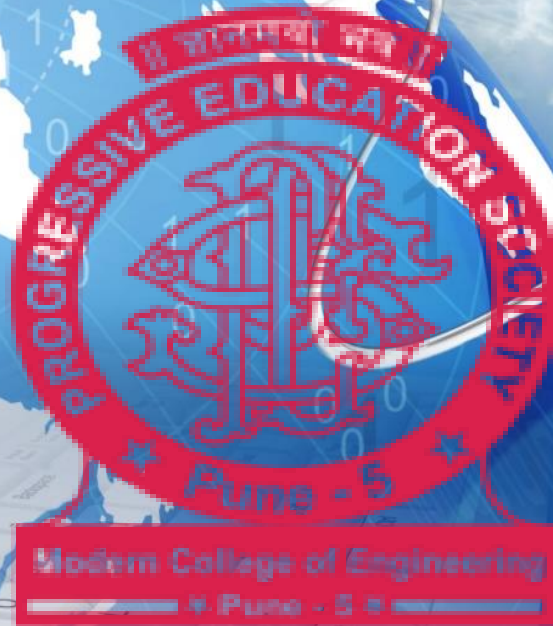


**Progressive Education Society's
Modern College of Engineering
Department of Computer Engineering**



Curriculum Booklet

**Fourth Year
2015 Pattern
Semester-I**



Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING



Progressive Education Society's
Modern College of Engineering, Shivajinagar, Pune-05.
Department of Computer Engineering

Curriculum Booklet

2015 Pattern
Semester: I
Class: BE (COMPUTER ENGINEERING)



Vision of the Institute:

- Creation of a collaborative academic environment to foster professional excellence and ethical values.

Mission of the Institute:

- To develop outstanding professionals with high ethical standards capable of creating and managing global enterprises.
- To foster innovation and research by providing a stimulating learning environment.
- To ensure equitable development of students of all ability levels and backgrounds.
- To be responsive to changes in technology, socio-economic and environmental conditions.
- To foster and maintain mutually beneficial partnerships with alumni and industry.

Vision of the Department

- To achieve excellence in the field of computing through quality education.

Mission of the Department

- To develop promising professionals in the field of computing.
- To provide exposure to emerging technologies and inculcate ethics.
- To strengthen association with alumni and industry.



Objectives of the Institute:

- To develop infrastructure appropriate for delivering quality education
- To develop the overall personality of students who will be innovators and future leaders capable of prospering in their work environment.
- To inculcate ethical standards and make students aware of their social responsibilities.
- Promote close interaction among industry, faculty and students to enrich the learning process and enhance career opportunities.
- Encourage faculty in continuous professional growth through quality enhancement programs and research and development activities.
- Foster a healthy work environment which allows for freedom of expression and protection of the rights of all stakeholders through open channels of communication.

PEO

The graduates of Computer Engineering Department will be,

PEO1: Capable of solving real world problems.

PEO2: Capable of working with multidisciplinary projects.

PEO3: Capable to adapt to changing technologies and life management skills.

PEO4: Able to exhibit professional and ethical responsibilities

Program Specific Outcomes

Graduate of computer engineering programme will demonstrate

- The ability to understand, analyze, develop and evaluate system based on various algorithmic approaches.
- The ability to pursue career in IT industries, to become an entrepreneur and have zest for higher studies.
- The ability to solve problems using engineering principles, tools and techniques.



Program Outcome

- 1. Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
- 2. Problem analysis:** Identify, formulate, research literature, and analyse complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
- 3. Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
- 4. Conduct investigations of complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
- 5. Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modelling to complex engineering activities with an understanding of the limitations.
- 6. The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
- 7. Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
- 8. Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
- 9. Individual and teamwork:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
- 10. Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
- 11. Project management and finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
- 12. Life-long learning:** Recognize the need for and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Course Structure

Faculty of Engineering

Savitribai Phule Pune University

Savitribai Phule Pune University												
Fourth Year of Computer Engineering (2015 Course) (with effect from 2018-19)												
<u>Semester I</u>												
Course Code	Course	Teaching Scheme Hours / Week		Examination Scheme and Marks						Credit		
		Theory	Practical	In-Sem	End-Sem	TW	PR	OR/ *PRE	Total	TH/ TUT	PR	
410241	High Performance Computing	04	--	30	70	--	--	--	100	04	--	
410242	Artificial Intelligence and Robotics	03	--	30	70	--	--	--	100	03	--	
410243	Data Analytics	03	--	30	70	--	--	--	100	03	--	
410244	Elective I	03	--	30	70	--	--	--	100	03	--	
410245	Elective II	03	--	30	70	--	--	--	100	03	--	
410246	Laboratory Practice I	--	04	--	--	50	50	--	100	--	02	
410247	Laboratory Practice II	--	04	--	--	50	--	*50	100	--	02	
410248	Project Work Stage I	--	02	--	--	--	--	*50	50	--	02	
Total Credit										16	06	
Total		16	10	150	350	100	50	100	750	22		
410249	Audit Course 5										Grade	
Elective I				Elective II								
410244 (A) Digital Signal Processing				410245 (A) Distributed Systems								
410244 (B) Software Architecture and Design				410245 (B) Software Testing and Quality Assurance								
410244 (C) Pervasive and Ubiquitous Computing				410245 (C) Operations Research								
410244 (D) Data Mining and Warehousing				410245 (D) Mobile Communication								

410249-Audit Course 5 (AC5) Options:

- | | |
|--|---|
| AC5-I Entrepreneurship Development | AC5-IV: Industrial Safety and Environment Consciousness |
| AC5-II: Botnet of Things | AC5-V: Emotional Intelligence |
| AC5-III: 3D Printing | AC5-VI: MOOC- Learn New Skills |

Abbreviations:

- TW:** Term Work **TH:** Theory **OR:** Oral **PR:** Practical
Sem: Semester ***PRE:** Project/ Mini-Project Presentation



Curriculum

Name of the Subject –410241: High Performance Computing (HPC)

Weekly Workload (in Hrs)	Lecture	Tutorial	Practical
	4	-	-

Online/ In-sem	Theory	Practical	Oral	Term-work	Total Marks	Credit
30	70	-	-	-	100	3

Course Objectives:

- To study parallel computing hardware and programming models
- To be conversant with performance analysis and modeling of parallel programs
- To understand the options available to parallelize the programs
- To know the operating system requirements to qualify in handling the parallelization

Course Outcome:

- C401.1 Explain parallel processing architecture.
- C401.2 Develop an efficient parallel algorithm to solve a problem.
- C401.3 Analyze and measure performance of parallel programs.
- C401.4 Explain the logic to parallelize the sorting and graph algorithms.



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Syllabus

Unit	<u>Course Contents</u>	<u>Hours</u>
I	Introduction	09
	Motivating Parallelism, Scope of Parallel Computing, Parallel Programming Platforms: Implicit Parallelism, Trends in Microprocessor and Architectures, Limitations of Memory, System Performance, Dichotomy of Parallel Computing Platforms, Physical Organization of Parallel Platforms, Communication Costs in Parallel Machines, Scalable design principles, Architectures: N-wide superscalar architectures, multi-core architecture.	
II	Parallel Programming	09
	Principles of Parallel Algorithm Design: Preliminaries, Decomposition Techniques, Characteristics of Tasks and Interactions, Mapping Techniques for Load Balancing, Methods for Containing Interaction Overheads, Parallel Algorithm Models, The Age of Parallel Processing, the Rise of GPU Computing, A Brief History of GPUs, Early GPU.	
III	Basic Communication	09
	Operations- One-to-All Broadcast and All-to-One Reduction, All-to-All Broadcast and Reduction, All-Reduce and Prefix-Sum Operations, Scatter and Gather, All-to-All Personalized Communication, Circular Shift, Improving the Speed of Some Communication Operations.	
IV	Analytical Models of Parallel Programs	09
	Analytical Models: Sources of overhead in Parallel Programs, Performance Metrics for Parallel Systems, and the effect of Granularity on Performance, Scalability of Parallel Systems, Minimum execution time and minimum cost, optimal execution time. Dense Matrix Algorithms: Matrix- Vector Multiplication, Matrix-Matrix Multiplication.	
V	Parallel Algorithms- Sorting and Graph	09
	Issues in Sorting on Parallel Computers, Bubble Sort and its Variants, Parallelizing Quick sort, All-Pairs Shortest Paths, Algorithm for sparse graph, Parallel Depth-First Search, Parallel Best- First Search.	
VI	CUDA Architecture	09
	CUDA Architecture, Using the CUDA Architecture, Applications of CUDA Introduction to CUDA C-Write and launch CUDA C kernels, Manage GPU memory, Manage communication and synchronization, Parallel programming in CUDA- C.	



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Textbooks

1. Ananth Grama, Anshul Gupta, George Karypis, and Vipin Kumar, "Introduction to Parallel Computing", 2nd edition, Addison-Wesley, 2003, ISBN: 0-201-64865-2
2. Jason Sanders, Edward Kandrot, "Cuda by Example", Addison-Wesley, ISBN-13: 978-0-13-138768-3

Reference Books:

- R1. Kai Hwang, "Scalable Parallel Computing", McGraw Hill 1998, ISBN:0070317984
R2. Shane Cook, "CUDA Programming: A Developer's Guide to Parallel Computing with GPUs", Morgan Kaufmann Publishers Inc. San Francisco, CA, USA 2013 ISBN: 9780124159884
R3. David Culler Jaswinder Pal Singh, "Parallel Computer Architecture: A Hardware/Software Approach", Morgan Kaufmann, 1999, ISBN 978-1-55860-343-1
R4. Rod Stephens, "Essential Algorithms", Wiley, ISBN: 978-1-118-61210-1

Reference Links:

https://onlinecourses.nptel.ac.in/noc18_cs50/preview

<https://www.top500.org/resources/>

<https://insidehpc.com/>



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Teaching Plan

Sr. No.	Unit	Broad Topic to be covered	Books Referred	Total Lectures Planned
1	Parallel Processing Concepts	Parallel Computing Platforms, Physical Organization of Parallel Platforms	T1, R1, R3	9
2	Parallel Programming	Principles of Parallel Algorithm Design	T1, R3	8
3	Basic Communication	Improving the Speed of Some Communication Operations.	T1, R4	8
4	Analytical Models of Parallel Programs	Performance Metrics for Parallel Systems	T1, R4	9
5	Parallel Algorithms- Sorting and Graph	Sorting on Parallel Computers, Algorithm for sparse graph	T1, R4	9
6	CUDA Architecture	Parallel programming in CUDA- C.	T2, R2	9

Question Bank

**Unit I
Introduction**

1. Explain pipelining and superscalar execution with example.
2. Explain the impact of strided access with example
3. Explain multithreading for latency hiding with example
4. Compare SIMD and MIMD
5. Explain static network topologies connecting p nodes with respect to cost and arc connectivity.
6. Describe invalidate and update protocol.
7. Explain Dynamic network topologies connecting p nodes with respect to cost and bisection width.



8. Explain simple 3-state protocol with diagram.
9. Explain snoopy bus cache coherence system with diagram.
10. Compare centralized and distributed directory based schemes.
11. Compare three types of routing with respect to communication cost.
12. List the overheads of applying communication costs to shared-address-space machines.
13. What is superscalar execution?
14. What is pipelining execution?
15. Explain the issues the superscalar execution
16. What is dynamic instruction issue?
17. Define with example vertical waste and horizontal waste.
18. What is VLIW?
19. Define following terms w. r. t. memory-i) latency ii) Bandwidth iii) hit ratio iv) memory bound v) temporal locality of reference vi) cache line vii) spacial locality viii) strided access ix) tiling x) prefetching xi) multithreading

Unit II

Parallel Programming

1. List the tasks in designing Non-trivial Parallel Algorithm.
2. What is decomposition of computation?
3. Define Tasks
4. What is task-dependency graph? Explain with example.
5. What is granularity of decomposition?
6. Define 1) Fine grained decomposition 2) coarse-grained decomposition
7. Explain Degree of Concurrency
8. Explain task interaction graph with example.
9. Explain task-process mapping with an example.
10. Explain recursive decomposition technique with an example.
11. Explain Data Decomposition technique with an example.
12. What is owner computes rule?
13. Explain exploratory decomposition with an example.
14. Explain speculative decomposition with an example.
15. Explain hybrid decomposition with an example.
16. Describe various characteristics of Tasks.
17. Describe various characteristics of interactions.
18. Explain different ways of mapping based on data partitioning
19. Explain mapping based on Task partitioning
20. Explain schemes for Dynamic Mapping.
21. What are different parallel algorithmic Models
22. Explain different techniques to reduce interaction overheads incurred by parallel programs



Unit III
Basic Communication

1. Group communication operations are built using which primitives?
2. ___ can be performed in an identical fashion by inverting the process.
3. Broadcast and reduction operations on a mesh is performed
4. Cost Analysis on a ring is
5. Cost Analysis on a mesh is
6. Communication between two directly link nodes
7. All-to-one communication (reduction) is the dual of _____ broadcast.
8. Which is known as Reduction?
9. Which is known as Broadcast?
10. The dual of all-to-all broadcast is
11. All-to-all broadcast algorithm for the 2D mesh is based on the
12. In the first phase of 2D Mesh All to All, the message size is ___
13. In the second phase of 2D Mesh All to All, the message size is ___
14. In All to All on Hypercube, the size of the message to be transmitted at the next step is ___ by concatenating the received message with their current data
15. The all-to-all broadcast on Hypercube needs ___ steps
16. One-to-All Personalized Communication operation is commonly called ___
17. The dual of the scatter operation is the
18. In Scatter Operation on Hypercube, on each step, the size of the messages communicated is ___
19. Which is also called "Total Exchange"?
20. All-to-all personalized communication can be used in ___
21. In collective communication operations, collective means
22. efficiency of data parallel algorithm depends on the
23. All processes participate in a single _____ interaction operation.
24. subsets of processes in _____ interaction.
25. Goal of good algorithm is to implement commonly used _____ pattern.
26. Reduction can be used to find the sum, product, maximum, minimum of _____ of numbers.
27. all processes that have the data can send it again is
28. accumulate results and send with the same pattern is...
29. every node on the linear array has the data and broadcast on the columns with the linear array algorithm in _____
30. using different links every time and forwarding in parallel again is
31. In a balanced binary tree processing nodes is equal to
32. In one -to- all broadcast there is
33. For sake of simplicity, the number of nodes is a power of
34. Nides with zero in i least significant bits participate in _____
35. every node has to know when to communicate that is
36. the procedure is disturbed and require only point-to-point _____
37. Renaming relative to the source is _____ the source.



Unit IV

Analytical Models of Parallel Programs

- 1 mathematically efficiency is _____
- 2 Cost of a parallel system is sometimes referred to ____ of product
- 3 Scaling Characteristics of Parallel Programs T_s is
- 4 Speedup tends to saturate and efficiency ____ because of Amdahl's law.
Speedup obtained when the problem size is _____ linearly
- 5 with the number of processing elements.
The $n \times n$ matrix is partitioned among n processors, with each processor storing
- 6 complete ____ of the matrix.
- 7 cost-optimal parallel systems have an efficiency of ____
The $n \times n$ matrix is partitioned among n^2 processors such that each processor owns a
- 8 _____ element.
- 9 how many basic communication operations are used in matrix vector multiplication
- 10 In DNS algorithm of matrix multiplication, it used
- 11 In the Pipelined Execution, steps contain
the cost of the parallel algorithm is higher than the sequential run time by a factor of
- 12 _____
The load imbalance problem in Parallel Gaussian Elimination: can be alleviated by
- 13 using a ____ mapping
- 14 A parallel algorithm is evaluated by its runtime in function of
- 15 For a problem consisting of W units of work, ____processors can be used optimally.
- 16 $C(W) \propto \Theta(W)$ for optimality (necessary condition).
- 17 many interactions in practical parallel programs occur in _____ pattern
- 18 efficient implementation of basic communication operation can improve
- 19 efficient use of basic communication operations can reduce
- 20 Group communication operations are built using _____ Messaging primitives.
- 21 one processor has a piece of data and it need to send to everyone is
- 22 the dual of one -to-all is
- 23 Data items must be combined piecewise, and the result made available at
- 24 simplest way to send $p-1$ messages from source to the other $p-1$ processors
- 25 In a eight node ring, node ____ is source of broadcast
- 26 The processors compute _____ product of the vector element and the local matrix
- 27 one to all broadcast use
- 28 In a broadcast and reduction on a balanced binary tree reduction is done in _____
- 29 if "X" is the message to broadcast it initially resides at the source node
- 30 logical operators used in algorithm are
- 31 Generalization of broadcast in Which each processor is
- 32 The algorithm terminates in _____ steps
- 33 Each node first sends to one of its neighbours the data it needs to....
- 34 The second communication phase is a column wise _____ broadcast of consolidated
- 35 All nodes collect _____ message corresponding to \sqrt{p} nodes to their respectively
- 36 It is not possible to port ____ for higher dimensional network
- 37 If we port algorithm to higher dimensional network, it will cause _____
- 38 In the scatter operation ____ node send message to every other node
- 39 The gather Operation is exactly the inverse of _____
- 40 Similar communication pattern to all-to-all broadcast except in the _____



Unit V

Parallel Algorithms-Sorting and Graphs

1. What are different issues in sorting on parallel computers.
2. Explain sorting network with suitable example.
3. What is the advantage of mapping bitonic sort to a Hypercube and a mesh
4. Explain odd-even transposition algorithm and parallel runtime for the same.
5. Explain Shell sort algorithm and parallel runtime for the same.
6. Explain CRCW PRAM parallel quicksort formulation
7. Explain shared address space parallel quick sort formulation.
8. Explain message passing parallel quick sort formulation.
9. Explain Cannon's Algorithm for Matrix-matrix multiplication
10. Explain the effect of pivot selection in the performance of quick sort.
11. Analyse the sample sort parallel algorithm on message passing computer.
12. Explain single source shortest path algorithm with suitable example.

Unit VI

CUDA Architecture

- 1 A CUDA program is comprised of two primary components: a host and a _____.
- 2 The kernel code is identified by the _____qualifier with void return type
- 3 The kernel code is only callable by the host(T/F)
- 4 The kernel code is executable on the device and host(T/F)
- 5 Calling a kernel is typically referred to as _____.
- 6 Host codes in a CUDA application can Initialize a device(T/F)
- 7 Host codes in a CUDA application can Allocate GPU memory(T/F)
- 8 A CUDA program is comprised of two primary components: a host and a _____.
- 9 Draw cuda architecture and explain.
- 10 Explain various types of memories in cuda.
- 11 Host codes in a CUDA application cannot Invoke kernels(T/F)
- 12 CUDA offers the Chevron Syntax to configure and execute a kernel. (T/F)
- 13 the BlockPerGrid and ThreadPerBlock parameters are related to the _____ model supported by CUDA.
- 14 _____ is Callable from the device only
- 15 _____ is Callable from the host
- 16 _____ is Callable from the host
- 17 CUDA supports _____ in which code in a single thread is executed by all other threads.
- 18 In CUDA, a single invoked kernel is referred to as a _____.
- 19 A grid is comprised of _____ of threads.
- 20 A block is comprised of multiple _____.
- 21 A solution of the problem in representing the parallel is min algorithm is _____
- 22 _____ is Callable from the host
- 23 In CUDA, a single invoked kernel is referred to as a _____.
- 24 the BlockPerGrid and ThreadPerBlock parameters are related to the _____ model supported by CUDA.
- 25 Host codes in a CUDA application can Transfer data to and from the device(T/F)
- 26 Host codes in a CUDA application cannot Deallocate memory on the GPU(T/F)
- 27 Host codes in a CUDA application cannot Reset a device(T/F)
- 28 Calling a kernel is typically referred to as _____.



Subject – 410242: Artificial Intelligence and Robotics

Weekly Work Load (in Hrs)	Lecture	Tutorial	Practical
	03	-	-

Online/ In-sem	Theory	Practical	Oral	Term-work	Total Marks	Credit
30	70	-	-	-	100	3

Course Objective

1. To teach various techniques of problem solving and game playing.
2. To introduce the concepts and application areas of Artificial Intelligence.
3. To explain importance of knowledge representation in certainty.
4. To give an overview of Planning and how robots perceive and act.

Course Outcome

- C402.1 Analyze the problem and apply Artificial Intelligence techniques for problem solving.
- C402.2 Interpret and solve problems related to planning and constraint satisfaction.
- C402.3 Represent various real-life problems using logic-based techniques and draw inference.
- C402.4 State and explain the applications of Artificial Intelligence in the real world.



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

- C402.5 Explain various components, functions, and applications of robots in real world.

Syllabus

Unit	<u>Course Contents</u>	<u>Hou rs</u>
I	Introduction	07
	Artificial Intelligence: Introduction, Typical Applications. State Space Search: Depth Bounded DFS, Depth First Iterative Deepening. Heuristic Search: Heuristic Functions, Best First Search, Hill Climbing, Variable Neighborhood Descent, Beam Search, Tabu Search. Optimal Search: A* algorithm, Iterative Deepening A*, Recursive Best First Search, Pruning the CLOSED and OPEN Lists.	
II	Problem Decomposition and Planning	07
	Problem Decomposition: Goal Trees, Rule Based Systems, Rule Based Expert Systems. Planning: STRIPS, Forward and Backward State Space Planning, Goal Stack Planning, Plan Space Planning, A Unified Framework For Planning. Constraint Satisfaction : N-Queens, Constraint Propagation, Scene Labeling, Higher order and Directional Consistencies, Backtracking and Look ahead Strategies.	
III	Logic and Reasoning	07
	Knowledge Based Reasoning: Agents, Facets of Knowledge. Logic and Inferences : Formal Logic, Propositional and First Order Logic, Resolution in Propositional and First Order Logic, Deductive Retrieval, Backward Chaining, Second order Logic Knowledge Representation: Conceptual Dependency, Frames, and Semantic nets.	
IV	Natural Language Processing and ANN	07
	Natural Language Processing: Introduction, Stages in natural language Processing, Application of NLP in Machine Translation, Information Retrieval and Big Data Information Retrieval. Learning: Supervised, Unsupervised and Reinforcement learning. Artificial Neural Networks (ANNs) : Concept, Feed forward and Feedback ANNs, Error Back Propagation, Boltzmann Machine.	
V	Robotics	07



**Progressive Education Society's
Modern College of Engineering**

DEPARTMENT OF COMPUTER ENGINEERING

	Robotics: Fundamentals, path Planning for Point Robot, Sensing and mapping for Point Robot, Mobile Robot Hardware, Non Visual Sensors like: Contact Sensors, Inertial Sensors, Infrared Sensors, Sonar, Radar, laser Rangefinders, Biological Sensing. Robot System Control: Horizontal and Vertical Decomposition, Hybrid Control Architectures, Middleware, High-Level Control, Human-Robot Interface.	
VI	Robots in Practice	07
	Robot Pose Maintenance and Localization: Simple Landmark Measurement, Servo Control, Recursive Filtering, Global Localization. Mapping: Sensorial Maps, Topological Maps, Geometric Maps, Exploration. Robots in Practice: Delivery Robots, Intelligent Vehicles, Mining Automation, Space Robotics, Autonomous Aircrafts, Agriculture, Forestry, Domestic Robots.	

Text Books

Sr.No		Text Books
1	T1	Deepak Khemani, "A First Course in Artificial Intelligence", McGraw Hill Education(India), 2013, ISBN : 978-1-25-902998-1
2	T2	Elaine Rich, Kevin Knight and Nair, "Artificial Intelligence", TMH, ISBN-978-0-07-008770-5
3	T3	Stuart Russell and Peter Norvig, "Artificial Intelligence: A Modern Approach", Third edition, Pearson, 2003, ISBN :10: 0136042597
4	T4	Michael Jenkin, Gregory, " Computational Principals of Mobile Robotics", Cambridge University Press, 2010, ISBN : 978-0-52-187157-0



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Reference Books

Sr.No		Text Books
1	R1	Nilsson Nils J, "Artificial Intelligence: A new Synthesis, Morgan Kaufmann Publishers Inc. San Francisco, CA, ISBN: 978-1-55-860467-4
2	R2	Patrick Henry Winston, "Artificial Intelligence", Addison-Wesley Publishing Company, ISBN: 0-201-53377-4
3	R3	Andries P. Engelbrecht-Computational Intelligence: An Introduction, 2nd Edition-Wiley India- ISBN: 978-0-470-51250-0

**Reference Web Links/ Research Paper/ Referred Book other than
Mention in Syllabus:**

- http://ictanddevelopment.blogspot.com/2009/06/ict-and-e-governance_15.html

Teaching Plan

Sr. No.	Unit	Broad Topic to be covered	Books Referred	Total Lectures Planned
1	I : Introduction	Artificial Intelligence, Heuristic Search, Optimal Search	4.a.1	7
2	II : Problem Decomposition and Planning	Problem Decomposition, Planning, Constraint Satisfaction	4.a.1	8
3	III : Logic and Reasoning	Knowledge Based Reasoning, Logic and Inferences, Second order Logic Knowledge Representation	4.a.1 4.b.1 4.c.1	7
4	IV : Natural Language Processing and ANN	Natural Language Processing and ANN, Artificial Neural Networks	4.b.1 4.a.2	8



**Progressive Education Society's
Modern College of Engineering**

DEPARTMENT OF COMPUTER ENGINEERING

5	V : Robotics	Robotics, Non Visual Sensors, Robot System Control	4.b.1 4.a.2	7
6	VI : Robots in Practice	Robot Pose Maintenance and Localization, Mapping, Robots in Practice	4.b.1 4.a.2	8

Question Bank

Unit 1: Introduction

Q. No	Questions
1	What are the typical Applications of Artificial Intelligence?
2	Define Artificial Intelligence.
3	What is State Space Search: Depth Bounded DFS?
4	What is Depth First Iterative Deepening?
5	Explain Best First Search
6	Explain Hill Climbing algorithm
7	Explain Variable Neighborhood Descent
8	Explain A* algorithm
9	Explain Iterative Deepening A*
10	Explain Recursive Best First Search
11	Explain Pruning the CLOSED and OPEN Lists
12	Explain below search algorithm in brief: <ul style="list-style-type: none"> a. Beam Search b. Tabu Search c. Heuristic Search d. Optimal Search
13	What are intelligent agents? Explain the architecture of typical agent and give at least two examples where agents are used.
14	What are attributes of agent design(hint:PEAS)
15	Which are the types of Agent? Explain any 2 in detail.



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

16	What mean by AI? Explain problem solving with AI?
17	Best First search uses both an OPEN list and a CLOSED list. Describe the purpose of each for the Best-First algorithm. Explain with suitable example.
18	Define search problem. Solve 8 queens as a state-space-search problem.

Unit 2: Problem Decomposition and Planning

Q. No	Questions
1	Explain Goal Trees along with neat diagram
2	Explain Rule Based Systems in detail
3	What is STRIPS?
4	Explain Forward and Backward State Space Planning
5	What is Goal Stack Planning?
6	Explain Plan Space Planning
7	Explain A Unified Framework For Planning
8	Explain N-Queens problem with neat diagrams
9	Describe the essence of a constraint satisfaction problem. What are some of the major applications of constraint satisfaction search?
10	What is Scene Labeling?
11	What is Higher order and Directional Consistencies?
12	Explain Backtracking and Look ahead Strategies
13	What is Hill climbing? Explain in detail with the problem of Local Maxima, Plateau and ridge



Unit 3: Logic and Reasoning

Q. No	Questions
1	Explain Knowledge Based Reasoning and Agents
2	What are Facets of Knowledge?
3	Explain Logic and Inferences
4	What is Propositional and First Order Logic?
5	Explain Resolution in Propositional and First Order Logic
6	What is Deductive Retrieval?
7	Explain Backward Chaining
8	What is Second order Logic?
9	Explain Knowledge Representation : Conceptual Dependency
10	What is Frames?
11	What are Semantic nets?
12	What are the applications of Knowledge Based systems?
13	Describe the PEAS for Wumpus world problem
14	Describe representation of knowledge using rules.
15	What is predicate logic? Describe the advantages of predicate logic over propositional logic.
16	Define Forward Chaining, Backward Chaining.
17	What is Algorithms for Planning as State-Space Search?
18	Write a note on goal stack planning
19	What is blocks world problem? Give suitable example.
20	Represent a suitable problem using STRIPS
21	What is a Logic? How it is represented in prolog?
22	Explain regression planner and progression planner.



Unit 4: Natural Language Processing and ANN

Q. No	Questions
1	What are the Stages in natural language Processing
2	What are application of NLP in Machine Translation
3	What is Information Retrieval and Big Data Information Retrieval
4	What is Artificial Neural Networks(ANNs)? Explain working of ANN
5	Explain Feed forward and Feedback ANNs,
6	Explain Error Back Propagation,
7	Explain Boltzman Machine.
8	What are application of Artificial Neural Networks(ANNs):
9	Explain below Learning methods: a. Supervised b. Unsupervised c. Reinforcement

Unit 5: Robotics

Q. No	Questions
1	Write short note on Mobile Robot Hardware
2	Explain path Planning for Point Robot
3	Explain Sensing and mapping for Point Robot
4	Explain Non Visual Sensors: a. Contact Sensors b. Inertial Sensors c. Infrared Sensors
5	Explain Sonar and Radar in brief
6	Explain laser Rangefinders and Biological Sensing
7	Explain Hybrid Control Architectures
8	Explain High-Level Control
9	Explain Human-Robot Interface
11	Explain Horizontal and Vertical Decomposition



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Unit 6: Robots in Practice

Q. No	Questions
1	Explain Robot Pose Maintenance and Localization
2	Explain Simple Landmark Measurement
3	What is Servo Control?
4	Recursive Filtering
5	What is Global Localization
6	Explain below Mapping: a. Sensorial Maps b. Topological Maps c. Geometric Maps
7	Write short note on Delivery Robots
9	Write short note on Space Robotics
11	Write short note on Domestic Robots



Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING

Subject – 410243: Data Analytics

Weekly Work Load (in Hrs)	Lecture	Tutorial	Practical
	03	-	-

Online/ In-sem	Theory	Practical	Oral	Term-work	Total Marks	Credit
30	70	-	-	-	100	3

Course Objectives

1. To develop problem solving abilities using Mathematics.
2. To apply algorithmic strategies while solving problems.
3. To develop time and space efficient algorithms.
4. To study algorithmic examples in distributed, Concurrent/Parallel environment.

Course Outcome

C403.1 Apply knowledge for analysis of data and select suitable methods for data pre-processing.

C403.2 Identify and apply analytical methods on data.

C403.3 Apply classification techniques on data to solve problems.

C403.4 Explain data visualization techniques.

C403.5 Make use of Hadoop ecosystem for analysis of unstructured data.



Syllabus

UNITWISE SYLLABUS AND OUTCOMES

UNIT – I(Introduction and Life Cycle)		
Topics:-		
Introduction: Big data overview, state of the practice in Analytics- BI Vs Data Science, Current Analytical Architecture, drivers of Big Data, Emerging Big Data Ecosystem and new approach. Data Analytic Life Cycle: Overview, phase 1- Discovery, Phase 2- Data preparation, Phase 3- Model Planning, Phase 4- Model Building, Phase 5- Communicate Results, Phase 6- Operationalize. Case Study: GINA		
Outcomes – At the end of this unit students will be able to -		No. of Lectures – 08
Sr. No.	Outcome	Bloom's Level
1	Define Big Data.	L1
2	Compare BI Vs Data Science.	L2
3	Explain Data Analytics Life Cycle.	L2
UNIT – II(Basic Data Analytic Methods)		
Topics –		
Statistical Methods for Evaluation- Hypothesis testing, difference of means, Wilcoxon rank-sum test, type 1 type 2 errors, power and sample size, ANNOVA. Advanced Analytical Theory and Methods: Clustering- Overview, K means- Use cases, Overview of methods, determining number of clusters, diagnostics, reasons to choose and cautions.		
Outcomes – At the end of this unit students will be able to -		No. of Lectures – 08
Sr. No.	Outcome	Bloom's Level
1	Make use of Hypothesis Testing for Statistical evaluation of Sample Data.	L3
2	Select the Appropriate analytical techniques for the given problem.	L3
3	Identify Suitable statistical method for evaluation of sample data	L3
UNIT – III(Association Rules and Regression)		
Topics –		



Advanced Analytical Theory and Methods: Association Rules- Overview, a-priori algorithm, evaluation of candidate rules, case study-transactions in grocery store, validation and testing, diagnostics. **Regression-** linear, logistics, reasons to choose and cautions, additional regression models.

Outcomes – At the end of this unit students will be able to - **No. of Lectures** – 08

Sr. No.	Outcome	Bloom's Level
1	Apply Apriori algorithm for data analysis.	L3
2	Define concept of Regression.	L1
3	Explain types of Regression	L2

UNIT – IV(Classification)

Topics –

Decision trees- Overview, general algorithm, decision tree algorithm, evaluating a decision tree. **Naïve Bayes** – Bayes' Algorithm, Naïve Bayes Classifier, smoothing, diagnostics. Diagnostics of classifiers, additional classification methods.

Outcomes – At the end of this unit students will be able to - **No. of Lectures** – 08

Sr. No.	Outcome	Bloom's Level
1	Explain basic concept of decision tree.	L1
2	Analyze the different techniques of classification of data.	L4
3	Make Use of Naïve Bayes Algorithm to solve problem.	L3

UNIT – V(Big Data Visualization)

Topics –

Introduction to Data visualization, Challenges to Big data visualization, Conventional data visualization tools, Techniques for visual data representations, Types of data visualization, Visualizing Big Data, Tools used in data visualization, Analytical techniques used in Big data visualization.

Outcomes – At the end of this unit students will be able to - **No. of Lectures** – 08

Sr. No.	Outcome	Bloom's Level
1	Tell what is data visualization.	L1
2	List challenges to big data visualization.	L1



Progressive Education Society's

Modern College of Engineering

DEPARTMENT OF COMPUTER ENGINEERING

3	Explain different types of data visualization & tool used for data visualization.	L2
UNIT – VI(Advanced Analytics-Technology and Tools)		
Topics –		
Analytics for unstructured data- Use cases, Map Reduce, Apache Hadoop. The Hadoop Ecosystem- Pig, HIVE, HBase, Mahout, NoSQL. An Analytics Project-Communicating, operationalizing, creating final deliverables.		
Outcomes – At the end of this unit students will be able to -		No. of Lectures – 08
Sr. No.	Outcome	Bloom's Level
1	Explain concept of Hadoop.	L2
2	Explain the Hadoop Ecosystem.	L2
3	Illustrate with example working of Map Reduce Algorithm.	L2

Text Books

Sr.No	Text Books
1	David Dietrich, Barry Hiller, “Data Science and Big Data Analytics”, EMC education services, Wiley publications, 2012, ISBN0-07-120413-X.
2	Ashutosh Nandeshwar , “Tableau Data Visualization Codebook”, Packt Publishing, ISBN 978-1-84968-978- 6.

Reference Books

Sr.No	Text Books
1	Maheshwari Anil, Rakshit, Acharya, “Data Analytics”, McGraw Hill, ISBN: 789353160258.
2	Mark Gardner, “Beginning R: The Statistical Programming Language”, Wrox Publication, ISBN: 978-1-118-16430-3
3	Luís Torgo, “Data Mining with R, Learning with Case Studies”, CRC Press, Talay and Francis Group, ISBN9781482234893.
4	.Carlo Vercellis, “Business Intelligence - Data Mining and Optimization for Decision Making”, Wiley Publications, ISBN: 9780470753866



Reference Web Links/ Research Paper/ Referred Book other than Mention in Syllabus:

1]Youtube Channel:

1. <https://www.youtube.com/watch?v=YZf5q-ICf8Y>
2. <https://www.youtube.com/watch?v=oQmxQh92Nh8>
3. <https://www.youtube.com/watch?v=THODdNXOjRw>

2]NPTEL PORTAL:

<https://nptel.ac.in/courses/110/106/110106072/>

3] ANALYTICS VIDHYA:

<https://www.analyticsvidhya.com/>

4] Slide Share:

<https://www.slideshare.net/krishnasingh320/1-data-analyticsintroduction>



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Teaching Plan

Sr. No.	Unit	Broad Topic to be covered	Books Referred	Total Lectures Planned
1	I	Introduction and Life Cycle	T1	8
2	II	Basic Data Analytic Methods	T1	8
3	III	Association Rules and Regression	T1	8
4	IV	Classification	T1	8
5	V	Big Data Visualization	T1	8
6	VI	Advanced Analytics - Technology and Tools	T1	8



Question Bank

UNIT_NO:1

- Q.No.1 What is Big Data?
- Q.NO.2 Explain Characteristic of Big Data.
- Q.No.3 Draw and Explain Data Analytics architecture.
- Q.No.4 Write Short note on Data Analytics life cycle.
- Q.No.5 Write Short note on GINA Case Study.
- Q.No.6 What is big data? Explain characteristics of Big data.
- Q.No.7 Explain different phases of data analytics life cycle.
- Q.No.8 What is big data? Explain 3V's of Big data.
- Q.No.9 Why communication is important in data analytics life cycle projects?
- Q.No.10 Compare BI Vs Data Science.

UNIT_NO:2

- Q.No.1 Explain Hypothesis Testing with example.
- Q.No.2 Explain Concept of Wilcoxon rank-sum test with example.
- Q.No.3 What is Clustering? Explain K-mean Clustering Technique.
- Q.No.4 Use the k-means algorithm and Euclidean distance to cluster the following 7 samples of A into 2 clusters: A1=(2,10), A2=(2,5), A3=(8,4), A4=(5,8), A5=(7,5) A6=(8,5) A7=(3,7).
- Q.No.5 Explain Null Hypothesis & Alternative Hypothesis.
- Q.No.6 Explain Type-I and Type-II Error.
- Q.No.7 How Wilcoxon Rank-sum Test works?
- Q.No.8 Explain ANOVA.
- Q.No.9 Use Below data and group them using k-mean clustering algorithm, show calculation of centroids.

Height	185	170	168	179	182	188	180	180	183	180	180	177
Weight	72	56	60	68	72	77	71	70	84	88	67	76



UNIT_NO:3

- Q.No.1 What is Regression? Explain Linear and Logistic Regression with Example.
Q.No.2 Define Support and Confidence.
Q.No.3 Explain Market Basket Analysis.
Q.No.4 Consider Database Consisting of 9 Transactions. Derive useful Association rule for given database transaction. Assume Min.Support count=2.

T_ID	List of Items
T100	I1,I2,I5
T101	I2,I4
T102	I2,I3
T103	I1,I2,I4
T104	I1,I3
T105	I2,I3
T106	I1,I3
T107	I1,I2,I3,I5
T108	I1,I2,I3

- Q.No.5 Explain Apriori association rule mining algorithm.
Q.No.6 Explain logistic regression? Explain use cases of logistic regression.
Q.No.7 Write Apriori Algorithms.
Q.No.8 Define Frequent Itemset,Lift.

UNIT_NO:4

- Q.No.1 Explain in brief Decision Tree Algorithms.
Q.No.2 Consider following database indicating patient previous history (Below are symptoms and Diagnosis).

Chills	Runny Nose	Headache	Fever	Flu
Y	N	MILD	Y	N
Y	Y	NO	N	Y
Y	N	STRONG	Y	Y
N	Y	MILD	Y	Y
N	N	NO	N	N
N	Y	STRONG	Y	Y
N	Y	STRONG	N	N
Y	Y	MILD	Y	Y

Do I believe that patient with following symptoms has flu? Solve using Naïve Bayes Classifier.

Chills	Runny Nose	Headache	Fever	Flu
---------------	-------------------	-----------------	--------------	------------



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Y	N	MILD	Y	?
---	---	------	---	---

- Q.No.3 Explain following decision tree algorithm a) ID3 Algorithm b) C4.5 c) CART
- Q.No.4 How Naïve Bayes classification works? Give its application.
- Q.No.5 Explain following terms i) Bagging ii) Boosting iii) Random forest
- Q.No.6 Explain Bayes theorem. Explain Naïve Bayes classifier.
- Q.No.7 Explain any three classification performance measures.
- Q.No.8 What is classification? List different classifiers.
- Q.No.9 What is decision tree? Explain how decision tree is constructed using ID3 algorithm.
- Q.No.10 Explain following with significance a) Entropy b) Information Gain c) Gain ratio

UNIT_NO:5

- Q.No.1 Explain types of data visualization.
- Q.No.2 List and Explain in brief tools used in visualization.
- Q.No.3 List Challenges to Big data visualization.
- Q.No.4 What is data visualization? Explain any four data visualization techniques.
- Q.No.5 Explain how data visualization done or visually represented, if 1-D, if data 2-D and data is 3-dimentional?
- Q.No.6 Explain analytical technique used in big data Visualization.
- Q.No.7 Explain data visualization tool-Tableau
- Q.No.8 Why it is difficult to visualize big data? Explain analytical techniques used in big data visualization.
- Q.No.9 Explain various tools to visualize big data.(Any Four)

UNIT_NO:6

- Q.No.1 Explain concept of Map Reduce with example.
- Q.No.2 Write Short note on HDFS.
- Q.No.3 Make a use of Map Reduce technique to process following input file.

I am Learning Data Analytics. I am Learning R Programming for Data Analytics.
--



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Q.No.4 Explain Hadoop Ecosystem.

Q.No.5 Write Short note on NOSQL.

Q.No.6 What is Map-Reduce? Explain working of Map reduce with example.

Q.No.7 Explain HDFS with respect to NameNode, DataNode, SecondaryNode with example.

Q.No.8 Explain Bigdata Ecosystem.

Q.No.9 What are four major categories of NOSQL tool(Store)?

Q.No.10 Explain Hadoop Ecosystem in detail with Pig, Hive, Hbase and Mathout.



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Subject -410244: Data Mining & Warehousing

Weekly Work Load (in Hrs)	Lecture	Tutorial	Practical
	03	-	08

Online/ In-sem	Theory	Practical	Oral	Term-work	Total Marks	Credit
30	70	-	-	-	100	3

Course Objectives

1. To explain fundamentals of Data Mining.
2. To explain the preprocessing, mining and post processing techniques of the data.
3. To elaborate different data mining techniques for decision making.
4. To illustrate different algorithms, techniques, and tools for data Mining applications.

Course Outcome

C404D.1 Analyze the problem and apply suitable pre-processing technique.

C404D.2 Construct a data warehouse model and identify OLAP operations to extract hidden patterns.

C404D.3 Apply basic, intermediate, and advanced techniques to mine the data.

C404D.4 Draw inference from the output generated to predict the patterns.

C404D.5 Compare and apply appropriate data mining technique to solve real world problems.



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Syllabus

Unit	<u>Course Contents</u>	<u>Hours</u>
I	Basics of Data Mining	08
	Data Mining, Data Mining Task Primitives, Data: Data, Information and Knowledge ; Attribute Types: Nominal, Binary, Ordinal and Numeric attributes, Discrete versus Continuous Attributes; Introduction to Data Preprocessing, Data Cleaning: Missing values, Noisy data; Data integration: Correlation analysis; transformation: Min-max normalization, zscore normalization and decimal scaling; data reduction: Data Cube Aggregation, Attribute Subset Selection, sampling; and Data Discretization: Binning, Histogram Analysis.	
II	Data Warehouse	08
	Data Warehouse, Operational Database Systems and Data Warehouses(OLTP Vs OLAP), A Multidimensional Data Model: Data Cubes, Stars, Snowflakes, and Fact Constellations Schemas; OLAP Operations in the Multidimensional Data Model, Concept Hierarchies, Data Warehouse Architecture, The Process of Data Warehouse Design, A three-tier data warehousing architecture, Types of OLAP Servers: ROLAP versus MOLAP versus HOLAP.	
III	Measuring Data Similarity and Dissimilarity	08
	Measuring Data Similarity and Dissimilarity, Proximity Measures for Nominal Attributes and Binary Attributes, interval scaled; Dissimilarity of Numeric Data: Minkowski Distance, Euclidean distance and Manhattan distance; Proximity Measures for Categorical, Ordinal Attributes, Ratio scaled variables; Dissimilarity for Attributes of Mixed Types, Cosine Similarity.	
IV	Association Rule Mining	08
	Market basket Analysis, Frequent item set, Closed item set, Association Rules, a-priori Algorithm, Generating Association Rules from Frequent Item sets, Improving the Efficiency of a-priori, Mining Frequent Item sets without Candidate Generation: FP Growth Algorithm; Mining Various Kinds of Association Rules: Mining multilevel association rules, constraint based association rule mining, Meta rule-Guided Mining of Association Rules.	
V	Classifier Algorithms	08
	Introduction to: Classification and Regression for Predictive Analysis, Decision Tree Induction, Rule-Based Classification: using IF-THEN Rules for Classification, Rule Induction Using a Sequential Covering Algorithm. Bayesian Belief Networks, Training Bayesian Belief Networks, Classification Using Frequent Patterns, Associative Classification, Lazy Learners-k-Nearest Neighbor Classifiers, Case-Based Reasoning.	



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

VI	Multiclass Classification	08
	Multiclass Classification, Semi-Supervised Classification, Reinforcement learning, Systematic Learning, holistic learning and multi-perspective learning. Metrics for Evaluating Classifier Performance: Accuracy, Error Rate, precision, Recall, Sensitivity, Specificity; Evaluating the Accuracy of a Classifier: Holdout Method, Random Sub sampling and Cross Validation.	

1.3 Text Books

Sr.No		Text Books
1	T1	1. Han, Jiawei Kamber, Micheline Pei and Jian, "Data Mining: Concepts and Techniques", Elsevier Publishers, ISBN:9780123814791, 9780123814807. 2. Parag Kulkarni, "Reinforcement and Systemic Machine Learning for Decision Making" by Wiley-IEEE Press, ISBN: 978-0-470-91999-6
2	T2	1. Han, Jiawei Kamber, Micheline Pei and Jian, "Data Mining: Concepts and Techniques", Elsevier Publishers, ISBN:9780123814791, 9780123814807. 2. Parag Kulkarni, "Reinforcement and Systemic Machine Learning for Decision Making" by Wiley-IEEE Press, ISBN: 978-0-470-91999-6

Reference Books

Sr.No		Text Books
1	R1	1. Matthew A. Russell, "Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More" , Shroff Publishers, 2nd Edition, ISBN: 9780596006068 2. Maksim Tsvetovat, Alexander Kouznetsov, "Social Network Analysis for Startups: Finding connections on the social web", Shroff Publishers , ISBN: 10: 1449306462
2	R2	



		<ol style="list-style-type: none">1. Matthew A. Russell, "Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More" , Shroff Publishers, 2nd Edition, ISBN: 97805960060682. Maksim Tsvetovat, Alexander Kouznetsov, "Social Network Analysis for Startups: Finding connections on the social web", Shroff Publishers , ISBN: 10: 1449306462
--	--	---

**Reference Web Links/ Research Paper/ Referred Book other than
Mention in Syllabus:**

1. Rapid Miner online tutorial

<https://academy.rapidminer.com/learning-paths/get-started-with-rapidminer-and-machine-learning>

1. Youtube video Channel Getting Started with RapidMiner

<https://www.youtube.com/watch?v=Gg01mmR3j&list=PLssWC2d9JhOZLbQNZ80uOxLypglgWqbJA>

2. Jiawei Han, Micheline Kamber and Jian Pei official blog for study material

https://hanj.cs.illinois.edu/bk3/bk3_slidesindex.htm



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Teaching Plan

Sr. No.	Unit	Broad Topic to be covered	Books Referred	Total Lectures Planned
1	I	Introduction to Data Mining, Data Preprocessing Techniques	T1, R1	8
2	II	Data Warehouse Architecture, Multidimensional Data Model ,OLAP operations,	T1, R1	8
3	III	Data Similarity and Dissimilarity techniques	T1, R1	8
4	IV	Association rule mining, Apriori Algorithm.	T1, R1	8
5	V	Introduction to Classification, Bayesian Belief Networks	T1, T2	8
6	VI	Multiclass Classification	T1, T2	8



Question Bank

UNIT - I

- what is data Mining?
- Explain the differences between Knowledge discovery and data mining.
- Explain different data mining tasks.
- what are the application areas of data Mining?
- what is the relation between data warehousing and data mining?
- List out different sources of information.
- what are the key issues in data mining?
- what do you mean by Data Processing?
- Explain data cleaning.
- Describe different data cleaning approaches.
- How can we handle missing values?
- Explain Noisy Data.
- Explain various normalization Techniques.
- Give Brief description of following:
 - a. Binning
 - b. Regression
 - c. Clustering
 - d. Smoothing
 - e. Generalization
 - f. Aggregation
- Briefly describe the four stages of knowledge discovery (KDD)?

UNIT - II

- Discuss the components of data warehouse.
- List out the differences between OLTP and OLAP.
- Discuss the various schematic representations in multidimensional model.
- Explain the OLAP operations In multidimensional model.
- Explain the design and construction of a data warehouse.
- Explain the three-tier data warehouse architecture.
- Explain indexing.
- Write notes on metadata repository.
- How does a snowflake schema differ from a star schema ? Name two advantages and two disadvantages of the snowflake schema.
- What is meant by slice and dice? Give an example.
- What are the essential differences between the MOLAP and ROLAP models? Also list a few similarities.



- Why is the entity-relationship modeling technique not suitable for the data warehouse?
- what are the similarities and differences between data warehouse & Database?
- What is ETL process.
- What is the STAR Schema? What are the Fact tables?

UNIT - III

- A data set for analysis includes only one attribute X:
 $X = \{7, 12, 5, 8, 5, 9, 13, 12, 19, 7, 12, 12, 13, 3, 4, 5, 13, 8, 7, 6\}$
What is the mean of the data set X?
What is the median?
Find the standard deviation for X.
- Define Clustering.
- What do you mean by Cluster Analysis?
- What are the fields in which clustering techniques are used?
- What are the requirements of cluster analysis?
- What is interval scaled variables?
- Define Binary variables? And what are the two types of binary variables?
- Define nominal, ordinal and ratio scaled variables?
- What do u mean by partitioning method?
- Differentiate Agglomerative and Divisive Hierarchical Clustering?

UNIT – IV

- Define Frequent sets, confidence, support and association rule.
- What do you mean by Market Basket analysis and how it can help in a supermarket?
- Explain whether association rule mining is supervised or unsupervised type of learning.
- Name some variants of Apriori Algorithm.
- Discuss the importance of Association Rule Mining.
- Consider the Data set D. Given the minimum support², apply apriori algorithm on this dataset.

• Transaction ID	• Items
• 100	• A, C, D
• 200	• B, C, E
• 300	• A, B, C,E
• 400	• B, E

- How to generate association rules from frequent item sets?
- Give few techniques to improve the efficiency of Apriori algorithm?



- What are the things suffering the performance of Apriori candidate generation technique?
- Describe the method of generating frequent item sets without candidate generation. Mention few approaches to mining Multilevel Association Rules? What are multidimensional association rules?
- Explain constraint-based association mining?
- Explain FP growth algorithm?
- Describe the different classifications of Association rule mining?

UNIT – V

- Classification is supervised learning. Justify.
- Define classification. Explain the purposes of using a classification model?
- Explain different classification Techniques.
- Entropy is an important concept in information theory. Explain its significance in mining context.
- Explain Naive Baye's Classification.
- Describe the essential features of decision trees in context of classification.
- What are the advantages and disadvantages of decision trees over other classification methods?

UNIT – VI

- Explain ID3 Algorithm.
- Explain classification by Decision tree induction?
- Explain Bayesian classification?
- Explain the general approach for building a classification model?
- What are the various characteristics of decision tree induction?
- Explain the rule-based classifier with an example?
- Explain the Nearest-Neighbor classifier?
- Discuss the k-nearest neighbor classification algorithm?
- Explain the characteristics of Nearest-Neighbor classifiers?



Name of the Subject –410245 (B): Software Testing and Quality Assurance (STQA)

Weekly Work Load(in Hrs)	Lecture	Tutorial	Practical
	03	-	-

Online/ In-sem	Theory	Practical	Oral	Term- work	Total Marks	Credit
30	70	-	-	-	100	3

Course Objective

1. Introduce basic concepts of software testing.
2. Understand white box, block box, object oriented, web based and cloud testing.
3. Know in details automation testing and tools used for automation testing.
4. Understand the importance of software quality and assurance software systems development.

Course Outcome

C405B.1 Identify type of software testing for a given scenario.

C405B.2 Design test plan, test cases and test using sample data.

C405B.3 Make use of automation tools for testing software.

C405B.4 Explain quality management, assurance, and quality standards for software systems.

C405B.5 Explain Software Quality Tools.



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Syllabus

Unit	<u>Course Contents</u>	<u>Hours</u>
I	Introduction Software Quality	07
	<p>Topics – Introduction, historical perspective, Definition, Core Components, Quality View, Financial Aspect, Customer’s suppliers and process, Total Quality Management (TQM), Quality practices of TQM, Quality Management through- Statistical process Control, Cultural Changes, Continual Improvement cycle, quality in different areas, Benchmarking and metrics, Problem Solving Techniques, Problem Solving Software Tools.</p> <p>Software Quality- Introduction, Constraints of Software product Quality assessment, Customer is a King, Quality and Productivity Relationship, Requirements of Product, Organization Culture, Characteristics of Software, Software Development Process, Types of Product, Criticality Definitions, Problematic areas of SDLC, Software Quality Management, Why Software has defects, Processes related to Software Quality, Quality Management System’s Structure, Pillars of Quality Management System, Important aspects of quality management</p>	
II	Test Planning and Management	07
	<p>Topics – Review of Fundamentals of Software Testing, Testing during development life cycle, Requirement Traceability matrix, essentials, Work bench, Important Features of Testing Process, Misconceptions, Principles, salient and policy of Software testing, Test Strategy, Test Planning, Testing Process and number of defects found, Test team efficiency, Mutation testing, challenges, test team approach, Process problem faced, Cost aspect, establishing testing policy, methods, structured approach, categories of defect, Defect/ error/ mistake in software, Developing Test Strategy and Plan, Testing process, Attitude towards testing, approaches, challenges, Raising management awareness for testing, skills required by tester..</p>	
III	Software Test Automation	07
	<p>Topics – What is Test Automation, Terms used in automation, Skills needed for automation, What to automate, scope of automation, Design and Architecture of automation, Generic requirement for Test Tool, Process Model for Automation, Selecting Test Tool, Automation for XP/Agile model, Challenges in Automation, Data-driven Testing. Automation Tools like JUnit, Jmeter</p>	
IV	Selenium Tool	07



**Progressive Education Society's
Modern College of Engineering**

DEPARTMENT OF COMPUTER ENGINEERING

	Topics – Introducing Selenium, Brief History of The Selenium Project, Selenium’s Tool Suite, Selenium- IDE, Selenium RC, Selenium Webdriver, Selenium Grid, Test Design Considerations.	
V	Quality Management	07
	Topics – Software Quality, Software Quality Dilemma, Achieving Software Quality, Software Quality Assurance. Elements of SQA, SQA Tasks, Goals, and Metrics, Formal Approaches to SQA, Statistical Software Quality Assurance, Six Sigma for Software Engineering, ISO 9000 Quality Standards, SQA Plan.	
VI	Software Quality Tools	07
	Topics – Total Quality Management, Product Quality Metrics, In process Quality Metrics, Software maintenance, Ishikawa's 7 basic tools, Checklists, Pareto diagrams, Histogram, Run Charts, Scatter diagrams, Control chart, Cause Effect diagram. Defect Removal Effectiveness and Process Maturity Level. Books	



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Text Books

Sr.No		Text Books
1	T1	M G Limaye, “Software Testing Principles, Techniques and Tools”, Tata McGraw Hill, ISBN: 9780070139909 0070139903
2	T2	Srinivasan Desikan, Gopalswamy Ramesh, “Software Testing Principles and Practices”, Pearson, ISBN-10: 817758121X

Reference Books

Sr.No		Text Books
1	R1	Srinivasan Desikan, Gopalswamy Ramesh, “Software Testing Principles and Practices”, Pearson, ISBN-10: 817758121X
2	R2	Stephen Kan, “Metrics and Models in Software Quality Engineering”, Pearson, ISBN-10: 0133988082; ISBN-13: 978-0133988086

**Reference Web Links/ Research Paper/ Referred Book other than
Mention in Syllabus:**

1. <https://github.com/SeleniumHQ/selenium>
2. <https://www.selenium.dev/>



Teaching Plan

Sr. No.	Unit	Broad Topic to be covered	Books Referred	Total Lectures Planned
1	I : Introduction	Artificial Intelligence, Heuristic Search, Optimal Search	4.a.1	7
I :	Introduction , Software Quality	T1,T2	8	8
II :	Test Planning and Management	T1,R1	8	7
III :	Software Test Automation	T1,T2	8	8
IV :	Selenium Tool	Web Link	8	7
V :	Quality Management	T1,R2	8	8
VI :	Software Quality Tools	T1,R2	8	7



Question Bank
Unit 1: Introduction

Q. No	Questions
1	1. Explain 'Quality' in terms of the generic expectations from any product.
2	2. Differentiate between continuous improvement and continual improvement.
3	3. Define the stakeholders for successful projects at micro level and for successful organizations at macro level.
4	4. Define 'quality' as viewed by different stakeholders of software development and usage.
5	5. Explain customers view of quality.
6	6. Explain manufacturers view of quality.
7	7. Define 'User's Gap' and 'Producer's Gap' and explain how these gaps can be closed effectively.
8	8. Describe various definitions of quality as per international standards.
9	9. Describe definition of quality as per Dr. Deming, Dr. Juran and Philip Crosby.
10	10. Describe 'Total Quality Management' principles of continual improvement.
11	11. Describe cultural change requirement for quality improvement.
12	12. Differentiate between tools and techniques.
13	What are the constraints of software requirement specifications?
14	2. Explain relationship between quality and productivity.
15	3. Explain the concept of 'q' organizations and 'Q' organizations.
16	4. Explain different development models.
17	5. How products are classified depending upon their criticality.
18	6. What are the different types of requirements?



Unit 2: Test Planning and Management

Q. No	Questions
1	1. Explain the evolution of software testing from debugging to prevention based testing.
2	2. Explain why independent testing is required.
3	3. Explain big bang approach of software testing.
4	4. Explain total quality management approach of software testing.
5	5. Explain concept of TQM cost perspective.
6	6. Explain testing as a process of software certification.
7	7. Explain the basic principles on which testing is based.
8	8. Explain the concept of test team's defect finding efficiency.
9	9. Explain test case's defect finding efficiency.
10	10. What are the challenges faced by testers.
11	11. Explain the process of developing test strategy.
12	12. Explain the process of developing test methodology.
13	13. Which skills are expected in a good tester.



Unit 3: Software Test Automation

Questions
1. What do you know about Selenium?
2. What are the technical challenges with selenium?
3. What are the test types supported by Selenium?
4. What are the capabilities of Selenium IDE?
5. What are the challenges with Selenium IDE?
6. Which are the browsers supported by Selenium IDE?
7. How to execute a single line command from Selenium IDE?
8. How to insert a start point in Selenium IDE?
9. How to insert a comment in Selenium IDE?
10. How to insert a break point in Selenium IDE?
11. How to debug the tests in Selenium IDE?
12. How to export the tests from Selenium IDE to Selenium RC in different languages?
13. How to export Selenium IDE Test Suite to Selenium RC Suites?
14. Which is the command used for displaying the values of a variable into the output console or log?
15. Which are the browsers supported by Selenium RC?
16. Which are the Operating Systems supported by Selenium?
17. What is Selenium RC? 18. Why Selenium RC is used?
18. Which are the languages supported by Selenium RC?
19. What is Selenium Grid?
20. What is Selenium WebDriver or Google WebDriver or Selenium 2.0?
21. What are the capabilities of Selenium WebDriver or Google WebDriver or Selenium 2.0?
22. What is the architecture of Selenium RC?



Unit 4: Selenium Tool

Q. No	Questions
1	What is Automation Testing?
2	What are the advantages/benefits of Automation Testing?
3	How many Test cases have you automated per day?
4	What is Selenium?
5	What are the different Selenium suite Components?
6	Why should I use Selenium?

Unit 5: Quality Management

Q. No	Questions
1	When should the process Manage Quality be conducted? ...
2	What is the best description of a Pareto diagram? ...
3	Who defines quality, according to the modern quality management approach?



Name of the Subject –Laboratory Practice I

Weekly Work Load(in Hrs)	Lecture	Tutorial	Practical
	-	-	04

Online/ In-sem	Theory	Practical	Oral	Term- work	Total Marks	Credit
-	-	50	-	50	100	2

Course Objectives

1. To develop problem solving abilities using Mathematical Modeling.
2. To develop problem solving abilities using HPC.
3. To develop problem solving abilities using Artificial Intelligence and Robotics.
4. To develop problem solving abilities using Data Analytics.

Course Outcomes:

- C406.1 Develop and analyze various parallel programs.
- C406.2 Apply AI algorithmic strategies for solving various problems.
- C406.3 Analyze datasets using analytical techniques and tools.
- C406.4 Apply data visualization technique for datasets.



ASSIGNMENT LIST

Course No.	Sr. No.	Batch	Title	Bloom's Level
C401	LA1	B1	Write a CUDA program that, given an N-element vector, find- The maximum element in the vector using parallel reduction.	L3,L4
		B2	Write a CUDA program that, given an N-element vector, find- The minimum element in the vector using parallel reduction.	
		B3	Write a CUDA program that, given an N-element vector, find- The arithmetic mean of the vector using parallel reduction.	
		B4	Write a CUDA program that, given an N-element vector, find- The standard deviation of values in the vector using parallel reduction.	
	LA2	B1	Vector and Matrix Operations- Design parallel algorithm - Multiply two $N \times N$ arrays using $n \times n$ processors	L3,L4
		B2	Vector and Matrix Operations- Design parallel algorithm- Multiply two $N \times N$ arrays using $n \times n$ processors	
		B3	Vector and Matrix Operations- Design parallel algorithm to - Multiply Vector and Matrix	
		B4	Vector and Matrix Operations- Design parallel algorithm to - Add two large vectors	
	LA3	B1,B3	Parallel Sorting Algorithms-For Bubble Sort based on existing sequential algorithms, design and implement parallel algorithm utilizing all resources available.	L3,L4
		B2,B4	Parallel Sorting Algorithms-For Merge Sort, based on existing sequential algorithms, design and implement parallel algorithm utilizing all resources available.	
	LA4	B1	Parallel Search Algorithm-Design and implement parallel algorithm utilizing all resources available for Best-First Search that (traversal of graph to reach a target in the shortest possible path).	L3,L4
		B2	Parallel Search Algorithm-Design and implement parallel algorithm utilizing all resources available for Breadth-First	



**Progressive Education Society's
Modern College of Engineering**

DEPARTMENT OF COMPUTER ENGINEERING

			Bound and Backtracking)	
	LA9	ALL	<p>Implement Tic-Tac-Toe using A* algorithm.</p> <p>OR</p> <p>Define the operators for controlling domestic robot; use these operators to plan an activity to be executed by the robot. For example, transferring two/three objects one over the other from one place to another. Use Means-Ends analysis with all the steps revealed.</p>	L3
C403	LA10	B2,B4	<p>Download Pima Indians Diabetes dataset. Use Naive Bayes“ Algorithm for classification</p> <ul style="list-style-type: none"> • Load the data from CSV file and split it into training and test datasets. • Summarize the properties in the training dataset so that we can calculate probabilities and make predictions. • Classify samples from a test dataset and a summarized training dataset. 	L4
		B3,B4	<p>Download the Iris flower dataset or any other dataset into a Data Frame. (eg https://archive.ics.uci.edu/ml/datasets/Iris) Use Python/R and Perform following –</p> <ul style="list-style-type: none"> • How many features are there and what are their types (e.g., numeric, nominal)? • Compute and display summary statistics for each feature available in the dataset. (eg. minimum value, maximum value, mean, range, standard deviation, variance and percentiles • Data Visualization-Create a histogram for each feature in the dataset to illustrate the feature distributions. Plot each histogram. • Create a boxplot for each feature in the dataset. All of the boxplots should be combined into a single plot. Compare distributions and identify outliers. 	
	LA11	B1,B3	Use Movies Dataset. Write the map and reduce methods to determine the average ratings of movies.	L4
		B2,B4	Write a Hadoop program that counts the number of occurrences of each word in a text file.	
	LA12	B1,B2	Trip History Analysis: Use trip history dataset that is from a bike sharing service in the United States. The data is provided quarter-wise from 2010 (Q4) onwards. Each file has 7 columns. Predict the class of user. Sample Test data set available here https://www.capitalbikeshare.com/trip-history-data	L4



**Progressive Education Society's
Modern College of Engineering**

DEPARTMENT OF COMPUTER ENGINEERING

		B3,B4	Twitter Data Analysis: Use Twitter data for sentiment analysis. The dataset is 3MB in size and has 31,962 tweets. Identify the tweets which are hate tweets and which are not. Sample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problem-twitter-sentiment-analysis/	
	LA13	B1,B3	Bigmart Sales Analysis: For data comprising of transaction records of a sales store. The data has 8523 rows of 12 variables. Predict the sales of a store. Sample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problem-big-mart-sales-iii/	L4
		B2,B4	Time Series Analysis: Use time series and forecast traffic on a mode of transportation. Sample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problem-time-series-2/	



Name of the Subject –Laboratory Practice II

Weekly Work Load(in Hrs)	Lecture	Tutorial	Practical
	-	-	04

Online/ In-sem	Theory	Practical	Oral	Term- work	Total Marks	Credit
-	-	-	50	50	100	2

Course Objectives

1. To develop problem solving abilities using Mathematical Modeling.
2. To apply algorithmic strategies, Software Engineering and testing while solving problems.
3. To develop time and space efficient algorithms.

Course Outcomes

C407.1 Make use of data mining Tools.

C407.2 Analyze and compare classification techniques.

C407.3 Develop an application and Apply manual testing.

C407.4 Develop web based application & apply automation testing using automation tool.



**Progressive Education Society's
Modern College of Engineering
DEPARTMENT OF COMPUTER ENGINEERING**

Assignment List

Course No.	Sr. No.	Title	Bloom's Level
C407.1	LA1	For an organization of your choice, choose a set of business processes. Design star / snow flake schemas for analyzing these processes. Create a fact constellation schema by combining them. Extract data from different data sources, apply suitable transformations and load into destination tables using an ETL tool. For Example: Business Origination: Sales, Order, and Marketing Process.	L3
C407.2	LA2	Consider a suitable dataset. For clustering of data instances in different groups, apply different clustering techniques (minimum 2). Visualize the clusters using suitable tool.	L3
	LA3	Apply a-priori algorithm to find frequently occurring items from given data and generate strong association rules using support and confidence thresholds. For Example: Market Basket Analysis	L3
	LA4	Consider a suitable text dataset. Remove stop words, apply stemming and feature selection techniques to represent documents as vectors. Classify documents and evaluate precision, recall	L3
	MP1	Mini project on classification: Consider a labeled dataset belonging to an application domain. Apply suitable data preprocessing steps such as handling of null values, data reduction, and discretization. For prediction of class labels of given data instances, build classifier models using different techniques (minimum 3), analyze the confusion matrix and compare these models. Also apply cross validation while preparing the training and testing datasets. For Example: Health Care Domain for predicting disease	L3,L4
C407.3	MP2	Mini-Project 1: Create a small application by selecting relevant system environment / platform and programming languages. Narrate concise Test Plan consisting features to be tested and bug taxonomy. Prepare Test Cases inclusive of Test Procedures for identified Test Scenarios. Perform selective Black-box and White-box testing covering Unit and Integration test by using suitable Testing tools. Prepare Test Reports based on Test	L3,L6



**Progressive Education Society's
Modern College of Engineering**

DEPARTMENT OF COMPUTER ENGINEERING

		Pass/Fail Criteria and judge the acceptance of application developed.	
C407.4	MP3	Mini-Project 2: Create a small web-based application by selecting relevant system environment / platform and programming languages. Narrate concise Test Plan consisting features to be tested and bug taxonomy. Narrate scripts in order to perform regression tests. Identify the bugs using Selenium Web Driver and IDE and generate test reports encompassing exploratory testing.	L3,L6