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 2020-2021]

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
Theory						
1	P19MDS101	Statistical Computing	2	1	0	3
2	P19MDS102	Matrix and Scientific Computing	2	1	0	3
3	P19MDS103	Advanced Data Structures and Algorithms	3	0	0	3
4	P19MDS104	Data Science and Big Data Analytics	3	0	0	3
5	P19GE101	Research Methodology and IPR	2	0	0	2
6	P19GE702	**Audit Course:** Stress Management by YOGA	2	0	0	0
Practical						
7	P19MDS105	Advanced Data Structures and Algorithms Laboratory	0	0	4	2
8	P19MDS106	Big Data Management and Data Analytics Laboratory	0	0	4	2
					Total Credits	**18**

Approved by

Chairperson, Computer Science and Engineering BOS
Dr.B.Sathiyabhama

Member Secretary, Academic Council
Dr.R.Shivakumar

Chairperson, Academic Council & Principal
Dr.S.R.R.Senthil Kumar

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

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	P19MDS509	**Elective**-Natural Language Processing	3	0	0	3	45
4	P19MDS513	**Elective** -Data Warehousing and Data Mining	3	0	0	3	45
5	P19MDS514	**Elective** -Information Retrieval Techniques	3	0	0	3	45
6	P19GE701	**Audit Course**-English for Research Paper Writing	2	0	0	0	30
Practical							
7	P19MDS203	Machine Learning Laboratory	0	0	4	2	60
8	P19MDS204	Data Visualization Laboratory	0	0	4	2	60
					Total Credits	**19**	

Approved by

Chairperson, Computer Science and Engineering BOS
Dr.B.Sathiyabhama

Member Secretary, Academic Council
Dr.R.Shivakumar

Chairperson, Academic Council & Principal
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	P19MDS504	**Elective-** Deep Learning	3	0	0	3	45
2	P19MDS505	**Elective-** Cloud Computing	3	0	0	3	45
3	P19CEM601	**Open Elective-** Disaster Mitigation and Management	3	0	0	3	45
Practical							
4	P19MDS301	Project Work Phase - I	0	0	16	8	240
					Total Credits	**17**	

Approved by

Chairperson, Computer Science and Engineering BOS
Dr.B.Sathiyabhama

Member Secretary, Academic Council
Dr.R.Shivakumar

Chairperson, Academic Council & Principal
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
Dr.B.Sathiyabhama

Member Secretary, Academic Council
Dr.R.Shivakumar

Chairperson, Academic Council & Principal
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
Theory						
1	P19MDS101	Statistical Computing	2	1	0	3
2	P19MDS102	Matrix and Scientific Computing	2	1	0	3
3	P19MDS103	Advanced Data Structures and Algorithms	3	0	0	3
4	P19MDS104	Data Science and Big Data Analytics	3	0	0	3
5	P19GE101	Research Methodology and IPR	2	0	0	2
6	P19GE702	**Audit Course:** Stress Management by YOGA	2	0	0	0
Practical						
7	P19MDS105	Advanced Data Structures and Algorithms Laboratory	0	0	4	2
8	P19MDS106	Big Data Management and Data Analytics Laboratory	0	0	4	2
					Total Credits	**18**

Approved by

Chairperson, Computer Science and Engineering BOS
Dr.B.Sathiyabhama

Member Secretary, Academic Council
Dr.R.Shivakumar

Chairperson, Academic Council & Principal
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 χ2-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** Rank of a matrix – Solution of linear system of equations by matrix method, Gauss elimination and Gauss-Jordan, Gauss-Jacobi and Gauss-Seidel methods	**9**
UNIT II	**VECTOR SPACES** 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).	**9**
UNIT III	**MATRIX DECOMPOSITION** Lower-Upper (LU) decomposition – Cholesky's factorization – QR factorization – House Holder transformation – Singular value decomposition – Pseudo inverse	**9**
UNIT IV	**CURVE FITTINGS** Principle of least squares – Fitting a straight line – Fitting a parabola – Fitting an exponential curve – Fitting a curve of the form $y = ax^b$.	**9**
UNIT V	**INTERPOLATION AND APPROXIMATION** Interpolation – Newton forward and backward difference formulae – Lagrange's interpolation formula – Inverse Lagrange's interpolation formula.	**9**

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	L	T	P	C	Marks
		3	0	0	3	100

COURSE OUTCOMES

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

2. 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

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

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

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** Algorithms using Map **9**

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

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

P19MDS105	ADVANCED DATA STRUCTURES AND ALGORITHMS LABORATORY	L	T	P	C	Marks
		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

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	L	T	P	C	Marks
		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

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 files 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.
- Find 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)

Output:
- 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.

P19GE101 RESEARCH METHODOLOGY AND IPR 2 0 0 2

COURSE OUTCOMES

At the end of the course, the student 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

UNIT 1 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 2 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 3 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 4 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 5 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

THEORY: 30 Hours TUTORIAL: - PRACTICAL: - TOTAL: 30 Hours

TEXT 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.

REFERENCE BOOKS

1. Panneerselvam, R., Research Methodology, Second Edition, Prentice-Hall of India, New Delhi, 2013.
2. Ranjith Kumar, Research Methodology – A step by step Guide for Begineers, 4th edition, Sage publisher, 2014.
3. D Llewelyn & T Aplin W Cornish, "Intellectual Property: Patents, Copyright, Trade Marks and Allied Rights", Sweet and Maxwell, 1st Edition, 2016.
4. Ananth Padmanabhan, "Intellectual Property Rights-Infringement and Remedies", Lexis Nexis, 1st Edition, 2012.
5. Ramakrishna B and Anil Kumar H.S, "Fundamentals of Intellectual Property Rights: For Students, Industrialist and Patent Lawyers", Notion Press, 1st Edition, 2017.
6. M.Ashok Kumar and Mohd.Iqbal Ali :"Intellectual Property Rights" Serials Pub

P19GE702 **Stress Management by Yoga**

2000

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	P19MDS509	**Elective**-Natural Language Processing	3	0	0	3	45
4	P19MDS513	**Elective** -Data Warehousing and Data Mining	3	0	0	3	45
5	P19MDS514	**Elective** -Information Retrieval Techniques	3	0	0	3	45
6	P19GE701	**Audit Course**-English for Research Paper Writing	2	0	0	0	30
Practical							
7	P19MDS203	Machine Learning Laboratory	0	0	4	2	60
8	P19MDS204	Data Visualization Laboratory	0	0	4	2	60
					Total Credits	**19**	

Approved by

Chairperson, Computer Science and Engineering BOS
Dr.B.Sathiyabhama

Member Secretary, Academic Council
Dr.R.Shivakumar

Chairperson, Academic Council & Principal
Dr.S.R.R.Senthil Kumar

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

P19MDS201	MACHINE LEARNING	L	T	P	C	Marks
		3	0	0	3	100

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

UNIT I **9**

INTRODUCTION

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 **9**

SUPERVISED LEARNING

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 **9**

UNSUPERVISED LEARNING

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 **9**

PROBABILISTIC GRAPHICAL MODELS

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 **9**

ADVANCED LEARNING

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

P19MDS202	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

UNIT I **9**

CORE SKILLS FOR VISUAL ANALYSIS

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 **9**

TIME-SERIES, RANKING, AND DEVIATION ANALYSIS

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 **9**

DISTRIBUTION, CORRELATION, AND MULTIVARIATE ANALYSIS

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 **9**

INFORMATION DASHBOARD DESIGN

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

UNIT V **9**

DASHBOARD DESIGN MEDIA

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

COURSE OUTCOMES

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

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

Total: 60

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

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

Total: 60

P19MDS509	NATURAL LANGUAGE PROCESSING	L	T	P	C	Marks
		3	0	0	3	100

COURSE OUTCOMES

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

- Compare the various models in morphology
- Analyze the various techniques used in syntactic analysis
- Compare the various parsing techniques in context free grammar
- Analyze the various the semantic analysis techniques
- Analyze the various techniques for language generation and disclosure analysis

UNIT I

OVERVIEW AND MORPHOLOGY **9**

Introduction – Models -and Algorithms - -Regular Expressions Basic Regular Expression Patterns – Finite State Automata. Morphology -Inflectional Morphology - Derivational Morphology. Finite-State Morphological Parsing --Porter Stemmer

UNIT II

WORD LEVEL AND SYNTACTIC ANALYSIS **9**

N-grams Models of Syntax - Counting Words - Unsmoothed N-grams. Smoothing- Backoff Deleted Interpolation – Entropy - English Word Classes - Tagsets for English. Part of Speech Tagging-Rule Based Part of Speech Tagging - Stochastic Part of Speech Tagging - Transformation-Based Tagging

UNIT III

CONTEXT FREE GRAMMARS **9**

Context Free Grammars for English Syntax- Context-Free Rules and Trees. Sentence- Level Constructions– Agreement – Sub Categorization. Parsing – Top-down – Earley Parsing -feature Structures – Probabilistic Context-Free Grammars

UNIT IV

SEMANTIC ANALYSIS **9**

Representing Meaning - Meaning Structure of Language - First Order Predicate Calculus. Representing Linguistically Relevant Concepts -Syntax-Driven Semantic Analysis - Semantic Attachments -Syntax-Driven Analyzer. Robust Analysis - Lexemes and Their Senses - Internal Structure -Word Sense Disambiguation - Information Retrieval

UNIT V **9**

LANGUAGE GENERATION AND DISCOURSE ANALYSIS

Discourse -Reference Resolution - Text Coherence -Discourse Structure – Coherence. Dialog and Conversational Agents - Dialog Acts – Interpretation -Conversational Agents. Language Generation – Architecture -Surface Realizations - Discourse Planning. Machine Translation -Transfer Metaphor–Interlingua – Statistical Approaches

Total: 45

REFERENCE BOOKS

1. Gerald J. Kowalski , Mark T. Maybury , " Information Storage And Retrieval Systems Theory and Implementation", Second Edition , Kluwer Academic Publishers
2. Daniel Jurafsky and James H Martin, "Speech and Language Processing: An introduction to Natural Language Processing, Computational Linguistics and Speech Recognition", Prentice Hall, 2nd Edition, 2008
3. C. Manning and H. Schutze, "Foundations of Statistical Natural Language Processing", MIT Press. Cambridge, MA:,1999
4. Tomek Strzalkowski " Natural Language Information Retrieval ", Kluwer academic Publishers,1999

P19MDS513 DATA WAREHOUSING AND DATA MINING

L	T	P	C	Marks
3	0	0	3	100

COURSE OUTCOMES

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

- Describe the role of statistics in data mining and identify a suitable mining technique to solve the given problem
- Identify and apply various data preprocessing techniques to improve data quality
- Analyze various classifications and clustering methods
- Apply OLAP operations to query processing in data mining
- Apply various mining techniques to developing areas-Web mining, Text mining and social networks and time series data

UNIT I **9**

INTRODUCTION

Relation to Statistics, Databases- Data Mining Functionalities-Steps in Data Mining Process-Architecture of a Typical Data Mining Systems- Classification of Data Mining Systems - Overview of Data Mining Techniques-Issues

UNIT II **9**

DATA PREPROCESSING AND ASSOCIATION RULES

Data Preprocessing-Data Cleaning, Integration, Transformation, Reduction, and Discretization Concept Hierarchies- Concept Description: Data Generalization And Summarization Based Characterization- Mining Association Rules In Large Databases – Analysis of Attribute Relevance- Exploratory Data Analysis Using tools (Python, Weka and R).

UNIT III **9**

PREDICTIVE MODELING

Classification and Prediction: Issues Regarding Classification and Prediction-Classification By Decision Tree Induction-Bayesian Classification-Classification by Back Propagation - Other Classification Methods-Prediction- Clusters Analysis: Types Of Data In Cluster Analysis- Categorization Of Major Clustering Methods: Partitioning Methods –Hierarchical Methods – Density Based Methods – Grid Based – Model Based – Outlier Analysis- Case Studies using tool (Python, Weka & R).

UNIT IV **9**

DATA WAREHOUSING

Data Warehousing Components -Multi Dimensional Data Model- Data Warehouse Architecture-Data Warehouse Implementation- -Mapping the Data Warehouse to Multiprocessor Architecture- OLAP Need-Categorization of OLAP Tools – OLAP Operations in Multidimensional Data Model

UNIT V **9**

APPLICATIONS

Recent trends in Distributed Warehousing and Data Mining, Class Imbalance Problem; Graph Mining; Social Network Analysis-Web Mining- Sentimental Analysis-Mining Multimedia data on the Web, Automatic classification of Web documents- Mining Time Series data and Sequential Pattern Mining

Total: 45

REFERENCE BOOKS

1. Jiawei Han, Micheline Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, 2002
2. Alex Berson,Stephen J. Smith, “Data Warehousing, Data Mining,& OLAP”, Tata Mcgraw- Hill, 2004.
3. Usama M.Fayyad, Gregory Piatetsky - Shapiro, Padhrai Smyth And Ramasamy Uthurusamy, "Advances In Knowledge Discovery And Data Mining", The M.I.T Press, 1996.
4. Ralph Kimball, "The Data Warehouse Life Cycle Toolkit", John Wiley & Sons Inc., 1998.
5. Sean Kelly, "Data Warehousing In Action", John Wiley & Sons Inc., 1997
6. Vipin Kumar, Introduction to Data Mining - Pang-Ning Tan, Michael Steinbach, Addison Wesley, 2006.
7. Jure Leskovec, Anand Rajaraman and Jeffrey D. Ullman, ”Mining of Massive Datasets”, Cambridge University Press, Second Edition,2016
8. G Dong and J Pei, Sequence Data Mining, Springer, 2007.

P19MDS514 INFORMATION RETRIEVAL TECHNIQUES

L	T	P	C	Marks
3	0	0	3	100

COURSE OUTCOMES

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

- Describe the concepts of Information Retrieval system
- Analyze various models of retrieval methods
- Identify and design the various components of an Information Retrieval system
- Apply machine learning techniques to text classification and clustering which is used for efficient Information Retrieval
- Design an efficient search engine and analyze the Web content structure

UNIT I **9**

INTRODUCTION: MOTIVATION

Basic Concepts – Practical Issues - Retrieval Process – Architecture - Boolean Retrieval – Retrieval Evaluation – Open Source IR Systems–History of Web Search – Web Characteristics– The impact of the web on IR —IR Versus Web Search–Components of a Search engine

UNIT II **9**

MODELING

Taxonomy and Characterization of IR Models – Boolean Model – Vector Model - Term Weighting – Scoring and Ranking –Language Models – Set Theoretic Models - Probabilistic Models – Algebraic Models – Structured Text Retrieval Models – Models for Browsing

UNIT III **9**

INDEXING

Static and Dynamic Inverted Indices – Index Construction and Index Compression. Searching - Sequential Searching and Pattern Matching. Query Operations -Query Languages – Query Processing - Relevance Feedback and Query Expansion - Automatic Local and Global Analysis – Measuring Effectiveness and Efficiency

UNIT IV **9**

CLASSIFICATION AND CLUSTERING

Text Classification and Naïve Bayes – Vector Space Classification – Support vector machines and Machine learning on documents. Flat Clustering – Hierarchical Clustering –Matrix decompositions and latent semantic indexing – Fusion and Meta learning

UNIT V **9**

SEARCHING THE WEB

Searching the Web –Structure of the Web –IR and web search – Static and Dynamic Ranking – Web Crawling and Indexing – Link Analysis - XML Retrieval Multimedia IR: Models and Languages – Indexing and Searching Parallel and Distributed IR – Digital Libraries

Total: 45

REFERENCES

1. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schutze, ―Introduction to Information Retrieval‖, Cambridge University Press, First South Asian Edition, 2008
2. Implementing and Evaluating Search Engines‖, The MIT Press, Cambridge, Massachusetts London, England, 2010
3. Ricardo Baeza – Yates, Berthier Ribeiro – Neto, ―Modern Information Retrieval: The concepts and Technology behind Search‖ (ACM Press Books), Second Edition, 2011
4. Stefan Buttcher, Charles L. A. Clarke, Gordon V. Cormack, ―Information Retrieval

Audit Course

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

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	P19MDS504	**Elective-** Deep Learning	3	0	0	3	45
2	P19MDS505	**Elective-** Cloud Computing	3	0	0	3	45
3	P19CEM601	**Open Elective-** Disaster Mitigation and Management	3	0	0	3	45
Practical							
4	P19MDS301	Project Work Phase - I	0	0	16	8	240
					Total Credits	**17**	

Approved by

Chairperson, Computer Science and Engineering BOS
Dr.B.Sathiyabhama

Member Secretary, Academic Council
Dr.R.Shivakumar

Chairperson, Academic Council & Principal
Dr.S.R.R.Senthil Kumar

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

P19MDS504 **DEEP LEARNING** **L T P C Marks**

3 0 0 3 100

COURSE OUTCOMES

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

- Describe the basics of deep learning
- Implement various deep learning network models
- Realign high dimensional data using reduction techniques
- Analyze optimization and generalization in deep learning
- Explore the deep learning 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	2	3	1	2	1	3	1	2	2	2	1	2	2	2	3
CO2	3	3	3	3	3	2	2	3	3	3	2	3	3	3	3
CO3	3	3	3	3	3	2	3	3	3	3	3	3	3	3	3
CO4	3	3	3	3	3	2	2	2	3	3	3	3	3	3	3
CO5	3	3	3	3	3	3	2	3	3	3	3	3	3	3	3

UNIT I INTRODUCTION 9

Biological Neuron- Idea of computational units-McCulloch–Pitts unit and Thresholding logic-Linear Perceptron- Perceptron Learning Algorithm- Linear separability- Convergence theorem for Perceptron Learning Algorithm-Multilayer Perceptron- Gradient Descent-Backpropagation-Vanishing Gradient-Activation Functions.

UNIT II DEEP NETWORKS 9

Difficulty of training deep neural networks-Greedy layerwise training-Better Training of Neural Networks: Newer optimization methods for neural networks :Adagrad, adadelta, rmsprop, adam, NAG- second order methods for training- Saddle point problem in neural networks- Regularization methods :dropout, drop connect, batch normalization. - Learning Vectorial Representations Of Words

UNIT III DEEP NETWORK ARCHITECTURES 9

Recurrent Neural Networks: Back propagation through time-Problem of Exploding Gradient and Vanishing Gradient-Long Short Term Memory- Gated Recurrent Units- Bidirectional LSTMs- Bidirectional RNNs

Convolutional Neural Networks: Architecture Overview-ConvNet Layers: Convolutional Layer, Pooling Layer, Normalization Layer, Fully Connected Layer, Converting fully connected layer to Convolutional Layer - Case Studies: LeNet, AlexNet.

UNIT IV GENERATIVE MODELS 9

Restrictive Boltzmann Machines (RBMs)- Introduction to MCMC and Gibbs Sampling- gradient computations in RBMs- Deep Boltzmann Machines .

Recent trends: Variational Auto encoders - Generative Adversarial Networks- Multi-task Deep Learning - Multi-view Deep Learning

UNIT V TOOLS AND APPLICATIONS 9

Introduction to Keras and Tensorflow-Deep learning for computer vision, Deep Learning Applications at the Enterprise Scale, Deep Learning Models for Healthcare Applications- Semantic parsing of Speech using Recurrent Net- LSTM network for sentiment analysis

TEXT BOOKS

1. Deep Learning, Ian Goodfellow and Yoshua Bengio and Aaron Courville, MIT Press, 2016

REFERENCE BOOKS

1. Cosma Rohilla Shalizi, Advanced Data Analysis from an Elementary Point of View, 2015
2. Deng & Yu, Deep Learning: Methods and Applications, Now Publishers, 2013
3. Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, MIT Press, 2016
4. Michael Nielsen, Neural Networks and Deep Learning, Determination Press, 2015.es of deep learning techniques

P19MDS505 CLOUD COMPUTING

L	T	P	C	Marks
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	2	2	1	2	3	2	1	3	2	2	3	2	2	2	3
CO2	3	3	3	3	3	2	2	3	3	3	3	3	3	3	3
CO3	3	3	2	3	3	2	3	2	3	3	3	3	3	2	3
CO4	3	3	3	3	3	2	3	3	3	3	3	3	3	3	3
CO5	2	3	3	3	3	2	3	3	3	3	3	3	3	3	2

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: 45

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

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
Dr.B.Sathiyabhama

Member Secretary, Academic Council
Dr.R.Shivakumar

Chairperson, Academic Council & Principal
Dr.S.R.R.Senthil Kumar

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