Rajiv Gandhi ProudyogikiVishwavidyalaya Bhopal
M.Tech (Artificial Intelligence & Data Science)
First Semester Syllabus

MTAD 101 - DATA STRUCTURES AND ALGORITHMS

UNIT- 1

INTRODUCTION: Basic concepts of OOPs – Templates – Algorithm Analysis – ADT - List (Singly, Doubly and Circular) Implementation - Array, Pointer, Cursor Implementation.

UNIT-2

BASIC DATA STRUCTURES: Stacks and Queues – ADT, Implementation and Applications - Trees – General, Binary, Binary Search, Expression Search, AVL, Splay, B-Trees – Implementations - Tree Traversals.

UNIT-3

ADVANCED DATA STRUCTURES: Set-Implementation, Basic operations on set –Priority Queue Implementation, Graphs, Directed Graphs, Shortest Path Problem, Undirected Graph, Spanning Trees, Graph Traversals.

UNIT-4

MEMORY MANAGEMENT: Issues - Managing Equal Sized Blocks, Garbage Collection Algorithms for Equal Sized Blocks, Storage Allocation for Objects with Mixed Sizes, Buddy Systems, Storage Compaction

UNIT-5

SEARCHING, SORTING AND DESIGN TECHNIQUES: Searching Techniques, Sorting – Internal Sorting – Bubble Sort, Insertion Sort, Quick Sort, Heap Sort, Bin Sort, Radix Sort – External Sorting – Merge Sort, Multi-way Merge Sort, Polyphase Sorting - Design Techniques - Divide and Conquer – Dynamic Programming – Greedy Algorithm – Backtracking - Local Search Algorithms.

REFERENCE BOOKS:

1. Mark Allen Weiss, "Data Structures and Algorithm Analysis in C++", Pearson Pub
2. Aho, Hopcroft, Ullman, "Data Structures and Algorithms", Pearson Education Pub.
3. Drozdek, Data Structures and algorithm in Jawa, Cengage (Thomson)
4. Gilberg, Data structures Using C++, Cengage
5. Horowitz, Sahni, Rajasekaran, "Computer Algorithms", Galgotia,
6. Tanenbaum A.S., Langram Y, Augestien M.J., "Data Structures using C & C++", Prentice Hall of India, 2002

MTAD 102 - Artificial Intelligence

UNIT-I

INTRODUCTION TO AI: AI application, Recent trends and Al techniques. Problem Definition, Problem characteristics, Production systems, Production system characteristics Control Strategies. Search strategies. Problem solving methods - Problem graphs, Matching, Indexing and Heuristic functions - Hill Climbing-Depth first and Breath first, A*, AO*. Constraints Satisfaction - Related algorithms.

UNIT-II

REPRESENTATION OF KNOWLEDGE: Knowledge representation, Knowledge representation using Predicate logic, Introduction to predicate calculus, Resolution, Use of predicate calculus, Knowledge representation using other logic-Structured representation of knowledge.

UNIT III

KNOWLEDGE INFERENCE: Knowledge representation - Production based system. Frame Based system. Inference - Backward chaining, Forward chaining, Rule value approach, Fuzzy Reasoning - Certainty factors, Bayesian Theory-Bayesian Network -Dempster - Shafer theory.

UNIT- IV

GAME PLAYING AND PLANNING: Game playing, Min-Max search procedure. Alpha Beta cut-offs, Basic plan generation systems - Strips ,Advanced plan generation systems - k strips.

UNIT V

EXPERT SYSTEMS: Expert systems - Architecture of expert systems, Roles of expert systems - Knowledge Acquisition -Meta knowledge ,Heuristics, Typical expert systems-MYCIN.

References

1. Kevin Night and Elaine Rich, Nair B ,"Artificial Intelligence Mc Graw Hill- 2008.
2. Dan W. Patterson, "Introduction to Al and ES" Pearson Education, 2007.
3.Stuart Russel and Peter Norvig Al - A Modern Approach 2nd Edition Pearson Education
4. Deepak Khemani "Artificial Intelligence" Tata Mc Graw Hill Education 2013

MTAD 103 - Mathematical Foundation of Computer Science

Unit 1-
Probability mass density and cumulative distribution functions. Parametric families of distributions, Expected value. Variance, conditional expectation, Applications of the univariate and multivariate Central Limit Theorem, Probabilistic inequalities, Markov chains.

Unit 2
Queuing system, transient and steady state ,traffic intensity, distribution queuing system, concepts of queuing models (M/M/1 :Infinity/ Infinity/ FC FS), (M/M/1: N/ Infinity/ FC FS), (M/W/S: Infinity/ Infinity/ FCFS)

Unit 3
Statistical inference, Introduction to multivariate statistical models: regression and classification problems, principal components analysis, the problem of over fitting model assessment.

Unit4
Graph Theory: Isomorphism, Planar graphs, graph coloring ,Hamilton circuits and Euler cycles. Permutations and Combinations with and without repetition. Specialized techniques to solve Combinatorial enumeration problems.

Unit 5
Random samples ,sampling distributions of estimators, Methods of Moments and Maximum Likelihood
Elementary properties of FT,DFT, WFT, Discrete Wavelet Transform (DWT). Haar transform.

Reference Books:
1. John Vince, Foundation Mathematics for computer Science, Springer.
2. K Trivedi, Probability and Statistics with Reliability , Queuing and computer Science Applications, Wiley.
3. M. Mitzenmacher and E. Upfal, Probability and Computing: Randomized Algorithms and Probabilistic Analysis.
4. Alan Tucker, Applied combinatorics , Wiley.

MTAD 104 - **Data Science**

Unit 1: Introduction to core concepts and technologies: Introduction Terminology, data science process,
Data science toolkit, Types of data, Example applications.

Unit 2: Data collection and management: Introduction, Sources of data, Data collection and APIs.
Exploring and fixing data. Data storage and management, Using multiple data sources.

Unit 3. Data analysis: Introduction , Terminology and concepts. Introduction to statistics Variance ,Distribution properties and arithmetic Samples/CLT, Basic machine learning algorithms ,Linear regression ,SVM, Naive Bayes.

Unit 4: Data Visualization: Introduction ,Types of data visualization, Data for visualization, Data types,
Data encodings, Retinal variables, Mapping variables to encodings. Visual encodings.

Unit 5: Applications of Data Science Technologies for visualization, Bokeh (Python) Recent trends in
various data collection and analysis techniques various visualization techniques, application development methods of used in data science.

REFERENCE BOOKS

1. Cathy O'Neil and Rachel schutt ,Dong Data Science, Straight Talk from the Frontline. O'Reilly.
2. Jure Leskovek, Anand Rajaraman and Jeffrey Ullman . Mining of Massive Datasets. V2.1. Cambridge University Press

MTAD 105 (A) Programming System

UNIT-I

Introduction, use in solving various business problems, information systems, transaction processing systems, MIS, ERP, decision support systems, EIS.

UNIT-II

Modeling & Design: OMT, methodologies, models, tools & techniques, SDLC, Unified Process life cycle: phases & iterations, Use cases, activity diagrams, UMI diagrams, System Design.

UNIT-III

Aspects of Compilation, overview of the various phases of compiler, Scanning, Syntax error handling, Symbol table conceptual design, Intermediate Code conceptual Design, Intermediate code interfaces, Dynamic storage allocation techniques, Dynamic Programming code generation algorithm, Principal sources of optimization, Approaches to compiler development. Register allocation techniques.

UNIT-IV

Operating system processes: Implementation oriented concepts related to inter process communication, process scheduling, process data structure, bootstrapping, system initialization, interrupt handling.

UNIT-V Procedural Paradigms of programming, Object Oriented Paradigm for programming, Procedural vs. Object Oriented Programming, Principles of OOP, Benefits and applications of OOP. OOP Concepts: Data Abstraction, Encapsulation, Inheritance and Polymorphism.

Reference Books:

1. Sebesta,"Concept of programming Language", Pearson Edu.
2. Louden, "Programming Languages: Principles and Practices", Cengage Learning
3. Tucker, "Programming Languages: Principles and paradigms", Tata McGraw —Hill
4. Terrance W Pratt, "Programming Languages: Design and Implementation", Pearson Edu.
5. G. Booch, Object-Oriented Analysis and Design, Pearson Education.
6. J. Rumbaugh, Object-Oriented Modeling and Design, Pearson Education.
7. Enterprise Resource Planning – A Managerial Perspective by D P Goyal, Tata McGraw Hill Education, 2011

MTAD 105 (B) - DATA WAREHOUSE AND DATA MINING

UNIT 1

Introduction : Data Mining: Definitions, KDD v/s Data Mining, DBMS v/s Data Mining , DM techniques, Mining problems, Issues and Challenges in DM, DM Application areas. Association Rules & Clustering Techniques: Introduction, Various association algorithms like A Priori, Partition, Pincer search etc., Generalized association rules.

UNIT 2

Clustering paradigms; Partitioning algorithms like K-Medioid, CLARA, CLARANS; Hierarchical clustering, DBSCAN, BIRCH, CURE; categorical clustering algorithms, STIRR, ROCK, CACTUS.

UNIT 3

Other DM techniques & Web Mining: Application of Neural Network, AI, Fuzzy logic and Genetic algorithm, Decision tree in DM. Web Mining, Web content mining, Web structure Mining, Web Usage Mining.

UNIT 4

Temporal and spatial DM: Temporal association rules, Sequence Mining, GSP, SPADE, SPIRIT, and WUM algorithms, Episode Discovery, Event prediction, Time series analysis.

UNIT 5

Spatial Mining, Spatial Mining tasks, Spatial clustering, Spatial Trends, Data Mining of Image and Video: A case study. Image and Video representation techniques, feature extraction, motion analysis, content based image and video retrieval, clustering and association paradigm, knowledge discovery.

Reference Books:

1. Data Mining Techniques; Arun K.Pujari ; University Press.
2. Data Mining; Adriaans & Zantinge; Pearson education.
3. Mastering Data Mining; Berry Linoff; Wiley.
4. Data Mining; Dunham; Pearson education.
5. Text Mining Applications, Konchandy, Cengage

MTAD 105 (C) Data Preparation and Analysis

UNIT- I

Data Gathering and Preparation: Data formats, Parsing and transformation, Scalability and real-time issues.

UNIT -II

Data Cleaning: Consistency checking, Heterogeneous and missing data, Data Transformation and Segmentation.

UNIT -III

Exploratory Analysis: Descriptive and comparative statistics, Clustering and association, Hypothesis Generation.

UNIT- IV

Visualization: Designing visualizations, Time series, Geo located data, Correlations and Connections, Hierarchies and networks, interactivity.

UNIT- V
Statistics : Descriptive statistics, Central tendency, Variation , Shape, Inferential statistics Confidence intervals, Hypothesis tests, Chi-square, One-way analysis of variance, Comparative statistics

References:

1. Making sense of Data : A practical Guide to Exploratory Data Analysis and Data Mining, by Glenn J. Myatt.,Wiley.

MTAD 105 (D) Information Retrieval

UNIT-I

Introduction - History of IR- Components of IR - Issues -Open source Search engine Frameworks - The Impact of the web on IR - The role of artificial intelligence (AI) in IR – IR Versus Web Search - Components of a search engine, Characterizing the web.

UNIT -II

Boolean and Vector space retrieval models- Term weighting - TF-IDF weighting- cosine similarity - Preprocessing - Inverted indices - efficient processing with sparse vectors Language Model based IR - Probabilistic IR -Latent Semantic indexing - Relevance feedback and query expansion.

UNIT- III

Web search overview, web structure the user paid placement search engine optimization, Web Search   Architectures - crawling - meta-crawlers, Focused Crawling - web indexes - Near-duplicate detection - Index Compression - XML retrieval.

UNIT -IV

Link Analysis -hubs and authorities - Page Rank and HITS algorithms -Searching and Ranking - Relevance Scoring and ranking for Web - Similarity - Hadoop & Map Reduce - Evaluation - Personalized search - Collaborative filtering and content-based recommendation of documents And products - handling invisible Web - Snippet generation, Summarization. Question Answering, Cross-Lingual Retrieval.

UNIT -V

Information filtering: organization and relevance feedback - Text Mining- Text classification and clustering - Categorization algorithms ,naive Bayes, decision trees and nearest neighbor - Clustering algorithms: agglomerative clustering, k-means, expectation maximization (EM).

References:

1. C. Manning, P. Raghvan and H Schutze: Introduction to Information Retrieval, Cambridge University Press, 2008.
2. Ricardo Baeza -Yates and Berthier Ribeiro –Neto, Modern Information Retrieval The Concepts and Technology behind Search 2nd Edition, ACM Press Books 2011.
3. Bruce Croft, Donald Metzler and Trevor Strohman Search Engines Information Retrieval in Practice 1st Edition Addison Wesley, 2009
4.Mark Levene, An Introduction to Search Engines and Web Navigation, 2nd Edition Wiley 2010.