OP JINDAL UNIVERSITY

Raigarh-Chhattisgarh

Scheme and Syllabus

Of

M. Tech.(01PG021)

Department of

Computer Science and Engineering

School of Engineering

SESSION: 2023- 2025

Approved scheme of teaching and syllabus for M. Tech. (Department of Computer Science and Engineering) by the members of the Board of Studies

Scheme & Syllabus of M. Tech (CSE) Programme

Semester I

S. No.	Subject Code	Subject	Periods Per Week			Scheme of Examination and Marks				Credit L+[T+P]/2
						PRE**		ESE*	Total Marks	
			L	T	P	Mid Sem	TA			
1.	SOE-M-CSE101	Advanced Mathematics	3	1	0	30	20	50	100	4
2.	SOE-M-CSE102	Advanced Data Structures and Algorithms	3	0	0	30	20	50	100	3
3.	SOE-M-CSE103	Fundamental of Data Analytics	3	0	0	30	20	50	100	3
4.	SOE-M-CSE104	Data Mining & Data Warehousing	3	0	0	30	20	50	100	3
5.	SOE-M-CSE105	Elective 1	3	0	0	30	20	50	100	3
6.	SOE-M-CSE106	Data Structures and Algorithms Lab	0	0	4	0	30	20	50	2
7.	SOE-M-CSE107	Data Mining & Data Ware Housing Lab	0	0	4	0	30	20	50	2
8.	SOE-M-CSE108	Fundamental of Data Analytics Lab	0	0	4	0	30	20	50	2
9.	SOE-M-CSE109	Elective Lab 1	0	0	4	0	30	20	50	2
Total			15	1	16	150	220	330	700	24

L: Lecture, **T:** Tutorial, **P:** Practical, **ESE:** End Semester Examination, **T.A:** Teacher's Assessment.

Elective-I

S. No	Subject Code	Subject Name
2.	SOE-M-CSE105(1)	Software Engineering
3.	SOE-M-CSE105(2)	IoT Technology & Applications
4.	SOE-M-CSE105(3)	Big Data & Cloud Computing

Elective Lab-I

S. No	Subject Code	Subject Name
1.	SOE-M-CSE109(1)	Software Engineering Lab
2.	SOE-M-CSE109(2)	Internet of Things (IoT) Lab
3	SOE-M-CSE109(3)	Big Data & Cloud Computing Lab

Programme	:	**M. Tech**	**Semester** :	**I**
Name of the Course	:	**Advanced Mathematics**	**Course Code:**	**SOE-M-CSE101**
Credits	:	**4**	**No of Hours :**	**4 Hrs./week**
Max Marks	:	**100**		

Course Description:

The course has been designed to understand the basic concepts of algebra. This course includes the study of algebra, probability theory, optimization techniques, Fourier series and transform and their applications. The concepts introduced has application in machine learning, data science and image processing.

Course Outcomes:

After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Apply the concept of vector space over real and complex fields
CO2	Apply probability theory in real life applications.
CO3	Apply optimization techniques in in real life applications.
CO4	Apply Fourier series and Transform in the area of signal processing.
CO5	Apply mathematical techniques in the research

Syllabus:

UNIT-I: Linear Algebra I

Vectors and geometry in two and three space dimensions. Algebraic properties. Dot products and the norm of a vector. Important inequalities. Vector spaces, subspaces and vector space axioms. Complex vector spaces, Eigenvalues and eigenvectors.

UNIT-II: Linear Algebra II

Examples of linear systems. Geometry of linear equations. Gaussian elimination. Row echelon form. Homogeneous and nonhomogeneous systems of linear equations. Application to the intersection of lines and planes, Properties and composition of linear transformations. Rotations, reflections and stretches. Translations using homogeneous coordinates. One-to-one and onto transformations

UNIT-III: Probability Theory

Review of basic probability theory. Definitions of random variables and probability distributions, probability mass and density functions, expectation, moments, central limit theorem, Poisson, Gaussian and Erlang distributions-examples.

UNIT-IV: Optimization Technique

Calculus of several variables, Implicit function theorem, Nature of singular points, Necessary and sufficient conditions for optimization, Constrained Optimization, Lagrange multipliers, Gradient method – steepest descent method.

UNIT-V: Fourier series and Transform

Fourier series and Transform: Fourier series, integrals and transforms and their properties. One dimensional Fourier transform, Convolution theorem, Parseval's formula, Introduction to 2-dimensional Fourier transform. (8L)

Text Book

1. J. Defranza and D. Gagliardi, Introduction to Linear Algebra with Applications, McGraw-Hill
2. Scott L. Miller, Donald G. Childers: "Probability and Random Process with application to Signal Processing", Elsevier Academic Press.
3. Kreyzig, 'Advanced Engineering Mathematics'

Reference Book

1. Schaum's outlines of Linear Algebra by Seymour Lipschutz, Marc Lipson, McGraw-Hill Education (India) Private Limited, New Delhi
2. T. Veerarajan "Probability, Statistics and Random Pr ocess", 3 rd Edition, Tata Mc-Graw Hill Co.
3. Elsgolts, L. "Differential Equations and Calculus of Variations", MIR Publications.

CO-PO&PSO Correlation

Course Name: Advanced Mathematics								
	Program Outcomes					**PSOs**		
Course Outcomes	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**
CO1:	**1**	**1**					**2**	
CO2:	**1**	**1**					**1**	**1**
CO3:	**1**	**2**					**2**	
CO4:	**2**	**2**				**1**	**1**	**1**
CO5:	**1**	**2**		**1**	**1**	**1**	**2**	**1**

Note: 1: Low 2.: Moderate 3: High

Programme	:	**M. Tech.**	**Semester**	:	**I**
Name of the Course:		**Advanced Data Structures and Algorithms**	**Course Code:**		**SOE-M-CSE102**
Credits	:	**3**	**No of Hours**	:	**3 Hrs./week**
Max Marks	:	**100**			

Course Description:

Engineering Algorithms & Data Structures deals with the fundamental means to approach the design and analysis of algorithms in an effective and methodologically correct manner. The student will acquire knowledge about general techniques for the design and analysis of algorithms as well as a collection of significant examples of solutions to representative problems. Furthermore, the student will have the opportunity to supplement the theory by writing actual programs in the C language during laboratory sessions.

Course Outcomes:

After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Argue the correctness of algorithms using inductive proofs and invariants.
CO2	Analyse worst-case running times of algorithms using asymptotic analysis.
CO3	Describe the divide-and-conquer paradigm and explain when an algorithmic design situation calls for it. Recite algorithms that employ this paradigm. Synthesize divide-and-conquer algorithms. Derive and solve recurrences describing the performance of divide-and-conquer algorithms.
CO4	Describe the dynamic-programming paradigm and explain when an algorithmic design situation calls for it. Recite algorithms that employ this paradigm. Synthesize dynamic-programming algorithms, and analyse them.
CO5	Describe the greedy paradigm and explain when an algorithmic design situation calls for it. Recite algorithms that employ this paradigm.

Syllabus:

UNIT–I: Introduction

Abstract Data Types - Time and Space Analysis of Algorithms - Big Oh and Theta Notations - Average, best and worst-case analysis - Simple recurrence relations, Array and Linked Structure Implementation of Lists, Stacks and Queues - Applications -Array of Nodes and Dynamic Pointer Implementation of Linked Structures.

UNIT–II: Linear and Non-linear Data Structure

Cursors –Sets, Priority Queues –Definition and applications, Max Priority Queue ADT-Implementation-Max Heap-Definition, Insertion into a Max Heap, Deletion from a Max Heap, Sorting techniques, Priority Queues, Trees - Binary trees - Search trees - Balanced trees - Advanced tree structures- B trees - AVL trees, 2-3 trees, Splay trees – applications of trees.

UNIT–III: Graph

Graphs - Directed -Shortest path - Undirected graph - Minimal spanning tree Hashing – Dictionary - Applications of Graphs, Graph based data structure: Network based

UNIT–IV: Analysis and Design Methodology of Algorithms

Algorithms Analysis - Sorting - Searching - Design Techniques - Greedy Methods- Dynamic Programming - Divide and Conquer - Back Tracking – Applications.

UNIT-V: Advance Algorithms

Parallel Algorithms: - Basic Techniques- Work & amp; Efficiency – Distributed Computation - Heuristic & Approximation Approaches.

Text Books:

1. Alfred. V. Aho, John. E. Hopcroft, Jeffrey. D. Ullman, & quot; Data Structures and Algorithms, Addison-Wesley Publications., 1985.

Reference Books:

1. Mark Allen Weiss, & quot; Data Structures and Algorithm Analysis in C & quot; Second Edition, Pearson Education, Asia
2. Jean-Paul Tremblay, Paul.G. Sorenson, & quot; An Introduction to Data Structures with Applications & quot;, Tata Mc Graw Hill second edition, 1991.
3. Thomas. H. Cormen, Charles.E. Leiserson, Ronald. L. Rivest, & quot; Introduction to Algorithms & quot;, PHI 1998.
4. Ellis Horowitz, Sartaj Sahni, Songuthevan Rajasekaran, Fundamentals of Computer Algorithms, Galgotial Publications Pvt. Ltd.

CO-PO & PSO Correlation

Course Name: Advanced Data Structures and Algorithms								
	Program Outcomes					**PSOs**		
Course Outcomes	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**
CO1:	**2**	**1**					**2**	
CO2:	**1**	**1**					**1**	**1**
CO3:	**2**	**2**					**2**	
CO4:	**2**	**2**				**1**	**1**	**1**
CO5:	**2**	**2**				**1**	**2**	**1**

Note: 1: Low 2.: Moderate 3: High

Programme	:	**M. Tech.**	**Semester**	:	**I**
Name of the Course	:	**Fundamental of Data Analytics**	**Course Code:**		**SOE-M-CSE103**
Credits	:	**3**	**No of Hours**	:	**3 Hrs./week**
Max Marks	:	**100**			

Course Descriptions:

This course will cover fundamental algorithms and techniques used in Data Analytics. The statistical foundations will be covered first, followed by various machine learning and data mining algorithms. Technological aspects and visualization will also be covered. This course will provide exposure to theory as well as practical systems and software used in data analytics.

Course Outcomes:

After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Find a meaningful pattern in data
CO2	Graphically interpret data
CO3	Implement the analytic algorithms
CO4	Handle large scale analytics projects from various domains
CO5	Develop intelligent decision support systems

Syllabus:

UNIT-I: Introduction

Data Definitions and Analysis Techniques: Elements, Variables, and Data categorization, Levels of Measurement, Data management and indexing, Introduction to statistical learning.

UNIT-II: Descriptive Statistics

Measures of central tendency, Introduction to Probability, Probability Distributions, Sampling and Sampling Distribution, Distribution of Sample Means, population, Measures of location of dispersions.

UNIT-III: Basic analysis techniques

Statistical hypothesis generation and testing, Hypothesis testing with two sample test Errors in Hypothesis Testing, Chi-Square test, t-Test, Analysis of variance, Correlation analysis, Maximum likelihood test.

UNIT-IV: Data Analysis techniques

Data analysis techniques: Regression analysis, Classification techniques, Clustering, Association rules analysis, Measures of attribute selection, Confusion

matrix and ROC, Different Plots (Line Plot, Bar Chart, Histogram Plot, Box and Whisker Plot, Scatter Plot.).

UNIT-V: Case studies and projects

Understanding business scenarios, Feature engineering and visualization, Scalable and parallel with Hadoop and Map-Reduce, Sensitivity Analysis.

Text book:

1. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer.
2. V. K. Jain "Data Science and Analytics".

References book:

1. Mize Edward "Data Analytics".
2. Tom White "Hadoop: The Definitive Guide" Third Edition, O'reilly Media.
3. Chris Eaton, Dirk DeRoos, Tom Deutsch, George Lapis, Paul Zikopoulos, "Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data", McGraw Hill Publishing.
4. Jiawei Han, Micheline Kamber "Data Mining Concepts and Techniques", 2nd Edition, Elsevier.

CO-PO&PSO Correlation

Course Name: Fundamental of Data Analytics								
	Program Outcomes					**PSOs**		
Course Outcomes	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**
CO1:	**2**					**1**		**1**
CO2:		**2**					**1**	**1**
CO3:	**2**	**2**	**1**					
CO4:	**2**					**1**	**1**	**1**
CO5:		**2**			**1**	**1**	**2**	**1**

Note: 1: Low 2.: Moderate 3: High

Programme	**:**	**M. Tech.**	**Semester**	**:**	**I**
Name of the Course:		**Data Mining & Data Warehousing**	**Course Code:**		**SOE-M-CSE104**
Credits	**:**	**3**	**No of Hours**	**:**	**3 Hrs./week**
Max Marks	**:**	**100**			

Course Description:

This course provides the student with in depth knowledge of Data Warehousing principles, Data Warehouse techniques, and Business Intelligence systems. The course introduces the topics of Data Warehouse design, Extract-Transform-Load (ETL), Data Cubes, and Data Marts. Students will create Business Intelligence using Data Warehouses with several OLAP and analytical tools.

Course Outcomes:

After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Understand fundamentals of data mining and Data warehousing techniques
CO2	Understand and apply data preprocessing techniques for data consistency
CO3	Understand features and their selection and reduction
CO4	Understand different algorithms for data mining
CO5	Understand and apply classification algorithm for classification of data

Syllabus:

UNIT–I: Introduction

Introduction to Data Science, data mining, machine learning, target applications, Knowledge Discovery, Data Mining Functionalities, Data Mining Techniques, Data Mining System categorization and its Issues.

Mathematical Background: Mean, Median, mode, standard deviation, correlation, covariance, likelihood, data: nominal, ordinal, ratio, interval, factor, levels. Interquartile range, Sampling, probability.

UNIT–II: Data Pre-processing

Data Cleaning, Data Integration and Transformation: standardization, normalization, smoothing, aggregation, generalization. Data Reduction:-Data Cube Aggregation, Dimensionality reduction, Data Discretization and Concept, encoding and decoding, Representing input data and output knowledge Visualization techniques, Guidelines for Successful Data Mining.

UNIT–III: Features

Introduction, importance, selection, extraction engineering Dimensionality Reduction: Principal Components Analysis, Sketching PCA, Applying PCA, Limitations of PCA.

UNIT–IV: Association Analysis

Basic Concepts and Algorithms, Problem Definition, Frequent Item set Generation, Rule Generation, Interesting measures: support and confidence, Apriori Algorithms, Improving the efficiency of the Apriori Algorithm, Tree Based Algorithms etc.

UNIT–V: Classification

Basic Concepts, linear vs nonlinear, Multiclass, class imbalance, Model Overfitting, linear classifier with examples, measuring classifier accuracy, clustering techniques.

Text Books:

1. J. Han & M. Kamber, "Data Mining: Concepts and Techniques", 2nd Ed, 2006.
2. Arun K Pujari "Data Mining Techniques", University Press.

Reference Books:

1. W. H. Inmon, "Building the Data Warehouse", 3rd edition.
2. Anahory and Murray, Data warehousing in the real world, Pearson Education/Addison Wesley.
3. Margaret Dunham, "Data Mining: Introductory and Advanced Topics", Prentice Hall.

CO-PO & PSO Correlation

Course Name: Data Mining & Data Warehousing								
	Program Outcomes					**PSOs**		
Course Outcomes	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**
CO1:	**1**					**1**	**2**	**1**
CO2:	**2**		**3**					**1**
CO3:		**2**					**2**	
CO4:	**2**	**3**				**1**	**1**	**1**
CO5:	**2**	**2**						**1**

Note: 1: Low 2.: Moderate 3: High

Programme	**:**	**M. Tech.**	**Semester**	**:**	**I**
Name of the Course:		**Software Engineering**	**Course Code:**		**SOE-M-CSE105(1)**
Credits	**:**	**3**	**No of Hours**	**:**	**3 Hrs./week**
Max Marks	**:**	**100**			

Course Description:

This course introduces the concepts and methods required for the construction of software interactive system. It will also include quality measures and maintenance of software. It also focuses on the Reliability Techniques for software development.

Course Outcomes:

Upon successful completion of this course, the student will be able to:

CO Number	Course Outcome
CO1	Understand the issues affecting the organization, planning and control of software based system development.
CO2	Analyze and design of a small software intensive system.
CO3	Manage risk and measure the software quality and reliability.

Syllabus:

UNIT–I:

Introduction, software life-cycle models, CMMI

UNIT-II:

SRS, SDS, formal requirements specification and verification - axiomatic and algebraic specifications Requirement Engineering Processes.

UNIT-III:

Software Design: UML; computer-aided software engineering (CASE), Verification and Validation: Software Testing, Quality assurance, Maintenance.

UNIT–IV:

Project Management: activities, planning, scheduling, Risk Management.

UNIT–V:

Reliability Techniques, Models of concurrency, Static analysis, Security vulnerabilities/attacks, Vulnerability detection.

Text Books:

1. Sommerville "Software Engineering, Pearson Education Publication", 7th edition.
2. M.Ben-Ari "Principles of concurrent and distributed programming", Addison-Wesley, 2006 "Handbook of model checking", Springer, 2014.

Reference Books:

1. R. S. Pressman: Software Engineering: A Practiioners Approach, 5th Edn., TMA, New Delhi.
2. M. Ben-Ari, "Principles of concurrent and distributed programming", Addison-Wesley, 2006 "Handbook of model checking", Springer, 2014.
3. Brian Chess and Jacob West, "Secure programming with static analysis", Addison Wesley, 2007 Additional research papers.

CO-PO&PSO Correlation

Course Name: Software Engineering								
	Program Outcomes					PSOs		
Course	**1**	**2**	**3**	**4**	**5**	**1**	**2**	
CO1:	**1**					**2**		**1**
CO2:		**1**	**1**	**1**	**1**		**1**	
CO3:		**1**		**1**				**2**

Note: 1.: Low 2.: Moderate 3.: High

Programme	:	**M. Tech.**	**Semester**	:	**I**
Name of the Course	:	**IoT Technology & Applications**	**Course Code:**		**SOE-M-CSE105(2)**
Credits	:	**3**	**No of Hours**	:	**3 Hrs./week**
Max Marks	:	**100**			

Course Description:

Internet of Things (IoT) is presently a hot technology worldwide. Government, academia, and industry are involved in different aspects of research, implementation, and business with IoT. IoT cuts across different application domain verticals ranging from civilian to defense sectors. These domains include agriculture, space, healthcare, manufacturing, construction, water, and mining, which are presently transitioning their legacy infrastructure to support IoT. Today it is possible to envision pervasive connectivity, storage, and computation, which, in turn, gives rise to building different IoT solutions. IoT-based applications such as innovative shopping system, infrastructure management in both urban and rural areas, remote health monitoring and emergency notification systems, and transportation systems, are gradually relying on IoT based systems. Therefore, it is very important to learn the fundamentals of this emerging technology.

Course Outcomes:

After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Understand the vision of IoT from a global context.
CO2	Determine the Market perspective of IoT.
CO3	Building state of the art architecture in IoT.
CO4	Application of IoT in Industrial and Commercial Building Automation and Real World Design Constraints.
CO5	Understand different IoT tools and its implementation.

Syllabus:

UNIT-I: IoT & Web Technology

The Internet of Things Today, Time for Convergence, Towards the IoT Universe, Internet of Things Vision, IoT Strategic Research and Innovation Directions, IoT Applications, Future Internet Technologies, Infrastructure, Networks and Communication, Processes, Data Management, Security, Privacy & Trust, Device Level Energy Issues, IoT Related Standardization, Recommendations on Research Topics.

UNIT-II: M2M to IoT – A Basic Perspective

Introduction, Some Definitions, M2M Value Chains, IoT Value Chains, an emerging industrial structure for IoT, the international driven global value chain and global information monopolies.

M2M to IoT-An Architectural Overview– Building an architecture, Main design principles and needed capabilities, An IoT architecture outline, standards considerations.

UNIT-III: IoT Architecture-State of the Art

Introduction, State of the art, Architecture Reference Model- Introduction, Reference Model and architecture, IoT reference Model, IoT Reference Architecture-Introduction, Functional View, Information View, Deployment and Operational View, Other Relevant architectural views.

UNIT-IV: IoT Applications for Value Creations

Introduction, IoT applications for industry: Future Factory Concepts, Brownfield IoT, Smart Objects, Smart Applications, Four Aspects in your Business to Master IoT, Value Creation from Big Data and Serialization, IoT for Retailing Industry, IoT for Oil and Gas Industry, Opinions on IoT Application and Value for Industry, Home Management, eHealth.

UNIT-V: IoT Privacy, Security, Governance & IoT solutions

Introduction, Overview of Governance, Privacy and Security Issues, Contribution from FP7 Projects, Security, Privacy and Trust in IoT-Data-Platforms for Smart Cities, First Steps Towards a Secure Platform, Smartie Approach. Data Aggregation for the IoT in Smart Cities, Security & privacy issues in IoT

Developing IoT Solutions

Introduction to Python, Introduction to different IoT tools, Introduction to Arduino and Raspberry Pi Implementation of IoT with Arduino and Raspberry, Cloud Computing, Fog Computing, Connected Vehicles.

Text Books:

1. Vijay Madisetti and Arshdeep Bahga, "Internet of Things (A Hands-on-Approach)", 1 st Edition, VPT, 2014
2. Francis daCosta, "Rethinking the Internet of Things: A Scalable Approach to Connecting Everything", 1 st Edition, Apress Publications, 2013

Reference Books:

1. Cuno Pfister, Getting Started with the Internet of Things, O"Reilly Media, 2011, ISBN: 978-1-4493- 9357-1

2. Jan Holler, Vlasios Tsiatsis, Catherine Mulligan, Stefan Avesand, Stamatis Karnouskos, David Boyle, "From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence", 1st Edition, Academic Press, 2014.

CO-PO & PSO Correlation

Course Name: IoT Technology & Applications								
	Program Outcomes					PSOs		
Course Outcomes	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**
CO1:	**1**	**2**	**3**			**1**	**2**	**3**
CO2:	**2**	**1**				**1**	**2**	**1**
CO3:	**2**						**1**	**1**
CO4:	**2**	**2**					**2**	
CO5:	**2**	**2**				**1**	**1**	**1**

Note: 1: Low 2.: Moderate 3: High

Programme	:	**M. Tech.**	**Semester**	:	**I**
Name of the Course	:	**Big Data & Cloud Computing**	**Course Code:**		**SOE-M-CSE105(3)**
Credits	:	**3**	**No of Hours**	:	**3 Hrs./week**
Max Marks	:	**100**			

Course Description:

This course is designed to introduce the concepts of Cloud Computing and Big Data as a new computing paradigm. The course will expose students to different views of understanding the Cloud Computing and Big Data such as theoretical, technical and commercial aspects.

Course Outcomes:

After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Differentiate different computing techniques
CO2	Compare various cloud computing providers/ Software.
CO3	Identify the appropriate cloud services for a given application
CO4	Compare various cloud computing providers/ Software.
CO5	Handle Open-Source Cloud Implementation and Administration.
CO6	Analyze authentication, confidentiality and privacy issues in Cloud computing environment.

Syllabus:

UNIT-I: Introduction of Computing

New Computing Paradigms & Services: Cloud computing - Edge computing, Grid computing - Utility computing - Cloud Computing Architectural Framework, Cloud Deployment Models - Virtualization in Cloud Computing, Parallelization in Cloud Computing, Security for Cloud Computing - Cloud Economics.

UNIT-II: Big Data

Introduction, Challenges, 5 V's- Ecosystem- Google's Solution Vs Hadoop- Hadoop: Ecosystem, Architecture- Cluster; Map Reduce- Information retrieval through Map Reduce- Hadoop File System, GFS- Page Ranking using Map Reduce.

UNIT-III: Big Data Analytics

Classification of analytics - Data Science - Terminologies in Big Data - CAP Theorem - BASE Concept. NoSQL: Types of Databases – Advantages – NewSQL - SQL vs. NOSQL vs NewSQL. Linear Regression

UNIT-IV: No SQL Databases

Mongo DB: Introduction – Features - Data types - Mongo DB Query language - CRUD operations – Arrays - Functions: Count – Sort – Limit – Skip – Aggregate - Map Reduce. Cursors – Indexes - Mongo Import – Mongo Export. Cassandra: Introduction – Features - Data types – CQLSH - Key spaces - CRUD operations – Collections – Counter – TTL - Alter commands - Import and Export - Querying System tables.

UNIT-V:

Case study- Apache Spark, Machine Learning, VMware.

Text Books:

1. EMC Education Services "Data Science and Big Data Analytics"
2. Viktor Mayer-Schonberger, Kenneth Cukier "Big Data"

Reference Books:

1. Subhashini Chellappan Seema Acharya "Big Data and Analytics"
2. George Reese "Cloud Application Architectures", O'Reillly Publications

CO-PO&PSO Correlation

Course Name: Big Data & Cloud Computing								
	Program Outcomes					PSOs		
Course Outcomes	1	2	3	4	5	1	2	3
CO1:	1	2	3	4	5	1	2	3
CO2:	2	1				1	2	1
CO3:	2		2				1	1
CO4:	1	2					2	
CO5:	2	2	2			1	1	1
CO6:	1	2				2	1	2

Note: 1: Low 2.: Moderate 3: High

Programme	:	**M. Tech.**	**Semester**	:	**I**
Name of the Course	:	**Data Structures and Algorithm Lab**	**Course Code:**		**SOE-M-CSE106**
Credits	:	**2**	**No of Hours**	:	**4 Hrs./week**
Max Marks	:	**50**			

Course Description:
In this course the student will have the opportunity to supplement the theory subject by writing actual programs in the C language during laboratory sessions.

Course Outcomes:
After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Implement and compare correctness of algorithms using inductive proofs and invariants.
CO2	Implement and Analyse worst-case running times of algorithms using asymptotic analysis.
CO3	Implement and compare different divide-and-conquer algorithms
CO4	Implement and analyses dynamic-programming algorithms
CO5	Implement and compare different greedy algorithms

The following concepts will be covered in the lab:

- Implementation of Sorting algorithm like Quick sort, Heap Sort, Merge sort etc. and computation of its time complexity. Run the program for varied values of n> 5000, and record the time taken to sort. Plot a graph of the time taken versus non graph sheet. The elements can be read from a file or can be generated using the random number generator.
- Implementation of divide-and-conquer method works along with its time complexity analysis: worst case, average case and best case.
- Implementation of ,0/1 Knapsack problem using (a) Dynamic Programming method (b) Greedy method.
- Implementation of Dijkstra's algorithm to find shortest paths from a given vertex in a weighted connected graph.
- Implementation of finding Minimum Cost Spanning Tree of a given connected undirected graph using Kruskal's algorithm.
- Implementation of Prim's algorithm to find Minimum Cost Spanning Tree of a given connected undirected graph.
- Implementation of All-Pairs Shortest Paths problem using Floyd's algorithm.
- Implementation of Travelling Sales Person problem using Dynamic programming.
- Implementation of finding all Hamiltonian Cycles in a connected undirected Graph G of n vertices using backtracking principle.

Text Books:

- Alfred. V. Aho, John. E. Hopcroft, Jeffrey. D. Ullman, & quot; Data Structures and Algorithms, Addison-Wesley Publications., 1985.

Reference Books:

- M. T. Goodrich and R. Tomassia, John Wiley and sons "Algorithm Design: Foundations, Analysis and Internet examples"
- S. Sridhar "Design and Analysis of Algorithms", Oxford Univ. Press.
- Aho, Ullman and Hopcroft "Design and Analysis of algorithms", Pearson Education.

CO-PO&PSO Correlation

Course Name: Data Structures and Algorithm Lab								
	Program Outcomes					PSOs		
Course Outcomes	1	2	3	4	5	1	2	3
CO1:	1	1	2			1	1	2
CO2:	2	2	2			2	2	2
CO3:	1	2	1					1
CO4:	1	2	1			1		
CO5:	1	2	2			1	1	1

Note: 1: Low 2.: Moderate 3: High

Programme	:	**M. Tech.**	**Semester**	:	**I**
Name of the Course	:	**Data Mining and Data Warehousing Lab**	**Course Code:**		**SOE-M-CSE107**
Credits	:	**2**	**No of Hours**	:	**4 Hrs./week**
Max Marks	:	**50**			

Course Description:

In this course students will implement the Data Warehouse techniques, and Business Intelligence systems. The concepts of Data Warehouse design and Business Intelligence using Data Warehouses with several OLAP and analytical tools will be explored.

Course Outcomes:

After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Design and evaluate the different models of OLAP
CO2	Implement various algorithms used in information analysis of Data Mining Techniques
CO3	Design and evaluate the different techniques for data pre-processing.
CO4	Implement Knowledge retrieval using data mining techniques
CO5	Implement classification models using datasets

The following concepts will be covered in the lab:

- Explore machine learning tool "WEKA"
  - Explore WEKA Data Mining/Machine Learning Toolkit Downloading and/or installation of WEKA data mining toolkit, Understand the features of WEKA toolkit such as Explorer, Knowledge Flow interface, Experimenter, command-line interface. Navigate the options available in the WEKA (ex. Select attributes panel, Preprocess panel, classify panel, Cluster panel, Associate panel and Visualize panel)
  - Study the arff file format Explore the available data sets in WEKA. Load a data set (ex. Weather dataset, Iris dataset, etc.) Load each dataset and observe the following:
    - List the attribute names and they types Number of records in each dataset
    - Identify the class attribute (if any).
    - Plot Histogram Determine the number of records for each class. Visualize the data in various dimensions.
- Perform data preprocessing tasks and Demonstrate performing association rule mining on data sets

- Explore various options available in Weka for preprocessing data and apply unsupervised filters like Discretization, Resample filter, etc. on each dataset.
- Load weather. nominal, Iris, Glass datasets into Weka and run Apriori algorithm with different support and confidence values. Study the rules generated.

- Apply different discretization filters on numerical attributes and run the Apriori association rule algorithm. Study the rules generated. Derive interesting insights and observe the effect of discretization in the rule generation process.

1. Demonstrate performing classification on data sets Load each dataset into Weka and run 1d3, J48 classification algorithm.
   - Study the classifier output. Compute entropy values, Kappa statistic.
   - Extract if-then rules from the decision tree generated by the classifier, Observe the confusion matrix.
   - Load each dataset into Weka and perform Naïve-bayes classification and k-Nearest Neighbor classification. Interpret the results obtained.
   - Plot RoC Curves Compare classification results of ID3, J48, Naïve-Bayes and k-NN classifiers for each dataset, and deduce which classifier is performing best and poor for each dataset and justify.
2. Demonstrate performing clustering of data sets Load each dataset into Weka and run simple k-means clustering algorithm with different values of k (number of desired clusters).
   - Study the clusters formed. Observe the sum of squared errors and centroids, and derive insights.
   - Explore other clustering techniques available in Weka.
   - Explore visualization features of Weka to visualize the clusters.
   - Derive interesting insights and explain.
3. Demonstrate knowledge flow application on data sets
   - Develop a knowledge flow layout for finding strong association rules by using Apriority.
   - FP Growth Algorithms Set up the knowledge flow to load an ARFF (batch mode) and perform a cross validation using J48 algorithm.
   - Demonstrate plotting multiple ROC curves in the same plot window by using j48 and Random forest tree.

Text Books:

- P. Adriaans & D. Zantinge, Data Mining, Addison Wesley, 1996.
- R. Mattison, Data Warehousing: Strategies, Tools and Techniques, McGraw Hill, 1996.
- P. Ponniah, Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals, Wiley, 2001.
- Soman K P, “Insight into Data Mining: Theory & Practice”, Prentice hall of India
- M.H. Dunham, “Data Mining Introductory and Advanced Topics”, Pearson

Education.
- Ralph Kimball, "The Data Warehouse Lifecycle toolkit", John Wiley.

CO-PO & PSO Correlation

Course Name: Data Mining and Data Warehousing Lab								
	Program Outcomes					PSOs		
Course Outcomes	1	2	3	4	5	1	2	
CO1:	3	2	1			3		
CO2:		2	1		1			2
CO3:	2	2	2		1		1	2
CO4:	2	2	2		1		1	1
CO5	3	1	2			1		1

Note: 1.: Low 2.: Moderate 3.: High

Programme	**:**	**M. Tech.**	**Semester**	**:**	**I**
Name of the Course:		**Fundamental of Data Analytics Lab**	**Course Code:**		**SOE-M-CSE108**
Credits	**:**	**2**	**No of Hours**	**:**	**4 Hrs./week**
Max Marks	**:**	**50**			

Course Descriptions:

This course will cover implementation of fundamental algorithms and techniques used in Data Analytics.

Course Outcomes:

After Completion of the course Students will be able to:

CO Number	Course Outcome
CO1	Find a meaningful pattern in data
CO2	Implement and analyze Graphically interpret data
CO3	Implement the analytic algorithms
CO4	Analyze large scale analytics projects from various domains
CO5	Design intelligent decision support systems

The following concepts will be covered in the lab:

- Text classification.
- Preprocessing techniques
- Designing tools to remove stop words in dictionary
- Designing of training and testing data set for text classification
- Machine learning based sentiment analysis.

Text Books:

- Breck Baldwin, ―Language Processing with Java and LingPipe Cookbook, Atlantic Publisher, 2015.
- Richard M Reese, ―Natural Language Processing with Java‖, O_Reilly Media, 2015.
- Nitin Indurkhya and Fred J. Damerau, ―Handbook of Natural Language Processing, Second Edition, Chapman and Hall/CRC Press, 2010.
- Tanveer Siddiqui, U.S. Tiwary, ―Natural Language Processing and Information Retrieval‖, Oxford University Press, 2008.

CO-PO&PSO Correlation

Course Name: Fundamental of Data Analytics Lab								
	Program Outcomes					**PSOs**		
Course Outcomes	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**
CO1:	**1**	**1**				**1**		
CO2:	**3**	**3**			**1**		**1**	**1**
CO3:	**2**	**1**			**1**	**1**	**2**	
CO4:	**2**	**2**			**1**	**1**	**2**	**1**
CO5:	**2**	**2**			**1**	**1**	**2**	**1**

Note: 1: Low 2.: Moderate 3: High

Programme	**:**	**M.Tech.**	**Semester :**	**I**
Name of the Course:		**Software Engineering Lab**	**Course Code:**	**SOE-M-CSE109(1)**
Credits	**:**	**2**	**No of Hours :**	**4 Hrs/Week**
Max Marks	**:**	**50**		

Course Descriptions:
This lab deals with the analysis and design of a software problem. The tool used in a lab is rational rose. this tool is used for an object oriented design of a problem. We draw a UML diagram in a rational rose which deals with the objects and classes in a system. The Unified Modeling Language or UML is a mostly graphical modelling language that is used to express designs. It is a standardized language in which to specify the artefacts and components of a software system. It is important to understand that the UML describes a notation and not a process. It does not put forth a single method or process of design, but rather is a standardized tool that can be used in a design process.

Course Outcomes:
At the end of the course, a student will be able to:

CO Number	Course Outcome
CO1	Provide users with a ready-to-use, expressive visual modeling language so they can develop and exchange meaningful models
CO2	Ability to generate a high-level design of the system from the software requirements
CO3	Will have experience and/or awareness of testing problems and will be able to develop a simple testing report
CO4	Ability to translate end-user requirements into system and software requirements

The following concepts will be covered in the lab:

1. **Introduction to Software Engineering-LAB.**
2. **Data flow diagram:**
   - **a.** What processes make up a system?
   - **b.** What data are used in each process?

c. What data are stored?
d. What data enter and leave the system?

3. **Sample Design:**
   a. Class Diagram
   b. Sequence Diagram
   c. State Chart Diagram
   d. Use-Case Diagram
4. **Project**:
   a. Write down the problem statement for a suggested system of relevance.
   b. Do requirement analysis and develop Software Requirement Specification Sheet (SRS) for suggested system.
   c. Perform the Data Flow Diagram (DFD).
   d. Performa the Sequence Diagram.
   e. Perform the State Chart Diagram.
   f. Perform The Use-Case Diagram.
   g. Perform the ER Diagram (If Database applicable).
   h. Prepare time line chart/Gantt Chart/PERT Chart for selected project.

Text Books :

- The unified modeling language user guide Grady Booch, James Rambaugh, Ivar Jacobson, Pearson Education, 2nd Edition, 2005.

CO-PO & PSO Correlation

Course Name: Software Engineering Lab												
	Program Outcomes								PSOs			
Course Outcomes	1	2	3	4	5	6	7	8	1	2	3	4
CO1:	1	3	1			1			2	2	1	1
CO2:	1	2							1		1	
CO3:	2	2				1			1		1	
CO4:	1	2	1			2			1	2	2	2

Note: 1.: Low 2.: Moderate 3.: High

Programme	:	**M. Tech**	**Semester**	**:**	**I**
Name of the Course	:	**IoT Technology and Applications Lab**	**Course Code:**		**SOE-M-CSE109(2)**
Credits	**:**	**2**	**No of Hours :**		**4 Hrs. / Week**
Max Marks	**:**	**50**			

Course Descriptions:

This course will describe the market around the Internet of Things (IoT), the technology used to build these kinds of devices, how they communicate, how they store data, and the kinds of distributed systems needed to support them. Divided into four modules, we will learn by doing. We will start with simple examples and integrate the techniques we learn into a class project in which we design and build an actual IoT system. The client will run in an emulated ARM environment, communicating using common IoT protocols with a cloud enabled backend system.

Course Outcomes:

CO Number	Course Outcome
CO1	Understand the importance of internet of things in present scenario
CO2	Describe the interfacing of IoT with arduino.
CO3	Design of direct and alternating type of electrical instruments using Arduino.
CO4	Analyze the protection schemes of induction motor against over current and under voltage.

The following concepts will be covered in the lab:

Design a Digital DC Voltmeter and Ammeter to measure the voltage and current in DC electrical circuits using Arduino and display the values in LCD display, design a Digital AC Voltmeter and Ammeter to measure the voltage and current in AC electrical circuits using Arduino and display the values in LCD display, Digital frequency meter to measure the frequency in any AC electrical circuit using Arduino and display the values in LCD display, Measure the power and energy in electrical circuit using Arduino and display the values in LCD display, Measure the phase shift and power factor in an electrical circuit for different loads using Arduino and display the value in LCD display, Design an over current relay for distribution system and displaying the tripping status of the relay in substation through IOT, Design a system to protect home appliances from over and under voltages using Arduino, Design a system for protecting the three phase induction motor from over voltages, over currents, temperature and displaying the status of the motor at remote location using IOT, Design a traffic control system using IOT.

Additional Experiments

Design a railway gate control using stepper motor using IOT, Control the speed and direction of a DC motor using Arduino and display the status of the motor at the remote location using IOT.

Reference Books:

- Mark torvalds, "Arduino Programming: Step-by-step guide to mastering arduino hardware and software (Arduino, Arduino projects, Arduinouno, Arduino starter kit, Arduino ide, Arduinoyun, Arduino mega, Arduinonano) Kindle", 2nd Edition, 2001
- Michael J Pont, "Embedded C", 2nd Edition, Pearson Education, 2008.

CO-PO & PSO Correlation

Course Name: Internet of Things (IoT) Lab								
	Program Outcomes					PSOs		
Course Outcomes	1	2	3	4	5	1	2	3
CO1:	1	1	1	2		1	2	2
CO2:	2	2	2	3		2	2	2
CO3:	1		2			1		3
CO4:	1						1	2

Note: 1.: Low 2.: Moderate 3.: High

Programme	**:**	**M.Tech.**	**Semester :**	**I**
Name of the Course:		**Big Data and Cloud Computing Lab**	**Course Code:**	**SOE-M-CSE109(3)**
Credits	**:**	**2**	**No of Hours :**	**4 Hrs/Week**
Max Marks	**:**	**50**		

Course Descriptions:
Configure various virtualization tools such as Virtual Box, VMware workstation. Design and deploy a web application in a PaaS environment. Learn how to simulate a cloud environment to implement new schedulers. Install and use a generic cloud environment that can be used as a private cloud. Manipulate large data sets in a parallel environment.

Course Outcomes:
At the end of the course, a student will be able to:

CO Number	Course Outcome
CO1	Configure various virtualization tools such as Virtual Box, VMware workstation.
CO2	Design and deploy a web application in a PaaS environment.
CO3	Learn how to simulate a cloud environment to implement new schedulers.
CO4	Install and use a generic cloud environment that can be used as a private cloud.

The following concepts will be covered in the lab:

- Install Virtualbox/VMware Workstation with different flavours of linux or windows OS on top of windows7 or 8.
- Install a C compiler in the virtual machine created using virtual box and execute Simple Programs
- Install Google App Engine. Create hello world app and other simple web applications using python/java.
- Use GAE launcher to launch the web applications.
- Simulate a cloud scenario using CloudSim and run a scheduling algorithm that is not present in CloudSim.

- Find a procedure to transfer the files from one virtual machine to another virtual machine.
- Find a procedure to launch virtual machine using trystack (Online Openstack Demo Version)
- Install Hadoop single node cluster and run simple applications like wordcount.

Text Books:

- Cloud computing a practical approach - Anthony T.Velte , Toby J. Velte Robert Elsenpeter, TATA McGraw- Hill , New Delhi – 2010
- John W. itinghouse james F.Ransome, "Cloud Computing Implementation, Management and Security", CRC Press.

Reference Books:

- Cloud Computing (Principles and Paradigms), Edited by Rajkumar Buyya, James Broberg, Andrzej Goscinski, John Wiley & Sons, Inc. 2011
- Cloud Security: A Comprehensive Guide to secure Cloud Computing, Ronald L. Krutz, Russell Dean Vines, Wiley

CO-PO & PSO Correlation

Course Name : Big Data and Cloud Computing Lab												
	Program Outcomes								PSOs			
Course Outcomes	1	2	3	4	5	6	7	8	1	2	3	4
CO1:	1								3			
CO2:		3	1		1						3	1
CO3:		3			1					1	2	
CO4:	1	3			1					1	2	

Note: 1.: Low 2.: Moderate 3.: High