SONA COLLEGE OF TECHNOLOGY, SALEM-5

(An Autonomous Institution)

M.E-Computer Science and Engineering

(M.Tech Data Science)

CURRICULUM and SYLLABI

[For students admitted in 2021-2022]

M.E / M.Tech Regulation 2019

Approved by BOS and Academic Council meetings

Sona College of Technology, Salem
(An Autonomous Institution)
Courses of Study for ME I Semester under Regulations 2019
Computer Science and Engineering
Branch: M.Tech Data Science

S. No	Course Code	Course Title	Lecture	Tutorial	Practical	Credit	Total Contact Hours
Theory							
1	P19MDS101	Statistical Computing	2	1	0	3	45
2	P19MDS102	Matrix and Scientific Computing	2	1	0	3	45
3	P19MDS103	Advanced Data Structures and Algorithms	3	0	0	3	45
4	P19MDS104	Data Science and Big Data Analytics	3	0	0	3	45
5	P19GE101	Research Methodology and IPR	2	0	0	2	30
6	P19GE702	**Audit Course:** Stress Management by YOGA	2	0	0	0	30
Practical							
7	P19MDS105	Advanced Data Structures and Algorithms Laboratory	0	0	4	2	60
8	P19MDS106	Big Data Management and Data Analytics Laboratory	0	0	4	2	60
					Total Credits	**18**	

Approved by

Chairperson, Computer Science and Engineering BOS   **Member Secretary, Academic Council**   **Chairperson, Academic Council & Principal**
Dr.B.Sathiyabhama   **Dr.R.Shivakumar**   **Dr.S.R.R.Senthil Kumar**

Copy to:-
     HOD/CSE, First Semester M.Tech DS Students and Staff, COE

S. No	Course Code	Course Title	Lecture	Tutorial	Practical	Credit	Total Contact Hours
Theory							
1	P19MDS201	Machine Learning	3	0	0	3	45
2	P19MDS202	Data Visualization	3	0	0	3	45
3	P19MDS511	**ELECTIVE**- Block Chain Techniques	3	0	0	3	45
4	P19MDS515	**ELECTIVE**-Business Intelligence	3	0	0	3	45
5	P19MDS516	**ELECTIVE**-Artificial Intelligence	3	0	0	3	45
6	P19GE701	**Audit Course**-English for Research Paper Writing	2	0	0	0	30
Practical							
6	P19MDS203	Machine Learning Laboratory	0	0	4	2	60
7	P19MDS204	Data Visualization Laboratory	0	0	4	2	60
					Total Credits	19	

Approved by

Chairperson, Computer Science and Engineering BOS   Member Secretary, Academic Council   Chairperson, Academic Council & Principal
Dr.B.Sathiyabhama   Dr.R.Shivakumar   Dr.S.R.R.Senthil Kumar

Copy to:-
HOD/CSE, Second Semester M.Tech DS Students and Staff, COE

Sona College of Technology, Salem
(An Autonomous Institution)
Courses of Study for ME III Semester under Regulations 2019
Computer Science and Engineering
Branch: M.Tech Data Science

S. No	Course Code	Course Title	Lecture	Tutorial	Practical	Credit	Total Contact Hours
Theory							
1	P19MDS512	**Professional Elective-** Internet of Things	3	0	0	3	45
2	P19MDS505	**Professional Elective-** Cloud Computing	3	0	0	3	45
3	P19WMC601	**Open Elective-** Mobile Technology and Network	3	0	0	3	45
Practical							
4	P19MDS301	Project Work Phase-1	0	0	16	8	240
					Total Credits	17	

Approved by

Chairperson, Computer Science and Engineering BOS   **Member Secretary, Academic Council**   **Chairperson, Academic Council & Principal**
      Dr.B.Sathiyabhama   **Dr.R.Shivakumar**   **Dr.S.R.R.Senthil Kumar**

Copy to:-
      HOD/CSE, Third Semester M.Tech DS Students and Staff, COE

Sona College of Technology, Salem
(An Autonomous Institution)
Courses of Study for ME IV Semester under Regulations 2019
Computer Science and Engineering
Branch: M.Tech Data Science

S. No	Course Code	Course Title	Lecture	Tutorial	Practical	Credit	Total Contact Hours
		Practical					
1	P19MDS401	Project Work Phase – II	0	0	28	14	420
					Total Credits	**14**	

Approved by

Chairperson, Computer Science and Engineering BOS   Member Secretary, Academic Council   Chairperson, Academic Council & Principal
Dr.B.Sathiyabhama   Dr.R.Shivakumar   Dr.S.R.R.Senthil Kumar

Copy to:-
    HOD/CSE, Fourth Semester M.Tech DS Students and Staff, COE

Sona College of Technology, Salem
(An Autonomous Institution)
Courses of Study for ME I Semester under Regulations 2019
Computer Science and Engineering
Branch: M.Tech Data Science

S. No	Course Code	Course Title	Lecture	Tutorial	Practical	Credit	Total Contact Hours
Theory							
1	P19MDS101	Statistical Computing	2	1	0	3	45
2	P19MDS102	Matrix and Scientific Computing	2	1	0	3	45
3	P19MDS103	Advanced Data Structures and Algorithms	3	0	0	3	45
4	P19MDS104	Data Science and Big Data Analytics	3	0	0	3	45
5	P19GE101	Research Methodology and IPR	2	0	0	2	30
6	P19GE702	**Audit Course:** Stress Management by YOGA	2	0	0	0	30
Practical							
7	P19MDS105	Advanced Data Structures and Algorithms Laboratory	0	0	4	2	60
8	P19MDS106	Big Data Management and Data Analytics Laboratory	0	0	4	2	60
					Total Credits	**18**	

Approved by

Chairperson, Computer Science and Engineering BOS   **Member Secretary, Academic Council**   **Chairperson, Academic Council & Principal**
Dr.B.Sathiyabhama   **Dr.R.Shivakumar**   **Dr.S.R.R.Senthil Kumar**

Copy to:-
     HOD/CSE, First Semester M.Tech DS Students and Staff, COE

P19MDS101   **STATISTICAL COMPUTING**   L T P C Marks
                                                          2 1 0 3 100

COURSE OUTCOMES
At the end of the course, the students will be able to

1. Apply the concepts of probability and random variable, moments, moment generating function and its properties to solve the problems.
2. Analyse the characteristics of the estimators, find the estimate of the parameters using maximum likelihood estimation and method of moment
3. Test the hypothesis about the population using Z,, and □2-test statistics.
4. Analyse the variances of several variables using standard designs
5. Apply the multivariate analysis concept to analyse the given set of data which involves more than one variable.

UNIT I   PROBABILITY AND RANDOM VARIABLE   9
Axioms of probability – Conditional probability – Total probability – Baye's theorem – Random variable – Probability mass function, probability density function, moment generating function and their properties

UNIT II   ESTIMATION THEORY   9
Estimators – Unbiasedness, consistency, efficiency and sufficiency (definitions and simple problems only) – Maximum likelihood estimation – Method of moments.

UNIT III   TESTING OF SIGNIFICANCE   9
Parameter and statistic – Null and alternative hypothesis – Errors in sampling, critical region and level of significance – One tailed and two tailed tests – Large sample tests for proportions, mean difference between means, standard deviation – t- test for single mean, difference between means and paired t-test- χ^2-test- independence of attributes, goodness of fit – F-test.

UNIT IV   DESIGN OF EXPERIMENTS   9
Analysis of variance – One way classification – Completely randomised design – Two way classification – Randomised block design.

UNIT V   MULTIVARIATE ANALYSIS   9
Random vectors and matrices – Mean vectors and covariance matrices – Multivariate normal density and its properties – Principal components: population components from standardized variables
                                        Theory :30 hours Tutorial :15 hours Total: 45

REFERENCE BOOKS
1. S. C. Gupta, V. K. Kapoor, "Fundamentals of Mathematical Statistics", Sultan Chand and Publishers, 11th Edition, Reprint, 2019
2. R. A. Johnson and D. W. Wichern, "Applied Multivariate Statistical Analysis", Pearson Publishers, 6th Edition, 2015.
3. J. L. Devore, "Probability and Statistics for Engineering and the Sciences", Thomson and Duxbury Publishers, 9th Edition, 2015
4. R. A. Johnson and C. B. Gupta, "Miller and Freund's, Probability and Statistics for Engineers", Pearson Publishers, 9th Edition, 2018.

P19MDS102   MATRIX AND SCIENTIFIC COMPUTING   L   T   P   C   Marks

                                                                    2   1   0   3   100

COURSE OUTCOMES

At the end of the course, the students will be able to

1. Solve linear system of equations by direct and indirect methods.
2. Apply the concepts of vector spaces and linear transformations in real world applications.
3. Apply the various matrix factorization techniques to decompose the given matrix.
4. Apply the principle of least square to represent a set of data by an equation.
5. Apply various numerical methods to find the intermediate value from a given set of data.

UNIT I   LINEAR SYSTEM OF EQUATIONS   9

Rank of a matrix – Solution of linear system of equations by matrix method, Gauss elimination and Gauss-Jordan, Gauss-Jacobi and Gauss-Seidel methods

UNIT II   VECTOR SPACES   9

Vector Space – Linear independence and dependence of vectors – Basis – Dimension – Linear transformations (maps) – Matrix associated with a linear map – Range and kernel of a linear map – Rank-nullity theorem (without proof).

UNIT III   MATRIX DECOMPOSITION   9

Lower-Upper (LU) decomposition – Cholesky's factorization – QR factorization – House Holder transformation – Singular value decomposition – Pseudo inverse

UNIT IV   CURVE FITTINGS   9

Principle of least squares – Fitting a straight line – Fitting a parabola – Fitting an exponential curve – Fitting a curve of the form $y = ax^b$.

UNIT V   INTERPOLATION AND APPROXIMATION   9

Interpolation – Newton forward and backward difference formulae – Lagrange's interpolation formula – Inverse Lagrange's interpolation formula.

                                   Theory :30 hours Tutorial :15 hours Total: 45

REFERENCE BOOKS

1. G. H. Golub and C. F. Van Loan, "Matrix Computations", Johns Hopkins University Press, 4th Edition, 2013.
2. T. Veerarajan, "Numerical Methods", McGraw Hill Publishers, Revised Edition, 2019.
3. D. W. Lewis, "Matrix Theory", Allied Publishers, First Indian Reprint, 1995.
4. S. Lipschutz and M. L. Lipson, "Linear Algebra", McGraw Hill Publishers, 6th Edition, 2018.
5. R. Bronson, "Matrix Operations", McGraw Hill Publishers, New York, Reprint, 2011.

P19MDS103   ADVANCED DATA STRUCTURES AND ALGORITHMS   LT P C Marks
3 0 0 3 100

<u>**COURSE OUTCOMES**</u>
At the end of the course, the students will be able to

1. Design algorithms to solve real time problems
2. Design and develop algorithms using various hierarchical data structures
3. Develop Graph algorithms to solve real-life problems
4. Apply suitable design strategy for problem solving
5. Analyse various NP hard and NP complete problems

COs	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	P09	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	3	3	3	2	2	2	2	1	2	2	3	3	3	3
CO2	3	3	3	3	2	2	2	2	2	2	2	3	3	3	3
CO3	3	3	3	3	2	2	2	2	2	2	1	3	3	3	3
CO4	3	3	3	3	2	2	2	2	2	2	2	3	3	3	3
CO5	3	3	3	3	2	2	2	2	2	2	2	3	3	3	3

UNIT I   ROLE OF ALGORITHMS IN COMPUTING   9
Algorithms – Algorithms as a Technology- Insertion Sort – Analyzing Algorithms – Designing Algorithms-Growth of Functions: Asymptotic Notation – Standard Notations and Common Functions- Recurrences: The Substitution Method – The Recursion-Tree Method

UNIT II   HIERARCHICAL DATA STRUCTURES   9
Binary Search Trees: Basics – Querying a Binary search tree – Insertion and Deletion- Red-Black trees: Properties of Red-Black Trees – Rotations – Insertion – Deletion -B-Trees: Definition of Btrees – Basic operations on B-Trees – Deleting a key from a B-Tree- Fibonacci Heaps: structure – Mergeable-heap operations- Decreasing a key and deleting a node-Bounding the maximum degree.

UNIT III   GRAPHS   9
Elementary Graph Algorithms: Representations of Graphs – Breadth-First Search – Depth-First Search – Topological Sort – Strongly Connected Components- Minimum Spanning Trees: Growing a Minimum Spanning Tree – Kruskal and Prim- Single-Source Shortest Paths: The Bellman-Ford algorithm – Single-Source Shortest paths in Directed Acyclic Graphs – Dijkstra's Algorithm; All-Pairs Shortest Paths: Shortest Paths and Matrix Multiplication – The FloydWarshall Algorithm;

UNIT IV   ALGORITHM DESIGN TECHNIQUES   9

Dynamic Programming: Matrix-Chain Multiplication – Elements of Dynamic Programming – Longest Common Subsequence- Greedy Algorithms: An Activity-Selection Problem – Elements of the Greedy Strategy- Huffman Codes.

UNIT V   NP COMPLETE AND NP HARD   9

NP-Completeness: Polynomial Time – Polynomial-Time Verification – NP- Completeness and Reducability – NP-Completeness Proofs – NP-Complete Problems

Total: 45 H

REFERENCE BOOKS

1. Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, Data Structures and Algorithms, Pearson Education, Reprint 2006.
2. S.Sridhar,Design and Analysis of Algorithms‖, First Edition, Oxford University Press, 2014.
3. Robert Sedgewick and Kevin Wayne, ―ALGORITHMS, Fourth Edition, Pearson Education, 2011
4. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, ―Introduction to Algorithms‖, Third Edition, Prentice-Hall, 2011.
5. Anany Levitin, Introduction to the Design and Analysis of Algorithms, Pearson Education, Third Edtion 2017.
6. Ellis Horowitz, Sartaj Sahni, Susan Anderson-Freed, Fundamentals of Data Structures in C, Universities Press; Second edition, 2008.

P19MDS104   DATA SCIENCE AND BIG DATA ANALYTICS

	L	T	P	C	Marks
	3	**0**	**0**	**3**	**100**

COURSE OUTCOMES
At the end of the course, the students will be able to
1. Comprehend Data science process
2. Apply various classifiers for real-time problems and analyze the results
3. Design and develop simple applications using R
4. Configure the Hadoop architecture
5. Process the big data using Mapreduce

COs	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	P09	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	3	3	3	3	3	1	1	2	2	1	3	2	3	3
CO2	3	3	3	3	3	2	1	1	1	1	1	3	1	3	3
CO3	3	3	3	3	3	3	2	1	1	1	1	3	1	3	3
CO4	3	3	3	3	3	2	2	1	1	2	1	3	1	3	3
CO5	3	3	3	3	3	3	3	1	1	1	1	3	2	3	3

UNIT I   INTRODUCTION TO DATA SCIENCE   9
Data science process – roles, stages in data science project – working with data from files – working with relational databases – exploring data – managing data – cleaning and sampling for modeling and validation – introduction to NoSQL.

UNIT II   MODELING METHODS   9
Choosing and evaluating models – mapping problems to machine learning, evaluating clustering models, validating models – cluster analysis – K-means algorithm, Naïve Bayes – Memorization Methods – Linear and logistic regression – unsupervised methods.

UNIT III   INTRODUCTION TO R   9
Reading and getting data into R – ordered and unordered factors – arrays and matrices – lists and data frames – reading data from files – probability distributions – statistical models in R - manipulating objects – data distribution. Introduction to graphical analysis – plot() function – displaying multivariate data – matrix plots – multiple plots in one window - exporting graph - using graphics parameters

UNIT IV   BIG DATA AND HADOOP DISTRIBUTED FILE SYSTEM ARCHITECTURE
                                                                                        9

Introduction of big data-Characteristics of big data-Data in the warehouse and data in Hadoop- Importance of Big data Hadoop: components of Hadoop-Application Development in Hadoop-Getting your data in Hadoop-other Hadoop Components

HDFS Architecture – HDFS Concepts – Blocks – NameNode – Secondary NameNode – DataNode – HDFS Federation – Basic File System Operations – Data Flow – Anatomy of File Read – Anatomy of File Write

UNIT V   PROCESSING YOUR DATA WITH MAPREDUCE   9
Algorithms using Map reduce, Matrix-Vector Multiplication by Map Reduce – Hadoop - Understanding the Map Reduce architecture - Writing Hadoop MapReduce Programs - Loading data into HDFS - Executing the Map phase - Shuffling and sorting - Reducing phase execution. Hadoop Word Count Implementation. Case studies.

                                                                                Total: 45 H

REFERENCE BOOKS

1. Tony Ojeda, Sean Patrick Murphy, Benjamin Bengfort, Abhijit Dasgupta, "Practical Data Science Cookbook", Packt Publishing Ltd., 2014.
2. Noreen Burlingame and Lars Nielsen, "A Simple Introduction to DATA SCIENCE", 2012.
3. Paul Zikopoulos, Chris Eaton, Dirk DeRoos, Tom Deutsch, George Lapis, Understanding Big Data: Analytics for Enterprise Class Hadoop and streaming Data,The McGraw-Hill Companies, 2012
4. Nina Zumel, John Mount, "Practical Data Science with R", Manning Publications, 2014.
5. Mark Gardener, "Beginning R - The Statistical Pr ogramming Language", John Wiley & Sons, Inc., 2012.
6. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, "Mining of Massive Datasets", Cambridge University Press, 2014.
7. Boris lublinsky, Kevin t. Smith, Alexey Yakubovich, "Professional Hadoop Solutions", Wiley, ISBN: 9788126551071, 2015
8. http://www.johndcook.com/R_language_for_programmers.html
9. http://bigdatauniversity.com/
10. http://home.ubalt.edu/ntsbarsh/stat-data/topics.htm#rintroduction

COURSE OUTCOMES

At the end of the course, the students will be able to

1. Review the literature of the research problem
2. Choose appropriate data collection and sampling method according to the research problem.
3. Interpret the results of research and communicate effectively with their peers
4. Explain the Importance of intellectual property rights
5. Evaluate trade mark, develop and register patents

CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak															
COs	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	P09	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	3	3	3	3	3	2	1	1	3	2	3	2	3	3
CO2	3	3	3	3	3	3	1	2	1	3	1	3	1	3	3
CO3	3	3	3	3	3	3	2	2	1	3	2	3	1	3	3
CO4	3	3	3	3	3	2	2	2	1	2	1	3	1	3	3
CO5	3	3	3	3	3	3	2	2	1	2	2	3	2	3	3

UNIT I   INTRODUCTION TO RESEARCH METHODS   6

Definition and Objective of Research, Various steps in Scientific Research, Types of Research, Criteria for Good Research, Defining Research Problem, Research Design , Case Study Collection of Primary and Secondary Data, Collection Methods: Observation, Interview, Questionnaires, Schedules

UNIT II   SAMPLING DESIGN AND HYPOTHESIS TESTING   6

Steps in Sampling Design, Types of Sample Designs, Measurements and Scaling Techniques - Testing of hypotheses concerning means (one mean and difference between two means -one tailed and two tailed tests), concerning variance – one tailed Chi-square test.

UNIT III   INTERPRETATION AND REPORT WRITING   6

Techniques of Interpretation, Precaution in Interpretation, Layout of Research Report, Types of Reports, Oral Presentation, Mechanics of Writing Research Report

UNIT IV   INTRODUCTION TO INTELLECTUAL PROPERTY   6

Introduction, types of intellectual property, international organizations, agencies and treaties, importance of intellectual property rights, Innovations and Inventions trade related intellectual property rights.

UNIT V   TRADE MARKS, COPY RIGHTS AND PATENTS   6

Purpose and function of trade marks, acquisition of trade mark rights, trade mark registration processes, trademark claims –trademark Litigations- International trademark law

Fundamental of copy right law, originality of material, rights of reproduction, rights to perform the work publicly, copy right ownership issues, copy right registration, notice of copy right, international copy right law.

Law of patents: Foundation of patent law, patent searching process, ownership rights and transfer

Total: 30 H

REFERENCE BOOKS

1. C.R. Kothari, Gaurav Garg, Research Methodology Methods and Techniques ,4th Edition, New Age International Publishers, 2019.
2. Deborah E. Bouchoux, "Intellectual Property: The Law of Trademarks, Copyrights, Patents, and Trade Secrets", Delmar Cengage Learning, 4th Edition, 2012.
3. Prabuddha Ganguli, "Intellectual Property Rights: Unleashing the Knowledge Economy", Tata Mc Graw Hill Education, 1st Edition, 2008.
4. Panneerselvam, R., Research Methodology, Second Edition, Prentice-Hall of India, New Delhi, 2013.
5. Ranjith Kumar, Research Methodology – A step by step Guide for Begineers, 4th edition, Sage publisher, 2014.
6. D Llewelyn & T Aplin W Cornish, "Intellectual Property: Patents, Copyright, Trade Marks and Allied Rights", Sweet and Maxwell, 1st Edition, 2016.
7. Ananth Padmanabhan, "Intellectual Property Rights-Infringement and Remedies", Lexis Nexis, 1st Edition, 2012.
8. Ramakrishna B and Anil Kumar H.S, "Fundamentals of Intellectual Property Rights: For Students, Industrialist and Patent Lawyers", Notion Press, 1st Edition, 2017
9. M.Ashok Kumar and Mohd.Iqbal Ali :"Intellectual Property Rights" Serials Pub

P19GE702   STRESS MANAGEMENT BY YOGA　　　**L　T　P　C　Marks**
　　　　　　　　　　　　　　　　　　　　　　　　　　　　2　0　0　0　100

<u>**COURSE OUTCOMES**</u>
At the end of the course, the students will be able to

1. Develop physical and mental health thus improving social health
2. Increase immunity power of the body and prevent diseases
3. Accelerate memory power
4. Achieve the set goal with confidence and determination
5. Improve stability of mind, pleasing personality and work with awakened wisdom

CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
COs	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)													
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	P09	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	2	2	1	2	0	2	2	3	2	2	0	3	2	1	3
CO2	2	3	1	2	0	2	2	2	2	1	0	3	1	1	3
CO3	2	0	1	2	0	2	3	2	1	1	0	3	1	3	3
CO4	2	2	1	2	0	2	2	3	2	2	0	3	1	2	3
CO5	2	2	1	2	0	3	3	1	2	1	0	3	2	2	3

UNIT I　　　　　　　　　　　　　　　　　　　　　　　　　　　　　**6**
Yoga-Introduction - Astanga Yoga- 8 parts-Yam and Niyam etc.- Do`s and Don'ts in life-Benefits of Yoga and Asana- Yoga Exercise- and benefits- Pranayam Yoga- Nadisuthi, Practice and Spinal Sclearance Practice-Regularization of breathing techniques and its effects-Practice and kapalapathy practice.

UNIT II　　　　　　　　　　　　　　　　　　　　　　　　　　　　　**6**

Neuromuscular breathing exercise and Practice- Magarasa Yoga, 14 points Acupressure techniques and practice- Body relaxation practice and its benefits- Raja Yoga- 1.Agna –explanation and practice- Activation of Pituitary- Raja Yoga-2. Santhi Yoga-Practice-Balancing of physical and mental power.

UNIT III　　　　　　　　　　　　　　　　　　　　　　　　　　　　**6**

Raja Yoga-3.Sagasrathara yoga –practice- Activation of dormant brain cells-Kayakalpa-theory- Kayakalpa –practice-Yogic exercise to improve physical and mental health and practice-Asanas –explanation-Practice-benefits

UNIT IV　　　　　　　　　　　　　　　　　　　　　　　　　　　　**6**

Sun namaskar- 12 poses-explanation and practice-Yoga –Asana-Padmasana, vajrasana,chakrasana, viruchasanaetc-Stress management with Yoga-Role of women and Yoga Equality, nonviolence, Humanity,Self-control- Food and yoga Aware of self-destructive habits
Avoid fault thinking (thought analysis-Practice)-Yoga Free from ANGER (Neutralization of anger)& practice

Moralisation of Desire & practice- Punctuality-Love-Kindness-CompassionEradication ofworries-Practice - Personality development, positive thinking-Good characters to lead a moral life
How to clear the polluted mind- Benefits of blessing- Five- fold culture –explanation- Karma Yoga Practice In Geetha- Sense of duty-Devotion, self- reliance, confidence, concentration, truthfulness, cleanliness.

Total: 30 H

REFERENCE BOOKS

1. 'Yogic Asanas for Group Tarining-Part-I" Janardan Swami YogabhyasiMandal, Nagpur
2. "Rajayoga or conquering the Internal Nature" by Swami Vivekananda, AdvaitaAshrama

P19MDS105   ADVANCED DATA STRUCTURES AND ALGORITHMS LABORATORY

L T P C M
0 0 4 2 100

COURSE OUTCOMES

At the end of course, the students will be able to

1. Design and implement basic and advanced data structures for real applications+
2. Design algorithms using graph structures
3. Implement for real applications using design techniques

COs	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	P09	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	3	3	3	3	2	2	2	3	3	2	2	3	3	3
CO2	3	3	3	3	2	2	2	2	3	3	2	2	3	3	3
CO3	3	3	3	3	2	2	2	2	3	3	3	2	3	3	3

Total: 60 H

List of Experiments:

1. Implementation of Merge Sort and Quick Sort-Algorithms
2. Implementation of a Binary Search Tree
3. Red-Black Tree Implementation
4. Heap Implementation
5. Fibonacci Heap Implementation
6. Graph Traversals
7. Spanning Tree Implementation
8. Shortest Path Algorithms (Dijkstra's algorithm, Bellmann Ford Algorithm)
9. Implementation of Matrix Chain Multiplication
10. Activity Selection and Huffman Coding Implementation.

P19MDS106   BIG DATA MANAGEMENT AND DATA ANALYTICS LABORATORY   LT P CM
                                                                     0 0 4 2 100

COURSE OUTCOMES

At the end of experiments, the students will be able to

1. Perform for data summarization, queries, and interpret the results
2. Design and develop data modeling techniques to large data sets
3. Creating and building a complete business data analytics solution

	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
CO / PO, PSO Mapping															
(3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak															
COs	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**P09**	**PO10**	**PO11**	**PO12**	**PSO1**	**PSO2**	**PSO3**
CO1	3	3	3	3	3	2	2	2	1	3	2	2	3	2	3
CO2	3	3	3	3	2	2	2	2	1	3	2	2	3	2	3
CO3	3	3	3	3	2	2	2	2	3	3	3	2	3	3	3

Total: 60 H

List of Experiments:

1. (i)Perform setting up and Installing Hadoop in its two operating modes:
   - Pseudo distributed
   - Fully distributed
   (ii) Use web based tools to monitor your Hadoop setup.
2. i)Implement the following file management tasks in Hadoop:
   - Adding file and directories
   - Retrieving files
   - Deleting files
   ii) Benchmark and stress test an Apache Hadoop cluster
3. Run a basic Word Count Map Reduce program to understand Map Reduce Paradigm.
   - Fine the number of occurrence of each word appearing in the input file(s)
   - Performing a MapReduce Job for word search count (look for specific keywords in a file)
4. Stop word elimination problem:
   Input:
   - A Large textual file containing one sentence per line
   - A small file containing a set of stop words (One stop word per line)
   - A textual file containing the same sentences of the large input file without the words appearing in the small file.

5. Write a Map Reduce program that mines weather data. Weather sensors collecting data every hour at many locations across the globe gather large volume of log data, which is a good candidate for analysis with MapReduce, since it is semi structured and record-oriented. Data available at: https://github.com/tomwhite/hadoopbook/tree/master/input/ncdc/all

- Find average, max and min temperature for each year in NCDC data set?

- Filter the readings of a set based on value of the measurement, Output the line of input files associated with a temperature value greater than 30.0 and store it in a separate file.

6. Purchases.txt Dataset
   - Instead of breaking the sales down by store, give us a sales breakdown by product category across all of our stores
   I. What is the value of total sales for the following categories?
- Toys
- Consumer Electronics
   II. Find the monetary value for the highest individual sale for each separate store
   III. What are the values for the following stores?
     - Reno
     - Toledo
     - Chandler

Find the total sales value across all the stores, and the total number of sales.

7. Install and Run Pig then write Pig Latin scripts to sort, group, join, project, and filter your data.
8. Write a Pig Latin scripts for finding TF-IDF value for book dataset (A corpus of eBooks available at: Project Gutenberg)
9. Install and Run Hive then use Hive to create, alter, and drop databases, tables, views,functions, and indexes
10. Install, Deploy & configure Apache Spark Cluster. Run apache spark applications using Scala.
11. Data analytics using Apache Spark on Amazon food dataset, find all the pairs of items frequently reviewed together
    - Write a single Spark application that:
    - Transposes the original Amazon food dataset, obtaining a PairRDD of the type:

<user_id> → <list of the product_ids reviewed by user_id>
    - Counts the frequencies of all the pairs of products reviewed together;

Writes on the output folder all the pairs of products that appear more than once and their frequencies. The pairs of products must be sorted by frequency

P19GE702　　　　　　**Stress Management by Yoga**

Course Outcomes:

At the end of completion of this course, students will be able to
1. Develop physical and mental health thus improving social health
2. Increase immunity power of the body and prevent diseases
3. Accelerate memory power
4. Achieve the set goal with confidence and determination
5. Improve stability of mind, pleasing personality and work with awakened wisdom

UNIT – I　　　　　　　　　　　　　　　　　　　　　　　　　　6
Yoga-Introduction - Astanga Yoga- 8 parts-Yam and Niyam etc.- Do`s and Don'ts in life-Benefits of Yoga and Asana- Yoga Exercise- and benefits- Pranayam Yoga- Nadi suthi, Practice and Spinal Sclearance Practice- Regularization of breathing techniques and its effects-Practice and kapalapathy practice.

UNIT – II　　　　　　　　　　　　　　　　　　　　　　　　　6
Neuromuscular breathing exercise and Practice- Magarasa Yoga, 14 points Acupressure techniques and practice- Body relaxation practice and its benefits- Raja Yoga- 1.Agna – explanation and practice- Activation of Pituitary- Raja Yoga- 2. Santhi Yoga-Practice-Balancing of physical and mental power.

UNIT – III　　　　　　　　　　　　　　　　　　　　　　　　6
Raja Yoga- 3. Sagasrathara yoga –practice- Activation of dormant brain cells-Kayakalpa-theory- Kayakalpa –practice-Yogic exercise to improve physical and mental health and practice-Asanas –explanation-Practice-benefits

UNIT –IV　　　　　　　　　　　　　　　　　　　　　　　　　6
Sun namaskar- 12 poses-explanation and practice-Yoga –Asana-Padmasana, vajrasana,chakrasana, viruchasana etc-Stress management with Yoga-Role of women and Yoga
Equality, nonviolence, Humanity, Self- control- Food and yoga Aware of self-destructive habits
Avoid fault thinking (thought analysis-Practice)-Yoga Free from ANGER (Neutralization of anger)& practice
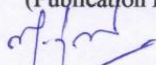
UNIT – V　　　　　　　　　　　　　　　　　　　　　　　　　6
Moralisation of Desire & practice- Punctuality-Love-Kindness-Compassion Eradication of worries-Practice -Personality development, positive thinking-Good characters to lead a moral life
How to clear the polluted mind- Benefits of blessing- Five- fold culture –explanation- Karma Yoga Practice In Geetha- Sense of duty-Devotion, self- reliance, confidence, concentration, truthfulness, cleanliness.

Reference Books
1. 'Yogic Asanas for Group Tarining-Part-I" Janardan Swami Yogabhyasi Mandal, Nagpur
2. "Rajayoga or conquering the Internal Nature" by Swami Vivekananda, AdvaitaAshrama
   (Publication Department), Kolkata

Total: 30 hours

Dr. M. Renuga
BoS – Chairperson,
Science & Humanities
HOD / H&L

Sona College of Technology, Salem
(An Autonomous Institution)
Courses of Study for ME II Semester under Regulations 2019
Computer Science and Engineering
Branch: M.Tech Data Science

S. No	Course Code	Course Title	Lecture	Tutorial	Practical	Credit	Total Contact Hours
-Theory							
1	P19MDS201	Machine Learning	3	0	0	3	45
2	P19MDS202	Data Visualization	3	0	0	3	45
3	P19MDS511	**ELECTIVE**- Block Chain Techniques	3	0	0	3	45
4	P19MDS515	**ELECTIVE**-Business Intelligence	3	0	0	3	45
5	P19MDS516	**ELECTIVE**-Artificial Intelligence	3	0	0	3	45
6	P19GE701	**Audit Course**-English for Research Paper Writing	2	0	0	0	30
-Practical							
6	P19MDS203	Machine Learning Laboratory	0	0	4	2	60
7	P19MDS204	Data Visualization Laboratory	0	0	4	2	60
					Total Credits	19	

Approved by

Chairperson, Computer Science and Engineering BOS　　**Member Secretary, Academic Council**　　**Chairperson, Academic Council & Principal**
Dr.B.Sathiyabhama　　　　　　　　　　　　　　**Dr.R.Shivakumar**　　　　　　　　　　**Dr.S.R.R.Senthil Kumar**

Copy to:-
　　HOD/CSE, Second Semester M.Tech DS Students and Staff, COE

30.05.2022　　　　　　　　　　　　　　　　　　　　Regulations-2019

COURSE OUTCOMES
At the end of the course, students will be able to

- Demonstrate the processes associated with Business Intelligence framework
- Solve business scenario, by identifying the metrics, indicators and make recommendations to achieve the business goals
- Develop analytical and critical thinking skills for the development of integrative plans for enterprise-wide systems that optimize enterprise performance
- Design an enterprise dashboard that depicts the key performance indicators which helps in decision making
- Apply business intelligence concepts in cloud computing, ERP systems

CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
COs	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)													
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	P0 9	PO1 0	PO1 1	PO1 2	PSO 1	PSO 2	PSO 3
CO 1	3	3	3	3	1	1	1	1	2	2	1	3	2	2	3
CO 2	3	3	3	2	1	1	1	1	1	1	1	3	1	3	3
CO 3	3	3	3	3	1	1	2	1	2	3	1	3	1	3	3
CO 4	3	3	3	2	2	1	2	1	1	2	1	3	1	3	3
CO 5	3	3	3	3	3	2	2	1	3	1	1	3	2	3	3

UNIT I   INTRODUCTION TO BUSINESS INTELLIGENCE   9
Introduction to digital data: Introduction, Types – structured, semi-structured and unstructured. **Introduction to OLTP and OLAP:** OLTP Vs OLAP - Architectures (MOLAP, ROLAP, HOLAP) - OLAP Operations - BI Definitions & Concepts: BI Framework-Data Warehousing concepts and its role in BI-BI Infrastructure Components – BI Process - BI Technology - BI Roles and Responsibilities - Business Applications of BI- BI best practices

UNIT II   BASICS OF DATA INTEGRATION   9

Data Integration: Concepts: needs and advantages of using data integration-introduction to common data integration approaches - Meta data - types and sources - Introduction to data quality-data profiling concepts and applications - Kettle Software: Introduction to ETL using Pentaho data Integration

UNIT III   INTRODUCTION TO MULTI-DIMENSIONAL DATA MODELING   9

Multidimensional data model: Introduction to data and dimension modeling-data modeling basics - Types - Technique - fact and dimension tables - Dimensional models - Measures and Metrics: Introduction to business metrics and KPIs - KPI usage in companies - Creating cubes using Microsoft Excel

UNIT IV   BASICS OF ENTERPRISE REPORTING   9

Reporting: A typical enterprise - Malcolm Baldrige - quality performance framework - balanced scorecard - Enterprise dashboard - balanced scorecard vs. enterprise dashboard - Best practices in the design of enterprise dashboards - Enterprise reporting using MS Access / MS Excel

UNIT V   BI APPLICATIONS AND CASE STUDIES   9

Applications: Understanding BI and mobility - BI and cloud computing - BI for ERP systems - Social CRM and BI - Case Study Briefs: Good Lift HealthCare group - Ten to Ten retail store

Total: 45 hours

TEXT BOOKS

1. RN Prasad and Seema Acharya, "Fundamental of Business Analytics", Wiley India Pvt. Ltd, 2nd Edition, 2017

REFERENCES

1. John Boyer, Bill Frank, Brian Green, Tracy Harris, and Kay Van De Vanter "Business Intelligence Strategy: A Practical Guide for Achieving BI Excellence", IBM Corporation, 2010
2. R. Sharda, D. Delen, & E. Turban .Business Intelligence and Analytics. Systems for Decision Support, 10th Edition. ; Pearson/Prentice Hall, 2015
3. Swain Scheps "Business Intelligence for Dummies", Wiley Publishing Inc, 2008
4. Cindi Howson " Successful Business Intelligence:Secrets to making BI a killer App", McGraw Hill, 2008
5. Elizabeth Vitt, Michael Luckevich, Stacia Misner "Business Intelligence: Making Better Decisions Faster", Microsoft Press, 2002
6. Elizabeth Vitt, Michael Luckevich, Stacia Misner "Business Intelligence: Making Better Decisions Faster", Microsoft Press, 2002

P19MDS511   BLOCKCHAIN TECHNIQUES   L   T   P   C   Marks

3   0   0   3   100

COURSE OUTCOMES

At the end of the course, the students will be able to

1. Discover the secure and efficient transactions using crypto-currencies
2. Experiment cryptocurrency trading and crypto exchanges
3. Apply bitcoin in real world application
4. Design and Develop private blockchain environment and a smart contract on Ethereum
5. Build the Hyperledger architecture and the consensus mechanism applied in the Hyperledger

	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
COs	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	P0 9	PO1 0	PO1 1	PO1 2	PSO 1	PSO 2	PSO 3
CO 1	3	3	3	3	1	1	1	1	2	2	1	3	2	2	3
CO 2	3	3	3	2	1	1	1	1	1	1	1	3	1	3	3
CO 3	3	3	3	3	1	1	2	1	2	3	1	3	1	3	3
CO 4	3	3	3	2	2	1	2	1	1	2	1	3	1	3	3
CO 5	3	3	3	3	3	2	2	1	3	1	1	3	2	3	3

UNIT I   CRYPTOCURRENCY AND BLOCKCHAIN- INTRODUCTION   9

Blockchain- An Introduction, Distinction between databases and blockchain, Distributed ledger. Blockchain ecosystem - Consensus Algorithms & Types, Blockchain structure, Distributed networks- Distributed Applications (DApps) – Web 3.0 - DApps Ecosystems. Working - Permissioned and permission-less Blockchain – Cross Chain Technologies. – IOT & Blockchain - Digital Disruption in Industries – Banking, Insurance, Supply Chain, Governments, IP rights, Creation of trustless Ecosystems – Block chain as a Service – Open Source Block chains

UNIT II   CRYPTO CURRENCIES   9

Crypto Currencies - Anonymity and Pseudonymity in Cryptocurrencies - Digital Signatures - Cryptocurrency Hash Codes -Need for Crypto Currencies – Crypto Markets – Explore Crypto Currency Ecosystems - ICOs – Crypto Tokens - Atomic Swaps – Crypto Currency Exchanges – Centralised and Decentralized Crypto exchanges – Regulations on Crypto Currencies & exchanges – Downside of non-regulated currencies – crypto Scams – Exchange hacks

UNIT III   BITCOIN   9

Bitcoin – history- Bitcoin- usage, storage, selling, transactions, working- Invalid Transactions- Parameters that invalidate the transactions- Scripting language in Bitcoin- Applications of Bitcoin script- Nodes and network of Bitcoin- Bitcoin ecosystem

UNIT IV   ETHEREUM   9

The Ethereum ecosystem, DApps and DAOs - Ethereum working- Solidity- Contract classes, functions, and conditionals- Inheritance & abstract contracts- Libraries- Types & optimization of Ether- Global variables- Debugging- Future of Ethereum- Smart Contracts on Ethereum- different stages of a contract deployment- Viewing Information about blocks in Blockchain- Developing smart contract on private Blockchain- Deploying contract from web and console

UNIT V   HYPERLEDGER   9

Hyperledger Architecture- Consensus- Consensus & its interaction with architectural layers- Application programming interface- Application model -Hyperledger frameworks- Hyperledger Fabric -Various ways to create Hyperledger Fabric Blockchain network- Creating and Deploying a business network on Hyperledger Composer Playground- Testing the business network definition- Transferring the commodity between the participants

Total: 45

REFERENCE BOOKS

1. Mastering Bitcoin: Unlocking Digital Cryptocurrencies, by Andreas M Antonopoulos 2018
2. Ethereum: Blockchains, Digital Assets, Smart Contracts, Decentralized Autonomous Organizations- 2016

OTHER ONLINE REFERENCES

1.https://www.coursera.org/learn/ibm-blockchain-essentials-for-developers

2. https://museblockchain.com/

3. https://www.provenance.org/

4. https://www.coursera.org/learn/blockchain-basics

5. https://steemit.com/

6. https://101blockchains.com/

7.https://followmyvote.com/

P19MDS516   ARTIFICIAL INTELLIGENCE   L   T   P   C   Marks
                                                 3   0   0   3   100

COURSE OUTCOMES

At the end of the course, the students will be able to
1. Design an intelligent agent by considering the nature of environment and applications
2. Solve the problems related to search application
3. Design knowledge base for any application using propositional/first order logic
4. Implement a suitable multi agent system for the given problem
5. Design a communicative agent for NLP application

COs	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak

COs	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	P09	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	3	3	3	1	1	1	1	2	2	1	3	2	2	3
CO2	3	3	3	2	1	1	1	1	1	1	1	3	1	3	3
CO3	3	3	3	3	1	1	2	1	2	3	1	3	1	3	3
CO4	3	3	3	2	2	1	2	1	1	2	1	3	1	3	3
CO5	3	3	3	3	3	2	2	1	3	1	1	3	2	3	3

UNIT I   INTRODUCTION   9
Introduction to Artificial Intelligence-The Foundations of Artificial Intelligence. The History of Artificial Intelligence-Intelligent Agents: Agents and Environments-The Concept of Rationality-The Nature of Environments-The Structure of Agents- Problem-Solving Agents-Example problems

UNIT II   PROBLEM SOLVING USING SERACH TECHNIQUES   9
Uninformed Search Strategies- Avoiding Repeated States- Searching with Partial Information- Informed Search and Exploration: Informed (Heuristic) Search Strategies- Heuristic Functions- Local Search Algorithms and Optimization Problems- Constraint Satisfaction problems-Adversarial search- minimax algorithm- Alpha-Beta pruning

UNIT III   KNOWLEDGE AND REASONING   9

Knowledge-Based agents – Logic –Propositional logic – First order logic- Representation – Syntax and semantics – Knowledge engineering – Inference in First order logic- Unification and lifting- Forward and backward chaining-Resolution

UNIT IV   SOFTWARE AGENTS   9

Architecture for Intelligent Agents – Agent communication – Negotiation and Bargaining – Argumentation among Agents – Trust and Reputation in Multi-agent systems.

UNIT V   COMMUNICATION AND APPLICATIONS OF AI   9

Communication: Phrase Structure Grammars - A Formal Grammar for a Fragment of English- Syntactic Analysis (Parsing) – Augmented Grammar and Semantic Interpretation - Machine translation –Speech recognition Tool for Artificial Intelligence -Tensor flow and IBM Watson

Total: 45

REFERENCE BOOKS

1. S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach", Third Edition, Prentice Hall, 2015
2. Gerhard Weiss, ―Multi Agent Systems‖, Second Edition, MIT Press, 2013
3. Nils J. Nilsson, "Artificial Intelligence: A New Synthesis", Harcourt Asia Pvt. Ltd., 2009
4. George F. Luger, "Artificial Intelligence-Structures and Strategies for Complex Problem Solving", Pearson Education, 2009.
5. Tom Mitchell, "Machine Learning", McGraw Hill, 2015.
6. P. Flach, "Machine Learning: The art and science of algorithms that make sense of data", Cambridge University Press, 2012
7. M. Mohri, A. Rostamizadeh, and A. Talwalkar, "Foundations of Machine Learning", MIT Press, 2012

COURSE OUTCOMES

At the end of the course, the students will be able to

- Comprehend machine learning basics
- Implement supervised learning algorithms for the given application and analyze the results
- Use tools to implement typical clustering algorithms for different types of applications
- Design and implement an HMM for a sequence model type of application
- Comprehend the advanced learning algorithms and identify the suitable applications for solving using these advanced learning techniques

	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
COs	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	P0 9	PO1 0	PO1 1	PO1 2	PSO 1	PSO 2	PSO 3
CO 1	3	3	3	3	1	1	1	1	2	2	1	3	2	2	3
CO 2	3	3	3	2	1	1	1	1	1	1	1	3	1	3	3
CO 3	3	3	3	3	1	1	2	1	2	3	1	3	1	3	3
CO 4	3	3	3	2	2	1	2	1	1	2	1	3	1	3	3
CO 5	3	3	3	3	3	2	2	1	3	1	1	3	2	3	3

UNIT I   INTRODUCTION   9

Machine Learning -Machine Learning Foundations –Overview –Design of a Learning system -Types of machine learning –Applications Mathematical foundations of machine learning -random variables and probabilities -Probability Theory –Probability distributions -Decision Theory-Bayes Decision Theory - Information Theory

UNIT II   SUPERVISED LEARNING   9

Linear Models for Regression -Linear Models for Classification –Naïve Bayes -Discriminant Functions - Probabilistic Generative Models -Probabilistic Discriminative Models -Bayesian Logistic Regression. Decision

Trees -Classification Trees-egression Trees -Pruning. Neural Networks -Feed-forward Network Functions -Back-propagation. Support vector machines -Ensemble methods-Bagging-Boosting

UNIT III   UNSUPERVISED LEARNING   9
Clustering-K-means -EM Algorithm-Mixtures of Gaussians. The Curse of Dimensionality -Dimensionality Reduction -Factor analysis -Principal Component Analysis -Probabilistic PCA-Independent components analysis

UNIT IV   PROBABILISTIC GRAPHICAL MODELS   9
Graphical Models -Undirected graphical models-Markov Random Fields -Directed Graphical Models -Bayesian Networks -Conditional independence properties -Inference –Learning-Generalization -Hidden Markov Models -Conditional random fields(CRFs).

UNIT V   ADVANCED LEARNING   9
Sampling –Basic sampling methods –Monte Carlo. Reinforcement Learning-K-Armed Bandit-Elements -Model-Based Learning-Value Iteration-Policy Iteration. Temporal Difference Learning-Exploration Strategies-Deterministic and Non-deterministic Rewards and Actions Computational Learning Theory -Mistake bound analysis, sample complexity analysis, VC dimension. Occam learning, accuracy and confidence boosting

Total: 45

REFERENCE BOOKS

1. Christopher Bishop, "Pattern Recognition and Machine Learning" Springer, 2007
2. Kevin P. Murphy, "Machine Learning: A Probabilistic Perspective", MIT Press, 2012.
3. EthemAlpaydin, "Introduction to Machine Learning", MIT Press, Third Edition, 2014
4. Jure Leskovec, Anand Rajaraman and Jeffrey D. Ullman,"Mining of Massive Datasets", Cambridge University Press, Second Edition.
5. 2016Tom Mitchell, "Machine Learning", McGraw-Hill, 1997
6. Trevor Hastie, Robert Tibshirani, Jerome Friedman, "The Elements of Statistical Learning", Springer, Second Edition, 2011.
7. Stephen Marsland, "Machine Learning -An Algorithmic Perspective", Chapman and Hall/CRC Press, Second Edition, 2014

DATAVISUALIZATION **L   T   P   C   Marks**
                                                      3   0   0   3   100

COURSE OUTCOMES
At the end of each unit, the students will be able to

- Comprehend the skills required for visual analysis
- Analyze various patterns in Data visualization
- Apply visualization techniques for various data analysis tasks
- Design information dashboard with required components
- Analyze critical design practices in real time application development

COs	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	P0 9	PO1 0	PO1 1	PO1 2	PSO 1	PSO 2	PSO 3
CO 1	3	3	3	3	1	1	1	1	2	2	1	3	2	2	3
CO 2	3	3	3	2	1	1	1	1	1	1	1	3	1	3	3
CO 3	3	3	3	3	1	1	2	1	2	3	1	3	1	3	3
CO 4	3	3	3	2	2	1	2	1	1	2	1	3	1	3	3
CO 5	3	3	3	3	3	2	2	1	3	1	1	3	2	3	3

UNIT I   CORE SKILLS FOR VISUAL ANALYSIS   9
Information visualization – effective data analysis – traits of meaningful data – visual perception –making abstract data visible – building blocks of information visualization – analytical interaction – analytical navigation – optimal quantitative scales – reference lines and regions – trellises and crosstabs – multiple concurrent views – focus and context – details on demand – over-plotting reduction – analytical patterns – pattern examples

UNIT II   TIME-SERIES, RANKING, AND DEVIATION ANALYSIS   9
Time-series analysis – time-series patterns – time-series displays – time-series best practices – part-to-whole and ranking patterns – part-to-whole and ranking displays – best practices – deviation analysis – deviation analysis displays – deviation analysis best practices.

UNIT III   DISTRIBUTION, CORRELATION, AND MULTIVARIATE ANALYSIS   9

Distribution analysis – describing distributions – distribution patterns – distribution displays – distribution analysis best practices – correlation analysis – describing correlations – correlation patterns – correlation displays – correlation analysis techniques and best practices – multivariate analysis – multivariate patterns – multivariate displays – multivariate analysis techniques and best practices-Exploratory data Analysis

UNIT IV   INFORMATION DASHBOARD DESIGN   9

Information dashboard – Introduction– dashboard design issues and assessment of needs – Considerations for designing dashboard-visual perception – Achieving eloquence

UNIT V   DASHBOARD DESIGN MEDIA   9

Advantages of Graphics _Library of Graphs – Designing Bullet Graphs – Designing Sparklines – Dashboard Display Media –Critical Design Practices – Putting it all together Unveiling the dashboard-Case Study-Tableau

Total: 45

REFERENCES

1. Ben Fry, "Visualizing data: Exploring and explaining data with the processing environment", O'Reilly, 2008
2. Edward R. Tufte, "The visual display of quantitative information", Second Edition, Graphics Press, 2001.
3. Evan Stubbs, "The value of business analytics: Identifying the path to profitability", Wiley, 2011
4. Gert H. N. Laursen and Jesper Thorlund, "Business Analytics for Managers: Taking business intelligence beyond reporting", Wiley, 2010.
5. Gert H. N. Laursen and Jesper Thorlund, "Business Analytics for Managers: Taking business intelligence beyond reporting", Wiley, 2010
6. Stephen Few, "Information dashboard design: Displaying data for at-a-glance monitoring", second edition, Analytics Press, 2013

P19MDS203   **MACHINE LEARNING LABORATORY**   **L   T   P   C   Marks**

                                                                  0   0   4   2   100

<u>**COURSE OUTCOMES**</u>

At the end of experiments, the students will be able to

- Understand the implementation procedures for the machine learning algorithms
- Solve the problems using machine learning techniques in image and language processing applications
- Choose appropriate algorithms/ techniques to solve computing problems in real-world

<table>
<tr><td colspan="16">CO / PO, PSO Mapping

(3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak</td></tr>
<tr><td rowspan="2">COs</td><td colspan="15">Programme Outcomes (POs) and Programme Specific Outcome (PSOs)</td></tr>
<tr><td>PO 1</td><td>PO 2</td><td>PO 3</td><td>PO 4</td><td>PO 5</td><td>PO 6</td><td>PO 7</td><td>PO 8</td><td>P0 9</td><td>PO1 0</td><td>PO1 1</td><td>PO1 2</td><td>PSO 1</td><td>PSO 2</td><td>PSO 3</td></tr>
<tr><td>CO1</td><td>3</td><td>3</td><td>3</td><td>3</td><td>3</td><td>2</td><td>2</td><td>2</td><td>1</td><td>3</td><td>2</td><td>2</td><td>3</td><td>2</td><td>3</td></tr>
<tr><td>CO2</td><td>3</td><td>3</td><td>3</td><td>3</td><td>2</td><td>2</td><td>2</td><td>2</td><td>1</td><td>3</td><td>2</td><td>2</td><td>3</td><td>2</td><td>3</td></tr>
<tr><td>CO3</td><td>3</td><td>3</td><td>3</td><td>3</td><td>2</td><td>2</td><td>2</td><td>2</td><td>3</td><td>3</td><td>3</td><td>2</td><td>3</td><td>3</td><td>3</td></tr>
</table>

List of Experiments:

1. Implement and demonstrate the FIND-S algorithm for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file
2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples
3. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample
4. Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets
5. Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets
6. Assuming a set of documents that need to be classified, use the naïve Bayesian Classifier model to perform this task. Built-in Java classes/API can be used to write the program. Calculate the accuracy, precision, and recall for your data set

---

7. Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set. You can use Java/Python ML library classes/API.

8. Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering. You can add Java/Python ML library classes/API in the program

9. Write a program to implement k-Nearest Neighbor algorithm to classify the iris data set. Print both correct and wrong predictions. Java/Python ML library classes can be used for this problem

10. Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs

11. Case Study on google Colab

P19MDS204   **DATA VISUALIZATION LABORATORY**   **L   T   P   C   Marks**

                                                            0   0   4   2   100

Course outcomes

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

- Present data with visual representations for your target audience, task, and data
- Identify appropriate data visualization techniques like deviation, correlation, ranking for the given particular requirements imposed by the data
- Analyze and design appropriate principles for creating multiple versions of digital visualizations using various software packages

COs	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak
	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	P0 9	PO1 0	PO1 1	PO1 2	PSO 1	PSO 2	PSO 3
CO1	3	3	3	3	3	2	2	2	1	3	2	2	3	2	3
CO2	3	3	3	3	2	2	2	2	1	3	2	2	3	2	3
CO3	3	3	3	3	2	2	2	2	3	3	3	2	3	3	3

1. **Visualization of Numerical Data :** Assigning data to appropriate chart elements, using glyphs, parallel coordinates, and streamgraphs, as well as implementing principles of design and color to make your visualizations more engaging and effective.
2. **Visualization of Non-Numerical Data:** Visualize graphs that depict relationships between data items and plot data using coordinates.
3. **Basic and Specialized Visualization Tools:** Learn and develop area plots, histograms, bar charts, pie charts, box plots and scatter plots and bubble plots using Matplotlib.
4. Visualization basics via linear regression graphing.
5. Visualization for time-series analysis.
6. Visualization for ranking analysis.
7. Visualization for deviation analysis.
8. Visualization for correlation analysis.
9. Visualization for multivariate analysis.

10. Visualization in R using ggplot.

11. Spatial Analysis with R.

12. The Visualization Dashboard.

Software : R, Python, Tableau

P19GE701   **English for Research Paper Writing**

2 0 0 0

Course Outcomes:

At the end of the course, the students will be able to
- Demonstrate research writing skills both for research articles and thesis
- Frame suitable title and captions as sub-headings for articles and thesis
- Write each section in a research paper and thesis coherently
- Use language appropriately and proficiently for effective written communication
- Exhibit professional proof-reading skills to make the writing error free

Unit – I   6
Planning and preparation, word order, breaking up long sentences, organising ideas into paragraphs and sentences, being concise and avoiding redundancy, ambiguity and vagueness

Unit – II   6

Interpreting research findings, understanding and avoiding plagiarism, paraphrasing sections of a paper/ abstract.

Unit- III   6

Key skills to frame a title, to draft an abstract, to give an introduction

Unit – IV   6

Skills required to organise review of literature, methods, results, discussion and conclusions

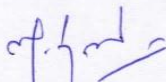**Unit – V**   6

Usage of appropriate phrases and key terms to make the writing effective - proof-reading to ensure error-free writing.

Text Books:

1. Adrian Wallwork , English for Writing Research Papers, Springer New York Dordrecht Heidelberg London, 2011
2. Highman N , Handbook of Writing for the Mathematical Sciences, SIAM. Highman's book, 1998.
3. Day R, How to Write and Publish a Scientific Paper, Cambridge University Press, 2006.
4. Goldbort R, Writing for Science, Yale University Press, 2006. (available on Google Books)

REFERENCES

Martin Cutts, Oxford Guide to Plain English, Oxford University Press, Second Edition, 2006

Total: 30 hours

Dr. M. Renuga
BoS – Chairperson,
Science & Humanities
HOD / H&L

S. No	Course Code	Course Title	Lecture	Tutorial	Practical	Credit	Total Contact Hours
Theory							
1	P19MDS512	**Professional Elective-** Internet of Things	3	0	0	3	45
2	P19MDS505	**Professional Elective-** Cloud Computing	3	0	0	3	45
3	P19WMC601	**Open Elective-** Mobile Technology and Network	3	0	0	3	45
Practical							
4	P19MDS301	Project Work Phase-1	0	0	16	8	240
				Total Credits		17	

Approved by

Chairperson, Computer Science and Engineering BOS   **Member Secretary, Academic Council**   **Chairperson, Academic Council & Principal**
Dr.B.Sathiyabhama   **Dr.R.Shivakumar**   **Dr.S.R.R.Senthil Kumar**

Copy to:-
        HOD/CSE, Third Semester M.Tech DS Students and Staff, COE

                                                          3   0   0   3   100

COURSE OUTCOMES

At the end of the course, the students will be able to

- Identify the architecture and infrastructure of cloud computing, including SaaS, PaaS, IaaS, public cloud, private cloud, hybrid cloud
- Comprehend the underlying principle of cloud virtualization, cloud storage and data management
- Analyze the trade-offs between deploying applications in the cloud and over the local infrastructure.
- Deploy applications over commercial cloud computing infrastructures such as Amazon Web Services, Windows Azure, and Google App Engine
- Implement cloud security for various real life applications

CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
COs	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)													
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	P09	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	2	3	3	2	2	1	1	2	2	1	3	2	2	3
CO2	3	2	3	3	3	3	3	2	1	3	1	3	3	3	3
CO3	3	3	3	3	2	1	2	2	2	3	1	3	3	3	3
CO4	3	3	3	3	3	2	2	1	2	3	2	3	3	3	3
CO5	3	3	2	3	3	2	3	2	3	3	3	3	3	3	3

UNIT I   CLOUD SYSTEM ARCHITECTURES AND MODELS   9

Technologies for Network-Based System – System Models for Distributed and Cloud Computing – NIST Cloud Computing Reference Architecture. Cloud Models - Characteristics – Cloud Services – Cloud computing delivery models (IaaS, PaaS, SaaS) – public, private and hybrid clouds –Cloud Solutions - Cloud ecosystem – Service management – Computing on demand.

UNIT II   VIRTUALIZATION   **9**

Basics of Virtualization - Types of Virtualization - Issues with virtualization- Implementation Levels of Virtualization - Virtualization Structures - Tools and Mechanisms - Virtualization of CPU, Memory, I/O Devices - Virtual Clusters and Resource management – Virtualization for Data-center Automation

UNIT III   CLOUD INFRASTRUCTURE   **9**

Architectural Design of Compute and Storage Clouds – Layered Cloud Architecture Development – Design Challenges - Inter Cloud Resource Management – Resource Provisioning and Platform Deployment – Global Exchange of Cloud Resources.

UNIT IV   CLOUD PROGRAMMING AND SOFTWARE ENVIRONMENTS   9

Parallel and Distributed Programming Paradigms – MapReduce , Twister and Iterative MapReduce – Hadoop Library from Apache – Mapping Applications - Programming Support - Google App Engine, Amazon AWS - Cloud Software Environments -Eucalyptus, Open Nebula, OpenStack, Aneka, CloudSim

UNIT V   CLOUD SECURITY   9

Security Overview – Cloud Security Challenges and Risks – Software-as-a-Service Security – Security Governance – Risk Management – Security Monitoring – Security Architecture Design – Data Security – Application Security – Virtual Machine Security – Identity Management and Access Control – Autonomic Security

Total: 45Hours

REFERENCE BOOKS

1. Rajkumar Buyya, J.Broberg, A. Goscinski, "Cloud Computing Principles and Paradigms", Wiley,2013
2. Sosinsky B., "Cloud Computing Bible", Wiley India Pvt Ltd, 2011
3. Resse G., "Cloud Application Architectures: Building Applications and Infrastructure in the Cloud ", First Edition ,O' Reilly.2009
4. Toby Velte, Anthony Velte, Robert Elsenpeter, "Cloud Computing, A Practical Approach", Tata Mcgraw Hill Education Private Limited, 2009
5. Shroff G., "Enterprise Cloud Computing", Cambridge University Press, 2010.
6. Tim Mather, "Cloud Security and Privacy", O'REILLY. 2009

COURSE OUTCOMES

At the end of the course, the students will be able to

- Comprehend the evolution and applications of IoT
- Identify suitable/appropriate sensors for an application and understand circuits
- Analyze different options for Embedded systems, connectivity and networking protocol and apply the appropriate one for a given application
- Write programs using Arduino and Raspberry Pi for simple applications
- Design and develop applications by considering the security challenges

COs	CO / PO, PSO Mapping (3/2/1 indicates strength of correlation) 3-Strong, 2-Medium, 1-Weak														
	Programme Outcomes (POs) and Programme Specific Outcome (PSOs)														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	P09	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	2	2	3	3	1	2	1	1	2	2	1	3	2	2	3
CO2	3	2	3	2	3	3	3	2	1	3	1	3	1	3	3
CO3	3	3	3	3	2	1	2	2	2	3	1	3	1	3	3
CO4	3	3	3	2	2	1	2	1	2	3	2	3	1	3	3
CO5	3	3	2	3	3	2	3	2	3	3	3	3	2	3	3

UNIT I   IOT – MOTIVATION AND APPLICATIONS   9

Importance of IoT. Motivating Applications of IoT:

Smart Cities- Smart Waste Management, Smart Street Lights, Smart Street Parking, Security without Surveillance, Connected Vehicles. Healthcare- Baby Monitoring, Elderly Monitoring, Mood Enhancing, Disease Treatment and Progression Monitoring, Enhance Adherence, Challenges. Agriculture- Precision Agriculture, Connected Livestock, Food Safety. Manufacturing and Logistics- Smart Manufacturing- Smart Packaging, Smart Label. Smart Electricity Grid- Managing Supply and Demand. Home Automation

UNIT II   SENSORS AND CIRCUITS   9

Sensor – Introduction, Terminology, Behavior, Selection,Circuits – Overview and Applications, Battery Issue and Energy Management, Wireless Link, Digital and Analog – Digital Computing, Analog to Digital Interfaces

UNIT III   EMBEDDED SYSTEMS, CONNECTIVITY AND NETWORKING 9

Embedded Systems – Overview, Technology Drivers, Energy, Microcontrollers, Software Connectivity and Networking – Introduction, Connectivity Challenges in IoT, Energy Harvesting Transmitters, Massive Multiple Access, Computation vs Communication

UNIT IV   ARCHITECTURE AND PROGRAMMING   9

IoT Architectures – embedded System, Gateway and Cloud (MGC) Architecture and other reference models and architectures Arduino vs Raspberry Pi vs Electric Imp – Key features and comparisons Arduino Interfaces – Arduino IDE – Programming

UNIT V   IOT CHALLENGES AND SECURITY STANDARDS   10

Technology Challenges – Security, Connectivity, Compatibility and Longevity, Standards, Intelligent Analysis and Actions .IoT Society Challenges – Privacy, Regulatory Standards -IoT security lifecycle - Cloud services and IoT – offerings related to IoT from cloud service providers – Cloud IoT security controls – An enterprise IoT cloud security architecture – New directions in cloud enabled IoT computing

Total: 45Hours

REFERENCE BOOKS

1. "The Internet of Things: Enabling Technologies, Platforms, and Use Cases", by Pethuru Raj and Anupama C. Raman (CRC Press)
2. Adrian McEwen and Hakim Cassimally, "Designing the Internet of Things", John Wiley & Sons Ltd., UK, 2014.
3. "Internet of Things: A Hands-on Approach", by ArshdeepBahga and Vijay Madisetti (Universities Press)
4. Dieter Uckelmann, Mark Harrison and Florian Michahelles, "Architecting the Internet of Things", Springer, NewYork, 2011
5. Olivier Hersent, David Boswarthick and OmarElloumi, "The Internet of Things: Key Applications and Protocols", John Wiley & Sons Ltd., UK, 2012.
6. B. Rusell and D. Van Duren, "Practical Internet of Things Security," Packet Publishing, 2016.

Sona College of Technology, Salem
(An Autonomous Institution)
Courses of Study for ME IV Semester under Regulations 2019
Computer Science and Engineering
Branch: M.Tech Data Science

S. No	Course Code	Course Title	Lecture	Tutorial	Practical	Credit	Total Contact Hours
		Practical					
1	P19MDS401	Project Work Phase – II	0	0	28	14	420
					Total Credits	**14**	

Approved by

Chairperson, Computer Science and Engineering BOS   Member Secretary, Academic Council   Chairperson, Academic Council & Principal
Dr.B.Sathiyabhama   Dr.R.Shivakumar   Dr.S.R.R.Senthil Kumar

Copy to:-
      HOD/CSE, Fourth Semester M.Tech DS Students and Staff, COE