**IT-701: Data Mining & Data Warehousing**

**Credits: 03**

**L-T-P::3-0-0**

**Unit-I**

Overview, Motivation(for Data Mining),Data Mining-Definition & Functionalities, Data Processing, Form of Data Preprocessing, Data Cleaning: Missing Values, Noisy Data,(Binning,

Clustering, Regression, Computer and Human inspection),Inconsistent Data, Data Integration

and Transformation. Data Reduction:-Data Cube Aggregation, Dimensionality reduction, Data

Compression, Numerosity Reduction, Clustering, Discretization and Concept hierarchy

generation

**Unit-II**

Concept Description:-

Definition, Data Generalization, Analytical Characterization, Analysis of attribute relevance, Mining Class comparisons, Statistical measures in large Databases. Measuring Central Tendency, Measuring Dispersion of Data, Graph Displays of Basic Statistical class Description, Mining Association Rules in Large Databases, Association rule mining, mining Single-Dimensional Boolean Association rules from Transactional Databases– Apriori Algorithm, Mining Multilevel Association rules from Transaction Databases and Mining Multi-Dimensional Association rules from Relational Databases

**Unit-III**

*Classification and Predictions:*

What is Classification & Prediction, Issues regarding Classification and prediction, Decision

tree, Bayesian Classification, Classification by Back propagation, Multilayer feed-forward

Neural Network, Back propagation Algorithm, Classification methods K-nearest neighbor

classifiers, Genetic Algorithm.

**Unit-IV**

*Cluster Analysis*:

Data types in cluster analysis, Categories of clustering methods, Partitioning methods.

Hierarchical Clustering- CURE and Chameleon, Density Based Methods-DBSCAN, OPTICS,

Grid Based Methods- STING, CLIQUE, Model Based Method –Statistical Approach, Neural

Network approach, Outlier Analysis

**Unit-V**

*Data Warehousing*:

Overview, Definition, Delivery Process, Difference between Database

System and Data Warehouse, Multi Dimensional Data Model, Data Cubes, Stars, Snow Flakes,

Fact Constellations, Concept hierarchy, Process Architecture, 3 Tier Architecture, Data Marting.

**Unit-VI**

Aggregation, Historical information, Query Facility, OLAP function and Tools. OLAP Servers,

ROLAP, MOLAP, HOLAP, Data Mining interface, Security, Backup and Recovery, Tuning

Data Warehouse, Testing Data Warehouse.

**Text Books / References:**

1. M.H.Dunham,”**DataMining:Introductory and Advanced Topics**” Pearson

Education

2. Jiawei Han, MichelineKamber, ”**Data Mining Concepts & Techniques**” Elsevier

3. Sam Anahory, Dennis Murray, “**Data Warehousing in the Real World : A Practical Guide**

**for Building Decision Support Systems**”, Pearson Education

4. Mallach,”**Data Warehousing System**”,McGraw –Hill

5. [Arun K. Pujari](http://www.google.co.in/search?tbo=p&tbm=bks&q=inauthor:%22Arun+K.+Pujari%22) “**Data Mining Techniques**”, Universities Press, 01-Jul-2001

**IT-702 Information Security and Cyber Laws**

**Credits: 04**

**L-T-P: 3-1-0**

**Unit-I:**

Security Attacks (Interruption, Interception, Modification and Fabrication), Security Services (Confidentiality, Authentication, Integrity, Non-repudiation, access Control and Availability) and Mechanisms, A model for Internetwork security, Internet Standards and RFCs, Buffer overflow & format string vulnerabilities, TCP session hijacking, ARP attacks, route table modification, UDP hijacking, and man-in-the-middle attacks.

**Unit-II:**

Conventional Encryption Principles, Conventional encryption algorithms, cipher block modes of operation, location of encryption devices, key distribution Approaches of Message authentication, Secure Hash Functions and HMAC.

**Unit-III:**

Public key cryptography principles, public key cryptography algorithms, digital signatures, digital Certificates, Certificate Authority and key management Kerberos, X.509 Directory Authentication Service.

**Unit-IV:**

Email privacy: Pretty Good Privacy (PGP) and S/MIME.

IP Security Overview, IP Security Architecture, Authentication Header, Encapsulating Security Payload, Combining Security Associations and Key Management.

**Unit-V:**

Web Security Requirements, Secure Socket Layer (SSL) and Transport Layer Security (TLS), Secure Electronic Transaction (SET).

**Unit-VI:**

Basic concepts of SNMP, SNMPv1 Community facility and SNMPv3. Intruders, Viruses and related threats.

**Unit-VII:**

Firewall Design principles, Trusted Systems. Intrusion Detection Systems.

Information Security & Law, IPR, Patent Law, Copyright Law, Legal Issues in Data mining Security, Building Security into Software Life Cycle Ethics- Ethical Issues, Issues in Data and Software Privacy Cyber Crime Types & overview of Cyber Crimes

**TEXT BOOKS / REFERENCES :**

1. **Network Security Essentials (Applications and Standards)**, William Stallings Pearson Education.

2. **Hack Proofing your network**, Ryan Russell, Dan Kaminsky, Rain Forest Puppy, Joe Grand, David Ahmad, Hal Flynn Ido Dubrawsky, Steve W.Manzuik and Ryan Permeh, wiley Dreamtech,

3.**Cyber Laws Simplified**, Sood, McGraw Hill

4. **Fundamentals of Network Security**, Eric Maiwald (Dreamtech press)

5. **Network Security - Private Communication in a Public World**, Charlie Kaufman, Radia Perlman and Mike Speciner, Pearson/PHI.

6. **Cryptography and network Security**, Third edition, Stallings, PHI/Pearson

7. **Principles of Information Security**, Whitman, Thomson.

8. **Network Security: The complete reference**, Robert Bragg, Mark Rhodes, TMH

9. **Introduction to Cryptography**, Buchmann, Springer.

**IT711:** **Software Quality Engineering**

**L-T-P:3-0-0**

**Credits :03**

**UNIT-I:**

*Introduction*

Defining Software Quality, Software Quality Attributes and Specification, Cost of Quality,

Defects, Faults, Failures, Defect Rate and Reliability, Defect Prevention, Reduction, and

Containment, Overview of Different Types of Software Review, Introduction to Measurement

and Inspection Process, Documents and Metrics.

**UNIT-II:**

*Software Quality Metrics*

Product Quality Metrics: Defect Density, Customer Problems Metric, Customer Satisfaction

Metrics, Function Points, In-Process Quality Metrics: Defect Arrival Pattern, Phase-Based

Defect Removal Pattern, Defect Removal Effectiveness, Metrics for Software Maintenance:

Backlog Management Index, Fix Response Time, Fix Quality, Software Quality Indicators.

**UNIT-III:**

*Software Quality Management and Models*

Modeling Process, Software Reliability Models: The Rayleigh Model, Exponential Distribution and Software Reliability Growth Models, Software Reliability Allocation Models, Criteria for

Model Evaluation, Software Quality Assessment Models: Hierarchical Model of Software Quality Assessment.

**UNIT-IV:**

*Software Quality Assurance*

Quality Planning and Control, Quality Improvement Process, Evolution of Software Quality

Assurance (SQA), Major SQA Activities, Major SQA Issues, Zero Defect Software, SQA

Techniques, Statistical Quality Assurance, Total Quality Management, Quality Standards and Processes.

**UNIT-V:**

*Software Verification, Validation & Testing:*

Verification and Validation, Evolutionary Nature of Verification and Validation, Impracticality of Testing all Data and Paths, Proof of Correctness, Software Testing, Functional, Structural and

Error-Oriented Analysis & Testing, Static and Dynamic Testing Tools, Characteristics of Modern Testing Tools.

**Books/References:**

1. Jeff Tian, **Software Quality Engineering (SQE)**, Wiley

2. Stephen H. Kan, **Metrics and Models in Software Quality Engineering**, Addison-Wesley

**IT712:** **Distributed System**

**L-T-P: 3-0-0**

**Credits: 03**

**Unit–I**

Characterization of Distributed Systems: 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.

**Unit-II**

Distributed Mutual Exclusion: 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.

**Unit–III**

Agreement Protocols: 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.

**Unit–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.

**Unit –V**

Transactions and Concurrency Control: 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.

**Text Books / References:**

1. Mukesh Singhal and Niranjan Shivaratri, **Advanced Concepts in Operating Systems**, McGraw-Hill.
2. Coulouris, Dollimore, Kindberg, **Distributed System: Concepts and Design**, Pearson
3. Education Andrew S. Tanenbaum, **Distributed Operating Systems**, ACM Press.
4. Jie Wu, Distributed Systems, CRC Press.
5. Hagit Attiya, Jennifer Welch, **Distributed Computing: Fundamentals, Simulations and Advanced Topics**, McGraw-Hill.
6. Sape Mullender (ed.), **Distributed Systems**, Addison-Wesley.

**IT713: IMAGE PROCESSING AND PATTERN RECOGNITION**

**L-T-P:3-0-0**

**Credits :03**

**UNIT I:**

*Basic Concepts:*

Pattern Recognition Systems, Fundamental Problems in pattern recognition system design, Design concepts and Methodologies: Character recognition, Speech recognition, Finger print Recognition. Pattern Recognition Model  
  
**UNIT II:**

*Decision Functions:*

Linear Decision functions, Distance functions. Minimum distance and Maximum distance classification, clustering concepts, Cluster seeking algorithms, K- means Algorithms.  
  
**UNIT III:**

*Baye's Classifier:*

Baye’s classified decision function for Baye’s classifier, Baye’s Classifier for normal patterns. Trainable pattern classifiers — deterministic approach, perception, and approach - reward— punishment concept  
  
**UNIT IV:**

*Gradient Approach:*

Gradient approach, Gradient Descent algorithms, LMSE Algorithms, Multi category classification.  
  
**UNIT V:**

*Trainable Pattern Classifiers*

Trainable pattern classifiers, statistical approach, stochastic approximation methods, Robbin Minro algorithms, increment correction algorithms, LMSE algorithms. Syntactic patter recognition, formulation — syntax directed recognition — picture descript.  
  
**UNIT VI:**

*Digital Image Fundamentals*

A simple image model, Sampling and Quantization, Imaging Geometry, Digital Geometry, Image Acquisition Systems, Different types of digital images.

**UNIT VII:**

*Bilevel Image Processing*

Basic concepts of digital distances, distance transform, medial axis transform, component labeling, thinning, morpho-logical processing, extension to grey scale morphology.

**UNIT VIII:**

*Binarization And Segmentation Of Grey Level Images*

Histogram of grey level images, Optimal thresholding using Bayesian classification, multilevel thresholding, Segmentation of grey level images, Water shade algorithm for segmenting grey level image.

**UNIT** **IX:**

*Detection Of Edges And Lines In 2d Images*

First order and second order edge operators, multi-scale edge detection, Canny's edge detection algorithm, Hough transform for detecting lines and curves, edge linking.

**UNIT X:**

*Images Enhancement*

Point processing, Spatial Filtering, Frequency domain filtering, multi-spectral image enhancement, image restoration.

**Text Books / References Books:**

1. R.C.Gonzalez and R.E.Wood, **Digital Image Processing**, Addison Wesley.

2. J.T. Tou, R.C. Gonzalez, **Pattern Recognition Principles**, Addison Wesley.

3. Anil Ku Jain, **Fundamentals of Digital Image Processing**, PHI.

4. B.Channda and D.Dutta, **Digital Image Processing and Analysis**, Prentice Hall.

5. Richard O. Duda, Peter E. Hart and David G. Stork, **Pattern Classification**, 2nd Edition,

John Wiley, 2006.

6. C. M. Bishop, **Pattern Recognition and Machine Learning**, Springer, 2009.

7. S. Theodoridis and K. Koutroumbas, **Pattern Recognition**, 4th Edition, Academic Press,

2009.demic Press, 2002

**IT714: Advanced Computer Architecture**

**L-T-P:3-0-0**

**Credits :03**

**UNIT** **I:**

Fundamentals of computer design, measuring and reporting performance.

Principles of linear pipelining; Instruction level parallelism and instruction pipelines, speedup, data dependency hazards, remedial measures, branch handling; Arithmetic pipelines; Pipeline control methods; Job sequencing, collision prevention and pipeline chaining; Case study of pipelined systems.

**UNIT** **II:**

Loop unrolling, software pipelining and trace scheduling techniques for exposing instruction level parallelism.

Dynamic scheduling algorithms, exploiting ILP using static scheduling and dynamic scheduling, hardware based speculation, multiple issues, and speculation.

**UNIT** **III:**

Vector processing characteristics and requirements, pipelined vector processing, vectorization methods, examples of vector processing.

Array processing, SIMD array processors, communication between PEs, SIMD interconnection networks, algorithms for array processing.

**UNIT** **IV:**

Data and control parallelism, concurrency, scalability, speedup and Amdahl’s law, PRAM model of parallel computation, parallel algorithms.

Multiprocessors and multi-computers; Processor organizations: mesh, binary tree, hypercube; Shared memory and message passing systems; Mapping and Scheduling: Embedding of task graphs in processor graphs, dilation and loading, load balancing, models for static

and dynamic scheduling.

**UNIT** **V:**

Overview of parallel programming using MPI and Open MP.

**Text Books/References:**

1. Hennessy, J. L. and Patterson, D. A., “**Computer Architecture**”, 4thEd., Morgan Kaufmann. 2007

2. Sima, D., Fountain, T. and Kacsuk, P., “**Advanced Computer Architecture: A Design Space**

**Approach**”, Pearson Education. 2007

3. Michael, J.Q., “**Parallel Computing: Theory and Practice**”, Tata McGraw-Hill.2002

4. Hwang, K., “**Advanced Computer Architecture**”, Tata McGraw-Hill. 2003

**IT715: ARTIFICIAL INTELLIGENCE**

**L-T-P:3-0-0**

**Credits :03**

**UNIT** **I:**

*Introduction:*

AI problems, foundation of AI and history of AI intelligent agents: Agents and Environments, the concept of rationality, the nature of environments, structure of agents, problem solving agents, problem formulation.

**UNIT** **II:**

*Searching:*

Searching for solutions, uniformed search strategies – Breadth first search, depth first search, Depth limited search, Iterative deepening depth first search bi-direction search - comparison. Search with partial information (Heuristic search) Greedy best first search, A\* search, Memory bounded heuristic search, Heuristic functions.

Local search Algorithms, Hill climbing, simulated annealing search, local beam search, and genetic algorithms.

**UNIT** **III:**

*Game Playing:*

Adversial search, Games, minimax, algorithm, optimal decisions in multiplayer games, Alpha-Beta pruning, Evaluation functions, cutting of search.

**UNIT** **IV:**

*Knowledge Representation & Reasons logical Agents:*

Knowledge – Based Agents, the Wumpus world, logic, propositional logic, Resolution patterns in propositional logic, Resolution, Forward & Backward Chaining.

First order logic. Inference in first order logic, propositional Vs. first order inference, unification & lifts forward chaining, Backward chaining, Resolution.

**UNIT** **V:**

*Planning:*

Classical planning problem, Language of planning problems, Expressiveness and extension, planning with state – space search, Forward states spare search, Backward states space search, Heuristics for stats space search. Planning search, planning with state space search, partial order planning Graphs.

**UNIT** V**I:**

*NLP:*

Introduction to Natural language processing and expert systems

**UNIT** **VII:**

*Learning:*

Forms of learning, Induction learning, Learning Decision Tree, Statistical learning methods, learning with complex data, learning with Hidden variables – The EM Algorithm, Instance Based learning, Neural Networks.

**TEXT BOOKS / REFERENCES :**

1. **Artificial Intelligence – A Modern Approach**. Second Edition, Stuart Russel, Peter Norvig, PHI/Pearson Education.

2. **Artificial Intelligence**, 3rd Edition, Patrick Henry Winston., Pearson Edition.

3**. Artificial Intelligence** , 2nd Edition, E.Rich and K.Knight (TMH).

4. **Artificial Intelligence and Expert Systems** – Patterson PHI.

5. **Expert Systems: Principles and Programming**- Fourth Edn, Giarrantana/ Riley, Thomson.

6. **PROLOG Programming for Artificial Intelligence**. Ivan Bratka- Third Edition – Pearson

**IT 716: PARALLEL COMPUTING**

**Credits: 03**

**L-T-P: 3-0-0**

**Unit I:**

Computational demands, advantages of parallel systems. Flynn’s classification, controlled parallelism and scalability. Topologies: Mesh, binary tree, Hyper tree, Cube Connected cycles, shuffle-Connected Exchange; Uniform Memory Access (UMA & Non uniform Memory Access (NUMA) Multi processor System. PARAM Model of Parallel Computation,

**Unit II:**

PARAM Algorithms; Parallel Reductions, Prefix sum, List Ranking, Merging of Two Sorted List.

**Unit III:**

Mapping and Scheduling; mapping of Data from Topology to other (Ring to 2-D Mesh, Binomial trees to 2-D mesh, Rings & mesh into 2-D Mesh, Ring & Mesh into Hypercubes), Load balancing, Static scheduling on UMA multi processor systems.

**Unit IV:**

Applications of parallel computing: Matrix Multiplication, Sorting (bitonic Merge sort, parallel quick sort, hyper quick sort), Searching a Graph (P-depth search, Breadth-Depth Search, Breath first search), parallel Brach and bound algorithms

**Text Books and References:**

1. Michel J. Quinn, ― **Parallel Computing: Theory and Practice**,‖ McGraw-Hill

2. Kai Hwang, ―**Advanced Computer Architecture**,‖ McGraw-Hill.

**IT 717: GRAPH THEORY**

**L-T-P:3-0-0**

**Credits :03**

**Unit I**

*Graph:*

Incidence and degree; Handshaking Lemma; Isomorphism; Subgraphs and Union of graphs; Connectedness; Walks, Paths and Circuits; Components and Connectedness; Walks, Walks, Paths and Circuits; Components and Connectedness algorithms; Shortest Path Algorithms, Eulerian graph, Fleury’s algorithm and Chinese postman problem; Hamiltonian graph - necessary and sufficient conditions; Traveling salesman; Bipartite graph.

**Unit II**

*Tree:*

Properties of trees; Pedant vertices in a tree; Center of a tree; Rooted binary trees; Spanning trees - Spanning tree algorithms; Fundamental circuits; Spanning trees of a weighted graph; cut-sets and cut-vertices; Fundamental cut-sets; Connectivity and separativity; network flow; max-flow min-cut theorem.

**Unit III**

*Planner graph:*

Combinatorial and geometric dual; Kuratowski’s graph; detection of planarity; Thickness and crossings.

**Unit IV**

*Matrix representations of graph:*

Incidence; Adjacency; matrices and their properties.

**Unit V**

*Colourings:*

Chromatic number: Chromatic polynomial; The six and five colour theorems; The four colour problem.

**Unit VI**

*Directed graphs:*

Binary relations; Directed graphs and connectedness; directed trees; Aborecence; Polish method; Tournaments.

**Unit VII**

*Counting of labeled trees:*

Cayley’s theorem; Counting methods; Polya theory.

**Text Books / References :**

1. Deo, N.: **Graph Theory with Applications to Engineering and Computer Science**.

2. Harary : **Graph Theory**, PHI (EEE)

**IT 718:** **Bioinformatics**

**L-T-P:3-0-0**

**Credits :03**

**Unit I:**

Bioinformatics objectives and overviews, Interdisciplinary nature of Bioinformatics, Data integration, Data analysis, Major Bioinformatics databases and tools. Metadata: Summary & reference systems, finding new type of data online. Molecular Biology and Bioinformatics: Systems approach in biology, Central dogma of molecular biology, problems in molecular approach and the bioinformatics approach, oerview of the bioinformatics applications.

**Unit II:**

Basic chemistry of nucleic acids, Structure of DNA, Structure of RNA, DNA Replication, -Transcription, -Translation, Genes- the functional elements in DNA, Analyzing DNA,DNA sequencing. Proteins: Amino acids, Protein structure, Secondary, Tertiary and Quaternary structure, Protein folding and function, Nucleic acid-Protein interaction.

**Unit III:**

Perl Basics, Perl applications for bioinformatics- Bioperl, Linux Operating System, mounting/unmounting files, tar, gzip / gunzip, telnet, ftp, developing applications on Linux OS, Understanding and Using Biological Databases, Overview of Java, CORBA, XML, Web deployment concepts.

**Unit IV:**

Genome, Genomic sequencing, expressed sequence tags, gene expression, transcription factor binding sites and single nucleotide polymorphism. Computational representations of molecular biological data storage techniques: databases (flat, relational and object oriented), and controlled vocabularies, general data retrieval techniques: indices, Boolean search, fuzzy search and neighboring, application to biological data warehouses.

**Unit V:**

Macromolecular structures, chemical compounds, generic variability and its connection to clinical data. Representation of patterns and relationships: sequence alignment algorithms, regular expressions, hierarchies and graphical models, Phylogenetics. BLAST.

**Text Books/ References:**

1. D E Krane& M L Raymer, ” **Fundamental Concepts of Bioinformatics**”, Perason Education.

2. Rastogi, Mendiratta, Rastogi, “**Bioinformatics Methods & Applications, Genomics,**

**Proteomics & Drug Discovery**” PHI, New Delhi

3. ShubhaGopal et.al.“ **Bioinformatics: with fundamentals of genomics and proteomics**”, Mc

Graw Hill.

4. O’Reilly, “ **Developing Bio-informatics Computer Skills**”, CBS

5. Forsdyke, “**Evolutionary Bioinformatics**”, Springer