontology Language	1 hours
3.9	OWL.	1 hours
4	**Module 4**	
4.1	Search Engine Architectures	1 hours
4.2	Cluster based Architecture	1 hours

4.3	Distributed Architectures	1 hours
4.4	Search Engine Ranking	1 hours
4.5	Link based Ranking	1 hours
4.6	Simple Ranking Functions	1 hours
4.7	Learning to Rank	1 hours
4.8	Evaluations — Search Engine Ranking	1 hours
4.9	Search Engine User Interaction – Browsing – Applications of a Web Crawler - Evaluation. INDRI SEARCH ENGINE.	1 hours
5	**Module 5**	
5.1	Recommender Systems Functions	1 hours
5.2	Data and Knowledge Sources	1 hours
5.3	Recommendation Techniques	1 hours
5.4	Basics of Content-based Recommender Systems	1 hours
5.5	High Level Architecture	1 hours
5.6	Advantages and Drawbacks of Content-based Filtering	1 hours
5.7	Collaborative Filtering	1 hours
5.8	Matrix factorization models – Neighborhood models	1 hours
5.9	Scout Portal Toolkit(SPT).	1 hours
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E105E

Course Name: INFORMATION RETRIEVAL

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Explain how the Boolean Model and Vector Model differ in their approach to information retrieval. Provide an example scenario where one model might be preferred over the other. (4 marks)

2. Describe how hash addressing and doubly chained trees can optimize the performance of information retrieval systems. Provide a practical example of their application. (4 marks)

3. How do precision and recall metrics complement each other in evaluating the effectiveness of an information retrieval system? Provide a scenario where a balance between these metrics is crucial. (4 marks)

4. Explain the key features and architecture of a web crawler. How does the URL frontier contribute to the efficiency of web crawling? (4 marks)

5. Discuss the role of anchor text and the web graph in link analysis. How do these elements impact the search engine's ranking algorithm and overall search quality? (4 marks)

PART B

Answer any five question from each module. Each question carries 8 marks.

6. Compare and contrast the Boolean, Vector, and Probabilistic models of information retrieval. (8 marks)

7. (a) Explain the underlying principles and assumptions of each model (5 marks)

   (b) Discuss the strengths and weaknesses of each model in the context of handling ambiguous search queries (3 marks)

8. Analyze the impact of advanced data structures on the efficiency of information retrieval. (8 marks)

9. Evaluate the importance of classification and clustering in information retrieval. (8 marks)

10. Examine the evaluation metrics used to assess information retrieval systems. (8 marks)

11. Discuss the concept and computation of PageRank in link analysis. (8 marks)

12. Discuss the various retrieval techniques such as Boolean search, matching functions, and cluster-based retrieval. (8 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1R106	**RESEARCH METHODOLOGY AND IPR**	ELECTIVE	2	0	0	2	2

Preamble : This course introduces the strategies and methods related to scientific research. It covers salient aspects of publication and patenting along with the crucial role of ethics in research. This course will equip students to define research problem and to adopt suitable methodologies for the solution of problem. The students are trained in the oral presentation with visual aids and writing technical thesis/reports/research papers.

Prerequisite : Computer Networks **Course Outcomes**

After the completion of the course the student will be able to:

CO 1	Approach research projects with enthusiasm and creativity.
CO 2	Conduct literature survey and define research problem.
CO 3	Adopt suitable methodologies for solution of the problem.
CO 4	Deliver well-structured technical presentations and write technical reports.
CO 5	Publish/Patent research outcome.

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	3	3	1		1	3
CO 2	3	3	1		1	3
CO 3	3	3	1		1	3
CO 4	3	3	1		1	3
CO 5	3	3	1		1	3

Assessment Pattern

Bloom's Category	**Continuous Internal Evaluation Tests**		**End Semester Examination (% Marks)**
	Test 1 (%Marks)	**Test 2 (%Marks)**	
Remember			
Understand	20%	20%	20%
Apply			
Analyse	40%	40%	40%
Evaluate	40%	40%	40%
Create			

Mark Distribution

Total Marks	**CIE Marks**	**ESE Marks**	**ESE Duration**
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar* : 10 marks

Course based task/Micro Project

Data collection and interpretation/Case study : 10 marks

Test paper 1 (Module 1 and Module 2) : 10 marks

Test paper 2 (Module 3 and Module 4) : 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination should be conducted by the college. The time duration will be for 3 Hrs and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

SYLLABUS

MODULE 1 (5 hours)

Meaning, and objective of research, Motivation for research: Motivational talks on research: "You and Your Research"- Richard Hamming, Types of research, Research Approaches, significance of research, Characteristics of good research, Research process. Thinking skills: Types and Levels of thinking - common-sense, scientific thinking, and logical thinking. Creativity: Some definitions, illustrations from day to day life, intelligence versus creativity, creative process, requirements for creativity.

MODULE 2 (4 hours)

Information gathering – reading, searching and documentation, Types of literature. Integration of research literature and identification of research gaps, Attributes and sources of research problems, problem formulation, Research question, multiple approaches to a problem, Problem solving strategies – reformulation or rephrasing, techniques of representation, Importance of graphical representation, examples.

MODULE 3 (6 hours)

Experimental and modelling skills Scientific method, role of hypothesis in experiment, dependent and independent variables, control in experiment, precision and accuracy, need for precision, definition, detection, estimation and reduction of random errors, statistical treatment of data, definition, detection and elimination of systematic errors. Design of experiments, experimental logic and documentation. Types of models, stages in modelling, curve fitting, the role of approximations, problem representation, logical reasoning, mathematical skills. Continuum/meso/micro scale approaches for numerical simulation, Case studies illustrating experimental and modelling skills.

MODULE 4 (5 hours)

Effective communication - oral and written Examples illustrating the importance of effective communication, stages and dimensions of a communication process. Oral communication –verbal and non-verbal, casual, formal and informal communication, interactive communication, listening, form, content and delivery, various contexts for speaking- conference, seminar etc. Guidelines for preparation of good presentation slides. Written communication – Rules of scientific writing, form, content and language, layout, typography and illustrations, nomenclature, reference and citation styles, contexts for writing – paper, thesis, reports etc. Tools for document preparation-LaTeX. Common errors in typing and documentation

MODULE 5 (5 hours)

Relative importance of various forms of publication, Choice of journal and reviewing process, Stages in the realization of a paper. Research metrics-Journal level, Article level and Author level, Plagiarism and research ethics. Introduction to IPR, Concepts of IPR, Types of IPR, Common rules of IPR practices, Types and Features of IPR Agreement, Trademark. Patents-Concept, Objectives and benefits, features, Patent process – steps and procedures.

Text Books

1. Panneerselvam, Research Methodology, Prentice Hall of India, New Delhi, 2012.

2. C. R. Kothari, Research Methodology, New Age International, 2004

Reference Books

1. E. M. Phillips and D. S. Pugh, "How to get a PhD - a handbook for PhD students and their supervisors", Viva books Pvt Ltd.

2. G. L. Squires, "Practical physics", Cambridge University Press

3. Antony Wilson, Jane Gregory, Steve Miller, Shirley Earl, Handbook of Science Communication, Overseas Press India Pvt Ltd, New Delhi, 1st edition 2005

4. Leedy P. D., Practical Research: Planning and Design, McMillan Publishing Co.

5. Day R. A., How to Write and Publish a Scientific Paper, Cambridge University Press, 1989.

6. William Strunk Jr., Elements of Style, Fingerprint Publishing, 2020

7. Peter Medawar, 'Advice to Young Scientist', Alfred P. Sloan Foundation Series, 1979.

8. E. O. Wilson, Letters to a Young Scientist, Liveright, 2014.

9. R. Hamming, You and Your Research, 1986 Talk at Bell Labs.

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	5 hours
1.1	Meaning, and objective of research, Motivation for research: Motivational talks on research: "You and Your Research"- Richard Hamming.	1 hours
1.2	Types of research, Research Approaches.	1 hours
1.3	Significance of research, Characteristics of good research, Research process.	1 hours
1.4	Thinking skills: Types and Levels of thinking - common-sense, scientific thinking, and logical thinking.	1 hours
1.5	Creativity: Some definitions, illustrations from day to day life, intelligence versus creativity, creative process, requirements for creativity.	1 hours
2	**Module 2**	4 hours

2.1	Information gathering – reading, searching and documentation, types of literature	1 hours
2.2	Integration of research literature and identification of research gaps	1 hours
2.3	Attributes and sources of research problems, problem formulation, Research question, multiple approaches to a problem	1 hours
2.4	Problem solving strategies – reformulation or rephrasing, techniques of representation, Importance of graphical representation, examples	1 hours
3	**Module 3**	5 hours
3.1	Scientific method, role of hypothesis in experiment, dependent and independent variables, control in experiment	1 hours
3.2	Precision and accuracy, need for precision, definition, detection, estimation and reduction of random errors, statistical treatment of data, definition, detection and elimination of systematic errors	1 hours
3.3	Design of experiments, experimental logic and documentation	1 hours
3.4	Types of models, stages in modelling, curve fitting, the role of approximations, problem representation, logical reasoning, mathematical skills	1 hours
3.5	Continuum/meso/micro scale approaches for numerical simulation, Case studies illustrating experimental and modelling skills.	1 hours
4	**Module 4**	5 hours

4.1	Examples illustrating the importance of effective communication, stages and dimensions of a communication process	1 hours
4.2	Oral communication –verbal and non-verbal, casual, formal and informal communication, interactive communication, listening, form, content and delivery, various contexts for speaking- conference, seminar etc.	1 hours
4.3	Guidelines for preparation of good presentation slides.	1 hours
4.4	Written communication – Rules of scientific writing, form, content and language, layout, typography and illustrations, nomenclature, reference and citation styles, contexts for writing – paper, thesis, reports etc. Tools for document preparation-LaTeX.	1 hours
4.5	Common errors in typing and documentation	1 hours
5	**Module 5**	5 hours
5.1	Relative importance of various forms of publication, Choice of journal and reviewing process, Stages in the realization of a paper.	1 hours
5.2	Research metrics-Journal level, Article level and Author level, Plagiarism and research ethics	1 hours
5.3	Introduction to IPR, Concepts of IPR, Types of IPR	1 hours
5.4	Common rules of IPR practices, Types and Features of IPR Agreement, Trademark	1 hours
5.5	Patents- Concept, Objectives and benefits, features, Patent process – steps and procedures	1 hours

MODEL QUESTION PAPER

QP CODE: **Pages: 1**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code : M24CS1R106

Course Name: RESEARCH METHODOLOGY AND IPR

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 12 marks.

1. (a) Discuss the salient recommendations for great research recommended by Richard Hamming in his famous talk "You and Your Research". (4marks)

   (b) Classify different types of researches. (8 marks)

2. (a) List out the different steps for identification of research gaps. (6 marks)

   (b) Classify various types of literature. (6 marks)

3. (a) Distinguish between continuum, meso-scale and micro scale approaches for numerical simulation. (6 marks)

   (b) Illustrate the role of approximations in research. (6 marks)

4. (a) Discuss any four rules of scientific writing. (4 marks)

(b) List out the Guidelines for preparation of good presentation slide (8 marks)

5. (a)) Examine the requirements for patentability? (6 marks)

   (b) Contrast between copyright and trademark protection. (6 marks)

6. (a) What are the characteristics of a good research question? Discuss with an example. (6 marks)

   (b) Explain the various sources of research problem. (6 marks)

7. (a) Discuss the various stages and dimensions of communication process. (6 marks)

   (b) Compare Journal level, Article level and Author level research metrics. (6 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1L107	**ADVANCED MACHINE LEARNING LAB**	CORE	4	0	0	4	4

Preamble : Study of the course enables the learners to make use of the machine learning concepts and algorithms to derive data insights. The course provides exposure to the design and implementation aspects of machine learning algorithms such as decision trees, regression, naive bayes algorithm, clustering algorithms and artificial neural network. This helps the students to develop machine learning based solutions to real world problems.

Prerequisites : M241T1CS03 ADVANCED MACHINE LEARNING

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Design and Develop various classification and regression models to solve real time problems and analyze their performance (Cognitive Knowledge Level: Apply)
CO 2	Understand the basic concepts of neural network model , implement and evaluate the performance (Cognitive Knowledge Level: Analyze)
CO 3	Develop and compare various clustering algorithms to solve real time problems (Cognitive Knowledge Level: Analyze)
CO 4	Understand and implement different dimensionality reduction techniques(Cognitive Knowledge Level: Apply)
CO 5	Implement various ensemble method and compare the performance of various models(Cognitive Knowledge Level: Analyze)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓	✓	✓	✓	✓	✓
CO 2	✓	✓	✓	✓	✓	✓
CO 3	✓	✓	✓	✓	✓	✓
CO 4	✓	✓	✓	✓	✓	✓
CO 5	✓	✓	✓	✓	✓	✓

Mark Distribution

Total Marks	CIE Marks	ESE Marks
100	60	40

Continuous Internal Evaluation Pattern

Lab work and Viva-voce : 60 marks

Final assessment Test and Viva voce : 40 marks

The laboratory courses will be having only Continuous Internal Evaluation and carries 100 marks. Final assessment shall be done by two examiners; one examiner will be a senior faculty from the same department.

SYLLABUS

Decision tree (ID3), Naïve bayesian classifier , Bayesian network, Expectation Maximization (EM) algorithm,K-means algorithm, K-nearest neighbor, Regression, Cross validation, Support Vector Machine (SVM), DBSCAN, Artificial neural network, Backpropagation algorithm, Principle Component Analysis (PCA), Google colab.

LIST OF EXPERIMENTS

1	Write a program to implement the Simple and Multiple Linear Regression for a sample training data set stored as a .CSV file. Compute Mean Square Error by considering few test data sets.
2	Write a program to implement the naïve bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.
3	Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.
4	Write a program to implement k-Nearest Neighbour algorithm to classify the iris data set. Print both correct and wrong predictions.
5	Implement a single layer neural network and for different logic gates.
6	Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets.
7	Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering.
8	Write a program to implement the DBSCAN clustering algorithm.
9	Write a program to implement Dimensionality reduction using Principle Component Analysis (PCA) method.
10	Write a program to implement Support Vector Machine algorithm to classify the iris data set. Print both correct and wrong predictions.
11	Write a program to implement 5-fold cross validation on a learning problem using real time dataset. Compare the accuracy, precision, recall, and F-score for different folds.
12	Write a program to implement Boosting ensemble method on a given dataset.

Reference Books

1. Jiawei Han, Micheline Kamber, Jian Pei. Data Mining Concepts and Techniques, Third Edition. Morgan Kaufmann.

2. Christopher M. Bishop. Pattern recognition and machine learning. Springer 2006.

3. Ethem Alpaydin, Introduction to Machine Learning, 2nd edition, MIT Press 2010.

4. Mohammed J. Zaki and Wagner Meira, Data Mining and Analysis: Fundamental Concepts and Algorithms, Cambridge University Press, First South Asia edition, 2016.

5. Goodfellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016. 6. Neural Networks and Deep Learning, Aggarwal, Charu C., c Springer International Publishing AG, part of Springer Nature 2018

SEMESTER II

M.TECH
COMPUTER SCIENCE AND ENGINEERING

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1T201	**ADVANCED DATA STRUCTURES AND ALGORITHMS**	CORE	4	0	0	4	4

Preamble : The course introduces advanced data structures and algorithms in different domains. The goal of this course is to provide a solid background in the design and analysis of the major classes of algorithms. The course helps the learners to develop their own versions for a given computational task and to compare and contrast their performance.

Prerequisites : Data Structures and Algorithms

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Analyze the relevance of amortized analysis and applications. (Cognitive Level: Apply)
CO 2	Illustrate string matching algorithms. (Cognitive Level: Apply)
CO 3	Illustrate advanced data structures like Binomial heap, Fibonacci heap, Disjoint set and string matching algorithms. (Cognitive Level: Apply)
CO 4	Illustrate network flow algorithms and applications. (Cognitive Level: Apply)
CO 5	Make use of probabilistic algorithms and approximation algorithms in computing. (Cognitive Level: Apply)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓		✓
CO 2	✓		✓	✓		✓
CO 3	✓		✓	✓		✓
CO 4	✓		✓	✓		✓
CO 5	✓		✓	✓		✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	30	30	30
Apply	40	40	40
Analyse	30	30	30
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Micro project/Course based project	: 10 marks
Course based task/Seminar/Quiz	: 10 marks
Test paper 1	: 10 marks
Test paper 2	: 10 marks

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (9 hours)

Overview of asymptotic notations and complexity analysis, Amortized analysis – aggregate analysis, accounting method, potential method. String matching – introduction, Rabin-Karp algorithm, Knuth-Morris-Pratt algorithm.

MODULE 2 (9 hours)

Overview of binary heap operations, Binomial tree and heap, Binomial heap operations, Fibonacci heap structure, Fibonacci heap operations, Disjoint set – overview, linked list representation, disjoint set forests.

MODULE 3 (9 hours)

Network flow properties, examples, residual network, augmenting path, cut of network, maxflow-mincut theorem, Ford-Fulkerson algorithm, Edmonds-Karp algorithm, maximum bipartite matching.

MODULE 4 (9 hours)

Introduction, types of probabilistic algorithms, Numerical algorithms – Numerical integration, Probabilistic counting, Monte-Carlo algorithms – Verifying matrix multiplication. Number theory fundamentals – modular arithmetic, modular exponentiation, Euler's Theorem and Fermat's Theorem, Primality testing – Miller-Rabin test. Las Vegas algorithms – Probabilistic selection and quick sort.

MODULE 5 (9 hours)

Introduction, Vertex-cover problem, Traveling-salesman problem, Set-covering problem, Subset-sum problem.

Text Books

1. T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, "Introduction to Algorithms", MIT Press, 3rd edition, 2009.

Reference Books

1. Gilles Brassard and Paul Bratley, "Fundamentals of algorithms", Prentice-hall of India Private Limited, 2001.
2. Rajeev Motwani, Prabhakar Raghavan, "Randomized Algorithms", Cambridge Universi Press, 2000.
3. Dexter C. Kozen, "The Design and Analysis of Algorithms", Springer.
4. Jon Kleinberg and Eva Tardos, "Algorithm Design", Pearson Education, 2006.

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	9 hours
1.1	Overview of asymptotic notations and complexity analysis	1 hour
1.2	Overview of asymptotic notations and complexity analysis	1 hour
1.3	Amortized analysis – aggregate analysis	1 hour
1.4	Accounting method	1 hour
1.5	Potential method	1 hour
1.6	String matching – introduction	1 hour
1.7	Rabin-Karp algorithm	1 hour
1.8	Knuth-Morris-Pratt algorithm (1)	1 hour
1.9	Knuth-Morris-Pratt algorithm (2)	1 hour
2	**Module 2**	9 hours
2.1	Overview of binary heap operations	1 hour
2.2	Binomial tree and heap	1 hour
2.3	Binomial heap operations (1)	1 hour
2.4	Binomial heap operations (2)	1 hour
2.5	Fibonacci heap structure	1 hour

2.6	Fibonacci heap operations (1)	1 hour
2.7	Fibonacci heap operations (2)	1 hour
2.8	Disjoint set – overview, linked list representation	1 hour
2.9	Disjoint set forests	1 hour
3	**Module 3**	9 hours
3.1	Network flow properties	1 hour
3.2	Network flow properties examples	1 hour
3.3	Residual network, augmenting path, cut of network	1 hour
3.4	Maxflow-mincut theorem	1 hour
3.5	Ford-Fulkerson algorithm	1 hour
3.6	Edmonds-Karp algorithm	1 hour
3.7	Edmonds-Karp algorithm	1 hour
3.8	Maximum bipartite matching	1 hour
3.9	Maximum bipartite matching	1 hour
4	**Module 4**	9 hours
4.1	Introduction, types of probabilistic algorithms	1 hour
4.2	Numerical algorithms – Numerical integration, Probabilistic counting)	1 hour
4.3	Monte-Carlo algorithms – Verifying matrix multiplication	1 hour

4.4	Number theory fundamentals – modular arithmetic, modular exponentiation	1 hour
4.5	Euler's Theorem	1 hour
4.6	Fermat's Theorem	1 hour
4.7	Primality testing – Miller-Rabin test (1)	1 hour
4.8	Primality testing – Miller-Rabin test (2)	1 hour
4.9	Las Vegas algorithms – Probabilistic selection and quick sort	1 hour
5	**Module 5**	10 hours
5.1	Introduction	1 hour
5.2	Vertex-cover problem	1 hour
5.3	Vertex-cover problem	1 hour
5.4	Traveling-salesman problem	1 hour
5.5	Traveling-salesman problem	1 hour
5.6	Set-covering problem	1 hour
5.7	Set-covering problem	1 hour
5.8	Subset-sum problem (1)	1 hour
5.9	Subset-sum problem (2)	1 hour
	Total	45 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1T201

Course Name: :ADVANCED DATA STRUCTURES AND ALGORITHMS

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Explain accounting method of amortized analysis with a suitable example.
2. Explain the algorithm for uniting two binomial heaps and analyze the running Time.
3. Maximum matching in a bipartite graph G corresponds to a maximum flow in its corresponding flow network G'. Comment on this statement. Explain how maximum flow problem can be used to solve maximum bipartite matching problem.
4. Explain the probabilistic algorithm for verifying matrix multiplication problem.
5. Explain the approximation algorithm for traveling salesperson problem.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. Describe Knuth-Morris-Pratt algorithm and illustrate using given text T = AABAACAADAABAABA and pattern P = AABA (8 marks)

7. (a) Using potential method, compute the amortized cost of incrementing a binary counter.

   (4 marks)

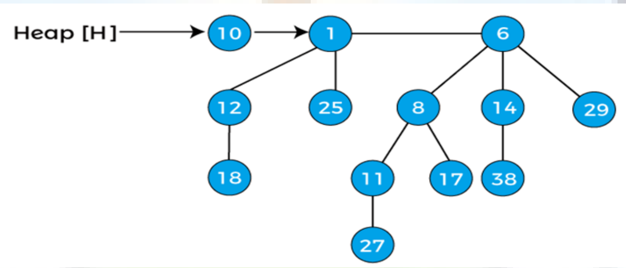
   (b) Suppose we perform a sequence of n operations on a data structure in which the ith operation costs i if i is an exact power of 2, and 1 otherwise. Use accounting method of amortized analysis to determine the amortized cost per operation.

   (4 marks)

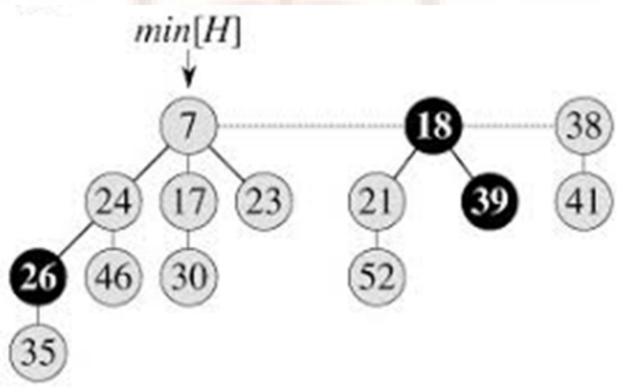
8. (a) Explain how disjoint set data structure is used to find connected components on an undirected graph. (4 marks)

   (b) Show the binomial heap that results when a node with key 11 is deleted from the binomial heap shown in figure.

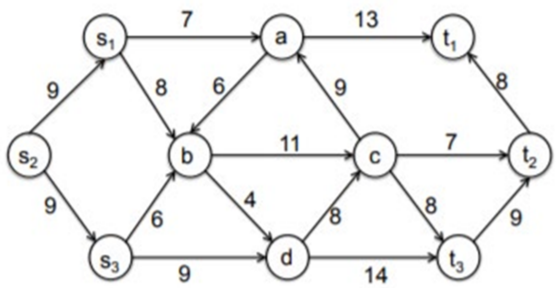
   (4 marks)



9. Explain the structure of Fibonacci heap. Apply extract minimum operation on the Fibonacci heap shown in figure and show the result. (8 marks)



10. Describe Ford-Fulkerson algorithm and apply on the following network. Also obtain minimum cut across the network. (8 marks)



11. Apply Miller-Rabin algorithm to test whether the number 341 is prime or not. (8 marks)

12. Explain probabilistic quick sort algorithm. (8 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1T202	**ADVANCED OPERATING SYSTEMS**	CORE	4	0	0	4	4

Preamble : Study of this course enables the learners to understand the configuration and functions of OS Kernel and have an overview on concepts implemented in modern operating systems. The course focuses on providing information on the design and implementation of the Linux kernel modules. This course will help the learners to suggest solutions/ modify the existing architectural features.

Prerequisites : Operating Systems

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Illustrate the concepts of process management and process scheduling mechanisms employed in the Linux operating system. (Cognitive Knowledge Level: Apply)
CO 2	Describe the set of interfaces by which the process running in user space can interact with the system and how the Kernel manages the various interrupts. (Cognitive Knowledge Level: Apply)
CO 3	Apply various synchronization methods to write race free code. (Cognitive Knowledge Level: Apply)
CO 4	Demonstrate how the kernel handles memory and implementation of the file system. (Cognitive Knowledge Level: Apply)
CO 5	Analyze how kernel manages block devices and their requests and identify the issues to be considered in writing portable codes (Cognitive Knowledge Level: Analyze)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	**Continuous Internal Evaluation Tests**		**End Semester Examination (% Marks)**
	Test 1 (%Marks)	**Test 2 (%Marks)**	
Remember			
Understand	30	30	30
Apply	40	40	40
Analyse	30	30	30
Evaluate			
Create			

Mark Distribution

Total Marks	**CIE Marks**	**ESE Marks**	**ESE Duration**
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Micro project/Course based project	:	10 marks
Course based task/Seminar/Quiz	:	10 marks
Test paper 1	:	10 marks
Test paper 2	:	10 marks

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (9 hours)

Process Management and Scheduling: Introduction to the Linux kernel, Process Management – Process, Process descriptor and the task structure, Process creation, The Linux implementation of threads, Process termination. Process Scheduling – Multitasking, Linux's process scheduler, Policy, Linux scheduling algorithm, Preemption and context switching, Real-time scheduling policies.

MODULE 2 (11 hours)

System calls and Interrupts: System Calls - Communicating with the Kernel, Syscalls, System call handler, System call implementation, System call context. Interrupts and Interrupt

Handlers – Interrupts, Interrupt handlers, Top halves versus bottom halves, Registering an interrupt handler, Writing an interrupt handler, Interrupt context, Interrupt control. Bottom Halves – Task queues, Softirqs, Tasklets, Work queues.

MODULE 3 (9 hours)

Kernel Synchronization: Kernel Synchronization – Critical regions and race conditions, Locking, Deadlocks, Contention and scalability. Kernel Synchronization Methods – Atomic operations, Spin locks, Semaphores, Mutexes, Completion variables, BKL: The Big Kernel Lock, Sequential locks, Preemption disabling.

MODULE 4 (8 hours)

Memory Management and Virtual File System: Memory Management – Pages, Zones, kmalloc(), kfree(), vmalloc(), Slab layer – design, PerCPU allocations. The Virtual File system – VFS objects, data structures, relationship and functionalities.

MODULE 5 (8 hours)

Block I/O Layer and Portability: The Block I/O Layer – Buffers and buffer heads, Request queues, I/O schedulers – Types, Scheduler selection. Portability – Word size and data types, Data alignment, Byte order, Time, Processor ordering.

Text Books

1. Robert Love, "Linux Kernel Development", 3/e, Addison-Wesley, 2010.

Reference Books

1. Daniel Bovet, Marco Cesati, "Understanding the Linux Kernel", 3/e, OReilly Media Inc., 2005.

2. Linux Kernel Architecture – Wolfgang Mauerer.

3. Reilly Christian Benvenuti, "Understanding Linux Network Internals", 1/e, OReilly Media Inc.,2005.

4. Jonathan Corbet, Alessandro Rubini, Greg Kroah-Hartman, "Linux Device Drivers", 3/e, OReilly Media Inc., 2005.

5. Operating Systems Concepts, 9th Edition- Silberschatz, Galvin, Gagne.

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	9 hours
1.1	Introduction to the Linux kernel , Process	1 hour
1.2	Process descriptor and the task structure	1 hour
1.3	Process creation	1 hour
1.4	The Linux implementation of threads	1 hour
1.5	Process termination, Multitasking	1 hour
1.6	Linux's process scheduler, Policy	1 hour
1.7	Linux scheduling algorithm	1 hour
1.8	Preemption and context switching	1 hour
1.9	Realtime scheduling policies	1 hour
2	**Module 2**	11 hours
2.1	Communicating with the Kernel, Syscalls	1 hour
2.2	System Call Handler	1 hour
2.3	System Call Implementation	1 hour

2.4	System call context	1 hour
2.5	Interrupts, Interrupt handlers, Top halves versus bottom halves	1 hour
2.6	Registering an Interrupt Handler	1 hour
2.7	Writing an Interrupt Handler	1 hour
2.8	Interrupt Context, Interrupt Control	1 hour
2.9	Bottom Halves – Task Queues, Softirqs	1 hour
2.10	Tasklets	1 hour
2.11	Work Queues	1 hour
3	**Module 3**	9 hours
3.1	Critical regions and race conditions	1 hour
3.2	Locking, Deadlocks	1 hour
3.3	Contention and scalability	1 hour
3.4	Kernel Synchronization Methods – Atomic operations	1 hour
3.5	Spin locks	1 hour
3.6	Semaphores	1 hour
3.7	Mutexes, Completion variables	1 hour
3.8	BKL: The Big Kernel Lock	1 hour
3.9	Sequential locks, Preemption disabling	1 hour
4	**Module 4**	8 hours

4.1	Pages, Zones	1 hour
4.2	kmalloc()	1 hour
4.3	kfree()	1 hour
4.4	vmalloc()	1 hour
4.5	Slab Layer - Design	1 hour
4.6	Per-CPU Allocations	1 hour
4.7	The Virtual File system – VFS objects	1 hour
4.8	Data structures, relationship and functionalities	1 hour
5	**Module 5**	8 hours
5.1	Buffers and buffer heads	1 hour
5.2	Request queues	1 hour
5.3	I/O Schedulers	1 hour
5.4	Types	1 hour
5.5	Scheduler Selection	1 hour
5.6	Portability – Word size and data types	1 hour
5.7	Data Alignment, Byte Order	1 hour
5.8	Time, Processor ordering	1 hour
	Total	45 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1T202

Course Name: :ADVANCED OPERATING SYSTEMS

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Discuss about Linux implementation of threads.
2. Describe how the work queue mechanism differs from other bottom half mechanisms.
3. Explain pre-emption disabling and list its advantages.
4. Specify the context in which kfree() is needed.
5. Elaborate the working of the deadline I/O scheduler and describe Linus Elevator.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) How does Linux process scheduler ensure a fair share of its computational power among the executing process? (4 marks)

(b) Consider a scenario with two runnable tasks: a text editor and a video encoder. Analyse the scheduling policy in action. (4 marks)

7. (a) Discuss how system calls are implemented and parameters are verified. (4 marks)

   (b) Describe the methods used in enabling and disabling interrupts? Justify the significance of disabling an interrupt line. (4 marks)

8. Explain the trouble that arises if a kernel code which has acquired a spin lock is interrupted by an interrupt handler trying to acquire the same spin lock? Is it possible to avoid such a situation? How? (8 marks)

9. (a) Explain data alignment. Describe the measures to avoid data alignment issues. (4 marks)

   (b) List the difference between big Endian and little Endian byte ordering. (4 marks)

10. (a) How does the TASK_INTERRUPTIBLE state differ from TASK_UNINTERRU-PTIBLE state? If a process is waiting for a specific event to happen, which state should be assigned? (4 marks)

    (b) Explain the situation where individual members of a task list are deallocated? (4 marks)

11. (a) Illustrate the importance of zones in Linux memory management. (4 marks)

    (b) Explain the design of slab layer and the advantage of slab layer allocation of kernel objects. (4 marks)

12. Explain the significance of the bottom half mechanism. In interrupt processing, if the deferred work needs to run in interrupt context and should guarantee that no two of same type run concurrently, which bottom half is preferred. Justify. (8 marks)

SEMESTER II

PROGRAM ELECTIVE III

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E203A	**SOFTWARE TESTING**	ELECTIVE	3	0	0	3	3

Preamble : This is a course in theoretical computer science that introduces the concepts and methods in software testing. It covers various techniques for test case design used to test software artifacts, including requirements, design, and code, the different techniques for test case design based on graphs, programming language syntaxes and symbolic execution using PEX tool. It enables the learners to follow a systematic software testing approaches while developing applications.

Prerequisites : Nil

Course Outcomes

After the completion of the course the student will be able to:

CO 1	List a range of different software testing techniques and be able to apply specific unit testing method to the projects using Junit.**(Cognitive Knowledge Level: Understand)**
CO 2	Illustrate using appropriate tools the mutation testing method for a given piece of code to identify hidden defects that can't be detected using other testing methods.**(Cognitive Knowledge Level:Apply)**
CO 3	Explain graph coverage criteria in terms of control flow graph and data flow graph for a given program. **(Cognitive Knowledge Level: Understand)**
CO 4	Demonstrate the importance of black-box testing techniques and apply them to test software applications.**(Cognitive Knowledge Level: Apply)**
CO 5	Analyze the effectiveness of different software testing tools and techniques in identifying and fixing software defects. **(Cognitive Knowledge Level: Analyze)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓	✓	✓			
CO 2	✓	✓	✓	✓	✓	
CO 3	✓	✓	✓			
CO 4	✓	✓	✓	✓		
CO 5	✓	✓	✓	✓		✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember	30	30	30
Understand	40	40	40
Apply	30	30	30
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Introduction to Software Testing: Software testing. Software Quality, Role of Testing. Testing Process - Level 0 thinking, Level 1 thinking, Level 2 thinking, Level 3 thinking, Level 4 thinking. Software Testing Terminologies - Verification, Validation and Testing, Faults, Error and Bug, Test cases, Coverage Criteria. Types of Testing- Unit testing, integration testing, System testing, Acceptance testing, Beta testing, Functional testing, Stress testing, Performance testing, Usability testing and Regres

MODULE 2 (7 hours)

Unit Testing: Concept of Unit testing. Static Unit testing. Dynamic Unit testing - Control Flow testing, Data Flow testing, Domain testing, Functional Program testing. Mutation testing - Mutation and Mutants, Mutation operators, Mutation score. Junit - Framework for Unit testing. Case Study - Mutation testing using Junit and Muclipse.

MODULE 3 (8 hours)

Unit Testing - White Box Approaches: Overview of Graph Coverage Criteria. Structural Graph Coverage Criteria - Node/vertex coverage, Edge coverage, Edge pair coverage, Path coverage, Complete path coverage, Prime path coverage, Complete round trip coverage, Simple round trip coverage. Data Flow Criteria - du paths, du pairs. Subsumption Relationships among Graph Coverage Criteria. Graph Coverage for Source Code - Control flow graphs for code, CFG: If statement, CFG: If statement with return, CFG: Switch-case, CFG: Loops, CFG: Exceptions (try-catch). Example program – Statistics. Graph Coverage for Design Elements - Call graphs and classes, Class inheritance testing: Coverage criteria, Coverage criteria on inheritance graph, Data flow at the design level, Inter-procedural DU pairs, Coupling du-pairs example. Example - Quadratic Root. Case Study - Graph Based testing using JUnit Framework.

MODULE 4 (7 hours)

Unit Testing - Black Box Approaches: Domain Testing / Input Space Partitioning - Partitions of a set. Input domain modelling - Interface-based approach, Functionality-based approach. Identifying values. Multiple partitions of the input domain - All Combinations Coverage (ACoC), Each Choice Coverage (ECC), Pair-wise Coverage, T-wise Coverage, Base Choice Coverage, Multiple Base Choices Coverage. TriTyp example. Functional Testing - Functional Testing Concepts of Howden. Functional testing - Important Steps. Types of Functional testing - Equivalence Class Partitioning, Boundary Value Analysis, Decision Tables, Random Testing. Case Study - Black Box testing approaches using JUnit.

MODULE 5 (7 hours)

Grey Box Testing Approaches: Introduction to Grey Box testing - Why Grey Box testing, Gray Box Methodology, Advantages and Disadvantages. Techniques of Grey Box Testing - Matrix Testing, Regression Testing, Orthogonal Array Testing or OAT, Pattern Testing. An Introduction to PEX - Parameterized Unit Testing, The Testing Problem. Symbolic

Execution – Example, Symbolic execution tree. PEX application Case Study – PEX.

Text Books

1. Paul Ammann and JeffOffutt , Introduction to Software Testing, Cambridge University Press
2. Kshirasagar Naik and Priyadarshi Tripathy, Software Testing And Quality Assurance: Theory And Practice, Wiley.

Reference Books

1. King, James C, "Symbolic Execution and Program Testing", Association for Computing Machinery, July 1976.

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	6 hours
1.1	Software testing.Software Quality, Role of Testing.	1 hour
1.2	Testing Process - Level 0 thinking, Level 1 thinking, Level 2 thinking, Level 3 thinking, Level 4 thinking.	1 hour
1.3	Software Testing Terminologies - Verification, Validation and Testing, Faults, Error and Bug, Test cases, Coverage Criteria	1 hour
1.4	Types of Testing- Unit testing, integration testing, System testing, Acceptance testing, Beta testing	1 hour

1.5	Functional testing, Stress testing, Performance testing, Usability testing and Regression testing	1 hour
1.6	Testing Methods - Black Box testing, White Box testing, Grey Box testing.	1 hour
2	**Module 2**	6 hours
2.1	Concept of Unit testing. Static Unit testing. Dynamic Unit testing	1 hour
2.2	Control Flow testing, Data Flow testing, Domain testing	1 hour
2.3	Functional Program testing. Mutation testing - Mutation and Mutants	1 hour
2.4	Mutation operators, Mutation score.	1 hour
2.5	Junit - Framework for Unit testing.	1 hour
2.6	Junit - Framework for Unit testing.Case Study - Mutation testing using Junit and Muclipse	1 hour
3	**Module 3**	8 hours
3.1	Overview of Graph Coverage Criteria. Structural Graph Coverage Criteria - Node/vertex coverage, Edge coverage	1 hour
3.2	Edge pair coverage, Path coverage, Complete path coverage, Prime path coverage, Complete round trip coverage, Simple round trip coverage	1 hour
3.3	Data Flow Criteria - du paths, du pairs. Subsumption Relationships among Graph Coverage Criteria	1 hour

3.4	Graph Coverage for Source Code - Control flow graphs for code, CFG: If statement, CFG: If statement with return, CFG: Switch-case, CFG: Loops,	1 hour
3.5	CFG: Exceptions (try-catch). Example program – Statistics. Graph Coverage for Design Elements - Call graphs and classes	1 hour
3.6	Class inheritance testing: Coverage criteria, Coverage criteria on inheritance graph, Data flow at the design level	1 hour
3.7	Inter-procedural DU pairs, Coupling du-pairs example. Example - Quadratic Root. Case Study - Graph Based testing using JUnit Framework.	1 hour
4	**Module 4**	8 hours
4.1	Domain Testing / Input Space Partitioning - Partitions of a set. Input domain modeling, Interface-based approach	1 hour
4.2	Functionality-based approach. Identifying values. Multiple partitions of the input domain	1 hour
4.3	All Combinations Coverage (ACoC), Each Choice Coverage (ECC), Pair-wise Coverage, T-wise Coverage	1 hour
4.4	Base Choice Coverage ,TriTyp example	1 hour
4.5	Functional Testing - Functional Testing Concepts of Howden	1 hour
4.6	Functional Testing - Functional Testing Concepts of Howden	1 hour
4.7	Functional testing - Important Steps. Types of Functional testing	1 hour

4.8	Equivalence Class Partitioning, Boundary Value Analysis, Decision Tables, Random Testing	1 hour
4.8	Case Study - Black Box testing approaches using JUnit.	1 hour
5	**Module 5**	8 hours
5.1	Introduction to Grey Box testing - Why Grey Box testing, Gray Box Methodology	1 hour
5.2	Advantages and Disadvantages. Techniques of Grey Box Testing - Matrix Testing	1 hour
5.3	Advantages and Disadvantages. Techniques of Grey Box Testing - Matrix Testing	1 hour
5.4	Regression Testing, Orthogonal Array Testing or OAT, Pattern Testing.	1 hour
5.5	Regression Testing, Orthogonal Array Testing or OAT, Pattern Testing.	1 hour
5.6	An Introduction to PEX - Parameterized Unit Testing	1 hour
5.7	The Testing Problem. Symbolic Execution – Example, Symbolic execution tree	1 hour
5.8	PEX application Case Study – PEX.	1 hour
5.8	PEX application Case Study – PEX.	1 hour
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E203A

Course Name: SOFTWARE TESTING

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Explain the differences between Validation and Verification?
2. Explain the differences between Fault, Error, and Bug?
3. Define Ground string, Mutation score, and Mutants?
4. What are the functions of Test driver and Test stubs in dynamic unit testing?
5. Define Node coverage, Edge coverage and Prime path coverage in a control flow graph?

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Explain the following types of testing

- Black Box testing
- White Box testing
- GreyBox testing
- Unit testing
- Integration testing
- System test
- Acceptance testing

(5 marks)

(b) Explain the following coverage criterias based on the code fragment given below?
(i) Functional coverage
(ii) Statement coverage
(iii)Conditional coverarage (3 marks)

7. (a) Write positive and negative test cases for an ATM Machine? (4 marks)

   (b) Explain Dynamic unit test environment with a neat figure. (4 marks)

8. (a) Explain the major difference between control flow testing and data flow testing (4 marks)

   (b) Explain seven types of mutation operators with neat examples? (4 marks)

9. (a) Explain touring, side trips and detours with a neat example (4 marks)

   (b) Draw CFG fragment for (i) Simple if (ii) Simple while loop (iii) Simple for loop. (4 marks)

9. (a) Explain the following concepts with examples?
   (i)Call graph (ii) Inheritance graph (iii) Coupling du-pairs (8 marks)

10. (a) Explain the importance of grey box testing, its advantages and disadvantages? (4 marks)

    (b) Explain the concept of symbolic execution tree? (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E203B	**CYBER FORENSICS AND INFORMATION SECURITY**	ELECTIVE	3	0	0	3	3

Preamble : This course provides an exposure to the concepts and techniques in Cyber Forensics Technologies. Different types of Computer Forensics systems, evidence of data gathered, network attacks, evidence data analysis and forensics tools are discussed in this course. This course helps the learners to design, and develop innovative ideas on different digital forensic investigation models

Prerequisite : Required knowledge in computer networks

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Apply forensics technologies, data recovery, evidence collection and handling different forensics issues. (Cognitive Knowledge Level: Apply)
CO 2	CO2 Choose the technique that provides the Quality of Service needs of a particular application. (Cognitive Knowledge Level: Apply)
CO 3	Investigate network intrusions and attacks. (Cognitive Knowledge Level: Analyze)
CO 4	Validating Forensics data and process crime, incident scene. (Cognitive Knowledge Level: Analyze)
CO 5	Exploring file structures and perform forensics investigation. (Cognitive Knowledge Level: Analyze)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	✓			✓	✓	✓	
CO 2	✓			✓	✓		
CO 3	✓			✓	✓	✓	
CO 4	✓		✓	✓	✓		
CO 5	✓		✓	✓	✓	✓	

Assessment Pattern

Course Name	Cyber Forensics and Information Security		
Bloom's Category	**Continuous Internal Evaluation Tests**		**End Semester Examination (% Marks)**
	Test 1 (%Marks)	**Test 2 (%Marks)**	
Remember		10%	10%
Understand	20%	20%	20%
Apply	30%	30%	30%
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar* : 10 marks

Course based task/Micro Project

Data collection and interpretation/Case study : 10 marks

Test paper 1 (Module 1 and Module 2) : 10 marks

Test paper 2 (Module 3 and Module 4) : 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Introduction to Computer Forensics: Computer Forensics Fundamentals, Types of Computer Forensics Technology – Types of Computer Forensics Systems-Data Recovery and Evidence Collection– Forensic duplication and preservation of Digital Evidence, Understanding Computer Investigation.

MODULE 2 (8 hours)

Evidence Data Gathering: Data Acquisition. - Data Recovery- Evidence Collection and Data Seizure - Duplication and Preservation of Digital Evidence.

MODULE 3 (6 hours)

Investigations: Network Traffic, Web Attacks, Router Forensics, DoS Attacks and Internet Crime.

MODULE 4 (7 hours)

Evidence Data Analysis: Discovery of Electronic Evidence - Identification of Data - Determining and Validating Forensics Data – Data Hiding Techniques – Performing Remote Acquisition– Cell Phone and Mobile Devices Forensics- Processing Crime and Incident Scenes.

MODULE 5 (8 hours)

Forensics Tools and Case Studies: Working with Windows and DOS Systems. - Understanding File systems, Exploring Microsoft file structures, Examining NTFS disks, Understanding whole disk encryption, windows registry, Microsoft start-up tasks.

Current Computer Forensics Tools: Software/ Hardware Tools. Computer forensics investigation – A case study

Reference Books

1. Man Young Rhee, “Internet Security: Cryptographic Principles”, “Algorithms and protocols”, Wiley Publications, 2003.

2. Nelson, Phillips, Enfinger, Steuart, “Computer Forensics and Investigations”, Cengage Learning, India Edition, 2008.

3. John R.Vacca, “Computer Forensics”, Cengage Learning, 2005.

4. Richard E.Smith, “Internet Cryptography”, 3rd Edition Pearson Education, 2008.

5. Marjie T.Britz, “Computer Forensics and Cyber Crime”: An Introduction”, 3rd Edition, Prentice Hall, 2013.

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1: Introduction to Computer Forensics**	7 hours
1.1	Computer Forensics Fundamentals	1 hours
1.2	Types of Computer Forensics Technology	1 hours
1.3	Types of Computer Forensics Systems	1 hours
1.4	Intrusion Detection Systems	1 hours
1.5	Data Recovery and Evidence Collection	1 hours
1.6	Forensic duplication and preservation of Digital Evidence	1 hours
1.7	Understanding Computer Investigation	1 hours
2	**Module 2 : Module 2: Evidence Data Gathering**	8 hours
2.1	Data Acquisition	1 hours
2.2	Data Recovery	1 hours
2.3	Evidence Collection	1 hours
2.4	MTypes of Evidence, The Rules of Evidence	1 hours
2.5	Volatile Evidence, General Procedure	1 hours
2.6	Collection and Archiving	1 hours
2.7	Data Seizure	1 hours

2.8	Duplication and preservation of Digital Evidence	1 hours
3	**Module 3: Investigations**	6 hours
3.1	Network Traffic: Investigating Network Intrusions	1 hours
3.2	Network Forensics and Investigating logs	1 hours
3.3	Web Attacks	1 hours
3.4	Router Forensics	1 hours
3.5	DoS Attacks	1 hours
3.6	Internet Crime	1 hours
4	**Module 4:Evidence Data Analysis**	7 hours
4.1	Discovery of Electronic Evidence	1 hours
4.2	Identification of Data	1 hours
4.3	Determining and Validating Forensics Data	1 hours
4.4	Data Hiding Techniques	1 hours
4.5	Performing Remote Acquisition	1 hours
4.6	Cell Phone and Mobile Devices Forensics	1 hours
4.7	Processing Crime and Incident Scenes	1hours
5	**Module 5:Forensics Tools and Case Studies**	8 hours
5.1	Working with Windows and DOS Systems	1 hours
5.2	Understanding File systems	1 hours

5.3	Exploring Microsoft file structures	1 hours
5.4	Examining NTFS disks	1 hours
5.5	Understanding whole disk encryption, windows registry	1 hours
5.6	Microsoft start-up tasks.	1 hours
5.7	Current Computer Forensics Software/Hardware Tools	1 hours
5.8	Case Study	1 hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code : M24CS1E204B

Course Name: Cyber Forensics and Information Security

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Analyse the procedure for corporate high-tech investigations. (4 marks)

2. Give specific instances to illustrate data seizure. (4 marks)

3. Explain the steps involved in investigating router attacks. (4 marks)

4. Illustrate the validation of forensics data. (4 marks)

5. Explain the different hardware tools for computer forensics. (4 marks)

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Analyse the procedure for corporate high-tech investigations (4 marks)
   (b) How is data seizure relevant in computer forensics and evidence collection. (4 marks)

7. Describe the different cyber forensic malicious software. (8 marks)

8. (a) Discuss Duplication and Preservation of Digital Evidence. (4 marks)
   (b) A bank suspected an employee of downloading sensitive files from the bank's computer network using his bank laptop computer from home while on leave of absence. The bank sent the computer for a computer forensic examination. How can this issue be addressed by the professionals? . (4 marks)

9. A new trainee joined a cyber-security firm. His trainer gave a talk about how to poison DNS. So what may be the techniques he has explained to him? (8 marks)

10. A cyber-attack happened in a Govt. Official server and the network has been breached. In this case how will the network forensics and investigation of the logs happen? (8 marks)

11. (a) Explain cyber surveillance and criminal tracking.. (8 marks)

12. (a) Discuss about any two software tools used for cyber forensics. (4 marks)
    (b) Develop a digital forensic investigation model with the idea of Venter that digital forensics investigation can be conducted by even non-technical persons. (4 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E203C	**DEEP LEARNING**	ELECTIVE	3	0	0	3	3

Preamble : Study of this course provides the learners an overview of the concepts and algorithms involved in deep learning. The course covers the basic concepts in deep learning, optimization techniques, regularization techniques, convolutional neural networks, recurrent neural networks, graphical models, deep generative models. This course helps the students to implement deep learning algorithms to solve real-world problems.

Course Outcomes

After the completion of the course the student will be able to:

CO1	Use the standard regularization and optimization techniques for the effective training of deep neural networks. **(Cognitive Knowledge Level: Apply)**
CO2	Build convolutional Neural Network (CNN) models for different use cases.**(Cognitive Knowledge Level: Apply)**
CO3	Apply the concepts of Recurrent Neural Network (RNN), Long Short Term Memory(LSTM), Gated Recurrent Unit (GRU) for solving problems. **(Cognitive Knowledge Level: Apply)**
CO4	Construct Bayesian networks, Markov networks and apply computational techniques to draw inferences.**(Cognitive Knowledge Level: Apply)**
CO5	Illustrate the concepts of auto encoder, sampling algorithms, deep generative models and transfer learning. **(Cognitive Knowledge Level: Apply)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	**Continuous Internal Evaluation Tests**		**End Semester Examination (% Marks)**
	Test 1 (%Marks)	**Test 2 (%Marks)**	
Remember			
Understand			
Apply	70	70	70
Analyse	30	30	30
Evaluate			
Create			

Mark Distribution

Total Marks	**CIE Marks**	**ESE Marks**	**ESE Duration**
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (8 hours)

Introduction to Deep Learning: Deep feed forward network, Training deep models - introduction, setup and initialization issues, Vanishing and exploding gradient problems, Optimization techniques - Gradient Descent (GD), Stochastic GD, GD with momentum, GD with Nesterov momentum, AdaGrad, RMSProp, Adam. Regularization Techniques - L1 and L2 regularization, Early stopping, Dataset augmentation, Parameter tying and sharing, Ensemble methods, Dropout.

MODULE 2 (7 hours)

Convolutional Neural Networks : Architecture, Convolution and Pooling operation, Motivation, Variants of convolution functions, Structured outputs, Data types, Efficient convolution algorithms, Training a Convolutional Network, Applications of Convolutional Networks, Case study of Convolutional Architectures – AlexNet

MODULE 3 (7 hours)

Recurrent Neural Networks : UComputational graphs, RNN design, encoder – decoder sequence to sequence architectures, deep recurrent networks, recursive neural networks, challenges of training Recurrent Networks, gated RNNs LSTM and GRU, Applications of RNNs.

MODULE 4 (5 hours)

Graphical Models and Sampling : Graphical models - Bayesian network, Markov networks, Inference on chains and factor graphs. Monte Carlo Methods – Basics of Monte Carlo Sampling, Importance sampling, Markov chain Monte Carlo methods(MCMC), Gibbs sampling.

MODULE 5 (9 hours)

Advanced Deep learning Topics : Autoencoders, Variational AutoEncoder , Deep generative models - Boltzmann machines, Restricted Boltzmann Machines, Deep Belief Networks, Deep Boltzmann Machines, ,Generative Adversarial Networks, Auto-Regressive Networks. Transfer Learning and Domain Adaptation.

Text Books

1. Goodfellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016
2. Neural Networks and Deep Learning, Aggarwal, Charu C., c Springer International Publishing AG, part of Springer Nature 2018.

Reference Books

1. Christopher M. Bishop. Pattern recognition and machine learning. Springer 2006.

2. David Foster. Generative Deep Learning - Teaching Machines to Paint, Write, Compose, and Play. O'Reilly Media, Inc., June 2019.

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	8 hours
1.1	Introduction to deep learning, Deep feed forward network . Training deep models - Introduction, setup and initialization issues	1 hours
1.2	Vanishing and exploding gradient problems	1 hours
1.3	Concepts of optimization, Gradient Descent (GD)	1 hours
1.4	Stochastic GD, GD with momentum, GD with Nesterov momentum	1 hours
1.5	AdaGrad, RMSProp, Adam	1 hours
1.6	Concepts of Regularization, L1 and L2 regularization	1 hours
1.7	Early stopping, Dataset augmentation	1 hours
1.8	Parameter tying and sharing, Ensemble methods, Dropout	1 hours
2	**Module 2**	7 hours
2.1	Convolutional Neural Networks, Architecture	1 hours

2.2	Convolution and Pooling operation with example	1 hours
2.3	Motivation, Variants of convolution functions	1 hours
2.4	Structured outputs, Data types, Efficient convolution algorithms	1 hours
2.5	Training a Convolutional Network	1 hours
2.6	Applications of Convolutional Networks	1 hours
2.7	Case study of Convolutional Architectures – AlexNet	1 hours
3	**Module 3**	7 hours
3.1	Recurrent neural networks – Computational graphs	1 hours
3.2	RNN design, Encoder – decoder sequence to sequence architectures	1 hours
3.3	Deep recurrent networks, Recursive neural networks	1 hours
3.4	Challenges of training Recurrent Networks	1 hours
3.5	LSTM	1 hours
3.6	GRU	1 hours
3.7	Applications of RNN	1 hours
4	**Module 4**	5 hours
4.1	Graphical models - Bayesian network	1 hours
4.2	Markov network	1 hours
4.3	Inference on chains and factor graphs	1 hours
4.4	Monte Carlo Methods – Basics of Monte Carlo Sampling, Importance sampling	1 hours
4.5	Markov chain Monte Carlo methods(MCMC), Gibbs sampling	1 hours

5	**Module 5**	9 hours
5.1	Autoencoders	1 hours
5.2	Variational Autoencoder	1 hours
5.3	Deep generative models - Boltzmann machines	1 hours
5.4	Restricted Boltzmann Machines	1 hours
5.5	Deep Belief Networks	1 hours
5.6	Deep Boltzmann Machines	1 hours
5.7	Generative Adversarial Networks	1 hours
5.8	Auto-Regressive Networks	1 hours
5.9	Transfer Learning and Domain Adaptation.	1 hours

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

SECOND SEMESTER M.TECH DEGREE EXAMINATION,DECEMBER 2024

Course Code: M24CS1E203C

Course Name: Deep Learning

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. A 2×2 image is represented by the following pixel value matrix.

$$\begin{bmatrix} 5 & 4 \\ 2 & 7 \end{bmatrix}$$

   This image is given to a 3-layer neural network, that is, two hidden layers and one output layer. Draw a schematic diagram of the network. Assuming all inter-connection weights having values 1, bias having value 0, the hidden layers having 3 neurons each, and a simple activation function of the form 11+x being used, compute output for one round of forward propagation.

2. In Convolutional Neural Networks, there is no need to perform feature extraction. Justify with an example.

3. Explain your understanding of unfolding a recursive or recurrent computation into a computational graph. List three differences between LSTM and GRU.

4. Sketch the core idea of the Monte Carlo method. What is a sample? What is a direct sampling method? Why can't it be used directly to do any inference? What is rejection sampling? What is its major disadvantage?

5. How does the variational auto-encoder(VAE) architecture allow it to generate new data points, compared to auto-encoder, which cannot generate new data points?

PART B

Answer any five question from each module. Each question carries 8 marks.

6. Differentiate gradient descent with and without momentum. Give equations for weight updation in GD with and without momentum. Illustrate plateaus, saddle points and slowly varying gradients. (8 marks)

7. (a) Consider an activation volume of size $13\times13\times64$ and a filter of size $3\times3\times64$. Discuss whether it is possible to perform convolutions with strides 2, 3 and 5. What happens if the stride of the convolutional layer increases? What can be the maximum stride? Justify your answer. (4 marks)

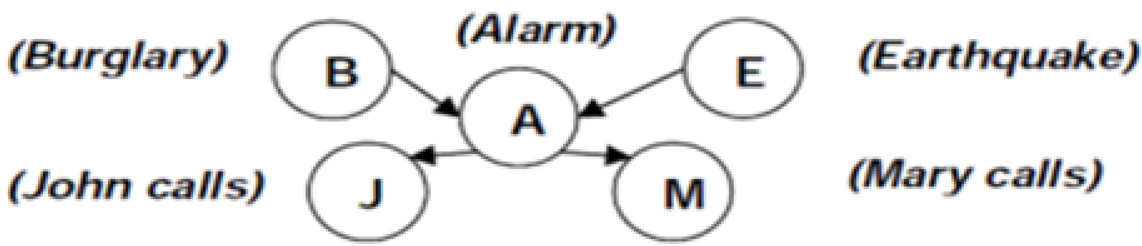
(b) How backpropagation is used to learn higher-order features in a convolutional (4 marks)

8. The vanishing gradient problem is more pronounced in RNN than in traditional neural networks. Give the reason. Discuss a solution for the problem (8 marks)

9. Shown below is the Bayesian network corresponding to the Burglar Alarm problem, P(J | A) P(M | A) P(A | B, E) P(B) P(E). The probability tables show the probability that variable is True, e.g., P(M) means P(M= t). Calculate

i $P(J \wedge M \wedge A \wedge \neg B \wedge \neg E)$

ii $P(J)$



P(E)
.002

A	P(M)
t	.70
f	.01

B	E	P(A)
t	t	.95
t	f	.94
f	t	.29
f	f	.001

P(B)
.001

A	P(J)
t	.90
f	.05

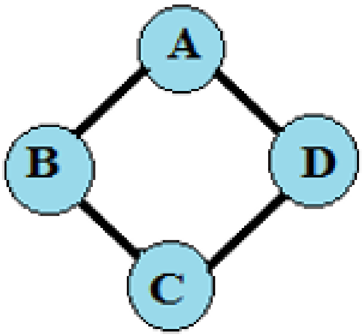
10. Compare Boltzmann Machine with Deep Belief Network. (8 marks)

11. (a) Training error of the deep learning model trained for the classification problem was found to be very low but generalization error was high. Identify the problem and suggest techniques to reduce this generalization error. (4 marks)

(b) Initializing the weights of a neural network with very small or large random numbers is not advisable. Justify (4 marks)

12. Consider the simple Markov network given below. Let A,B,C and D be binary random variables representing four people's beliefs as to whether the earth is round (1 for believes, 0 for does not believe). Determine the probability of only person D believes that the earth is round. (8 marks)

A	B	ϕ_{AB}
0	0	50
0	1	5
1	0	5
1	1	50



A	D	ϕ_{AD}
0	0	5
0	1	50
1	0	50
1	1	5

B	C	ϕ_{BC}
0	0	1
0	1	5
1	0	45
1	1	50

C	D	ϕ_{CD}
0	0	1
0	1	15
1	0	40
1	1	50

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E203D	**WIRELESS SENSOR NETWORK**	ELECTIVE	3	0	0	3	3

Preamble : To understand the fundamentals of wireless sensor networks and its application to critical real time scenarios**.**The course deals with hardware aspects like sensors with their built in transceivers and associated electronics as well as the software's needed to run them. The MAC layer with its myriad protocols, as well as routing layer protocols, are discussed in the course. Localization of sensor nodes, deployment and coverage, as well as security of wireless sensor networks also form part of the course. This course enables the learners to design and develop wireless sensor protocols and applications.

Prerequisite : NIL

Course Outcomes

After the completion of the course the student will be able to:

CO 1	List the applications, hardware and software components of wireless sensor networks, and the challenges faced in design of sensor networks. **(Cognitive knowledge: Analyze)**
CO 2	Apply design principles and formulate necessary service interfaces while designing a wireless sensor network **(Cognitive knowledge: Apply)**
CO 3	Design MAC protocols for wireless sensor networks taking into account the specific requirements of the network. (Cognitive knowledge: Create) Design MAC protocols for wireless sensor networks taking into account the specific requirements of the network. **(Cognitive knowledge: Create)**
CO 4	Explain localization techniques, coverage problem and security issues in wireless sensor networks **(Cognitive knowledge: Evaluate)**
CO 5	Design energy efficient routing protocols for wireless sensor networks **(Cognitive knowledge: Create)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓		✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand			
Apply	30%	30%	30%
Analyse	40%	40%	40%
Evaluate	30%	30%	
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 :Introduction to Wireless Sensor Networks (8 Hours)

Motivation, Challenges and Constraints, Applications. Node Architecture – Hardware elements, Sensors and Actuators, Power supply, EnergyConsumption of sensor nodes. Challenges in sensor network programming, Operating systems and execution environments-embedded OS, issues, Programming models.

MODULE 2 : Sensor Network Architecture(6 Hours)

Sensor network scenarios, Optimization goals and figures of merit,Design principles for WSNs, Service interfaces of WSNs. Gateway concepts, WSN- Internet Communication. Discrete time Markov Chain, Performance analysis of IEEE 802.11 DCF using Markov Chain

MODULE 3 : MAC Protocols for WSNs(7 Hours)

Characteristics and design goals, low duty cycle protocols and wakeup concepts.Contention based protocols: PAMAS, STEM, T-MAC,Schedule based protocols: SMACS, TRAMA, Hybrid MAC protocols: Z-MAC Case Studies: S-MAC, 802.15.4, 802.15.6

MODULE 4 : Routing Protocols for WSNs(7 Hours)

Introduction, Routing challenges and design issues, Routing Strategies Routing Techniques: Flooding and Gossiping, SPIN, LEACH, PEGASIS, Directed Diffusion, Rumour Routing, Geographic Routing- Forwarding Strategies.

MODULE 5 : Localization, Coverage and Security in WSNs (8 Hours)

Localization: approaches-proximity-trilateration and triangulation- scene analysis Coverage and deployment: sensing models, coverage measures, uniform random deployments,coverage determination Security: Security challenges in WSNs, Security attacks in WSNs.

Text Books

1. W. Dargie and C. Poellabauer, Fundamentals of Wireless Sensor Networks, Theory and Practice,Wiley, 2010
2. Holger Karl, Andreas Willig, “Protocols and architectures for wireless sensor networks”, John wiley & Sons.

Reference Books

1. K. Sohraby, D. Minoli and T. Znati, Wireless Sensor Networks, Technology, Protocols, and Applications,Wiley-Interscience, 2007
2. C. Siva Ram Murthy and B. S Manoj, Adhoc Wireless Networks Architectures and Protocols,Prentice Hall, 2004

3. Feng Zhao and Leonidas Guibas, Wireless Sensor Networks An Information Processing Approach, Morgan Kaufman, 2005

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1: Introduction to Wireless Sensor Networks g**	8 hours
1.1	Motivation, Challenges and Constraints	1 hours
1.2	Applications, Node Architecture	1 hours
1.3	Hardware elements	1 hours
1.4	Sensors and Actuators, Power supply	1 hours
1.5	Energy Consumption of sensor nodes	1 hours
1.6	Challenges in sensor network programming	1 hours
1.7	Operating systems and execution environments- embedded OS, issues	1 hours
1.8	programming models	1 hours
2	**Module 2: Sensor Network Architecture**	6 hours

2.1	Sensor network scenarios	1 hours
2.2	Optimization goals and figures of merit	1 hours
2.3	Design principles for WSNs	1 hours
2.4	Service interfaces of WSNs	1 hours
2.5	Gateway concepts, WSN- Internet Communication	1 hours
2.6	Discrete time Markov Chain, Performance analysis of IEEE 802.11 DCF using Markov Chain	1 hours
3	**Module 3 : MAC Protocols for WSNs**	7 hours
3.1	Characteristics and design goals	1 hours
3.2	Low duty cycle protocols and wakeup concepts	1 hours
3.3	Contention based protocols: PAMAS	1 hours
3.4	STEM, T-MAC	1 hours
3.5	Schedule based protocols: SMACS TRAMA	1 hours
3.6	Hybrid MAC protocols: Z-MAC	1 hours
3.7	Case Studies: S-MAC 1 , 802.15.4, 802.15.6	1 hours
4	**Module 4:Routing Protocols for WSNs**	7 hours
4.1	Introduction, Routing challenges and design issues	1 hours

4.2	Routing Strategies, Routing Techniques: Flooding and Gossiping, SPIN	1 hours
4.3	LEACH, PEGASIS	1 hours
4.4	Directed Diffusion	1 hours
4.5	Rumour Routing	1 hours
4.6	Geographic Routing	1 hours
4.7	Forwarding Strategies	1 hours
5	**Module 5 :Localization, Coverage and Security in WSNs**	7 hours
5.1	Localization: approaches-proximity-trilateration	1 hours
5.2	Triangulation, scene analysis	1 hours
5.3	Coverage and deployment: sensing models	1 hours
5.4	Coverage measures	1 hours
5.5	Uniform random deployments	1 hours
5.6	Coverage determination	1 hours
5.7	Security: Security challenges in wireless sensor networks	1 hours
5.8	Security attacks in wireless sensor networks	1 hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code : M24CS1E203D

Course Name: WIRELESS SENSOR NETWORKS

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Describe important applications of wireless sensor networks.
2. Differentiate between single hop and multihop wireless sensor networks.
3. How does T-MAC take care of the shortcomings of S-MAC? Comment on TMAC's ability to adapt to traffic density.
4. Evaluate security challenges in wireless sensor networks?
5. Explain any three ways in which gossiping overcomes the challenges of flooding.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Explain how to calculate the amount of energy saved if a sensor node is allowed to sleep (5 marks)

   (b) A cross layered approach is more suitable for wireless sensor networks than a layered approach. Explain (3 marks)

7. (a) Demonstrate the different types of in-network processing.

   (8marks)

8. (a) Show how the design of the MAC protocol affects the energy efficiency of a sensor node. (3 marks)

   (b) What are the advantages and disadvantages of TRAMA protocol? What is the purpose of NP component? (5 marks)

9. (a) Why is overhearing a problem in a wireless sensor network. How does PAMAS solve this problem ?

   (3 marks)

   (b) Define "early sleeping problem" and discuss the solution provided by TMAC.

   (5 marks)

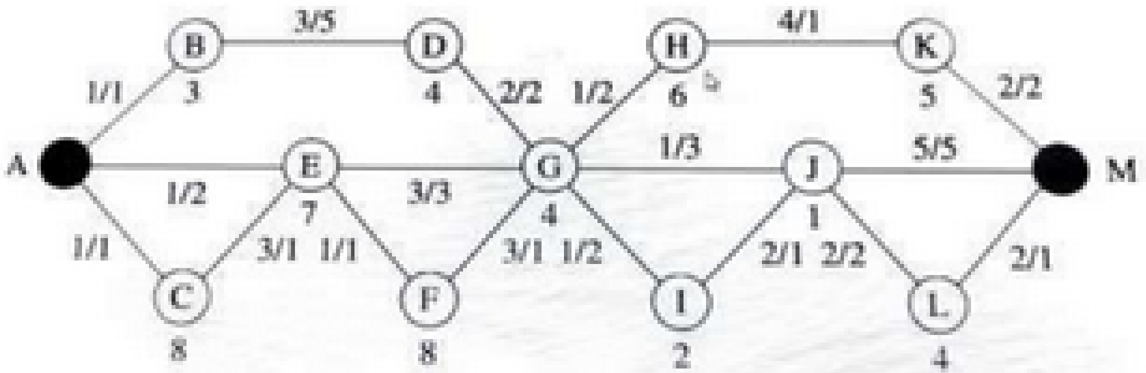
10. (a) For each link in the above figure, the ratio gives latency/energy cost for transmitting a single packet, while the number under each node gives the node's remaining energy capacity.

    Node G's routing table is given as below. G sends queries toward events E1, E2 using rumor routing. Show how it is made possible. (5 marks)

EVENT	DISTANCE	DIRECTION
E1	3	F
E2	4	I

11. (a) Why is localization needed in wireless sensor networks? Give two methods for localization in wireless sensor networks.

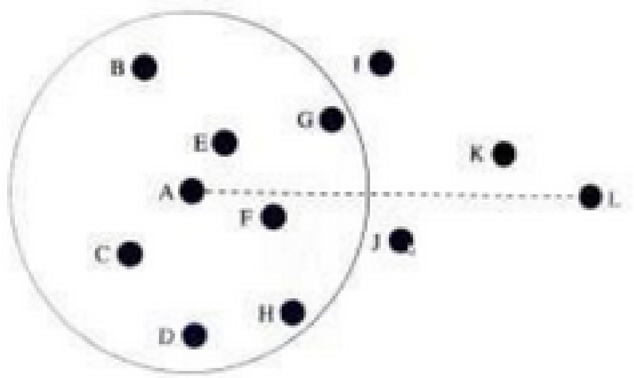
    (5 marks)



(b) Discuss a method to find the maximum breach path in a wireless sensor network.

12. (a) Enumerate the issues and challenges while designing a routing protocol for wireless sensor network (3 marks)

(b) In the below figure, if node A wants to forward a packet to node L using greedy forwarding, which neighbor should it choose? The communication range of node A is given by the circle (5 marks)



CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E203E	**MACHINE LEARNING IN COMPUTATIONAL BIOLOGY**	ELECTIVE	3	0	0	3	3

Preamble : This course is intended to provide the learners a outlook towards application of Machine learning algorithms in the field of computational biology. This course covers Machine learning methods likes clustering algorithms, dimensionality reduction, decision drees, Artificial Neural Network, Support Vector Machine and its application in computational biology. By the end of the course, students will have gained the skills to apply machine learning methods to complex problems in computational biology.

Prerequisites : Basic background in Bioinformatics and Machine Leaning

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Describe the basic concepts of Machine Leaning, Bioinformatics and apply classification algorithm on computational biology problems. **(Cognitive Knowledge Level : Apply)**
CO 2	Demonstrate the clustering algorithms and Artificial Neural Network on computational biology problems. **(Cognitive Knowledge Level: Apply)**
CO 3	Illustrate Dimensionality reduction, Feature Extraction and Pattern recognition in the domain of Computational Biology analysis communication.**(Cognitive Knowledge Level: Apply)**
CO 4	Apply Hidden Markov Models (HMMs), Genetic Algorithms (GAs), and Support Vector Machines (SVMs) to solve complex problems in computational biology.**(Cognitive Knowledge Level: Apply)**
CO 5	Explain the role and challenges of Machine Learning in Computational biology scenario.**(Cognitive Knowledge Level: Understand)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	30	30	30
Apply	70	70	70
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (8 hours)

Overview of Machine Learning and Bioinformatics: Supervised and unsupervised learning, Role of Machine Learning Technique, Introduction to Bioinformatics, Databases, Sequence Alignment, Proteins and peptides, Computational Biology and Bioinformatics.
Classification: Protein Classification, Decision Trees in Bioinformatics, Proteomic Mass Spectra Classification Using Decision Tree Technique, Drug Repositioning Using

Classification Approach.

MODULE 2 (8 hours)

Clustering: Hierarchical Clustering, Partition Clustering, Clustering for creating phylogenetic trees, Using Clustering Approach to Identify Patients' Subtypes, Application of clustering algorithms on gene expression data.
Artificial Neural Network: Artificial Neural Network (ANN) in Bioinformatics, Designing ANN for Bioinformatics, ANN in Protein Bioinformatics

MODULE 3 (8 hours)

Dimensionality reduction: Proteomics Dataset, Data Pre-processing Algorithms, Dimension and Feature Subset Selection, Principal Component Analysis (PCA), Partial Least Square (PLS), Linear Discriminant Analysis (LDA), Machine-Learning Algorithms for Feature Selection from Gene Expression Data, Feature Extraction and Pattern recognition from sequence data, measures of a Feature.

MODULE 4 (7 hours)

Markov Models: Hidden Markov Models, Forward and Backward Algorithms, Most probable state path: Viterbi algorithm, Parameter Estimation for HMMs:-Baum-Welch Algorithm, Applications of profile HMMs for multiple alignment of proteins and for finding genes in the DNA.
Genetic Algorithms (GA) in Bioinformatics, Support Vector Machine with Feature Elimination.

MODULE 5 (5 hours)

Issues Related to Machine Learning in Bioinformatics: Data Errors in Machine Learning, Mean Square Error, Generative versus Discriminative, Approximation Versus Explanation, Single Versus Multiple Methods, Future directions of Machine Learning in Computational Biology.

Text Books

1. Statistical Modelling and Machine Learning Principles for Bioinformatics Techniques,

Tools, and Applications. Germany, Springer Singapore, 2020.

2. Yang, ZhengRong. Machine Learning Approaches to Bioinformatics. Singapore, World Scientific Publishing Company, 2010.

Reference Books

1. Izadkhah, Habib. Deep Learning in Bioinformatics: Techniques and Applications in Practice. Netherlands, Elsevier Science, 2022.
2. Agapito, Giuseppe, et al. Artificial Intelligence in Bioinformatics: From Omics Analysis to Deep Learning and Network Mining. Netherlands, Elsevier Science, 2022.
3. Data Analytics in Bioinformatics: A Machine Learning Perspective. United States, Wiley, 2021.
4. Michailidis, George, et al. Introduction to Machine Learning and Bioinformatics. United Kingdom, CRC Press, 2008.
5. Zhang, Yanqing, and Rajapakse, Jagath C, Machine Learning in Bioinformatics, Germany, Wiley, 2009.
6. Baldi, Professor Pierre, et al. Bioinformatics, Second Edition: The Machine Learning Approach. India, Bradford, 2001.

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	8 hours
1.1	Supervised and unsupervised learning	1 hour

1.2	Role of Machine Learning Technique	1 hour
1.3	Introduction to Bioinformatics, Databases, Sequence Alignment	1 hour
1.4	Proteins and peptides, Computational Biology and Bioinformatics	1 hour
1.5	Protein Classification	1 hour
1.6	Decision Trees in Bioinformatics	1 hour
1.7	Proteomic Mass Spectra Classification Using Decision Tree Technique	1 hour
1.8	Drug Repositioning Using Classification Approach	1 hour
2	**Module 2**	8 hours
2.1	Hierarchical Clustering	1 hour
2.2	Partition Clustering	1 hour
2.3	Clustering for creating phylogenetic trees	1 hour
2.4	Using Clustering Approach to Identify Patients' Subtypes	1 hour
2.5	Application of clustering algorithms on gene expression data	1 hour
2.6	Artificial Neural Network (ANN) in Bioinformatics	1 hour
2.7	Designing ANN for Bioinformatics	1 hour
2.8	ANN in Protein Bioinformatics	1 hour
3	**Module 3**	8 hours
3.1	Proteomics Datasets	1 hour

3.2	Data Pre-processing Algorithms	1 hour
3.3	Dimension and Feature Subset Selection	1 hour
3.4	Principal Component Analysis (PCA)	1 hour
3.5	Partial Least Square (PLS), Linear Discriminant Analysis (LDA)	1 hour
3.6	Machine-Learning Algorithms for Feature Selection from Gene Expression Data	1 hour
3.7	Feature Extraction and Pattern recognition from sequence data	1 hour
3.8	measures of a Feature	1 hour
4	**Module 4**	7 hours
4.1	Hidden Markov Models	1 hour
4.2	Forward and Backward Algorithms	1 hours
4.3	Most probable state path: Viterbi algorithm,	1 hour
4.4	Parameter Estimation for HMMs:-Baum-Welch Algorithm	1 hour
4.5	Applications of profile HMMs for multiple alignment of proteins and for finding genes in the DNA	1 hour
4.6	Genetic Algorithms (GA) in Bioinformatics	1 hour
4.7	Support Vector Machine with Feature Elimination	1 hour
5	**Module 5**	5 hours
5.1	Data Errors in Machine Learning, Mean Square Error	1 hour

5.2	Generative versus Discriminative	1 hour
5.3	Approximation Versus Explanation	1 hours
5.4	Single Versus Multiple Methods	1 hour
5.5	Future directions of Machine Learning in Computational Biology	1 hour

MODEL QUESTION PAPER

QP CODE: **Pages: 3**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E203E

Course Name: MACHINE LEARNING IN COMPUTATIONAL BIOLOGY

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. What does the regression line equation tell you?
2. Illustrate the design of Artificial Neural Network for solving Computational Biology question.

3. Explain the process involved in feature extraction and pattern recognition from sequence data.

4. In what ways are profile HMMs superior to traditional sequence alignment methods for multiple sequence alignment?

5. Explain different kinds of Data Errors in Machine Learning that would happen in case of applying it in to the Computational Biology domain?

PART B

Answer any five question from each module. Each question carries 8 marks.

6. How do decision tree classifiers work? what types of problems can they solve in Computational Biology. (8 marks)

7. Illustrate with a toy example the application of clustering algorithms on gene expression data.

   (8 marks)

8. (a) 20 physicochemical properties of 100 set of proteins were given with the help of PCA, explain how will you reduce 20x100 in to Five properties (5x100) for the next level analysis. (5 marks)

   (b) What is the importance of using PCA before applying Machine learning method?

   (3 marks)

9. (a) How can Genetic Algorithms be integrated with other machine learning techniques in bioinformatics for enhanced data analysis.

   (4 marks)

   (b) What role does the kernel function play in the performance of SVMs when dealing with high-dimensional biological data? (4 marks)

10. The transformation of huge volume of data into knowledge is the biggest challenge faced in computational biology" How can machine learning techniques help in this?

    (8 marks)

11. Design and implement an ANN model for the prediction of relative solvent accessibility.

    (8 marks)

12. Explain how Linear Discriminant Analysis can be used for the dimensionality reduction with the help of a scenario in computational biology.

    (8 marks)

SEMESTER II

PROGRAM ELECTIVE IV

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E204A	**SOFTWARE ARCHITECTURE AND DESIGN PATTERNS**	ELECTIVE	3	0	0	3	3

Preamble : This course helps the students to learn the basic knowledge of software, software development process and the concepts of software design principles.It contain concepts of Software and the software design process, Process models, Importance of software architecture, Software design principles, Software Architectural styles, Software Architecture patterns, Evaluation of architectural design. Students will be able to impart knowledge on the different architectural styles and architectural patterns for the software.

Prerequisites : Nil

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Understand the concept of software architecture, its importance, and the role of a software architect in the development process.(Cognitive Level: Understand)
CO 2	Analyze trade-offs among various design principles to achieve optimal software design.(Cognitive Level: Apply)
CO 3	Design UML diagram for the software(Cognitive Level: Apply)
CO 4	Make use of design pattern for the development of software systems.(Cognitive Level: Apply)
CO 5	Explore behavioral design patterns like Chain of Responsibility, Command, Interpreter, Mediator, State, Template, and Observer.(Cognitive Level: Apply)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓		✓
CO 2	✓		✓	✓	✓	
CO 3	✓		✓	✓	✓	
CO 4	✓		✓	✓	✓	
CO 5	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	20	20	10
Apply	40	40	25
Analyse	40	40	25
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (8 hours)

The Nature of Software – Defining software, Software Application domains. Software Engineering. The Software Process, Process Models: - Water fall model, Incremental model, Prototyping model, RAD, Spiral model and Agile Development. (Brief introduction of Agile Development).Software Architecture, importance of software architecture, Role of software architect.

MODULE 2 (7 hours)

Software Design principles – Correctness and Robustness – Flexibility, Reusability and Efficiency – Tradeoffs among robustness, flexibility, reusability and efficiency

MODULE 3 (7 hours)

Introduction to UML diagrams – Use case diagrams, Class diagrams, Sequence diagrams, Activity diagrams, State Transition diagram, Deployment diagram.

Software architectural styles - pipes & filters, layered, eventbased, data-centered, interpreter, MVC, message dispatcher, multi-tier distributed

MODULE 4 (9 hours)

Design pattern. Creational patterns – Factory, Abstract Factory, Prototype and Singleton. Structural patterns – Composite, Decorator, Adapter, Façade and Flyweight.

MODULE 5 (6 hours)

Behavioral patterns- Chain of responsibility, Command, Interpreter, Mediator, State, Template and Observer. Evaluation of architectural design - ATAM

Reference Books

1. Eric J. Braude , Software Design, John Wiley and Sons.
2. Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides : Design Patterns: Elements of Reusable Object-Oriented Software, Addison – Wesley, 1994
3. James Rumbaugh, Object Oriented Modeling and Design, Prentice Hall India
4. Len Bass, Paul Clements, Rick Kazman, Software Architecture in Practice (2nd Ed.), Pearson
5. Mary Shaw & David Garlan, Software Architecture – Perspectives on an emerging discipline, Pearson, 1996
6. Roger S.Pressman, Software Engineering – A Practitioner's approach, 8th edition(2014), McGraw Hill Education

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	8 hours
1.1	The Nature of Software – Defining software, Software Application domains.	1 hours
1.2	Software Engineering.	1 hours
1.3	The Software Process, Process Models: - Water fall model, Incremental model, (Brief introduction of Agile Development).	1 hours
1.4	Prototyping model, RAD,	1 hours
1.5	Spiral model and Agile Development.(Brief introduction of Agile Development).	1 hours
1.6	Software Architecture	1 hours
1.7	Importance of software architecure	1 hours
1.8	Role of software architect.	1 hours
2	**Module 2**	7 hours
2.1	Software Design principles	1 hours
2.2	Software Design principles	1 hours
2.3	Correctness and Robustness	1 hours
2.4	Flexibility, Reusability and Efficiency	1 hours
2.5	Flexibility, Reusability and Efficiency	1 hours
2.6	Tradeoffs among robustness, flexibility, reusability and efficiency	1 hours

2.7	Tradeoffs among robustness, flexibility, reusability and efficiency	1 hours
3	**Module 3**	7 hours
3.1	Introduction to UML diagrams	1 hours
3.2	Use case diagrams, Class diagrams	1 hours
3.3	Sequence diagrams,Activity diagrams,	1 hours
3.4	State Transition diagram, Deployment diagram.	1 hours
3.5	Software architectural styles - pipes filters, layered,	1 hours
3.6	eventbased, data-centered, interpreter,	1 hours
3.7	MVC, message dispatcher, multi-tier distributed	1 hours
4	**Module 4**	8 hours
4.1	Design pattern,Creational patterns – Factory	1 hours
4.2	Abstract Factory	1 hours
4.3	Prototype and Singleton.	1 hours
4.4	Prototype and Singleton.	hours
4.5	Structural patterns	hours
4.6	Composite,	hours
4.7	Decorator, Adapter	hours
4.8	Façade and Flyweight.	1 hours
5	**Module 5**	6 hours
5.1	Behavioral patterns- Chain of responsibility	1 hours
5.2	Command, Interpreter	1 hours
5.3	Mediator, State	1 hours

5.4	Template and Observer.	1 hours
5.5	Evaluation of architectural design - ATAM	1 hours
5.6	Evaluation of architectural design - ATAM	1 hours
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E204A

Course Name: :SOFTWARE ARCHITECTURE AND DESIGN PATTERNS

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Explain about Umbrella Activities of software process.
2. Explain about different Process Flows.
3. List different use case relationships
4. Explain software architectural style .list some examples.
5. Explain sequence diagram of Prototype Pattern.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Describe Water Fall Model and Incremental Model. (4 marks)

   (b) Discuss about Spiral and RAD Model with neat diagram. (4 marks)

7. (a) Explain about Informal and formal approaches to correctness. (4 marks)

   (b) Prepare Sequence diagram for Two Party Mobile Phone Call. (4 marks)

8. (a) List different use case relationships (4 marks)

   (b) Explain event based styles and data-cantered styles. (4 marks)

9. (a) Construct an observer design pattern for converting a decimal value into Binary, Hexadecimal and octal values. (8 marks)

10. (a) Explain with example Façade Structural pattern. Consider a simple banking example that allows users to make deposit to customer accounts and obtain balance. (4 marks)

    (b) Prepare Sequence diagram for Two Party Mobile Phone Call. (4 marks)

11. (a) Explain Abstract Factory Method. Suppose that there are three different shapes rectangle, square and circle. The shapes are available in three different colours namely red, blue, green .Draw the Abstract Factory to represent to draw shape having any of above three colours. (8 marks)

12. (a) Explain chain of responsibility. Consider a situation in job portal application, where students update their profile. Prepare chain of responsibility class model. (8 marks)

12. (a) Construct an observer design pattern for converting a decimal value into Binary, Hexadecimal and octal values. (8 marks)

12. (a) Explain Mediator class model in Java API. (8 marks)

Code	Course Name	CATEGORY	L	T	P	S	CREDIT
M24CS1E204B	**DATA PRIVACY AND SECURITY**	ELECTIVE	3	0	0	3	3

Preamble : The course provides an in-depth understanding of data privacy and security in technology-enabled environments and it focuses on technological solutions, methods and practices for data protection in business organizations and peer to peer networks. This provides students with an understanding of strategic and regulatory issues of data privacy and security in modern organizing as well as an overview of mechanisms for privacy and security assurance in enterprise architectures and projects.

Prerequisites : Understanding of Database, Cryptography and Network Security.

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Understand the fundamental concepts of data security and privacy.**(Cognitive Level: Understand)**
CO 2	Analyze security and privacy threats to data assets, including the privacy level of various data release mechanisms, privacy-utility trade-offs, and statistical inference attacks to infer sensitive information.**(Cognitive Level: Apply)**
CO 3	Apply the basic concepts of hiding data in text and images.**(Cognitive Level: Apply)**
CO 4	Understand the concepts of privacy, authentication, web and email security.**(Cognitive Level: Understand)**
CO 5	Make use of different technology for preserving privacy.**(Cognitive Level: Apply)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1				✓		✓
CO 2	✓		✓	✓	✓	✓
CO 3				✓		✓
CO 4				✓		✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	40	40	40
Apply	60	60	60
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

Fundamentals of Data Privacy & Security: Databases and Exploratory Data Analysis, Data Representation and Storage, Authentication and Authorization, Database Security, Understanding Privacy: Social Aspects of Privacy, Legal Aspects of Privacy and Privacy Regulations, Effect of Database and Data Mining technologies on privacy.

MODULE 2 (7 hours)

Data Privacy Models and Anonymization: Linkage and re-identification attacks, k-anonymity, l-diversity, t-closeness, differential privacy Implementing anonymization, Anonymizing complex data, Privacy and anonymity in mobile environments

MODULE 3 (8 hours)

Data Hiding in Text: Basic Features, Applications of Data Hiding, Watermarking, Intuitive Methods, Simple Digital Methods, Data Hiding in Text, Innocuous Text, Mimic Functions. Data Hiding in Images: LSB Encoding, BPCS Steganography, Lossless Data Hiding, Spread Spectrum Steganography, Data Hiding by Quantization, Patchwork , Signature Casting in Images, Transform Domain Methods, Robust Data Hiding in JPEG Images, Robust Frequency Domain Watermarking, Detecting Malicious Tampering.

MODULE 4 (8 hours)

Privacy: Privacy Concepts, Privacy Principles and Policies, Authentication and Privacy, Data Mining, Privacy on the Web, E-Mail Security, Impacts on Emerging Technologies. Legal and Ethical Issues in Computer Security: Protecting Programs and Data, Information and the Law, Rights of Employees and employers, Redress for Software Failures, Computer Crime, Ethical Issues in Computer Security.

MODULE 5 (6 hours)

Using technology for preserving privacy: Statistical Database security, Inference Control, Secure Multiparty computation and Cryptography, Privacy preserving Data mining, Hippocratic databases

Reference Books

1. Data Privacy and Security by Salomon, David, Springer, 2003.

2. Security in Computing by Charles Pfleeger, Shari Lawrence Pfleeger, 5th Edition, PHI,2015.

3. Data Privacy Principles and Practice by Nataraj Venkataramanan, Ashwin Shriram, 2016.

4. Data Privacy: Foundations, New Developments and Big Data Challenge by VicencTorra, Springer, 2017.

5. Differential Privacy and Applications by TianqingZhu, GangLi, WanleiZhou, Philip S. Yu, Springer, 2017.

6. Security, Privacy, and Trust in Modern Data Management, by Milan Petkovic, Willem Jonker,2007

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	7 hours
1.1	Fundamentals of Data Privacy & Security: Databases and Exploratory Data Analysis	1 hour
1.2	Data Representation and Storage	1 hour
1.3	Authentication and Authorization	1 hour
1.4	Database Security	1 hour
1.5	Understanding Privacy: Social Aspects of Privacy	1 hour
1.6	Legal Aspects of Privacy and Privacy Regulations	1 hour
1.7	Effect of Database and Data Mining technologies on privacy	1 hour
2	**Module 2**	7 hours

2.1	Data Privacy Models and Anonymization: Linkage and re-identification attacks	1 hour
2.2	k-anonymity	1 hour
2.3	l-diversity, t-closeness	1 hour
2.4	Differential privacy	1 hour
2.5	Implementing anonymization	1 hour
2.6	Anonymizing complex data	1 hour
2.7	Privacy and anonymity in mobile environments	1 hour
3	**Module 3**	8 hours
3.1	Data Hiding in Text: Basic Features, Applications of Data Hiding	1 hour
3.2	Watermarking, Intuitive Methods, Simple Digital Methods	1 hour
3.3	Data Hiding in Text, Innocuous Text, Mimic Functions	1 hour
3.4	Data Hiding in Images: LSB Encoding, BPCS Steganography	1 hour
3.5	Lossless Data Hiding, Spread Spectrum Steganography	1 hour
3.6	Data Hiding by Quantization, Patchwork , Signature Casting in Images	1 hour
3.7	Transform Domain Methods, Robust Data Hiding in JPEG Images	1 hour

3.8	Robust Frequency Domain Watermarking, Detecting Malicious Tampering	1 hour
4	**Module 4**	8 hours
4.1	Privacy: Privacy Concepts, Privacy Principles and Policies	1 hour
4.2	Authentication and Privacy, Data Mining,	1 hour
4.3	Privacy on the Web, E-Mail Security	1 hour
4.4	Impacts on Emerging Technologies	1 hour
4.5	Legal and Ethical Issues in Computer Security: Protecting Programs and Data, Information and the Law	1 hour
4.6	Rights of Employees and employers	1 hour
4.7	Redress for Software Failures, Computer Crime	1 hour
4.8	Ethical Issues in Computer Security	1 hour
5	**Module 5**	6 hours
5.1	Using technology for preserving privacy: Statistical Database security	1 hour
5.2	Inference Control	1 hour
5.3	Secure Multi-party computation and Cryptography	1 hour
5.4	Privacy-preserving Data mining	2 hour
5.5	Hippocratic databases	1 hour
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E204B

Course Name: DATA PRIVACY AND SECURITY

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. What are the challenges of data privacy?
2. Explain the properties of Differential Privacy.
3. Describe steganographic algorithm with the help of a diagram.
4. Explain the privacy principles to describe the rights of individuals.
5. What is inference control in databases.

PART B

Answer any five question from each module. Each question carries 8 marks.

7. What is database security and why is it important? (8 marks)

8. Explain different privacy models used to represent data anonymization. (8 marks)

9. Explain the threats caused by linkage and reidentification attacks. What are the protection mechanisms against that. (8 marks)

10. Describe LSB encoding technique. (8 marks)

11. Explain the procedure for data hiding in BPCS steganography. (8 marks)

12. Explain the technologies which are frequently used to monitor web surfing activities without the user's knowledge and how we can take precautions against these activities. (8 marks)

13. (a) Differentiate between law and ethics. (3 marks)

    (b) Illustrate the ways of redress for software failures. (5 marks)

14. Describe the problems and techniques in privacy-preserving data mining (8 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E204C	**WEB MINING**	ELECTIVE	3	0	0	3	3

Preamble : This course introduces the web mining backgrounds, the concepts of Information retrieval, Structured Data Extraction in web structure and usage mining and Web search with special emphasis on Web Crawling. This course helps the learner to use various aspects of web usage mining.

Prerequisites : NIL

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Explain data mining process and techniques, specifically those that are relevant to Web mining. (Cognitive Knowledge Level: Understand)
CO 2	Identify the use of Social Networks Analysis in Web Mining. (Cognitive Knowledge Level: Apply)
CO 3	Describe the basics of Information retrieval and Web search with special emphasis on Web Crawling. (Cognitive Knowledge Level: Understand)
CO 4	Develop the role of Structured Data Extraction in web structure mining . (Cognitive Knowledge Level: Apply)
CO 5	Illustrate the various aspects of web usage mining.(Cognitive Knowledge Level: Understand)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓		✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓		✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand			
Apply	70	60	70-80
Analyse	30	40	30-40
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (6 hours)

Introduction – Web Mining – Theoretical background – Association rule mining – Sequential Pattern Mining -Information retrieval and Web search – Information retrieval Models-Relevance Feedback- Text and Web page Pre-processing – Inverted Index – Latent Semantic Indexing – Web Search – Meta-Search – Web Spamming.

MODULE 2 (8 hours)

Introduction -Social Networks Analysis- Co-Citation and Bibliographic Coupling- Page Rank: PageRank Algorithm, Link-Based Similarity Search, Enhanced Techniques for Page Ranking HITS: HITS Algorithm, Finding Other Eigenvectors-Community Discovery: Problem Definition, Bipartite Core Communities.

MODULE 3 (7 hours)

Web Crawling -A Basic Crawler Algorithm: Breadth-First Crawlers, Preferential Crawlers - Implementation Issues- Universal Crawlers- Focused Crawlers- Topical Crawlers -Evaluation - Crawler Ethics and Conflicts - New Developments.

MODULE 4 (8 hours)

Structured Data Extraction: Wrapper Generation – Preliminaries- Wrapper Induction- Instance Based Wrapper Learning - Automatic Wrapper Generation: Problems - String Matching and Tree Matching -Multiple Alignment - Building DOM Trees - Extraction Based on a Single List Page and Multiple pages- Introduction to Schema Matching – Schema Level Match -Domain and Instance Level Matching – Extracting and Analysing Web Social Networks.

MODULE 5 (7 hours)

Web Usage Mining - Data Collection and Pre-Processing: Sources and Types of Data, Key Elements of Web Usage Data - Data Modelling for Web Usage Mining - Discovery and Analysis of Web Usage Patterns – Applications- Recommender Systems and Collaborative Filtering – Query Log Mining

Text Book

1. Bing Liu, "Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data Centric Systems and Applications)", Springer; 2nd Edition 2009

Reference Books

1. Zdravko Markov, Daniel T. Larose, "Data Mining the Web: Uncovering Patterns in Web Content, Structure, and Usage", John Wiley & Sons, Inc., 2007

2. Guandong Xu, Yanchun Zhang, Lin Li, "Web Mining and Social Networking: Techniques and Applications", Springer; 1st Edition.2010

3. Soumen Chakrabarti, "Mining the Web: Discovering Knowledge from Hypertext Data", Morgan Kaufmann; edition 2002

4. Adam Schenker, "Graph-Theoretic Techniques for Web Content Mining", World Scientific Pub Co Inc, 2005

5. Min Song, Yi Fang and Brook Wu, Handbook of research on Text and Web mining technologies, IGI global, information Science Reference – imprint of: IGI publishing,2008

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	
1.1	Introduction – Web Mining, Theoretical background	1 hour
1.2	Association rule mining	1 hour
1.3	Sequential Pattern Mining -Information retrieval and Web search	1 hour
1.4	Information retrieval Models-Relevance Feedback	1 hour
1.5	Text and Web page Pre-processing	1 hour

1.6	Inverted Index – Latent Semantic Indexing	1 hour
1.7	Web Search – Meta-Search – Web Spamming	1 hour
2	**Module 2**	
2.1	Introduction -Social Networks Analysis	1 hour
2.2	Co-Citation and Bibliographic Coupling	1 hour
2.3	Page Rank: PageRank Algorithm	1 hour
2.4	Link-Based Similarity Search	1 hour
2.5	Enhanced Techniques for Page Ranking	1 hour
2.6	HITS: HITS Algorithm, Finding Other Eigenvectors	1 hour
2.7	Community Discovery, Problem Definition, Bipartite Core Communities	1 hour
3	**Module 3**	
3.1	Web Crawling	1 hour
3.2	A Basic Crawler Algorithm: Breadth-First Crawlers	1 hour
3.3	Preferential Crawlers	1 hour
3.4	Implementation Issues- Universal Crawlers	1 hour
3.5	Focused Crawlers- Topical Crawlers	1 hour
3.6	1 hour Evaluation - Crawler Ethics and Conflicts maxima	1 hour
3.7	New Developments	1 hour

4	**Module 4**	
4.1	Structured Data Extraction: Wrapper Generation	1 hour
4.2	Preliminaries- Wrapper Induction	1 hour
4.3	Instance Based Wrapper Learning - Automatic Wrapper Generation	1 hour
4.4	Problems - String Matching and Tree Matching -Multiple Alignment	1 hour
4.5	Building DOM Trees	1 hour
4.6	Extraction Based on a Single List Page and Multiple pages- Introduction to Schema Matching	1 hour
4.7	Schema Level Match -Domain and Instance Level Matching	1 hour
4.8	Extracting and Analysing Web Social Networks	1 hour
5	**Module 5**	
5.1	Web Usage Mining - Data Collection and Pre-Processing	1 hour
5.2	Sources and Types of Data, Key Elements of Web Usage Data	1 hour
5.3	Data Modelling for Web Usage Mining	1 hour
5.4	Discovery and Analysis of Web Usage Patterns	1 hour
5.5	Application of web usage mining	1 hour
5.6	Recommender Systems and Collaborative Filtering	1 hour

5.7	Query Log Mining	1 hour
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E204C

Course Name: :Web Mining

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Differentiate between Information retrieval and Web search in web mining
2. Describe the role of link analysis and random walks in the PageRank algorithm.
3. Compare breadth-first crawlers and preferential crawlers in terms of their crawling strategies and efficiency.
4. List out the techniques used in string matching and tree matching for automatic wrapper generation.
5. Explain Web Recommender systems based on User and Item.

PART B

Answer any five question from each module. Each question carries 8 marks.

6. (a) Explain the characteristics of Web search.

   4 marks

   (b) Use Latent Semantic Indexing (LSI) to rank these documents for the query gold silver truck for the following "documents":

   d1: Shipment of gold damaged in a fire.

   d2: Delivery of silver arrived in a silver truck

   4 marks

   d3: Shipment of gold arrived in a truck.

7. Describe how the HITS algorithm can be extended to find other eigenvectors beyond the top-ranked ones. 8 marks

8. Differentiate Schema Level Match-Domain and Instance Level Matching in Web mining. 8 marks

9. Discuss potential techniques to improve the performance and effectiveness of basic crawler algorithms. 8 marks

10. Make use of web usage mining how it can be applied to improve website recommendation systems and collaborative filtering. 8 marks

11. Describe Click stream Analysis and Web Server Log Files. 8 marks

12. Identify the techniques used in query log mining, such as query clustering, query categorization, and query intent analysis 8 marks

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E204D	**SEMANTIC WEB ARCHITECTURE**	ELECTIVE	3	0	0	3	3

Preamble : : This course helps the learners to understand the advanced concepts in emerging trends in Web architecture. The learner will be able to grasp the basics of metadata, OWL, RDF and ontologies. This course enables the learners to build an ontology for a given domain.

Prerequisites : NIL

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Distinguish between Traditional web Environment and Semantic web Environment and in the behaviour of a search engine. **(Cognitive Knowledge Level: Understand)**
CO 2	Recognize and Apply the concept of RDF and RDFS and the benefits of Ontology and Taxonomy **(Cognitive Knowledge Level: Apply)**
CO 3	Identify how to use OWL to rewrite the ontology and new features in OWL. **(Cognitive Knowledge Level: Analyze)**
CO 4	Analyze the Web services like UDDI and Real World Examples like Swoogle and FOAF. **(Cognitive Knowledge Level: Analyze)**
CO 5	Apply concepts OWL-S for Web service annotation and mapping OWL-S to UDDI **(Cognitive Knowledge Level: Apply)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓		✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓		✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand			
Apply	70	60	70-80
Analyse	30	40	30-40
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (6 hours)

Introduction to semantic web technology: Traditional web to semantic web-WWW- First Look at the Semantic Web – meta data- Search Engine for the Traditional Web- Search Engine for the Semantic Web

MODULE 2 (8 hours)

Resource Description Framework : Elements - Resource – Property- Statement -rules of RDF – tools- RDFS core elements- XML Schema and RDF Schema –Taxonomy and ontology concepts.

MODULE 3 (7 hours)

Web ontology language: OWL: Define classes- set operators –enumerations- defining properties- Symmetric Properties- Transitive Properties- Functional Properties- Inverse Property- Inverse Functional Property- Validating OWL ontology-.

MODULE 4 (8 hours)

Web services and Real world examples: Web services – web services standards – web services to semantic web services- UDDI. Swoogle- architecture and usage of meta data; FOAF – vocabulary – creating documents

MODULE 5 (7 hours)

Concept of OWL-S – building blocks of OWL-S- OWL-S Profile Ontology- OWL-S Process Ontology- OWL-S Grounding Ontology - mapping OWL-S to UDDI - WSDL

Text Book

1. Foundations and Technologies. Introduction to the Semantic Web Technologies. John Domingue, Dieter Fensel, James A. Hendler.

Reference Books

1. Liyang Yu, Introduction to the Semantic Web and Semantic web services. Chapman & Hall/CRC, Taylor & Francis group, 2007.

2. Johan Hjelm. Creating the Semantic Web with RDF. Wiley, 2001

3. Grigoris Antoniou and Frank van Harmelen. A Semantic Web Primer. MIT Press

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	
1.1	Traditional web to semantic web	1 hour
1.2	World Wide Web	1 hour
1.3	First Look at the Semantic Web	1 hour
1.4	Meta data	1 hour
1.5	Search Engine for the Traditional Web	1 hour
1.6	Search Engine for the Semantic Web	1 hour
2	**Module 2**	
2.1	Elements- Resource	1 hour
2.2	Property- Statement	1 hour
2.3	Rules of RDF	1 hour
2.4	Tools	1 hour
2.5	RDFS core elements	1 hour
2.6	XML Schema	1 hour

2.7	RDF Schema	1 hour
2.8	Taxonomy and ontology concepts	1 hour
3	**Module 3**	
3.1	OWL: define classes	1 hour
3.2	Set operators	1 hour
3.3	Enumerations	1 hour
3.4	Defining properties- Symmetric Properties	1 hour
3.5	Transitive Properties- Functional Properties	1 hour
3.6	Inverse Property- Inverse Functional Property	1 hour
3.7	Validating OWL ontology	1 hour
4	**Module 4**	
4.1	Web services	1 hour
4.2	Web services standards	1 hour
4.3	Web services to semantic web services	1 hour
4.4	UDDI	1 hour
4.5	Swoogle	1 hour
4.6	Architecture and usage of meta data	1 hour
4.7	FOAF – vocabulary –	1 hour

4.8	FOAF- creating documents	1 hour
5	**Module 5**	
5.1	Concept of OWL-S	1 hour
5.2	Building blocks of OWL-S-	1 hour
5.3	OWL-S Profile Ontology	1 hour
5.4	OWL-S Process Ontology	1 hour
5.5	OWL-S Grounding Ontology	1 hour
5.6	Mapping OWL-S to UDDI	1 hour
5.7	WSDL	1 hour
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: : M24CS1E203D

Course Name: :SEMANTIC WEB ARCHITECTURE

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. "T he Semantic Web is an extension of the current Web". Explain the statement.
2. Differentiate RDF and RDFS.
3. Compare and contrast OWL and RDF? Illustrate the need for OWL in Semantic Web?
4. Determine in what manner ranking of documents using metadata done in Swoogle
5. Discuss the issues of mapping OWL-S profile information into UDDI registry

PART B

Answer any five question from each module. Each question carries 8 marks.

6. Compare search engines work in traditional and Semantic web? Explain in detail.
7. A Set of RDF Statements (8 marks)

a.

Subject	Predicate	Object
mySLR:Nikon-D70	mySLR:weight	1.4 lb
mySLR:Nikon-D70	mySLR:pixel	6.1 M
mySLR:Nikon-D50	mySLR:weight	1.3 lb

8. What properties did we define in order to describe Nikon D70? Using the table Write a pseudocode to help the computer to answer the question? (8 marks)
9. OW L has created a new class called "owl:Class" to define classes in OWL documents; it is a subclass of "rdfs:Class". The relationship between all these top classes is summarized in Figure. Define all the classes in our camera ontology (8 marks)
10. Illustrate why the "Person class" is the core of the FOAF vocabulary. Explain with an Example (8 marks)

11. Swoogle is expected to be used more often by the researchers and developers in the Semantic Web community Why? (8 marks)

12. Determine the necessity about the concept of degree of matching? Explain the four degrees of match between two concepts. (8 marks)

13. "O WL-S is not about learning a new language; instead, it is about understanding and using three new ontologies." Justify the statement. (8 marks)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1E204E	**REINFORCEMENT LEARNING**	ELECTIVE	3	0	0	3	3

Preamble : Study of this course provides learners the fundamentals of Reinforcement learning. The course covers the basic concepts of reinforcement learning, Multi Armed bandits problem, Finite Markov Decision Process, Dynamic programming, Monte Carlo methods and Temporal-Difference Learning, Tabular Methods and Prediction with Approximation. This course helps the students to familiarize with applications and case studies of reinforcement learning and solve real-world problems.

Prerequisites : Basic knowledge in machine learning.

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Understand the Reinforcement Learning, Multi Armed Bandits. **(Blooms Level : Understand)**
CO 2	Apply concepts of Finite Markov Decision Processes and Dynamic Programming to model and solve decision-making problems. **(Blooms Level: Apply)**
CO 3	Apply Monte Carlo reinforcement learning algorithms for prediction. **(Blooms Level : Apply)**
CO 4	Analyze and evaluate the effectiveness of Temporal-Difference Learning methods and Tabular Planning techniques in optimizing decision-making processes. **(Blooms Level : Analyse)**
CO 5	Implement and analyze on-policy approximation techniques for action values and apply these methods to real-world problems such as game playing and job-shop scheduling. **(Blooms Level : Analyse)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	**PO 2**	**PO 3**	**PO 4**	**PO 5**	**PO 6**
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	**Continuous Internal Evaluation Tests**		**End Semester Examination (% Marks)**
	Test 1 (%Marks)	**Test 2 (%Marks)**	
Remember			
Understand			
Apply	70	70	70
Analyse	30	30	30
Evaluate			
Create			

Mark Distribution

Total Marks	**CIE Marks**	**ESE Marks**	**ESE Duration**
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar*	: 10 marks
Course based task/Micro Project	
Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

*Seminar should be conducted in addition to the theory hours. Topics for the seminar should be from recent technologies in the respective course

End Semester Examination Pattern

The end semester examination will be conducted by the college. There will be two parts; Part A and Part B. Part A contain 5 numerical questions (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students), with 1 question from each module, having 4 marks for each question. Students shall answer all questions. Part B contains 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student shall answer any five. Each question can carry 8 marks. Total duration of the examination will be 3 Hrs.

SYLLABUS

MODULE 1 (7 hours)

The Reinforcement Learning Problem: Reinforcement Learning, Examples, Elements of Reinforcement Learning, Limitations and Scope, An Extended Example: Tic-Tac-Toe.

Multi-arm Bandits: An n-Armed Bandit Problem, Action-Value Methods, Incremental Implementation, Tracking a Nonstationary Problem, Optimistic Initial Values, Upper-Confidence Bound Action Selection, Gradient Bandits, Associative Search (Contextual Bandits)

MODULE 2 (9 hours)

Finite Markov Decision Processes: The Agent–Environment Interface, Goals and Rewards, Returns, Unified Notation for Episodic and Continuing Tasks, The Markov Property, Markov Decision Processes, Value Functions, Optimal Value Functions, Optimality and Approximation.

Dynamic Programming: Policy Evaluation, Policy Improvement, Policy Iteration, Value Iteration, Asynchronous dynamic programming, Generalized Policy Iteration, Efficiency of dynamic programming.

MODULE 3 (7 hours)

RMonte Carlo Methods: Monte Carlo Prediction, Monte Carlo Estimation of Action Values, Monte Carlo Control, Monte Carlo Control without Exploring Starts, Off-policy Prediction via Importance Sampling, Incremental Implementation, Off-Policy Monte Carlo Control, Importance Sampling on Truncated Returns.

MODULE 4 (8 hours)

Temporal-Difference Learning: TD Prediction, Advantages of TD Prediction Methods, Optimality of TD(0), Sarsa: On-policy TD control, Q-learning Off-policy TD control.

Planning and Learning with Tabular Methods: Models and Planning, Integrated Planning, acting and learning, Prioritized Sweeping, Full vs. Sample Backups, Trajectory sampling, Heuristic search.

MODULE 5 (5 hours)

On-policy Approximation of Action Values: Value Prediction with Function Approximation, Gradient-Descent Methods , Linear Methods , Control with Function Approximation.

Applications and Case Studies: TD-Gammon, Samuel's Checkers Player, Job-Shop

Scheduling.

Reference Books

1. Richard S. Sutton and Andrew G. Barto, "Reinforcement Learning-An Introduction",2nd Edition,The MIT Press,2018
2. Marco Wiering , Martijn van Otterlo Reinforcement Learning: State-of-the-Art (Adaptation, Learning, and Optimization (12)) 2012th Edition
3. Vincent François-Lavet , Peter Henderson , Riashat Islam, An Introduction to Deep Reinforcement Learning (Foundations and Trends(r) in Machine Learning) , 2019

COURSE CONTENTS AND LECTURE SCHEDULE

(For 4 credit courses, the content can be for 45 hrs. and for 3 credit courses, the content can be for 36 hrs.)

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1**	
1.1	The Reinforcement Learning Problem: Reinforcement Learning, Examples	1
1.2	Elements of Reinforcement Learning, Limitations and Scope	1
1.3	An Extended Example: Tic-Tac-Toe	1
1.4	Multi-arm Bandits: An n-Armed Bandit Problem, Action-Value Methods	1
1.5	Incremental Implementation, Tracking a Nonstationary Problem, Optimistic Initial Values	1
1.6	Upper-Confidence-Bound Action Selection, Gradient Bandits	1

1.7	Associative Search (Contextual Bandits)	1
2	**Module 2**	
2.1	Finite Markov Decision Processes: The Agent–Environment Interface	1
2.2	Goals and Rewards, Returns	1
2.3	Unified Notation for Episodic and Continuing Tasks, The Markov Property	1
2.4	Markov Decision Processes, Value Functions	1
2.5	Optimal Value Functions, Optimality and Approximation.	1
2.6	Dynamic Programming: Policy Evaluation	1
2.7	Policy Improvement, Policy Iteration	1
2.8	Value Iteration, Asynchronous dynamic programming	1
2.9	Generalized Policy Iteration, Efficiency of dynamic programming	1
3	**Module 3**	
3.1	Monte Carlo Methods: Monte Carlo Prediction	1
3.2	Monte Carlo Estimation of Action Values, Monte Carlo Control	1
3.3	Monte Carlo Control without Exploring Starts	1
3.4	Off-policy Prediction via Importance Sampling	1
3.5	Incremental Implementation	1

3.6	Off-Policy Monte Carlo Control	1
3.7	Importance Sampling on Truncated Returns	1
4	**Module 4**	
4.1	Temporal-Difference Learning: TD Prediction, Advantages of TD Prediction Methods	1
4.2	Optimality of TD(0), Sarsa: On-policy TD control	1
4.3	Q-learning Off-policy TD control	1
4.4	Planning and Learning with Tabular Methods: Models and Planning	1
4.5	Integrated Planning, acting and learning	1
4.6	Prioritized Sweeping	1
4.7	Full vs. Sample Backups, Trajectory sampling	1
4.8	Heuristic search	1
5	**Module 5**	
5.1	On-policy Approximation of Action Values: Value Prediction with Function Approximation	1
5.2	Gradient-Descent Methods , Linear Methods	1
5.3	Control with Function Approximation	1
5.4	Applications and Case Studies: TD-Gammon	1
5.5	Samuel's Checkers Player, Job-Shop Scheduling	1
	Total	36 Hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: M24CS1E204E

Course Name: REINFORCEMENT LEARNING

Max. Marks: 60 Duration: 3 hours

PART A

Answer all questions. Each question carries 4 marks.

1. Using Tic-Tac-Toe as an example, explain how reinforcement learning can be applied to solve sequential decision-making problems. (4 marks)

2. Suppose that we are doing value iteration with $= 0$. How many iterations will it take for value iteration to converge to the optimal value function? (4 marks)

3. Why Monte Carlo methods for learning value functions require episodic tasks? How is it that n-step TD methods avoid this limitation and can work with continuing tasks? (4 marks)

4. Supose that a Q-learning agent always chooses the action which maximizes the Q-value. What is one potential problem with that approach? (4 marks)

5. Consider a five state random walk. There are five states, s1, s2, ..., s5, in a row with two actions each, left and right. There are two terminal states at each end, with a reward of $+1$ for terminating on the right, after s5 and a reward of 0 for all other

transitions, including the one terminating on the left after s1. In designing a linear function approximator, what is the least number of state features required to represent the value of the equi-probable random policy? (4 marks)

PART B

Answer any five question from each module. Each question carries 8 marks.

6. Illustrate how Optimistic Initial Values can influence the exploration-exploitation tradeoff in an n-Armed Bandit problem (8 marks)

7. Consider designing a recycling robot whose job is to collect empty bottles around the building. The robot has a sensor to detect when a bottle is in front of it, and a gripper to pick up the bottle. It also senses the level of its battery. The robot can navigate, as well as pick up a bottle and throw a bottle it is holding in the trash. There is a battery charger in the building, and the robot should not run out of battery.

   i. Describe this problem as an MDP. What are the states and actions?

   ii. Suppose that you want the robot to collect as many bottles as possible, while not running out of battery. Describe what rewards would enable it to achieve this task. (8 marks)

8. Consider an undiscounted Markov Reward Process with two states A and B. The transition matrix and reward function are unknown, but you have observed two sample episodes:

   A +3 –> A +2 –> B -4 –> A +4 –> B -3

   B -2 –> A +3 –> B -3

   i. Using first-visit Monte-Carlo evaluation, estimate the state-value function V(A),V(B).

   ii. Using every-visit Monte-Carlo evaluation, estimate the state-value function V(A),V(B).

   iii. Draw a diagram of the Markov Reward Process that best explains these two episodes. Show rewards and transition probabilities on your diagram.

   (8 marks)

9. (a) For Q-learning to converge we need to correctly manage the exploration vs. exploitation tradeoff. What property needs to be hold for the exploration strategy? (4 marks)

   (b) Compare full backups and sample backups, highlighting the scenarios where each would be most effective (4 marks)

10. (a) Value function based methods are oriented towards finding deterministic policies whereas policy search methods are geared towards finding stochastic policies. True or false? Justify. (4 marks)

    (b) Suppose that we are using a policy gradient method to solve a reinforcement learning problem and the policy returned by the method is not optimal. Give three plausible reasons for such an outcome? (4 marks)

11. Evaluate the efficiency of dynamic programming methods in solving MDPs. What are the primary computational challenges, and how do techniques like function approximation and hierarchical decomposition address these challenges? (8 marks)

12. Analyze the contributions of TD-Gammon to the field of artificial intelligence and game theory. How has it influenced the development of other AI systems in gaming and beyond? (8 marks)

SEMESTER II

INDUSTRY INTEGRATED COURSE

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1S205A	**CLOUD PRACTITIONER ESSENTIALS**	Industry Course	3	0	0	3	3

Preamble : Learners will gain a comprehensive understanding of AWS services, including their benefits, pricing models, and use cases. They will develop skills in managing AWS infrastructure, such as EC2, networking, storage, databases, and security. Additionally, they will learn to monitor, optimize costs, and plan effective cloud migration strategies.

Prerequisite : Basic Knowledge on Computer Networks

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Summarize the working definition of AWS and differentiate between on-premises, hybrid-cloud, and all-in cloud**(Cognitive Level: Understand)**
CO 2	Describe the basic global infrastructure of the AWS Cloud and describe and provide an example of the core AWS services, including compute, network, databases, and storage.**(Cognitive Level: Apply)**
CO 3	Identify an appropriate solution using AWS Cloud services with various use cases and understand the shared responsibility model.**(Cognitive Level: Analyse)**
CO 4	Describe the core security services within the AWS Cloud and describe the basics of AWS Cloud migration and articulate the financial benefits of the AWS Cloud for an organization's cost management.**(Cognitive Level: Evaluate)**
CO 5	Define the core billing, account management, and pricing models and explain how to use pricing tools to make cost-effective choices for AWS services. **(Cognitive Level: Evaluate)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	20	20	20
Apply	60	60	60
Analyse	20	20	20
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar	: 10 marks
Course based task/Seminar/Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

End Semester Examination Pattern

The examination will be conducted by the College with the question paper provided by the industry. The examination will be for 3 Hrs. and will contain 7 questions, with a minimum of one question from each module of which the student should answer any five. Each question can carry 12 marks. The evaluation of the answer scripts shall be done by the expert in the industry handling the course.

SYLLABUS

MODULE 1 (7 hours)

AWS-Benefits of AWS-Differences between on-demand delivery and cloud deployments–pay-as-you-go pricing model-Benefits of Amazon Elastic Compute Cloud (Amazon EC2) at a basic level-Identify the different Amazon EC2 instance types-Differentiate between the various billing options for Amazon EC2-Describe the benefits of Amazon EC2 Auto Scaling-Summarize the benefits of Elastic Load Balancing-Give an example of the uses for Elastic Load Balancing-Summarize the differences between Amazon Simple Notification Service (Amazon SNS) and Amazon Simple Queue Services (Amazon SQS)-Summarize additional AWS compute options.

MODULE 2 (6 hours)

Benefits of the AWS Global Infrastructure-Basic concept of Availability Zones-Benefits of Amazon CloudFront and Edge locations-Compare different methods for provisioning AWS services-Basic concepts of networking-Difference between public and private networking resources-Virtual private gateway using a real life scenario-Virtual private network (VPN) using a real life scenario-Benefit of AWS Direct Connect-Benefit of hybrid deployments-Layers of security used in an IT strategy-Services are used to interact with the AWS global network

MODULE 3 (6 hours)

Basic concept of storage and databases-Benefits of Amazon Elastic Block Store (Amazon EBS)-Benefits of Amazon Simple Storage Service (Amazon S3) Benefits of Amazon Elastic File System (Amazon EFS)-Various storage solutions-Benefits of Amazon Relational Database Service (Amazon RDS)-Benefits of Amazon DynamoDB-Various database services-Benefits of the shared responsibility model-Multi-factor authentication (MFA)- Differentiate between the AWS Identity and Access Management (IAM) security levels-Security policies at a basic level-Benefits of AWS Organizations-Benefits of compliance with AWS-Primary AWS security services at a basic level.

MODULE 4 (7 hours)

Summarize approaches to monitoring your AWS environment-Benefits of Amazon Cloud Watch-Benefits of AWS Cloud Trail-Benefits of AWS Trusted Advisor-Understand AWS pricing and support models-AWS Free Tier-Benefits of AWS Organizations and consolidated billing-Benefits of AWS Budgets-Benefits of AWS Cost Explorer-Primary benefits of the AWS Pricing Calculator-Distinguish between the various AWS Support Plans-Benefits of AWS Marketplace

MODULE 5 (8 hours)

Migration and innovation in the AWS Cloud-AWS Cloud Adoption Framework (AWS CAF)-Six key factors of a cloud migration strategy-Benefits of various AWS data migration solutions, such as AWS Snowcone, AWS Snowball, and AWS Snowmobile-Broad scope of innovative solutions that AWS offers-Six pillars of the AWS Well-Architected Framework-Six benefits of cloud computing.

Reference Books

1. John W. Rittinghouse,James F. Ransome, “Cloud Computing Implementation, Management, and Security”

2. Michael Miller Cloud Computing-Web Based Applications that change the way you work and Collaborate Online.

3. https://explore.skillbuilder.aws/learn/public/learning_plan/view/82/cloud-foundations-learning-plan?la=cta&cta=topbanner

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1 (7 Hours)**	
1.1	AWS-Benefits of AWS-Differences between on-demand delivery and cloud deployments	1 hours
1.2	pay-as-you-go pricing model-Benefits of Amazon Elastic Compute Cloud (Amazon EC2) at a basic level	1 hours
1.3	Identify the different Amazon EC2 instance types-Differentiate between the various billing options for Amazon EC2	1 hours
1.4	Describe the benefits of Amazon EC2 Auto Scaling-Summarize the benefits of Elastic Load Balancing	1 hours
1.5	Give an example of the uses for Elastic Load Balancing	1 hours
1.6	Differences between Amazon Simple Notification Service (Amazon SNS) and Amazon Simple Queue Services (Amazon SQS)	1 hours
1.7	Summarize additional AWS compute options.	1 hours
2	**Module 2**	

2.1	Benefits of the AWS Global Infrastructure-Basic concept of Availability Zones	1 hours
2.2	Benefits of Amazon CloudFront and Edge locations-Compare different methods for provisioning AWS services	1 hours
2.3	Basic concepts of networking-Difference between public and private networking resources	1 hours
2.4	Virtual private gateway using a real life scenario -Virtual private network (VPN) using a real life scenario	1 hours
2.5	Benefit of AWS Direct Connect-Benefit of hybrid deployments	1 hours
2.6	Layers of security used in an IT strategy-Services are used to interact with the AWS global network	1 hours
3	**Module 3**	
3.1	Basic concept of storage and databases-Benefits of Amazon Elastic Block Store (Amazon EBS)	1 hours
3.2	Benefits of Amazon Simple Storage Service (Amazon S3)	1 hours
3.3	Amazon Elastic File System (Amazon EFS)-Various storage solutions	1 hours
3.4	Benefits of Amazon Relational Database Service (Amazon RDS)	1 hours
3.5	Benefits of Amazon DynamoDB-Various database services-Benefits of the shared responsibility model	1 hours

3.6	Multi-factor authentication (MFA)-Differentiate between the AWS Identity and Access Management (IAM) security levels	1 hours
3.7	Security policies at a basic level-Benefits of AWS Organizations	1 hours
3.8	Benefits of compliance with AWS-Primary AWS security services at a basic level	1 hours
4	**Module 4**	
4.1	Approaches to monitoring your AWS environment	1 hours
4.2	Benefits of Amazon CloudWatch-Bbenefits of AWS CloudTrai	1 hours
4.3	Benefits of AWS Trusted Advisor Understand AWS pricing and support models	1 hours
4.4	AWS Free Tier-Benefits of AWS Organizations and consolidated billing	1 hours
4.5	Benefits of AWS Budgets-Benefits of AWS Cost Explorer	1 hours
4.6	Primary benefits of the AWS Pricing Calculator	1 hours
4.7	Distinguish between the various AWS Support Plans-Benefits of AWS Marketplace	1 hours
5	**Module 5**	
5.1	Migration and innovation in the AWS Cloud	1 hours
5.2	Migration and innovation in the AWS Cloud	1 hours

5.3	AWS Cloud Adoption Framework (AWS CAF)	1 hours
5.4	Six key factors of a cloud migration strategy	1 hours
5.5	Benefits of various AWS data migration solutions such as AWS Snowcone, AWS Snowball, and AWS Snowmobile	1 hours
5.6	Broad scope of innovative solutions that AWS offers	1 hours
5.7	Six pillars of the AWS Well-Architected Framework	1 hours
5.8	Six benefits of cloud computing	1 hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: M24CS1S205A

Course Name: CLOUD PRACTIONER ESSENTIALS

Max. Marks: 60 Duration: 3 hours

PART A

Answer any five questions. Each question carries 12 marks.

1. Your company is planning to migrate its existing on-premises web application infrastructure to Amazon Web Services (AWS) to improve scalability and cost-efficiency. Using your knowledge of AWS services, design a solution that leverages Amazon EC2, Auto Scaling, and Elastic Load Balancing to ensure high availability and optimal performance. Explain the types of EC2 instances you would choose, the billing options you would consider, and how Auto Scaling and Elastic Load Balancing would be configured to handle varying levels of web traffic.

2. A startup is developing a mobile application that requires reliable message delivery and coordination between different microservices. As the AWS architect, you need to choose between Amazon Simple Notification Service (Amazon SNS) and Amazon Simple Queue Service (Amazon SQS) for the messaging system. Describe a scenario where each service would be most beneficial, and explain how you would implement them within the application's architecture. Additionally, outline other AWS compute options that could support the backend processing needs of the mobile application

3. Design a solution for securely connecting an on-premises data center to an AWS environment for a hybrid cloud strategy. Incorporate a virtual private gateway and a virtual private network (VPN), and provide real-life scenarios for their use. Additionally, explain how AWS Direct Connect can enhance this hybrid deployment. Finally, describe the security layers required to protect sensitive financial data as it travels through the AWS global network, and identify the AWS services that would support these security measures.

4. You are tasked with enhancing the security and compliance of your organization's AWS environment. Using your understanding of the shared responsibility model, multi-factor authentication (MFA), and AWS Identity and Access Management (IAM) security levels, develop a comprehensive security strategy. Describe how you would implement security policies and leverage AWS Organizations for centralized management. Furthermore, explain the benefits of complying with AWS security standards and identify the primary AWS security services that would be utilized at a basic level to safeguard the organization's data and infrastructure.

5. You are responsible for monitoring and optimizing the performance and security of your company's AWS environment. Design a monitoring strategy that utilizes Amazon CloudWatch, AWS CloudTrail, and AWS Trusted Advisor. Describe how each service will be implemented to monitor different aspects of the environment, such as performance metrics, API activity, and best practice recommendations. Provide specific examples of how these services can proactively address potential issues.

6. A large enterprise is planning to migrate its on-premises infrastructure to the AWS Cloud. Using the AWS Cloud Adoption Framework (AWS CAF) and the six key factors of a cloud migration strategy, develop a comprehensive migration plan. Discuss how you would leverage AWS data migration solutions such as AWS Snowcone, AWS Snowball, and AWS Snowmobile to facilitate the migration. Additionally, explain how the six pillars of the AWS Well-Architected Framework would guide the design and implementation of the cloud architecture to ensure security, reliability, and efficiency.

7. A technology startup is looking to leverage AWS for rapid innovation and scalability. Identify and describe the six benefits of cloud computing and how they apply to the startup's needs. Additionally, explore the broad scope of innovative solutions that AWS offers, and propose a strategy for the startup to utilize these services to accelerate development and deployment. Integrate the principles of the AWS Well-Architected Framework into your strategy to ensure that the startup's cloud environment is well-designed and sustainable.

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1S205B	**CLOUD INFRASTRUCTURE ENGINEERING**	Industry Course	3	0	0	3	3

Preamble : Learners will gain a comprehensive understanding of cloud services, including their benefits, pricing models, and use cases. They will develop skills in managing Cloud infrastructure, such as docker, container and kubernets.Lerners will able to apply DevOps principles and methodologies to bridge the gap between development and operations teams.

Prerequisite : Nil

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Apply DevOps principles and methodologies to bridge the gap between development and operations teams. **(Blooms Level Apply)**
CO 2	Master the fundamentals of DevOps culture, including collaboration, communication, and shared responsibility. **(Blooms Level Apply)**
CO 3	Analyze Linux system basics, including essential commands, user management, file permissions, and package management. **((Blooms Level Analyze))**
CO 4	Analyze core technologies like Python scripting and Git version control for efficient development workflows. **(Blooms Level Analyze)**
CO 5	Apply containerization concepts, create and manage containerized applications using Docker, explore container orchestration with Kubernetes **(Blooms Level Create)**

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓		✓	✓	✓	✓
CO 2	✓		✓	✓	✓	✓
CO 3	✓		✓	✓	✓	✓
CO 4	✓		✓	✓	✓	✓
CO 5	✓		✓	✓	✓	✓

Assessment Pattern

Bloom's Category	Continuous Internal Evaluation Tests		End Semester Examination (% Marks)
	Test 1 (%Marks)	Test 2 (%Marks)	
Remember			
Understand	10	10	10
Apply	60	60	60
Analyse	30	30	30
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	40	60	3 Hours

Continuous Internal Evaluation Pattern

Seminar	: 10 marks
Course based task/Seminar/Data collection and interpretation/Case study	: 10 marks
Test paper 1 (Module 1 and Module 2)	: 10 marks
Test paper 2 (Module 3 and Module 4)	: 10 marks

End Semester Examination Pattern

The examination will be conducted by the College with the question paper provided by the industry. The examination will be for 3 Hrs. and will contain 7 questions, with a minimum of one question from each module of which the student should answer any five. Each question can carry 12 marks. The evaluation of the answer scripts shall be done by the expert in the industry handling the course.

SYLLABUS

MODULE 1 Fundamental of DevOps(6 hours)

DevOps Foundations - Agile, People, Process, Culture, Technology / Product. Linux -Basic Commands, Text Editors: Vim, Users and Groups, File permissions, Package Management, Services, Processes. Networking Basics - Basic Concepts, Networking Commands in Linux, IPs and ports, SSL & TLS basics. Python Basics, Shell Programming, Basic Git Operations

MODULE 2 Application Architecture and CICD(7 hours)

Monolithic/Microservices. Micro services communication - sync vs async (rest api, queue), service discovery, api gateway, event driven. Backend - Database management, Authentication/Authorization, Caching, server side languages and frameworks. Frontend - static hosting, CDN, Build packages. Why CI/CD? Fundamentals and sample setup using github actions. Scalability and Performance

MODULE 3 Containerization and Orchestration (7 hours)

Introduction to containers, need of containers, Containers vs Virtual Machines, Container implementation, Advantages. Docker Overview - Docker, Main Docker Commands, Docker networking, Docker volumes, Container Registry. Docker in Practice -Developing with Containers, Dockerfile - Building our own Docker image, pushing our built Docker Image into a private Registry (AWS/Self hosted), Debugging a Container, Persist data in Docker (volumes). Docker compose - Running multiple services, what is docker compose, how to use it - Create the Docker Compose File, Docker Networking in Docker Compose. Container Orchestration -What is orchestration, Different orchestration tools, Intro to Kubernetes

MODULE 4 Observability (7 hours)

Introduction - observability, Importance of observability in modern software development and operations. Observability Pillars - Exploring the pillars of observability: Monitoring, Logging, and Tracing, Real-world use cases and benefits. Monitoring & Logging Tools. Metrics and Alerting. Application Performance Monitoring (APM). Service Level Objectives (SLOs) and SLIs, Setting and measuring SLOs with observability metrics, Practical examples and case studies. Incident Response and Management

MODULE 5 Infrastructure as Code (IAC) and Cloud Computing (9 hours)

Introduction to IaC Principles, Overview of IaC- Understanding the concept of treating infrastructure as code. Benefits of IaC, Discussing the advantages of automating infrastructure provisioning and management. IaC Tools and Frameworks, Popular IaC Tools- Introducing common IaC tools such as Terraform, Ansible, Chef, and Puppet. Choosing the Right Tool, evaluating different tools based on use cases and infrastructure requirements. Leveraging Infrastructure Metadata, Tagging and Labeling, using metadata to enrich observability data. Standardized Naming Conventions, Establishing naming conventions for observability purposes. Cloud Computing - Cloud computing Saas, Paas, Iaas. Shared responsibility models. Could components - Networking, compute services, File storage, Database, Queue, Orchestration, IAM, Different providers - AWS, GCP, Azure etc . Comparison of components in each provider and understanding the names.

Reference Books

1. Del Toro V, “Electrical Engineering Fundamentals”, Pearson Education.
2. T. K. Nagsarkar, M. S. Sukhija, “Basic Electrical Engineering”, Oxford Higher Education.

COURSE CONTENTS AND LECTURE SCHEDULE

No	Topic	No of Lecture/ Tutorial Hours
1	**Module 1 Fundamental of DevOps**	6 Hours
1.1	TDevOps Foundations - Agile, People, Process, Culture, Technology / Product	2 hours
1.2	Linux -Basic Commands, Text Editors: Vim, Users and Groups, File permissions, Package Management, Services, Processes.	1 hours
1.3	Networking Basics - Basic Concepts, Networking Commands in Linux	2 hours
1.4	IPs and ports, SSL & TLS basics. Python Basics, Shell Programming,Basic Git Operations	1 hours
2	**Module 2 Application Architecture and CICD**	7 Hours
2.1	Monolithic/Microservices. Micro services communication	2 hours
2.2	sync vs async (rest api, queue), service discovery, api gateway, event driven	1 hours
2.3	Backend - Database management, Authentication/Authorization, Caching, server-side languages and frameworks.	1 hours
2.4	Frontend - static hosting, CDN, Build packages.	1 hours

2.5	Why CI/CD? Fundamentals and sample setup using github actions.	1 hours
2.6	Scalability and Performance	1 hours
3	**Module 3 Containerization and Orchestration**	7 Hours
3.1	Introduction to containers, need of containers, Containers vs Virtual Machines, How is it done, Container Advantages.	2 hours
3.2	Docker Overview - Docker, Main Docker Commands, Docker networking, Docker volumes, Container Registry.	1 hours
3.3	Docker in Practice -Developing with Containers, Dockerfile - Building our own Docker image, pushing our built Docker Image into a private Registry (AWS/Self hosted)	1 hours
3.4	Debugging a Container, Persist data in Docker (volumes). Docker compose - Running multiple services	1 hours
3.5	what is docker compose, how to use it - Create the Docker Compose File, Docker Networking in Docker Compose.	1 hours
3.6	Container Orchestration -What is orchestration, Different orchestration tools, Intro to Kubernetes	1 hours
4	**Module 4 Observability**	7 Hours
4.1	Introduction - observability, Importance of observability in modern software development and operations	2 hours
4.2	Observability Pillars - Exploring the pillars of observability: Monitoring, Logging, and Tracing, Real-world use cases and benefits	1 hours
4.3	Monitoring & Logging Tools. Metrics and Alerting	1 hours

4.4	Application Performance Monitoring (APM). Service Level Objectives (SLOs) and SLIs	1 hours
4.5	Setting and measuring SLOs with observability metrics, Practical examples and case studies. Incident Response and Management	2 hours
5	**Module 5 Infrastructure as Code (IAC) and Cloud Computing**	9 Hours
5.1	Introduction to IaC Principles, Overview of IaC - Understanding the concept of treating infrastructure as code	1 hours
5.2	Benefits of IaC, Discussing the advantages of automating infrastructure provisioning and management	1 hours
5.3	IaC Tools and Frameworks,Popular IaC Tools- Introducing common IaC tools such as Terraform, Ansible, Chef, and Puppet	1 hours
5.4	Choosing the Right Tool, evaluating different tools based on use cases and infrastructure requirements	1 hours
5.5	Leveraging Infrastructure Metadata, Tagging and Labeling, using metadata to enrich observability data	1 hours
5.6	Standardized Naming Conventions, Establishing naming conventions for observability purposes	1 hours
5.7	Cloud Computing - Cloud computing Saas, Paas, Iaas. Shared responsibility models.	1 hours
5.8	Could components - Networking, compute services, File storage, Database, Queue, Orches-tration, IAM, Different providers - AWS, GCP, Azure etc .	1 hours

5.9	Could components - Networking, compute services, File storage, Database, Queue, Orches- tration, IAM, Different providers - AWS, GCP, Azure etc .	1 hours
Total		36 hours

MODEL QUESTION PAPER

QP CODE: **Pages: 2**

Reg.No.: ....

Name: ....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS), KOTHAMANGALAM

FIRST SEMESTER M.TECH DEGREE EXAMINATION, DECEMBER 2024

Course Code: M24CS1S205B

Course Name: Cloud Infrastructure Engineering

Max. Marks: 60 Duration: 3 hours

PART A

Answer any five questions. Each question carries 12 marks.

1. You are responsible for maintaining a high-availability database cluster in a cloud environment:

   (a) Describe different strategies for achieving database redundancy and failover. (6)

   (b) Discuss how IaC tools can be used to automate the provisioning and configuration of your database cluster. (6)

2. (a) Describe the three pillars of observability: Monitoring, Logging, and Tracing. (6)
   (b) How can these tools be used to diagnose performance issues in a cloud-native application? (6)

3. Your team is developing a new API using a serverless architecture on AWS Lambda:
   (a) Discuss the advantages and disadvantages of serverless functions compared to traditional web server deployments. (6)
   (b) Explain how you would configure and manage security access for your Lambda functions. (6)

4. (a) Explain the concept of Git branching strategies like forking, rebasing, and cherry picking. (6)
   (b) How can Git be used to effectively manage code versions and collaboration within a DevOps team? (6)

5. (a) Describe the shared responsibility model in cloud computing. (6)
   (b) Compare and contrast the key cloud services (Compute, Storage, Database) across AWS, GCP, and Azure. (6)

6. (a) Discuss the differences between containerization and virtualization. (6)
   (b) Provide a step-by-step guide on how to containerize a simple Node.js application using Docker. (6)

7. (a) Explain the concept of cloud computing service models (SaaS, PaaS, IaaS). (6)
   (b) Compare the key components (Networking, Compute, Storage) among AWS, GCP, and Azure (6)

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1P206	**MINI PROJECT**	CORE	4	0	0	4	4

Preamble : Mini project can help to strengthen the understanding of student's fundamentals through application of theoretical concepts and to boost their skills and widen the horizon of their thinking. The ultimate aim of an engineering student is to resolve a problem by applying theoretical knowledge. Doing more projects increases problem- solving skills. The introduction of mini projects ensures preparedness of students to undertake dissertation. Students should identify a topic of interest in consultation with PG Programme Coordinator. Demonstrate the novelty of the project through the results and outputs.

Prerequisites : NIL

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Identify and define a relevant and significant problem or challenge in the relevant field
CO 2	Formulate research methodologies for the innovative and creative solutions
CO 3	Plan and execute tasks utilizing available resources within timelines, following ethical professional and financial norms
CO 4	Organize and communicate technical and scientific findings effectively in written reports, oral presentation, and visual aids

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓	✓	✓	✓	✓	✓
CO 2	✓	✓	✓	✓	✓	
CO 3	✓	✓	✓	✓	✓	
CO 4	✓	✓	✓	✓	✓	

Continuous Internal Evaluation Pattern

The progress of the mini project is evaluated based on three reviews, two interim reviews and a final review. A report is required at the end of the semester.

Total marks : 100 marks

Interim evaluation: (2 Evaluations)	60 (30 marks for each review)
Final evaluation by a Committee (will be evaluating the level of completion and demonstration of functionality/ specifications, clarity of presentation, oral examination, work knowledge and involvement)	25
Report (the committee will be evaluating for the technical content, adequacy of references, templates followed and permitted plagiarism level is not more than 25%)	10
Supervisor/Guide:	5
Total	100

CODE	COURSE NAME	CATEGORY	L	T	P	S	CREDIT
M24CS1L207	MODERN DATABASE LAB	CORE	4	0	0	4	4

Preamble : This course provides an in-hand experience related to various database management systems. Also equips them to design and implement a database application built over the concepts. This course helps the learners to develop applications that manage data efficiently with the help of suitable data models and techniques.

Prerequisites : Advanced Database Management Systems

Course Outcomes

After the completion of the course the student will be able to:

CO 1	Able to perform the distributed processing of large data sets across clusters of computers using simple programming models with the help of Hadoop. (Cognitive Knowledge Level: Apply)
CO 2	Perform data summarization and ad hoc querying using Hive. (Cognitive Knowledge Level: Create)
CO 3	Operates on document databases and techniques using DynamoDB. (Cognitive Knowledge Level: Apply)
CO 4	Capable of implementing XML and XML queries for data management. (Cognitive Knowledge Level: Create)
CO 5	Apply emerging technologies in column store along with Cassandra. (Cognitive Knowledge Level: Apply)

Mapping of Course Outcomes With Program Outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6
CO 1	✓	✓	✓	✓	✓	✓
CO 2	✓	✓	✓	✓	✓	
CO 3	✓	✓	✓	✓	✓	
CO 4	✓	✓	✓	✓	✓	✓
CO 5	✓	✓	✓	✓	✓	

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
100	60	40	3 Hours

Continuous Internal Evaluation Pattern

Lab work and Viva-voce : 60 marks

Final assessment Test and Viva voce : 40 marks

The laboratory courses will be having only Continuous Internal Evaluation and carries 100 marks. Final assessment shall be done by two examiners; one examiner will be a senior faculty from the same department.

SYLLABUS

Basic concepts of Big Data, Configuration of Hadoop, Parallel Database, Distributed Database, Semi-structured Data and XML Databases: XML Data Model – XML Schema-DTD-XSD – XPath and XQuery, JDOQL (Java Data Object-based Query Language), No SQL Databases: Key value stores – DynamoDB, Column Based- HBase, Cassandra, Web Page ranking algorithm.

LIST OF EXPERIMENTS

1	Study and Configure Hadoop for Big Data.
2	Study of NoSQL Databases such as Hive/HBase/Cassandra/DynamoDB.
3	Design Data Model using NoSQL Databases such as Hive/Hbase/Cassandra/DynamoDB
4	Implement any one Partitioning technique in Parallel Databases.

5	Implement Two Phase commit protocol in Distributed Databases.
6	Web page using XML simple tags
7	Web page using XML Schema with simple elements
8	Create XML, XML schemas, DTD for any database application and implement min 10 queries using Xquery FLOWR expression and Xpath.
9	Design database schemas and implement min 10 queries using Hive/ Hbase/ Cassandra column-based databases.
10	Develop the XSD for the XML document
11	Create an XML document and validate it against an XML Schema/DTD. Use XQuery to query and view the contents of the database.
12	Consider an application in which the results of football games are to be represented in XML,DTD and XQuery. For each game, we want to be able to represent the two teams involved ,which one was playing at home, which players scored goals(some of which may have been penalties)and the time when each was scored, and which players were shown yellow or red cards. You may use some attributes. You can check your solutions with the online demo of the Zorba XQueryengine.
13	Implement Web Page ranking algorithm

Reference Books

1. Tom White, Hadoop: The Definitive Guide, O'Reilly Media 4th Edition, April 2015.
2. Joe Fawcett, Danny Ayers, Liam R. E. Quin, Beginning XML, 5/e, John Wiley & Sons, 2012.
3. Jeff Carpenter, Eben Hewitt, "Cassandra: The Definitive Guide", 3/e, O'Reilly, 2020.
4. Tanmay Deshpande, "DynamoDB Cookbook", Packt Publishing, September 2015.