Institute of Engineering & Technology

Sitapur Road, Lucknow

Evaluation Scheme & Syllabus

for

M. Tech.

(ARTIFICIAL INTELLIGENCE & DATA SCIENCE)

Syllabus-Ist Year

(Effectiv from the Session: 2023-24 onwards)

M. Tech. (Artificial Intelligence and Data Science)

Objectives

M. Tech in Artificial Intelligence and Data Science program provides students with a state-of-the-art educational experience to become outstanding, socially responsible engineers of AI systems. Students will learn the theoretical foundations and gain practical skills to manage and engineer various data forms, to make predictive analyses, to derive actionable insights, and develop intelligent hardware and software systems and related cutting-edge AI technologies. This program also provides students with foundational knowledge, experience, and skills that are crucial for a data science career. It features a multidisciplinary curriculum in various application domain of data science such as data Science in health care, social networks, public policy and finance through applied math, statistics and machine learning. Its rigorous, balanced coursework provides students with strong mathematical, computing, and data skills, as well as project experience.

Graduates of this program will have the necessary theoretical knowledge, practical data analysis and coding skills for solving complex real-world problems based on massive amounts of data and also meet the qualifications for various data scientist positions in industry and government. They will graduate with necessary training in applied research in the area.

Program Learning Outcomes

1. Identify and apply fundamental concepts, algorithms, and methodologies to formulate, analyze, design, and evaluate artificial intelligence and machine learning systems.
2. Apply computer science knowledge and tools to assist in performing data science tasks. Summarize and evaluate statistical and machine learning concepts, models, and techniques.
3. Tackle complex machine learning and artificial intelligence problems using contemporary principles, algorithms, technologies, methodologies, and tools.
4. Participate in a team to develop AI and machine learning applications; while recognizing the ethical, economic and environmental implications of their work.
5. Sustain a process of life-long learning in engineering or other professional areas; and Identify ways in which data scientists can contribute to the cultural and economic well-being of diverse societies in local, national, and global scopes.

Program Outcomes (POs) and Program Specific Outcomes (PSO)

Program Outcomes (POs)

PO1: Independently carry out research /investigation and development work to solve practical problems related to Artificial intelligence and Data science.

PO2: Write and present a substantial technical report/document in the area of Artificial intelligence and Data science.

PO3: Demonstrate a degree of mastery over the area related to Artificial intelligence, Machine learning and Data Science.

Program Specific Outcomes (PSO)

PSO1: Analyze and solve the real world problems by applying the knowledge of artificial intelligence, machine learning and data Science.

Credit Summary for M.Tech. (AI & DS) Curriculum

Semester	Subjects	Credits	Marks
I	Core Subjects-02 (4 credit each) Research Methodology (4 credit) Elective Subject-01 (4 credits) Labs-02 (2 credit each)	20	600
II	Core Subjects-02 (4 credit each) Elective Subject-02 (4 credit each) Lab/Mini Project-01 (2 credit) Seminar-01 (2 credits)	20	600
III	Elective Subject-01 (4 credit) Seminar-01 (2 credit) Dissertation-I (10 credit)	16	500
IV	Dissertation-II (16 credit)	16	500
Total		**72**	**2200**

Semester– I

S. No.	Subjects Code	Subject	Period			Credit	Evaluation Scheme					Subject Total
							Theory			Practical		
			L	T	P		CT	TA	ESE	TA	ESE	
1	MADS-101	Mathematical Foundation of Data Science	4	0	0	4	20	10	70			100
2	MADS-102	Advanced Data Structures and Algorithms	4	0	0	4	20	10	70			100
3	MADS(011-017)*	Departmental Elective-I	4	0	0	4	20	10	70			100
5	MARM-101	Research Process and Methodology	4	0	0	4	20	10	70			100
5	MADS-151	Data Science Lab	0	0	3	2				50	50	100
6	MADS-152	Advanced Data Structure and Algorithms Lab	0	0	3	2				50	50	100
	TOTAL					20						600

Semester– II

S. No.	Subjects Code	Subject	Period			Credit	Evaluation Scheme					Subject Total
							Theory			Practical		
			L	T	P		CT	TA	ESE	TA	ESE	
1	MADS-201	Artificial Intelligence and Knowledge Discovery	4	0	0	4	20	10	70			100
2	MADS-202	Advanced Database Management System	4	0	0	4	20	10	70			100
3	MADS(021-025)*	Departmental Elective-II	4	0	0	4	20	10	70			100
4	MADS(031-034)*	Departmental Elective-III	4	0	0	4	20	10	70			100
5	MADS-251	Artificial Intelligence and Knowledge Discovery Lab	0	0	3	2				50	50	100
6	MADS-252	Seminar-I	0	0	3	2				100		100
	TOTAL					20						600

List of Departmental Elective

Departmental Elective-I*

Subject Code	Subject Name
MADS-011	Data Warehousing and Mining
MADS-012	Data Science Theory & Practices
MADS-013	Optimization Theory, Algorithms & Techniques
MADS-014	Soft Computing
MADS-015	Applied Matrix Analysis
MADS-016	Big Data Management
MADS-017 (added)	Distributed Computing
MADS-01X*	--------

Departmental Elective-II*

Subject Code	Subject Name
MADS-021	Data Analytics and Visualization
MADS-022	Deep Learning
MADS-023	Cloud Computing
MADS-024	Web and Big Data Mining
MADS-025 (added)	Natural language processing
MADS-02X*	-----

Departmental Elective-III*

Subject Code	Subject Name
MADS-031	Pattern Recognition Theory & Applications
MADS-032	Recommender Systems
MADS-033	Probability and Stochastic Process
MADS-034	Advanced Machine Learning
MADS-03X*	----

(*) The list of above Departmental Elective courses is open to include other relevant courses as per the expertise available at the commencement of the semester.

M. Tech. (AI & DS)

Semester-I

SYLLABUS

MADS-101 MATHEMATICAL FOUNDATION OF DATA SCIENCE		
Course Outcome (CO)		**Bloom's Level (KL)**
At the end of course, the student will be able to:		
CO1	Ability to understand the basic mathematical concept in the field of data science.	K2, K3, K4
CO2	Will be able to understand linear algebra and solve the system of equation using numerical methods, develop understanding of vectors space.	K2, K3, K4
CO3	Will have understanding of probability theory, random processes and discrete and continuous probability distributions and be able to apply to the problems.	K2, K3
CO4	Will be able to solve problems related to testing of hypotheses.	K2, K3
CO5	Will develop understanding of group theory, important theorems and be able to apply Fermat's theorem and Chinese remainder theorem in solving the problems.	K2, K3, K4
DETAILED SYLLABUS		
Unit	**Topics**	**Lectures**
I	Basics of Data Science: Introduction, Typology of problems, Importance of linear algebra, Fields, Vector-spaces, sub-spaces, linear-combination, linear-dependence and independence. Basis, dimensions, spanning sets. Statistics and optimization from a data science perspective; Structured thinking for solving data science problems.	**08**

II	Linear-transformation, definition and examples, matrix representation, similarity, range and kernel, rank-nullity theorem and its consequences, Singular and non singular linear transformations, Inner product spaces, definition and examples, Distance measures; Projections; Notion of hyper planes; half-planes. System of Linear algebraic equation: Gaussian Elimination, Cholesky Decomposition, Ill-conditioned equations, Iterative Methods, Convergence, Eigen Value problems, QR decomposition,	**08**
III	Probability, Statistics and Random Processes: Frequency Distribution Measures of central tendency, Probability theory, axioms, conditional probability, Bayes Theorem, Bernoulli Trials, Random variables, Probability distributions with discrete and continuous random variables:, geometric, Poisson, Normal, Exponential and Gaussian distributions, Expectations and Moments	**08**
IV	Sampling & Estimation Theory**:** Sampling and Testing of Hypothesis: Introduction To testing ofhypothesis - Tests of significance for large samples – t, F and Chi-square tests; Estimation Theory, goodness of fit,Confidence interval, Correlation Coefficient	**08**
V	Number Theory: Euclidean Algorithm, Cyclic Group Subgroup, Generators, Lagrange Theorem, Euler's Totient Function, Fermat's theorem, Chinese Remainder Theorem, Discrete Logarithm Problem	**08**

Textbook & References:

1. Gilbert Stang: Introduction to Linear Algebra.
2. Jain M.K.,Iyengar S.R.K.& Jain R.K.,Jain and Iyengar: Numerical Methods.
3. Johnson R.A.: Probability and Statistics for Engineers.
4. Papoulis A., Unnikrishna Pillai S.: Probability, random variable and stochastic Processes.

MADS-102 ADVANCED DATA STRUCTURES & ALGORITHMS		
Course Outcome (CO)		**Bloom's Level (KL)**
At the end of course, the student will be able to understand		
CO1	Understand concepts of advanced data structures.	K1, K2
CO2	Understand and apply divide and conquer algorithms, greedy methods, dynamic programming and backtracking approach for problem solving.	K2, K3
CO3	Discuss and understand genetic algorithms and artificial neural networks.	K1, K2
CO4	Understand and analyze complexity theory of algorithms.	K1, K2, K4
CO5	Understand Randomized and Approximation Algorithms and solve real world problems.	K1, K2, K5
DETAILED SYLLABUS		
Unit	**Topics**	**Lectures**
I	Advanced Data Structure**:** Array, Linked-List, Stack and Queue ADTs, Priority queue (abstract data type) ADT, implementation using Heaps Hashing-Hash functions, Collision Resolution methods-Open Addressing, Chaining. Heaps, Binary tree ADT, representations, recursive and non-recursive traversals, Graph ADT, representations, graph traversals/search methods.	**08**
II	Abstract Algorithms**:** Divide and Conquer, Greedy methods Dynamic Programming, Backtracking, Branch and Bound method, Elementary Graph Algorithms.	**08**
III	Complexity Theory: Introduction to Complexity theory,Diagonalization Theorem, Tractable and Non-Tractable problems; Few NP and NPC problems	**08**
IV	Advanced Topics**:** Randomized and Approximation Algorithms , Introduction to Computational Geometry	**08**
V	Natural Algorithms**:** Genetic Algorithms, Simulated Annealing, Artificial Neural Networks, Markov chains and random walks, Particle Swarm optimization (PSO), Ant Colony optimization	**08**

Textbook & References:

1. E. Horowitz, S.Sahni and Dinesh Mehta, "Fundamentals of Data structures in C++", University Press, 2007
2. E. Horowitz, S. Sahni and S. Rajasekaran, "Computer Algorithms/C++", Second Edition, University Press, 2007
3. Introduction to Algorithms by Cormen, Licerson, Steina and Rivest
4. Aho, Hopcroft and Ullman. "Design and Analysis of Algorithm".
5. The Design of Approximation Algorithms by Williamson and Shmoys.
6. Approximation Algorithms by Vazirani.
7. Randomized Algorithms by Motwania and Raghavan.

MADS-151 DATA SCIENCE LAB	
S. No.	**List of Experiment/Assignment**
1.	Write a program to demonstrate the basic applications of working using numpy arrays, panda data frames and plot using matplotlib.
2.	Write a program for Frequency distribution, Variability and Averages.
3.	Develop a python program for Normal Curves.
4.	Write a program for Correlation, Correlation coefficient and scatter plots
5.	Develop a python program for Simple Linear regression.
6.	Considering "US Crime Dataset in "MASS " package in R, use t-test to analyze the difference between southern group (So=1) Andno southern group (So=0) on the probability of imprisonment (Probe). Discuss hypothesis and results.
7.	Considering "Loblolly" dataset in R determine mean, median, quartiles of age and height. Determine the correlation between age and height. Also apply the chi-square test to determine whether there is an association between age and height of a tree. hypothesis and result.
8.	Considering the "cholesterol" dataset in the "multicomp" package in R, use ANOVA to analyze the difference between trt and response. Discuss the hypothesis and result
9.	Write a program to implement k-Nearest Neighbour algorithm to classify the iris data set. Print both correct and wrong predictions.

1. The instructor cans add/modify/tune experiments, wherever he/she feels in a justified manner.
2. The language/tools/environment to conduct the lab are R, Python, MATLAB/JAVA.

MADS-152 ADVANCED DATA STRUCTUE & ALOGIRTHM DESIGN LAB	
S.No.	**List of Experiment / Assignment**
1.	Implement dependent task scheduling via Knapsack / TSP / Genetic Algorithm. Discuss the efficiency of the implementation.
2.	Write a program to implement flow Min Cut problem using the Ford Fulkerson algorithm. The implementation must exhibit limitations of the problem. Also should explain whether the algorithm works for bipartite matching or not.
3.	Write a program to implement string matching using Rabin Karp algorithm, Knuth Morris algorithm, and Boyer Moore algorithm
4.	Write a program to find all pair shortest path using Johnson's algorithm for sparse graphs.
5.	Implementation of randomized quick sort algorithm. Also discuss its complexity in all cases.
6.	Implementation of Edmonds-Karp algorithm. Discuss its complexity.
7.	Write a program to implement Travelling salesperson using branch and bound, dynamic programming & genetic algorithm.
8.	Write a program to implement Graph Coloring using a backtracking method. Discuss its complexity.
9.	Write a program to implement Simple optimization problem using soft computing approaches (GA,PSO, ACO & NN).

1. The instructors can add/modify/tune experiments, wherever he/she feels in justified manner.
2. The language/tools/environment to conduct the lab are R, Python, MATLAB/JAVA.

Departmental Elective-I

MADS-011 DATA WAREHOUSING AND MINING		
Course Outcome (CO)		**Bloom's Level (KL)**
At the end of course, the student will be able to understand		
CO1	To understand the principles of data warehousing and data mining.	K1, K2
CO2	To be familiar with the data warehouse architecture and its implementation.	K2, K3
CO3	To know the architecture of a data mining system	K1, K2
CO4	To understand the various data preprocessing methods.	K1, K2, K4
CO5	To perform classification and prediction of data	K1, K2, K5

SYLLABUS		
Unit	**Topics**	**Lectures**
I	Data Warehousing and Business Analysis: - Data warehousing Components – Building a Data warehouse –Data Warehouse Architecture – DBMS Schemas for Decision Support – Data Extraction, Cleanup, and Transformation Tools – Metadata – reporting – Query tools and Applications – Online Analytical Processing (OLAP) – OLAP and Multidimensional Data Analysis.	**08**
II	Data Mining: - Data Mining Functionalities – Data Preprocessing – Data Cleaning – Data Integration and Transformation – Data Reduction – Data Discretization and Concept Hierarchy Generation- Architecture Of A Typical Data Mining Systems- Classification Of Data Mining Systems. Association Rule Mining: - Efficient and Scalable Frequent Item set Mining Methods – Mining Various Kinds of Association Rules – Association Mining to Correlation Analysis – Constraint-Based Association Mining.	**08**
III	Classification and Prediction: - Issues Regarding Classification and Prediction – Classification by Decision Tree Introduction – Bayesian Classification – Rule Based Classification – Classification by Back propagation – Support Vector Machines – Associative Classification – Lazy Learners – Other Classification Methods – Prediction – Accuracy and Error Measures – Evaluating the Accuracy of a Classifier or Predictor – Ensemble Methods – Model Section.	**08**
IV	Cluster Analysis: - Types of Data in Cluster Analysis – A Categorization of Major Clustering Methods – Partitioning Methods – Hierarchical methods – Density-Based Methods – Grid-Based Methods – Model-Based Clustering	**08**

	Methods – Clustering High-Dimensional Data – Constraint-Based Cluster.	
V	Mining Object, Spatial, Multimedia, Text and Web Data: Multidimensional Analysis and Descriptive Mining of Complex Data Objects – Spatial Data Mining – Multimedia Data Mining – Text Mining – Mining the World Wide Web	**08**

Textbook & References:

1. Alex Berson and Stephen J. Smith "Data Warehousing, Data Mining & OLAP", Tata McGraw – Hill Edition, Tenth Reprint 2007.
2. K.P. Soman, Shyam Diwakar and V. Ajay "Insight into Data mining Theory and Practice", Easter Economy Edition, Prentice Hall of India, 2006.
3. G. K. Gupta "Introduction to Data Mining with Case Studies", Easter Economy Edition, Prentice Hall of India, 2006.
4. Pang-Ning Tan, Michael Steinbach and Vipin Kumar "Introduction to Data Mining", Pearson Education, 2007.

MADS-012 DATA SCIENCE THEORY AND PRATCTICES		
Course Outcome (CO)		**Bloom's Level (KL)**
At the end of course, the student will be able to understand		
CO1	Understand and apply the, random number generation and density estimations to perform analysis of various kinds of data.	K1, K2
CO2	Understand and apply the basic concept of python.	K2, K3
CO3	Understand and apply the various method of optimization.	K1, K2
CO4	Understand and apply data processing and data validation.	K1, K2, K4
CO5	Perform statistical analysis on variety of data by using various libraries.	K1, K2, K5

SYLLABUS

Unit	Topics	Lectures
I	Background of Data Science, Selecting rows/observations, Rounding Number , Selecting columns/fields, Merging data, Data aggregation, Matrix Algebra and Random Vectors, Multivariate Analysis , Time Series Forecasting, Introduction to R.	**08**
II	Python basics, Strings, Lists, Tuples, Sets, Dictionaries, Functions, Classes and objects, File input/output.	**08**
III	Fundamentals of Optimisation: Single-variable and multi-variable optimisation - Optimality Criteria, Non-gradient and Gradient-based methods in single variable, Contour plots, Unidirectional search, Gradient-based approaches in multi-variable, Constrained Optimisation, Implementation of Optimisation.	**08**
IV	Data Preprocessing and Data Validations, Data Visualizations, Discriminant and Cluster Analysis, Multidimensional Scaling, Linear / Polynomial Regression, Logistic Regression (Classification), Regularization, Support vector machines, Clustering, Dimensionality reduction, Manifold learning, 2D/3D Convolution, Introduction to Neural Networks, Evaluation Metrics.	**08**
V	Fourier Analysis, Wavelets Analysis, Stochastic Analysis andpython libraries - Numpy (ndarray, indexing, slicing and other functions), Matplotlib, Pandas, Scikit-Learn, Implementation of ML.	**08**

Textbooks & References:

1. Let us Python, Yashavant Kanetkar and Aditya Kanetkar, First Edition, 2019, BPB Publications.
2. Optimization for Engineering Design, - Algorithms and Examples, Kalyanmoy Deb, Second Edition, 2016, PHI.
3. Machine Learning Refined: Foundations, Algorithms, and Applications, Jeremy Watt, Reza Borhani, Aggelos K. Katsaggelos, Cambridge University Press, 2nd Edition, 2020.
4. https://matplotlib.org/tutorials/introductory/pyplot.html.
5. Data science from scratch - First principles with Python, Joel Grus, OâReily, 2015.
6. Machine Learning, Tom Mitchell, McGrawhill, 1997

MADS-013
OPTIMIZATION THEORY, ALGORITHMS AND TECHNIQUES

Course Outcome (CO)		Bloom's Leve l (KL)
At the end of course, the student will be able to understand		
CO1	Discuss and explain various concepts of optimization methods.	K1, K2
CO2	Explain linear programming and analyses problems.	K1, K2, K3
CO3	Discuss and explain transportation and transshipment problem and assignment problem.	K1, K2
CO4	Discuss and analyze inventory models and dynamic programming problems.	K1, K2, K3
CO5	Discuss and apply Queuing theory and integer programming.	K2, K3, K5

DETAILED SYLLABUS

Unit	Topics	Lectures
I	Introduction to optimization techniques: Nature, scope, applications, Historical development, Role Indecision making. Introduction to linear programming: Graphical approach with algorithm, Simple Method with algorithm, Duality, Post optimal analysis.	**08**
II	Linear Programming: Transportation Problems With Algorithms-North West Method, Column minima method, Row Minima method, Matrix minima method, Vogel's Approximation method, Modified method, Transshipment Problem.	**08**
III	Assignment Problems: Hungarian algorithm and its application. Dynamic Programming: Bellman's principle of optimality of Dynamic Programming, Application.	**08**
IV	Inventory Models: Various Models, Applications, Economic Lot Sizes-price Replacement Problems: Concept, Applications.	**08**
V	Queuing Theory: Different models and their solution. Integer Programming: Concept, Applications, cutting plane method, Branch And Bound Methods.	**08**

Textbooks & References:

1. kalyanmoy Deb, "Optimization for Engineering Design Algorithms and Examples", Prentice Hall of India, New Delhi, 2004.
2. Edwin K. P. Chong and Stanislaw H. Zak, "An Introduction to Optimization", Second Edition, Wiley-Interscience Series in Discrete Mathematics and Optimization, 2004.
3. M. Asghar Bhatti, "Practical Optimization Methods: with Mathematics Applications", Springer Verlag Publishers, 2000.

MADS-014
SOFT COMPUTING

Course Outcome (CO)		Bloom's Level (KL)
At the end of course, the student will be able to understand		
CO1	Understand and apply fuzzy logic and its applications.	K1, K2
CO2	Understand artificial neural networks and its applications.	K2, K3
CO3	Apply and solving multi-objective optimization problems using evolutionary algorithms (MOEAs).	K1, K2
CO4	Applications of Soft computing to solve problems in varieties of application domains.	K1, K2, K4
CO5	Solving single-objective optimization problems using GAs.	K1, K2, K5

DETAILED SYLLABUS

Unit	Topics	Lectures
I	Introduction to Soft Computing: Concept of computing systems."Soft" computing versus "Hard" computing Characteristics of Soft computing Some applications of Soft computing techniques	**08**
II	Fuzzy logic: Introduction to Fuzzy logic: Fuzzy sets and membership functions. Operations on Fuzzy sets. Fuzzy relations, rules, propositions, implications and inferences. Defuzzification techniques. Fuzzy logic controller design. Some applications of Fuzzy logic.	**08**
III	Genetic Algorithms: Concept of "Genetics" and "Evolution" and its application to proablistic search techniques, Basic GA framework and different GA architectures. GA operators: Encoding, Crossover, Selection, Mutation, etc. Solving single-objective optimization problems using GAs.	**08**
IV	Multi-objective Optimization Problem Solving: Concept of multi-objective optimization problems (MOOPs) and issues of solving them. Multi-Objective Evolutionary Algorithm (MOEA). Non-Pareto approaches to solve MOOPs Pareto-based approaches to solve MOOPs. Some applications with MOEAs.	**08**
V	Artificila Neural Networks :Biological neurons and its working. Simulation of biolgical neurons to problem solving. Different ANNs architectures. Trainging techniques for ANNs. Applications of ANNs to solve some real life problems.	**08**

Textbooks & References:

1. Fuzzy Logic: A Pratical approach, F. Martin, Mcneill, and Ellen Thro, AP Professional, 2000.
2. Fuzzy Logic with Engineering Applications (3rd Edn.), Timothy J. Ross, Willey, 2010.
3. Foundations of Neural Networks, Fuzzy Systems, and Knowldge Engineering, Nikola K. Kasabov, MIT Press, 1998.
4. Fuzzy Logic for Embedded Systems Applications, Ahmed M. Ibrahim, Elesvier Press, 2004.
5. An Introduction to Genetic Algorithms, Melanie Mitchell, MIT Press, 2000.
6. Genetic Algorithms In Search, Optimization And Machine Learning, David E. Goldberg, Pearson Education, 2002.

MADS-015 APPLIED MATRIX ANALYSIS		
Course Outcome (CO)		**Bloom's Level (KL)**
At the end of course, the student will be able to understand		
CO1	Understand and describe different types of matrix and vector norms, and calculate or estimate.	K1, K2
CO2	Apply and analyse iterative algorithms for calculating eigenvalues and eigenvectors	K2, K3
CO3	Describe the properties of quadratic forms, projections, spectral theory and their use in quadratic optimization.	K1, K2
CO4	Analyse the matrix computations in geometrical terms of linear spaces, linear transformations and symmetries	K1, K2, K4
CO5	Describe in detail a freely chosen application area of matrix analysis	K1, K2, K5
DETAILED SYLLABUS		
Unit	**Topics**	**Lectures**
I	Elementary row operations, Echelon form of a matrix, Rank of a matrix, System of Linear Equations, Vector spaces and their properties, subspaces, basis and dimension, linear transformations	**08**
II	Eigen values and eigen vectors, Calyey Haminton theorem, diagonalization. Week Special matrices, Gerschgorin theorem, inner product spaces, matrix norms and Gram Schmidt Process.	**08**
III	Normal and Positive Definite matrices, Quadratic forms with applications , Evaluation of matrix functions, SVD and its application.	**08**
IV	Stationary and non-stationary iterative methods for linear system, Krylov subspace methods, analysis of positive and non-negative matrices, polar decomposition theorem	**08**
V	Matrix computations in terms of linear transformations and symmetries; applications of matrix analysis in engineering and Data Science.	**08**

Textbooks & References:

1. Hoffman, K. and Kunze, R., "Linear Algebra", 2nd edition, Pearson Education (Asia) Pvt. Ltd/ Prentice Hall of India (2004)
2. Leon, S.J., "Linear Algebra with Applications", 8th Edition, Pearson (2009)
3. Peter, J. Olever and Shakiban, C., "Applied Linear Algebra", 1st Edition , Prentice Hall (2005)
4. Strang, G., "Linear Algebra and its Applications", 3rd edition, Thomson Learning Asia Pvt Ltd (2003)

MADS-016
BIG DATA MANAGEMENT

Course Outcome (CO)		Bloom's Level (KL)
At the end of course, the student will be able to understand		
CO1	Understand Big data and its analytics in the real world.	K1, K2
CO2	Design of Algorithms to solve data intensive problems using map reduce paradigm.	K2, K3, K4
CO3	Analyze the big data framework like hadoop and HDFS to efficiently store and process big data to generate analytics.	K1, K2, K3
CO4	To study scientific and information visualization techniques with a focus on data compression, statistical and information theory technique.	K1, K2
CO5	To learn about extreme scale visual computing and state-of-the-art in data analysis techniques and conclude by discussing the future paradigms of the big data in Machine learning.	K1, K2, K3

DETAILED SYLLABUS

Unit	Topics	Lectures
I	Introduction to Big Data, Different types of data, Big data and its characteristics Big Data Architecture.	**08**
II	Integrated with Big Data Capabilities: storage management, processing, data integration, statistical analysis, large file system, distributed file system, and Map Reduce	**08**
III	HDFS and Hadoop, Data analysis using statistical methods, Data compression and visualization.	**08**
IV	Visual perception, Information analysis and visual variables, Data and task abstraction Computation predictive analysis of data, Mining of Big data.	**08**
V	Machine/Deep learning techniques for big data, Data exploration at extreme-scale, Future paradigm in extreme-scale data visualization	**08**

Textbooks & References:

1. Rajaraman, A., Ullman, J. D., Mining of Massive Datasets, Cambridge University Press, United Kingdom, 2012
2. Berman, J.J., Principles of Big Data: Preparing, Sharing and Analyzing Complex Information,Morgan Kaufmann, 2014
3. Barlow, M., Real-Time Big Data Analytics: Emerging Architecture, O Reilly, 2013
4. Schonberger, V.M. , Kenneth Cukier, K., Big Data, John Murray Publishers, 2013

Distributed Computing (MADS-017)		
Course Outcome (CO)		**Bloom's Knowledge Level (KL)**
At the end of course , the student will be able to understand		
CO 1	To provide hardware and software issues in modern distributed systems.	K1 , K2
CO 2	To get knowledge in distributed architecture, naming, synchronization, consistency and replication, fault tolerance, security, and distributed file systems.	K2
CO 3	To analyze the current popular distributed systems such as peer-to-peer (P2P) systems will also be analyzed.	K4
CO 4	To know about Shared Memory Techniques and have Sufficient knowledge about file access	K1
CO 5	Have knowledge of Synchronization and Deadlock.	K1
DETAILED SYLLABUS		**3-0-0**
Unit	**Topic**	**Proposed Lecture**
I	Introduction, Examples of distributed Systems, Resource sharing and the Web Challenges. Architectural models, Fundamental Models. Theoretical Foundation for Distributed System: Limitation of Distributed system, absence of global clock, shared memory, Logical clocks ,Lamport's & vectors logical clocks. Concepts in Message Passing Systems: causal order, total order, total causal order, Techniques for Message Ordering, Causal ordering of messages, global state, termination detection.	**08**
II	Classification of distributed mutual exclusion, requirement of mutual exclusion theorem, Token based and non token based algorithms, performance metric for distributed mutual exclusion algorithms. Distributed Deadlock Detection: system model, resource Vs communication deadlocks, deadlock prevention, avoidance, detection & resolution, centralized dead lock detection, distributed dead lock detection, path pushing algorithms, edge chasing algorithms.	**08**
III	Introduction, System models, classification of Agreement Problem, Byzantine agreement problem, Consensus problem, Interactive consistency Problem, Solution to Byzantine Agreement problem, Application of Agreement problem, Atomic Commit in Distributed Database system. Distributed Resource Management: Issues in distributed File Systems, Mechanism for building distributed file systems, Design issues in Distributed Shared Memory, Algorithm for Implementation of Distributed Shared Memory.	**08**
IV	Failure Recovery in Distributed Systems: Concepts in Backward and Forward recovery, Recovery in Concurrent systems, Obtaining consistent Checkpoints, Recovery in Distributed Database Systems. Fault Tolerance: Issues in Fault Tolerance, Commit Protocols, Voting protocols, Dynamic voting protocols	**08**
V	Transactions, Nested transactions, Locks, Optimistic Concurrency control, Timestamp ordering, Comparison of methods for concurrency control. Distributed Transactions: Flat and nested distributed transactions, Atomic Commit protocols, Concurrency control in distributed transactions, Distributed deadlocks, Transaction recovery. Replication: System model and group communication, Fault - tolerant services, highly available services, Transactions with replicated data.	**08**

Text books:

1. Singhal&Shivaratri, "Advanced Concept in Operating Systems", McGraw Hill
2. Ramakrishna,Gehrke," Database Management Systems", McGraw Hill
3. Vijay K.Garg Elements of Distributed Compuitng , Wiley
4. Coulouris, Dollimore, Kindberg, "Distributed System: Concepts and Design", Pearson Education
5. Tenanuanbaum, Steen," Distributed Systems", PHI

M. Tech. (AI & DS) Semester-II

SYLLABUS

MADS-201 ARTIFICIAL INTELLIGENCE & KNOWLEDGE DISCOVERY		
Course Outcome (CO)		**Bloom's Knowledge Level (KL)**
At the end of course, the student will be able to understand		
CO1	Understand the concepts of Artificial Intelligence and intelligent agents.	K1, K2
CO2	Understand and learn knowledge representation and reasoning for the problem-solving.	K2, K3
CO3	Apply basic search techniques for problem-solving.	K3
CO4	Understand and apply learning techniques.	K1, K2, K3
CO5	Apply and utilize AI knowledge for application in the real world.	K3, K5, K6
DETAILED SYLLABUS		
Unit	**Topic**	**Proposed Lecture**
I	Introduction to Artificial Intelligence, Foundation and history of Artificial Intelligence, Agents and Environments, Structure of intelligent agents, State-of-the-art applications of AI.	**08**
II	Knowledge Representation and Reasoning: Representation and Reasoning using predicate logic, Inference in first order logic, forward and backward chaining. Probabilistic reasoning, Bayesian networks, Dempster Shafer theory, Probabilistic Reasoning overtime: Hidden Markov Models.	**08**
III	Uninformed search and Informed search based on heuristics, Local search algorithms and optimization problems, Adversarial search: Games, Optimal decisions in games, Alpha-betapruning, Online search.	**08**
IV	Learning from examples, Forms of Learning, Inductive Learning, learning decision trees, learning in problem solving, Learning Probabilistic models, Bayesian learning, Learning in neural and belief networks. Learning with hidden variable.	**08**
V	AI Applications: Expert system, decision support systems, speech and vision, natural language processing, semantic web, robotics, Al-based programming tools.	**08**

Textbooks & References:

1. Artificial Intelligence: A Modern Approach, S Russel and P Norvig, 3rd Edition, 2015 Prentice Hall.
2. Introduction to Artificial Intelligence and Expert Systems, Dan W. Patterson, Pearson Education.
3. Artificial Intelligence and Expert Systems — Patterson PHI.

MADS-202
Advanced Database Management System

	Course Outcome (CO)	Bloom's Knowledge Level (KL)
At the end of course, the student will be able to understand		
CO1	Understand and implement the concepts of database systems in real world, relational model and ER diagram.	K1, K2, K3
CO2	Apply query processing techniques to automate their all-time problems of databases and understand and implement normalization.	K2, K3, K4
CO3	Understand the concepts of transactions and database management issues including data integrity, security and recovery.	K2, K3, K6
CO4	Understand the concepts of distributed systems and techniques forms sage ordering.	K1, K2, K3
CO5	Understand the concepts of distributed transactions and concurrency control in distributed environment.	K2, K3

DETAILEDSYLLABUS

Unit	Topic	Proposed Lecture
I	Data Modelling using the Entity Relationship Model, diagram, Relational data model concepts, integrity constraints, relational algebra, ERD, DFD using tools, e.g. Power designer	**08**
II	Database Systems Design and SQL: Querying Database Using SQL, Database Design Concepts, FDs, MVDs, JD and concepts in Normalization.	**08**
III	Transaction and Concurrency Control, Schedules, testing serializability of schedules, conflict & view serializable, recoverability, Recovery from transaction failures, log-based recovery, checkpoints, deadlock handling.	**08**
IV	Theoretical Foundation for Distributed System, Logical Clocks, Lamport's Vectors logical clocks. Message Passing Systems: causal order, total order, total causal order, Techniques for Message Ordering	**08**
V	Distributed Transactions, Commit protocols, Concurrency control in distributed transactions, Distributed Deadlocks, Transaction Recovery. Replication: System Model and Group Communication, Fault-tolerant services	**08**

Textbooks & References:

1. R. Ramakrishnan, J. Gehrke, Database Management Systems, McGraw Hill, 2004
2. A. Silberschatz, H. Korth, S. Sudarshan, Database system concepts, 5/e, McGraw Hill, 2008.
3. S Ceri and G Pelagatti, "Distributed databases principles and systems", 1st Edition, TMH, 2008.
4. Elmasri R, Navathe S B, Somayajulu D V L N, and Gupta S K, "Fundamentals of Database Systems", 5th Edition, Pearson Education, 2009.
5. C. J. Date, "Introduction to Database Systems", 8th Edition, Pearson Education, 2009.

MADS-251 ARTIFICAL INTELLIGENCE & KNOWLEDGE DISCOVERY LAB	
S. No.	**List of Experiments/Assignments**
1.	Create a SWI Prolog program to represent the family tree. The topmost, nodes are parents and bottom most nodes are children nodes. Nodes in the middleware parent or child or both. All children have two arrows going to its parents
2.	Translate the following text into Prolog Logic to answer the queries: Problem: A, B and C belong to the Himalayan club. Every member in the club is either a mountain climber or a skier or both. A likes whatever B dislikes and dislikes whatever B likes. A likes rain and snow. No mountain climber likes rain. Every skier likes snow. Query 1: Is there a member who is a mountain climber and not a skier? Query 2: Is there a member who is both a mountain climber and a skier? Query 3: Is there a member who likes both rain and snow?
3.	Eight queens problem is a constraint satisfaction problem (CSP).Write the required predicates to solve the Eight Queens placement problem.
4.	Medical Diagnostic Expert system design which will hypothesis the name of the disease by learning the symptoms the patient have. Medical Diagnostic system expert knowledge about symptoms and name of the disease. Write a prolog program that will represent this expert knowledge in terms of rules in its knowledge base
5.	Implementation of searching techniques in AI • A* search. • Uniformed Cost search. • Graph search. • Alpha-beta pruning.
6.	Chat bots are computer programs that simulate human conversation through text or voice interactions. Write a program to implement AI based chatbots.
7.	Implement a Web Crawler with the following functionality- ● The crawler collects a set of web pages and identify their properties related to the web structure ● Ranking of pages by using various techniques. ● Grouping of pages by similarity.
8.	Create a sentiment analysis model with IMDB dataset.

1. The instructors can add/modify/tune experiments, wherever he/she feels in justified manner.
2. The language/tools/environment to conduct the lab are R, Python, MATLAB/JAVA.

MADS-021
DATA ANALYTICS & VISUALIZATION

Course Outcome (CO)		Bloom's Knowledge level (KL)
At the end of course, the student will be able to understand		
CO1	Discuss various concepts of data analytics pipeline.	K1, K2
CO2	Discus sand apply various data analytic methods.	K1, K2, K3
CO3	Discuss and apply text and sentiment analysis.	K2, K3
CO4	Apply R tool for Data Analytics problem solving.	K3, K5, K6
CO5	Understand NoSQL and Data Visualization methods and implementation in tools.	K1, K2, K5

DETAILEDSYLLABUS

Unit	Topic	Proposed Lecture
I	Introduction : Data, Types of Data, Big Data, Big data Characteristics, Business Intelligence, Levels of measurement, Introduction to Statistical Learning, Mean, Median, Mode, Standard deviation. Life cycle of Data centric projects.	08
II	Basic Analysis Techniques: Chi-Square, t Test, Correlation Analysis, Analysis of Variance. Advanced Analytics Techniques: Regression, Clustering, Classification, Association Mining.	08
III	Text Analytics & Web Mining: Process of Text Analytics, Topic Modelling, Sentiment Analysis, Web Mining. Time Series Analysis: Overview of Time Series Analysis, Forecasting Models, ARMA and ARIMA Models	08
IV	R language: Introduction to R., Basic Syntax, Implementation basic and advanced Data analytic methods, Data visualization using R,Text Analysis Process in R.	08
V	NoSQL: Introduction to NoSQL, Principles of NoSQL Data Models, CAP, NoSQL Data Model. Data Visualization: Plots and graphs for data visualization.	08

Textbooks & References:

1. Michael Berthold, David J. Hand, Intelligent DataAnalysis, Springer
2. Anand Rajaraman and Jeffrey David Ullman, Mining of Massive Datasets, Cambridge University Press.
3. Michael Minelli, Michelle Chambers, and Ambika Dhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley
4. David Dietrich, Barry Heller, Beibei Yang, "Data Science and Big Data Analytics", EMC Education Series, John Wiley
5. Mark Gardner, "Beginning R: The Statistical Programming Language", Wrox Publication
6. GlennJ.Myatt,MakingSenseofData,JohnWiley&Sons
7. JiaweiHan,MichelineKamber"DataMiningConceptsandTechniques",SecondEdition,Elsevier
8. ScientificArticles published in InternationalJournals and Conferences relatedto DataAnalytics

MADS-022
Deep Learning

Course Outcome (CO)		Bloom's Knowledge Level (KL)
At the end of course, the student will be able to understand		
CO1	Understand and Analyze concepts of Bayesian Learning, Neural networks, Perceptron convergence theorem, Back propagation.	K2, K4
CO2	Understand and Apply Deep Learning, CNN and Transfer Learning	K2,K3
CO3	Understand and analyse PCA and Deep learning architectures.	K2, K4, K5
CO4	Explain and Understand Gradient Descent, Normalization & LSTM	K1, K2, K3
CO5	Implementation of Case Studies in Deep Learning.	K3, K4, K6

DETAILEDSYLLABUS

Unit	Topic	Proposed Lecture
I	Introduction to Bayesian Learning, Decision Surface, Linear Models (SVM, Perceptron, LogisticRegression), Optimization Techniques, Gradient Descent, Linear and non linear separable problems, Perceptron Convergence Theorem, Introduction to Neural Network.	08
II	History of Deep Learning, Introduction to Probability, Continuous and Discrete Distributions, Probabilistic Theory of Deep Learning, Deep Vs Shallow Networks, Multilayer Perceptron, Back Propagation Learning, Convolutional Neural Network (CNN), Building Blocks CNN, Transfer Learning	08
III	Principal Component Analysis (PCA), Auto encoders, Deep Learning Architectures-AlexNet, VGG, ResNet etc.	08
IV	Revisiting Gradient Descent, Momentum Optimizer, RMSPrOP, ADAM, Loss Functions, EffectiveTraining Deep Net-Early Stopping, Drop-Out, Batch Normalization, Instance Normalization, Group Normalization, Recurrent Networks, LSTM.	08
V	ImageNet, WaveNet, Generative Modelling with DL, Generative Adversarial Networks, Case studies of DL models with some real-life applications.	08

Textbooks & References:

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, MIT Press,2016.
2. Michael Nielsen, Neural Networks and Deep Learning, Determination Press, 2015.
3. Pattern Classification-Richard O. Duda, Peter E. Hart, David G. Stork, John Wiley & Sons Inc.
4. Cosma Rohilla Shalizi, Advanced Data Analysis from Elementary Point of View, 2015
5. Deng Yu, Deep Learning: Methods and Applications, Now Publishers, 2013

MADS-023
Cloud Computing

Course Outcome (CO)		Bloom's Knowledge Level (KL)
At the end of course, the student will be able to understand		
CO1	Describe architecture and underlying principles of cloud computing.	K2, K3
CO2	Identify security and privacy issues in cloud computing.	K1, K3, K4
CO3	Analyze the performance, scalability, and availability of the underlying cloud technologies and software.	K4, K5
CO4	Deploy applications over commercial cloud computing infrastructures such as Amazon Web Services, Windows Azure, and Google App Engine.	K2, K3, K4
CO5	Explain recent research results in cloud computing and identify their pros and cons.	K2, K3, K4

SYLLABUS

Unit	Topic	Proposed Lecture
I	Origins of Cloud computing – Cloud components - Essential characteristics – On-demand self-service, Broad network access, Comparing cloud providers with traditional IT service providers, Roots of cloud computing	08
II	Architectural influences – High-performance computing, Utility and Enterprise grid computing, Cloud scenarios – Benefits: scalability, simplicity, vendors, security, Limitations – Sensitive information	08
III	Layers in cloud architecture, Software as a Service (SaaS), features of SaaS and benefits, Platform as a Service (PaaS), features of PaaS and benefits, Infrastructure as a Service (IaaS), features of IaaS and benefits, Service providers, challenges and risks in cloud adoption. Cloud deployment model: Public clouds – Private clouds – Community clouds - Hybrid clouds - Advantages of Cloud computing.	08
IV	Introduction to Simulator, understanding CloudSim simulator, CloudSim Architecture (User code, CloudSim, GridSim, SimJava) Understanding Working platform for CloudSim, Introduction to GreenCloud	08
V	Basics of VMWare, advantages of VMware virtualization, using VMware workstation, creating virtual machines-understanding virtual machines, create a new virtual machine on local host, cloning virtual machines, virtualize a physical machine, starting and stopping a virtual machine.	08

Textbooks & References:

1. Cloud computing a practical approach - Anthony T. Velte, Toby J. Velte Robert Elsenpeter, TATA McGraw- Hill, New Delhi – 2010.
2. Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online - Michael Miller - Que 2008.
3. Cloud computing for dummies- Judith Hurwitz, Robin Bloor, Marcia Kaufman, Fern Halper, Wiley Publishing, Inc, 2010
4. Cloud Computing (Principles and Paradigms), Edited by Rajkumar Buyya, James Broberg, Andrzej Goscinski, John Wiley & Sons, Inc. 2011

MADS-024
WEB and BIG DATA MINING

Course Outcome (CO)		Bloom's Knowledge Level (KL)
Attend Course, the student will be able to understand		
CO1	Discuss and explain various concepts of Data Mining and Web Mining.	K1, K2, K3
CO2	Discuss and understand various methods of Supervised and Unsupervised learning methods.	K1, K2
CO3	Discuss and analyse Information Retrieval and Web Search.	K2, K3, K4
CO4	Understand Web crawling and implementation of Web Crawler and Page Rank Algorithm.	K3, K5, K6
CO5	Discuss and explain Sentiment Analysis and Opinion Mining.	K1, K2, K3

DETAILEDSYLLABUS

Unit	Topic	Proposed Lecture
I	Introduction to Web Data Mining and Data Mining Foundations, Introduction – World Wide Web (WWW), A Brief History of the Web and the Internet, Web Data Mining - Data Mining, Web Mining. Data Mining Foundations–Association Rules And Sequential Patterns – Basic Conceptsof Association Rules, Apriori Algorithm - Frequent Item set Generation, Association Rule Generation, Data Formats for Association Rule Mining, Mining with multiple minimum supports, Mining Sequential Patterns on GSP, Mining Sequential Patterns on Prefix Span, Generating Rules from Sequential Patterns.	08
II	Supervised and Unsupervised Learning Supervised Learning – Basic Concepts, Decision Tree Induction – Learning Algorithm, Impurity Function, Handling of Continuous Attributes, Classifier Evaluation, Rule Induction – Sequential Covering, Rule Learning, Classification Based on Associations, Naïve Bayesian Classification, Naïve Bayesian Text Classification – Probabilistic Framework, Naïve Bayesian Model Unsupervised Learning – Basic Concepts, K-means Clustering–K-means Algorithm, Representation of Clusters, Hierarchical Clustering – Single link method, Complete link Method, Average Link Method, Strength and Weakness.	08
III	Information Retrieval and Web Search: Basic Concepts of Information Retrieval. Information Retrieval Methods – Boolean Model, Vector Space Model and Statistical Language Model, Relevance Feedback, Evaluation Measures, Text and Web Page Preprocessing– Stopword Removal, Stemming, Inverted Index and Its Compression–Inverted Index, Search using Inverted Index, Index Construction, Index Compression, Latent Semantic Indexing–Singular Value Decomposition, Query and Retrieval, Web Search, Meta Search, WebSpamming.	08
IV	Analysis, Co-Citation and Bibliographic Coupling, PageRank Algorithm, HITS Algorithm, Community Discovery-Problem Definition, Bipartite Core Communities, Maximum Flow Communities, Email Communities. Web Crawling – A Basic Crawler Algorithm - Breadth First Crawlers, Preferential Crawlers, Implementation Issues–Fetching, Parsing, Stop word Removal, Link Extraction, Spider Traps, Page Repository, Universal Crawlers, Focused Crawlers, Topical Crawlers, Crawler Ethics and Conflicts	08
V	Sentiment Analysis and Opinion Mining – Sentiment Classification – Classification based on Sentiment Phrases, Classification Using Text Classification Methods, Feature based Opinion Mining and Summarization –Opinion Search and Opinion Spam. Aspect Based Sentiment Analysis, Supervised and Unsupervised Spam Detection, Group Spam Detection.	08

Textbooks & References:

1. Web Data Mining: Exploring Hyperlinks, Contents and usage data by Bing Liu (SpringerPublication)
2. Data Mining: Concepts and Techniques, Second Edition by Jiawei Han, MichelineKamber (Elsevier Publications)
3. Mining the Web: Discovering Knowledge from Hypertext Data by SoumenChakrabarti (Morgan Kaufmann Publishers)
4. ResearchArticles&JournalPapers

Natural Language Processing (MADS-025)		
Course Outcome (CO)		**Bloom's Knowledge Level (KL)**
At the end of course , the student will be able :		
CO 1	To learn the fundamentals of natural language processing.	K_1 , K_2
CO 2	To understand the use of CFG and PCFG in NLP.	K_1 , K_2
CO 3	To understand the role of semantics of sentences and pragmatic.	K_2
CO 4	To introduce speech production and related parameters of speech.	K_1 , K_2
CO 5	To show the computation and use of techniques such as short time fourier transform, linear predictive coefficients and other coefficients in the analysis of speech.	K_3, K_4
DETAILED SYLLABUS		**3-0-0**
Unit	**Topic**	**Proposed Lecture**
I	Origins and challenges of NLP – Language Modeling: Grammar-based LM, Statistical LM – Regular Expressions, Finite-State Automata – English Morphology, Transducers for lexicon and rules, Tokenization, Detecting and Correcting Spelling Errors, Minimum Edit Distance. Unsmoothed N-grams, Evaluating N-grams, Smoothing, Interpolation and Backoff – Word Classes, Part-of-Speech Tagging, Rule-based, Stochastic and Transformation-based tagging, Issues in PoS tagging – Hidden Markov and Maximum Entropy models.	**08**
II	Context Free Grammars, Grammar rules for English, Treebanks, Normal Forms for grammar – Dependency Grammar – Syntactic Parsing, Ambiguity, Dynamic Programming parsing – Shallow parsing – Probabilistic CFG, Probabilistic CYK, Probabilistic Lexicalized CFGs – Feature structures, Unification of feature structures.	**08**
III	Requirements for representation, First-Order Logic, Description Logics – Syntax-Driven Semantic analysis, Semantic attachments – Word Senses, Relations between Senses, Thematic Roles, selectional restrictions – Word Sense Disambiguation, WSD using Supervised, Dictionary & Thesaurus, Bootstrapping methods – Word Similarity using Thesaurus and Distributional methods.	**08**
IV	Speech Fundamentals: Articulatory Phonetics – Production And Classification Of Speech Sounds; Acoustic Phonetics – Acoustics Of Speech Production; Review Of Digital Signal Processing Concepts; Short-Time Fourier Transform, Filter-Bank And LPC Methods.	**08**
V	Features, Feature Extraction And Pattern Comparison Techniques: Speech Distortion Measures– Mathematical And Perceptual – Log– Spectral Distance, Cepstral Distances, Weighted Cepstral Distances And Filtering, Likelihood Distortions, Spectral Distortion Using A	

	Warped Frequency Scale, LPC, PLP And MFCC Coefficients, Time Alignment And Normalization – Dynamic Time Warping, Multiple Time – Alignment Paths. Hidden Markov Models: Markov Processes, HMMs – Evaluation, Optimal State Sequence – Viterbi Search, Baum-Welch Parameter Re-Estimation, Implementation Issues.	**08**

Text books:

1. Daniel Jurafsky, James H. Martin—Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech, Pearson Publication, 2014.
2. Steven Bird, Ewan Klein and Edward Loper, —Natural Language Processing with Python, First Edition, OReilly Media, 2009.
3. Lawrence Rabiner And Biing-Hwang Juang, "Fundamentals Of Speech Recognition", Pearson Education, 2003.
4. Daniel Jurafsky And James H Martin, "Speech And Language Processing – An Introduction To Natural Language Processing, Computational Linguistics, And Speech Recognition", Pearson Education, 2002.
5. Frederick Jelinek, "Statistical Methods Of Speech Recognition", MIT Press, 1997.
6. 1. Breck Baldwin, —Language Processing with Java and LingPipe Cookbook, Atlantic Publisher, 2015.
7. Richard M Reese, —Natural Language Processing with Java, OReilly Media, 2015.
8. Nitin Indurkhya and Fred J. Damerau, —Handbook of Natural Language Processing, Second Edition, Chapman and Hall/CRC Press, 2010.
9. Tanveer Siddiqui, U.S. Tiwary, —Natural Language Processing and Information Retrieval, Oxford University Press, 2008.

MADS-031
PATTERN RECOGNITION

Course Outcome (CO)		Bloom's Knowledge Level (KL)
At the end of course, the student will be able to understand		
CO1	Understand Bayes decision theory, Bayes inference, Bayes classifier, and Bayes Networks.	K1, K2
CO2	Discuss and understand Linear discriminant functions, logistic discriminant functions, SVM for separable and nonlinearly separable classes.	K1, K2, K3
CO3	Understand and apply Non-linear classifiers and their combinations.	K2, K3, K4
CO4	Apply host of Clustering algorithms for small and large data set.	K3, K4, K5
CO5	Discuss and understanding of theory to select the right approach to solve a given problem.	K3, K5, K6

DETAILED SYLLABUS

Unit	Topic	Proposed Lecture
I	Introduction, Features, Feature Vectors, Classifiers, Supervised, Unsupervised and Semi Supervised Learning. Different Paradigms of Pattern Recognition, Probability estimation, Proximity measures, Feature extraction.	**08**
II	Introduction, Bayes Decision Theory, Discriminant Functions, Bayes Classification for Normal Distributions, Estimation of Unknown Probability Distributions: ML Parameter Estimation, MAP Estimation, Bayesian Inference, Maximum Entropy Estimation, Mixture Models, Non Parametric Estimation, the Naïve-Bayes Classifier, the Nearest Neighbor Rule, Bayesian Networks	**08**
III	Introduction, Linear Discriminant Functions and Decisions, Hyper-planes, The Perceptron algorithm, Least Square Methods, Mean Square Estimation Revisited, Logistic Discrimination, Support Vector Machines for Separable Classes, SVM for Non-Separable Classes, SVM for Multiclass Case, 9-SVM	**08**
IV	XOR Problem, Two Layer Perceptron, Three-Layer Perceptrons, Algorithms based on Exact Classification of Training Set, The Back-Propagation Algorithm, Variation of BP Theme, Choice of Cost Function, Choice of Network Size, Generalized Linear Classifiers, Capacity of d-dimensional space in linear Dichotomies, Polynomial Classifiers, Radial Basis Function Networks, Universal Approximators, Probabilistic Neural Networks, SVM-Nonlinear Case, Beyond SVM Paradigm, Decision Trees, Combining Classifiers, Boosting, Class Imbalance Problem	**08**
V	Introduction, Proximity Measures, Number of Possible Clusterings, Categories of Clustering Algorithms, Sequential Clustering Algorithms, Agglomerative Algorithms, Divisive Algorithms, Hierarchical Algorithms for Large Datasets, Choice of the Best Number of Clusters, Hard Clustering Algorithms, Vector Quantization. Algorithms based on Graph Theory, Competitive Learning algorithms	**08**

Textbooks & References:

1. R. O. Duda, P. E. Hart and D. G. Stork Pattern Classification, Wiley Publications, 2001.
2. D. McKay, Information Theory, Inference, and Learning Algorithms, Cambridge University Press 2003.
3. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

MADS-032
RECOMMENDER SYSTEMS

Course Outcome (CO)		Bloom's Knowledge Level (KL)
At the end of course, the student will be able		
CO1	To understand basic techniques and problems in the field of recommender systems	K1, K2
CO2	To understand filtering algorithms and apply on recommendations..	K1, K2, K3
CO3	To introduce different approaches of recommender systems	K2, K3
CO4	To apply algorithms and techniques to develop Recommender Systems that are widely used in the Internet industry.	K3, K4
CO5	To explore various types of recommender systems and develop state-of-the-art recommender systems.	K3, K4, K5

DETAILED SYLLABUS

Unit	Topic	Proposed Lecture
I	Basic concepts for recommender systems: detailed taxonomy of recommender systems, Evaluation of recommender systems, Applications of recommendation systems, Issues with recommender system.	08
II	Collaborative filtering algorithms: User-based nearest neighbor recommendation, Item based nearest-neighbour recommendation, Model based and pre-processing based approaches, Attacks on collaborative recommender systems.	08
III	Content-based recommendation: High level architecture of content-based systems, Advantages and drawbacks of content based filtering, Item profiles, Discovering features of documents, Obtaining item features from tags, Representing item profiles, Methods for learning user profiles, Similarity based retrieval, Classification algorithms.	08
IV	Knowledge based recommendation: Knowledge representation and reasoning, Constraint based recommenders, Case based recommenders	08
V	Hybrid approaches: Opportunities for hybridization, Monolithic hybridization design: Feature combination, Feature augmentation, Parallelized hybridization design: Weighted, Switching, Mixed, Pipelined hybridization design: Cascade Meta-level, Limitations of hybridization strategies. Evaluating Recommender System: General properties of evaluation research, Evaluationdesigns, Evaluation on historical datasets.	08

Textbooks & References:

1. Charu Aggarwal “Recommender Systems: The Textbook,” First Edition, Springer.
2. Francesco Ricci, Lior Rokach, and Bracha Shapira “Recommender Systems Handbook,” First Edition, Springer.
3. Rounak Banik “Hands-On Recommendation Systems with Python,” First Edition, Packt Publishing.
4. Kim Falk “Practical Recommender Systems,” First Edition, Manning Publications.
5. Deepak Agarwal and Bee-Chung Chen “Statistical Methods for Recommender Systems,” First Edition, Cambridge University Press.

MADS-033
Probability and Stochastic Process

	Course Outcome (CO)	Bloom's Knowledge Level (KL)
At the end of course, the student will be able to understand		
CO1	Understand and apply fundamental probability concepts.	K1, K2
CO2	Understand and apply concept of random variables and stochastic process.	K1, K2, K3
CO3	Understand multiple random variable and operations performed on multiple random variables.	K2, K3, K4
CO4	To analyze random process and Transmission of random process.	K3, K4, K5
CO5	To understand concept of temporal & spectral characterstics of random process.	K2, K3, K4

SYLLABUS

Unit	Topic	Proposed Lecture
I	Probability Theory Refresher: Axiomatic construction of probability spaces, random variables and vectors, probability distributions, functions of random variables; mathematical expectations, transforms and generating functions, modes of convergence of sequences of random.	08
II	Introduction to Stochastic Processes (SPs) & Stationary Processes: Definition and examples of SPs, classification of random processes according to state space and parameter space, types of SPs, elementary problems. Weakly stationary and strongly stationary processes, moving average and auto regressive processes	08
III	Discrete-time Markov Chains (DTMCs) & Continuous-time Markov Chains (CTMCs): Definition and examples of MCs, transition probability matrix, Chapman-Kolmogorov equations; calculation of n-step transition probabilities, limiting probabilities, classification of states, ergodicity, stationary distribution, transient MC; random walk and gambler's ruin problem, applications. infinitesimal generator, Poisson process, birth-death process, stochastic Petri net, applications to queueing theory and communication networks	08
IV	Martingales & Brownian Motion: Conditional expectations, definition and examples of martingales. Wiener process as a limit of random walk; process derived from Brownian motion, stochastic differential equation, stochastic integral equation, Ito formula, Some important SDEs and their solutions, applications to finance	08
V	Renewal Processes & Branching Processes: Renewal function and its properties, renewal theorems, cost/rewards associated with renewals, Markov renewal and regenerative processes, non Markovian queues, applications of Markov regenerative processes: Definition and examples branching processes.	08

Textbooks & References:

1. J. Medhi, Stochastic Processes, 3rd Edition, New Age International, 2009.
2. H. Stark and J. W. Woods, Probability and Random Processes with Applications to Signal Processing, Prentice Hall, 2002.
3. A. Papoulis and S. U. Pillai, Probability, Random Variables and Stochastic Processes, 4th Edn., McGraw-Hill, 2002

MADS 034
Advanced Machine Learning

Course Outcome (CO)		Bloom's Knowledge Level (KL)
At the end of course, the student will be able to understand		
CO1	Discuss various concepts of machine learning with hypothesis is and version space.	K1, K2
CO2	Understand a wide variety of learning algorithms and how to evaluate models generated from data.	K1, K3
CO3	Understand and apply Bayesian method of machine learning.	K2, K3
CO4	Understand and apply Instance based learning and Case based learning.	K1, K2, K3
CO5	Apply and evaluate machine learning algorithms on real world.	K2, K3, K5

DETAILED SYLLABUS

Unit	Topic	Proposed Lecture
I	Introduction: Well defined learning problems, Designing a Learning System, Issues in Machine Learning, THE CONCEPT LEARNING TASK-General-to-specific ordering of hypotheses, Find-S, List then eliminate algorithm, Candidate elimination algorithm, Inductive bias.	**08**
II	Decision tree learning algorithm-Inductive Bias-Issues Decision tree learning; ARTIFICIAL NEURAL NETWORKS–Perceptrons, Gradient descent and the Delta Rule, Adaline, Multi layer Networks, Derivation of Back propagation Rule Back propagation Algorithm Convergence, Generalization.	**08**
III	Estimating Hypotheses Accuracy, Basics of sampling Theory, Comparing Learning Algorithms; Bayesian Learning: Bayes Theorem, Concept Learning, Bayes Optimal Classifier, Naïve Bayes Classifier, Bayesian Belief Networks, EM algorithm.	**08**
IV	Sample Complexity for Finite Hypothesis spaces, Sample Complexity for Infinite Hypothesis spaces, The Mistake Bound Model of Learning; INSTANCE-BASED LEARNING–k- Nearest Neighbour Learning, Locally Weight Regression, Radial basis function networks, Case-based learning.	**08**
V	An illustrative example, Hypothesis spaces search, Genetic Programming, Models of Evolution and Learning; Learning first order rules-sequential covering algorithms General to specific search-FOIL; REINFORCEMENT LEARNING- The Learning Task, Q Learning.	**08**

Textbooks & References:

1. TomM.Mitchell,—MachineLearning,McGraw-HillEducation(India)PrivateLimited,2013.
2. Ethem Alpaydin, Introduction to Machine Learning (Adaptive Computation and Machine Learning, The MITPress2004.
3. Stephen Marsland—Machine Learning: An Algorithmic Perspective, CRC Press, 2009.
4. Bishop C. Pattern Recognition and Machine Learning Berlin: Springer-Verlag.
5. Scientific Articles published in International Journals and Conferences related o Data Analytics.