



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

CURRICULUM AND CREDIT FRAMEWORK FOR M.SC. IN MATHEMATICS WITH ARTIFICIAL INTELLIGENCE & DATA SCIENCE (2YEARS)

Semester I			
COURSE CODE	TYPE	COURSE TITLE	CREDITS
B260701T	CORE	Linear Algebra	4
B260702T	CORE	Probability and Statistics	4
B260703T	CORE	Artificial Intelligence	4
B260704T	CORE	Mathematics for Data Science	4
B260705P	Practical-1	Lab based on the courses B260701T to B260704T	4
TOTAL			20

Semester II			
COURSE CODE	TYPE	COURSE TITLE	CREDITS
B260801T	CORE	Numerical Analysis	4
B260802T	CORE	Discrete Mathematics	4
B260803T	CORE	Operation Research	4
B260804T	CORE	Data Structure	4
B260805T B260806T B260807T B260808T	Elective-1	Any one from DSE Courses-I 1. Cyber Security 2. Graph Theory 3. Big Data Analytics	4
TOTAL			20

NOTE: Total credits after completion of 1 year=40

Deccan

Handwritten signature

Handwritten signature

Handwritten signature



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Semester III

COURSE CODE	TYPE	COURSE TITLE	CREDITS
B260901T	CORE	Database Management Systems	4
B260902T	CORE	Introduction to Machine Learning	4
B260903T	CORE	Design and Analysis of Algorithms	4
B260904P	Practical-2	Lab based on the courses B260901T to B260903T	4
B260905R	Project	Master's Dissertation	4
TOTAL			20

Semester IV

COURSE CODE	TYPE	COURSE TITLE	CREDITS
B261001T	CORE	Cryptography & Data Privacy	4
B261002T	CORE	Deep Learning & Neural Networks	4
B261003T	CORE	SWAYAM Course	4
B261004T	Elective-2	Any one from DSE Courses-II 1. Reinforcement Learning (Mathematical Approach) 2. AI in Cybersecurity & Adversarial Machine Learning 3. Natural Language Processing (NLP) & AI Applications 4. Data Mining & Warehousing	4
B261005T			
B261006T			
B261007T			
B261008R	Project	Master's Dissertation / Research Thesis	4
TOTAL			20
Total credits after completion of 2 nd year			40+40=80

Deeath

[Handwritten signature]

[Handwritten signature]

[Handwritten signature]



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Discipline Specific Elective Papers (DSE-I): (Credit 04 each) (1 paper to be selected for III semester)

1. Cyber Security (4)
2. Graph Theory (4)
3. Big Data Analytics (4)

Discipline Specific Elective Papers (DSE-II): (Credit 04 each) (1 paper to be selected for IV semester)

5. Reinforcement Learning (Mathematical Approach) (4)
6. AI in Cybersecurity & Adversarial Machine Learning (4)
7. Natural Language Processing (NLP) & AI Applications (4)
8. Data Mining & Warehousing (4)

Semester I

Course Code- B260701T

Course Name- Linear Algebra

(Credits: 04)

Course Objectives:

CO1: Find rank, nullity of linear transformation and its row space and column space.

CO2: Understand notion of dual space and dual of dual space.

CO3: Understand concepts of bilinear forms, adjoint operators and spectral theorem.

CO4: Find geometric and algebraic multiplicity of Eigen values and its relation with diagonalization of matrix.

CO5: Apply the above concepts to other disciplines.

Unit I:

Review of matrices and system of equations, Vector spaces, subspaces, linear dependence, basis, dimension, dual space, quotient space.

Unit II:

Algebra of linear Transformation, representation of linear transformations by matrices, eigen values and eigen vectors, invariant subspaces, annihilating polynomials, triangulation and diagonalization.

Unit III:

Primary decomposition theorem, rational and Jordan form, inner product spaces, orthonormal bases, Gram-Schmidt orthogonalization process.

Unit IV:



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Linear functionals, adjoint, self adjoint, normal and unitary operators, spectral theorem for normal operators.

Reference Books:

- K. Hoffman and R. Kunze, Linear Algebra, PHI, 1996.
- S. Axler, Linear Algebra Done Right, UTM, Springer 1997.
- G.C. Cullen, Linear Algebra with Applications, Addison Wesley 1997.
- P. R. Halmos, Finite dimensional vector spaces, Springer Verlag, New York, 1987.

Course Code- B260702T

Course Name- Probability and Statistics

(Credits: 04)

Course Objectives:

- CO1: Find the probability of real life problems.
CO2: Understand notion of random variables.
CO3: Understand concepts of Discrete and Continuous Distributions.
CO4: Find the Correlation, Regression and Calculation of covariance.
CO5: Apply the above concepts to other disciplines.

Unit I:

Probability Functions and Moment Generating Function Basic notions of probability, Conditional probability and independence, Baye's theorem.

Unit II:

Random variables - Discrete and continuous, Cumulative distribution function, Probability mass/density functions; Mathematical expectation, Moments, Moment generating function, Characteristic function.

Unit-III:

Univariate Discrete and Continuous Distributions Discrete distributions: Uniform, Bernoulli, Binomial, Negative binomial, Geometric and Poisson; Continuous distributions: Uniform, Gamma, Exponential, Chi-square, Beta and normal; Normal approximation to the binomial distribution.

Unit-IV:

Correlation, Regression and Central Limit Theorem, Correlation coefficient, Covariance, Calculation of covariance from joint moment generating function, Independent random variables, Linear regression for two variables, method of least squares, Bivariate normal distribution, Chebyshev's theorem, Strong law of large numbers, Central limit theorem and weak law of large numbers.

Reference Books:



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

- Robert V. Hogg, Joseph W. McKean & Allen T. Craig (2013). Introduction to Mathematical Statistics (7th edition), Pearson Education.
- Irwin Miller & Marylees Miller (2014). John E. Freund's Mathematical Statistics with Applications (8th edition). Pearson. Dorling Kindersley Pvt. Ltd. India.
- Jim Pitman (1993). Probability, Springer-Verlag.
- Sheldon M. Ross (2014). Introduction to Probability Models (11th edition). Elsevier.
- A. M. Yaglom and I. M. Yaglom (1983). Probability and Information. D. Reidel Publishing Company. Distributed by Hindustan Publishing Corporation (India) Delhi.

Course Code- B260703T

Course Name- Artificial Intelligence

(Credits: 04)

Course Objectives:

CO1: Learn the differences between optimal reasoning vs human like.

CO2: Understand basic principles of AI towards problem solving, inference perception, knowledge representation and learning.

CO3: Understand the notions of state space representation, heuristic search, time and space complexities.

CO4: Understand the applications of AI namely Intelligent Agents, Game Play, Expert Systems, Machine Learning and NLP.

CO5: Apply the above concepts to other disciplines.

Unit I:

Introduction: What is AI? Foundations of AI, History of AI, Agents and Environments, the nature of the Environment, Problem Solving Agents, Problem Formulation, Search Strategies.

Unit II:

Knowledge and Reasoning: Knowledge-based Agents, Representation, Reasoning and Logic, Propositional logic, First-order logic, Using First-order Logic, Inference in First-order logic, forward and Backward Chaining.

Unit III:

Learning: Learning from Observations, Forms of Learning, Inductive Learning, Learning Decision Trees, Why Learning Works, Learning in Neural and Belief networks.

Unit IV:



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Practical Natural Language Processing: Practical applications, Efficient parsing, scaling up the lexicon, Scaling up the Grammar, Ambiguity, Perception, Image formation, Image processing operations for Early vision, Speech recognition and Speech Synthesis.

Reference Books:

- Stuart Russell, Peter Norvig, —Artificial Intelligence: A Modern Approach, 2nd Edition, Pearson Education, 2007.
- B. Yagna Narayana, —Artificial Neural Networks, PHI, 2005.
- 3. E. Rich and K. Knight, —Artificial Intelligence, 3rd Edition, TMH, 2017.
- Dan W. Patterson, —Artificial Intelligence and Expert Systems, PHI, 2015.
- Giarrantana, Riley, —Expert Systems: Principles and Programming, 4th Edition, Course Technology Inc, 2004.
- Ivan Bratka, —PROLOG Programming for Artificial Intelligence, Pearson Education, 3rd Edition, 2012.

Course Code- B260704T

Course Name- Mathematics for Data Science

(Credits: 04)

Course Objectives:

CO1: Understanding of the Mathematical and Logical Basis to many Modern Techniques in the Technology like Data Science, Artificial Intelligence, Machine Learning and Programming Language Design etc.

CO2: Understand the basic concepts of sets, vector space, subspace, basis and dimension.

CO3: Understand important characteristics of Matrices, Eigen Values, Eigen Vectors and Vector Spaces etc.

CO4: Learn how to analyze and solve a linear system of equations.

CO5: Check linear dependency of vectors and identify Eigen values and Eigen Vectors
Derivative of Matrix, which will form the basis for Principal Component Analysis.

Unit I:

Set Theory: Binary Operations, Functions and Relations, Recurrence relations and Generating Functions, Permutations and Combinations: Pigeon Hole Principle, Propositional Calculus, Tautology & Contradiction, Boolean Algebra, Congruence, Euler's phi function.

Unit II:

Differentiation: Logarithmic Differentiation, Partial differentiation, Euler's Theorem for Homogenous Functions, Maxima and Minima. Integration: Double Integration, Change of Order of Integration.

Unit III:



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Vectors: Dot Product, Cross Product, Triple product, Scalar and Vector Fields, Vector Differentiation, Gradient, Divergence and Curl, Vector Integration.

Unit IV:

Matrices & Determinants: Matrix, Def., types, Addition, Subtraction, Multiplication of Matrices, Singular and Non-Singular Matrices, Rank of a Matrix, Solution of Simultaneous Equations, Cayley Hamilton Theorem, Eigen Values & Eigen Vectors.

Reference Books:

- Seymour Lipschutz, Marc Lipson, —Linear Algebra, 6th Edition, Schaum Series, 2018.
- Seymour Lipschutz, Marc Lipson, H. Patil, —Discrete Mathematics, 3rd Edition, Schaum Series, 2017.
- Elliott Mendelson, Frank Ayres, —Calculus, 6th Edition, Schaum Series, 2012.
- Philip N. Klein, —Coding the Matrix: Linear Algebra Through Applications to Computer Science, Newtonian Press, 2013.
- Sheldon Axler —Linear Algebra Done Right, 3rd Edition, Springer, 2015.

B260705P Practical - 1

Lab based on the courses B260701T to B260704T

Credits:04

Semester II

Course Code- B260801T

Course Name- Numerical Analysis

(Credits: 04)

Course Objectives:

CO1: To solve the algebraic and transcendental Equation.

CO2: To solve the linear system of equation using different numerical methods.

CO3: To understand the interpolation techniques.

CO4: To able to integrate the functions using Numerical Methods.

CO5: To solve the various differential equations using Numerical Methods.

Unit I:

Roots of transcendental equations and polynomial equations, Bisection method, Iteration based on first degree equations, Regula-Falsi methods, Rate of convergence, Generalized Newton- Raphson method, Computation of errors.

Unit II:

System of linear equation: Direct method-: Gauss Elimination method, Triangularization method, Iterative methods-: Jacobi's method, Gauss-Seidel method, SOR method,

Dr. Anshu

Madhvi

Dr. Anshu

Dr. Anshu



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Givens power method for Eigen value and Eigenvectors.

Unit III:

Interpolation and Approximation Lagrange's and Newton's divided difference, Finite difference operators, Hermite interpolation, piecewise & cubic spline interpolation, Least square approximation.

Unit IV:

Newton Cotes Quadrature formula, trapezoidal rule, Simpson's 1/3 rule, Simpson's 3/8 rule, Weddle's rule, Numerical Methods for ODE: Single step method -Euler's method, Taylor series method, Runge- Kutta method of 2nd and 4th order, Numerical methods for BVP, Multi step method-predictor- corrector method, Adams Bash forth method.

Reference Books:

- Gerald, C. Fand Wheatly, P.O, Applied Numerical Analysis", 6th edition, Wesley, 2002
- Jain, M.K, Iyengar, S.R.K and Jain, R.K, "Numerical methods for Scientific and Engineering computation, New Age Pvt. Pub, New-Delhi, 2000
- Introduction to Numerical Analysis by S.S Sastry, Prentice Hall, Flid
- Krishnamurthy, E.V & Sen, S.K, Applied Numerical analysis, East West Publication

Course Code- B260802T

Course Name- Discrete Mathematics

(Credits: 04)

Course Objectives:

CO1: Have the knowledge of Fibonacci sequence, linear recurrence relations with constant coefficients.

CO2: Construct generating function and study its application to counting and in solving recurrence relations

CO3: Simplify logic and Boolean circuits using K-maps.

CO4: Find principle disjunctive & conjunctive normal forms and application of inference theory.

Unit I:

Logic: Introduction to logic, Rules of Inference, Validity of arguments, Normal forms, Direct and Indirect proofs, Proof by contradiction, Recurrence relations with examples of Fibonacci numbers, Difference equation, generating function, solution of recurrence relation using generating functions.

Unit II:

Decalhi

[Handwritten signature]

[Handwritten signature]

[Handwritten signature]



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Definition and types of relations, representing relations using digraphs and matrices, closure of relations, paths in digraph, Transitive closure using Warshall's algorithm, Posets, Hasse diagram, Lattices.

Unit III:

Boolean algebra and Boolean functions, different representations of Boolean function, application to synthesis of circuits, circuit minimization and simplification, Karnaugh map.

Unit IV:

Automata theory, Finite state automaton, Types of automaton, Deterministic finite state automaton, Non-deterministic finite state automaton, Non-deterministic finite state automaton with, Equivalence of NFA and DFA, Equivalence of NFA and NFA-ε, Equivalence of NFA-ε and DFA, Finite state machines: Moore and Mealy machine and their conversion, Turing machine.

Reference Books:

- C.L Liu, Elements of Discrete Mathematics, Tata McGraw- Hill, 2000.
- Kenneth Rosen, WCB McGraw-Hill, 6th edition, 2004.
- J.P Tremblay and R.P Manohar, Discrete Mathematical structures with Application to Computer science, McGraw-Hill (1975).

Course Code- B260803T

Course Name- Operations Research

(Credits: 04)

Course Objectives:

CO1: Formulate and solve the LPP including those that lead to cycling and degeneracy.

CO2: Apply integer programming to the LPP's where integer solution is sought.

CO3: Solve transportation and assignment problems and their importance.

CO4: Apply the above concepts to real life problems.

CO5: Make models using Game Theory and apply in real world.

Unit I:

Origin of OR and its definition, Phases of OR problem approach, Formulation of Linear Programming problems, Graphical solution of LPP, Solution of LPP by Simplex method, Two phase method, Big-M method, Methods to solve degeneracy in LPP.

Unit II:

Concept of duality in LPP, Comparison of solutions of Dual and Primal, Dual Simplex method, Sensitivity Analysis, Integer Programming.

Deo

Mohit

Pratishtha



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Unit III:

Mathematical formulation of Transportation problem, Tabular representation, Methods to find initial basic feasible solution, Optimality test, Method of finding Optimal solution, Degeneracy in Transportation problem, Unbalanced Transportation problem, Mathematical formulation of Assignment problem, Hungarian Assignment method.

Unit IV:

Theory of Games: Introduction, Two-Person Zero-Sum Games, Saddle point, Maximin-Minimax Criteria for Optimal Strategy, Minimax Theorem, Principle of Dominance, Graphical Method, Arithmetic Method, Game without Saddle Points- Mixed Strategies, Solution of Games by LPP.

Reference Books:

- Rao, S.S, Optimization theory and applications, 2nd edition, Willey Eastern Ltd., New-Delhi.
- Hiller, F.S and Liberman, Introduction to Operations Research, 6th Ed. McGraw- Hill, International Edition, Industrial Engg. Series, 1995.
- Taha, H.A, Operations Research, An Introduction, 8thEd, Prentice Hall Publishers.
- Gupta, P.K, Hira, D.S, Operations Research, S. Chand & Company Pvt. Ltd.
- Sharma, S. D, Operations Research, Kedar Nath Ram Nath and Co. Meerut, 2002.

Course Code- B260903T

Course Name- Data Structures

(Credits: 04)

Course Objectives:

CO1: Identify and apply appropriate data structures to solve problems and improve their efficiency.

CO2: Analyze and understand the complexity of data structures and associated methods.

CO3: provide an overview of data structures and algorithms commonly used in Computer Science.

CO4: solve complex problems by applying appropriate Data Structures and algorithms.

CO5: Apply the above concepts to other disciplines.

Unit – I

Linear Data Structures: Introduction – Abstract Data Types (ADT) – Stack –Queue – Circular Queue – Double Ended Queue – Applications of stack –Evaluating Arithmetic Expressions-Other Applications- Applications of Queue

Linked Lists – Singly Linked List – Circularly Linked List – Doubly Linked lists, Applications of linked list.

Unit – II

Non-Linear Data Structures: Binary Tree – expression trees – Binary tree



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Traversals – Applications of Trees – Huffman Algorithm – Binary Search Tree –

Balanced Trees – AVL Tree – B-Tree – Splay Trees – Heap- Heap operations—Binomial Heaps
Fibonacci Heaps- Hash set.

Unit -III

Graphs: Graphs Terminology, Graph ADT, Representations, Graph Search Methods DFS and BFS, Applications of Graphs-Minimum Cost Spanning tree Using Kruskal's algorithm, Dijkstra's algorithm for Single Source Shortest Path Problem.

Unit – IV

Algorithm Analysis – Asymptotic Notations – Divide and Conquer – Merge Sort –Quick Sort – Binary Search – Greedy Algorithms – Knapsack Problem – Dynamic Programming– Optimal Binary Search Tree –Back Tracking- N Queen's Problem.

Reference Books:

- Mark Allen Weiss,— Data Structures and Algorithm Analysis in C++", 2nd Edition,Pearson, 2004.
- M T Goodrich, Roberto Tamassia, —Data Structures and Algorithms in Java, 6thEdition, Wiley, 2014.
- Sartaj Sahni, —Data Structures, Algorithms and Applications in Java, 2nd Edition,Universities Press, 2005.
- M T Goodrich, Roberto Tamassia, Algorithm Design, John Wiley, 2002.

DSE Courses-I (Any One)

(Choose any one from following 1-4)

1. Course Code- B260805T

(Credit 04 each)

Course Name- Cyber Security

Course Objectives:

CO1. understand the concept of security, types of attack experienced, encryption and authentication for deal with attacks

CO1. understand the concepts of Intrusion prevention, detection and firewall and attack detection and prevention.

CO3. Have the knowledge of plaintext, cipher text, RSA and other cryptographic algorithm, Key Distribution, Intrusion detection, attacks and their prevention.

CO4. Apply the above concepts to other disciplines.

Syllabus:



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Unit-I: Critical characteristics of Information - NSTISSC Security Model - Components of information System - SDLC - Information assurance - Security Threats and vulnerabilities - Overview of Security threats - Security Standards.

Unit-II: Classical cryptography-symmetric Cryptography-Asymmetric Cryptography- Modern Cryptography -Access Control - DRM - Steganography - Biometrics. Cyber laws.

Unit-III: Network Security - Intrusion Prevention, detection and Management - Firewall- Ecommerce Security - Computer Forensics - Security for VPN and Next Generation Networks.

Unit-IV: Host and Application security -Control hijacking, software architecture and a Simple buffer overflow - Common exploitable application bugs, shell code -Buffer Overflow - Side-channel attacks - Timing attacks.

Unit-V: Power analysis, cold-boot attacks, defenses - Malware - Viruses and worms, spyware, key loggers, and botnets; defenses auditing, policy - Defending weak applications- Isolation, sandboxing, virtual machines.

Reference Books:

1. William Stallings, Cryptography and Network Security: Principles and Practice, 6th Edition, PHI, 2014.
2. Michael E. Whitman and Herbert J Mattord, —Principles of Information Security, 6th Edition, Vikas Publishing House, 2017

2. Course Code- B260806T

(Credit 04 each)

Course Name- Graph Theory

Course Objectives:



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

CO1. Students will achieve command of the fundamental definitions and concepts of graph theory.

CO2. Students will understand and apply the core theorems and algorithms, generating examples as needed, and asking the next natural question.

CO3. Students will achieve proficiency in writing proofs, including those using basic graph theory proof techniques such as bijections, minimal counterexamples, and loaded induction.

CO4. Students will work on clearly expressing mathematical arguments, in discussions and in their writing

CO5. Students will become familiar with the major viewpoints and goals of graph theory: classification, extremality, optimization and sharpness, algorithms, and duality **Unit I:**

Syllabus

Unit I (8 Lectures)

Graph and its terminology, Directed and undirected graph, Multi graph, Simple graph, Complete graph, Weighted graph, Planar and non-planar graph, Regular graph, Graph isomorphism and homeomorphism, Euler's formula, Statement and applications of Kuratowski's theorem, Path factorization of a graph, representing graphs in computer system, Coloring of graph.

Unit II (8 Lectures)

Graph connectivity, Konigsberg bridge problem, Eulerian path and Eulerian circuit, Hamiltonian path and Hamiltonian circuit, Shortest path, Dijkstra's algorithm, Paths between the vertices, Path matrix, Warshall's algorithm, Cut point, bridge, cut sets and connectivity, Menger's theorem.

Unit III (8 Lectures)

Tree and related terminology, spanning tree, Finding minimum spanning tree by Kruskal's algorithm and Prim's algorithm, in order, preorder, and post order tree traversals, Binary tree.

Unit IV (8 Lectures)

Flow network, Feasible flows, Multiple sources and multiple sinks, Cut sets in flow network, Relation between flows and cuts, Max flow problem, Max flow min-cut theorem, Matching, Covering, Introduction to CPM and PERT.

Reference Books:

- Graph Theory, F. Harary, Addison- Wesley 1969.
- Introduction to Graph Theory, D. B. West, Prentice Hall 1996.
- Graph Theory and Its Applications, Jonathan Gross and Jay Yellan, CRC 1998.

3. Course Code- B260807T

(Credit 04 each)

Course Name- BIG DATA ANALYTICS

Course Objectives:



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

CO1. Understand Big Data and its analytics in the real world

CO2. Analyze the Big Data framework like Hadoop and

CO3. Design of Algorithms to solve Data-Intensive Problems using the Map Reduce Paradigm

CO4. Design and Implementation of Big Data Analytics using pig and spark to solve data-intensive problems and generate analytics

CO5. Implement Big Data Activities using Hive

SYLLABUS

Unit-1: Essentials of Big Data and Analytics: Data, Characteristics of data and Types of digital data, Sources of data, Working with unstructured data, Evolution, and Definition of big data, Characteristics and Need of big data, Challenges of big data; Overview of business intelligence, Data Science and Analytics, Meaning and Characteristics of big data analytics, Need of big data analytics, Classification of analytics, Challenges to big data analytics, Importance of big data analytics.

Unit-2: Hadoop: Introducing Hadoop, Need of Hadoop, limitations of RDBMS, RDBMS versus Hadoop, distributed computing challenges, History of Hadoop, Hadoop overview, Use case of Hadoop, Hadoop distributors, HDFS (Hadoop Distributed File System), Processing data with Hadoop, Managing resources and applications with Hadoop YARN (Yet Another Resource Negotiator), Interacting with Hadoop Ecosystem.

Unit-III: Mapreduce Programming: Introduction, Mapper, Reducer, Combiner, Partitioner, Searching, Sorting, Compression, Real-time applications using Map Reduce, Data serialization, and Working with common serialization formats, Big data serialization formats.

Unit-IV: HIVE: Introduction to Hive, Hive architecture, Hive data types, Hive file format, Hive Query Language (HQL), User-Defined Function (UDF) in Hive.

Unit-V: PIG: The anatomy of Pig, Pig on Hadoop, Pig Philosophy, Use case for Pig: ETL Processing, Pig Latin overview, Data types in Pig, Running Pig, Execution modes of Pig, HDFS commands, Relational operators, Piggy Bank, Word count example using Pig.

LAB WORK: Implementation of any 2-3 algorithms from each UNIT (selected by the instructor) using Python.

Reference Books:

1. Seema Acharya, Subhashini Chellappan, "Big Data Analytics", 1st Edition, Wiley, 2015.
2. Boris Lubinsky, Kevin I. Smith, Alexey Yakubovich, "Professional Hadoop Solutions", 1st Edition, Wrox, 2013.
3. Chris Eaton, Dirk Deroos et. al., "Understanding Big data", Indian Edition, McGraw Hill, 2015.
4. Tom White, "HADOOP: The Definitive Guide", 3rd Edition, O Reilly, 2012.
5. Vignesh Prajapati, "Big Data Analytics with R and Hadoop", 1st Edition, Packet Publishing Limited, 2013.



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Semester III

Course Code- B260901T

Course Name- DATABASE MANAGEMENT SYSTEMS

(Credits: 04)

Course Objectives:

CO1: Use the basic concepts of Database Systems in Database design.

CO2: Apply SQL queries to interact with Database.

CO3: Design a Database using ER modeling

CO4: Apply normalization on database design to eliminate anomalies.

CO5: Analyze database transactions and can control them by applying ACID properties.

Unit I:

INTRODUCTION: Introduction and applications of DBMS, Purpose of data base, Data, Independence, Database System architecture- Levels, Mappings, Database, users and DBA

DATABASE DESIGN: Database Design Process, ER Diagrams – Entities, Attributes, Relationships, Constraints, keys, extended ER features, Generalization, Specialization, Aggregation, Conceptual design with the E-R model.

Unit II:

THE RELATIONAL MODEL: Introduction to the relational model, Integrity constraints over Relations, Enforcing integrity constraints, Querying relational data, Logical database design: E-R to Relational, Introduction to views, Destroying/altering tables and views.

RELATIONAL ALGEBRA AND CALCULUS: Preliminaries, relational algebra operators, relational calculus – Tuple and domain relational calculus, expressive power of algebra and calculus.

Unit III:

SQL: Basics of SQL, DDL, DML, DCL, structure – creation, alteration, defining constraints – Primary Key, foreign key, unique, not null, check, IN operator, Functions – aggregate functions, Built-in functions – numeric, date, string functions, set operations, sub-queries, correlated sub-queries, Use of group by, having, order by, join and its types, Exist, Any, All, view and its types. Transaction control commands – Commit, Rollback, save point, cursors, stored procedures, Triggers

Unit IV:

TRANSACTIONS MANAGEMENT: Transaction concept, transaction state, implementation of atomicity and durability, concurrent executions, Serializability, recoverability, implementation of Isolation, transaction definition in SQL, testing for Serializability.

CONCURRENCY CONTROL AND RECOVERY SYSTEM: Concurrency control, lock based Protocols, time-stamp based protocols, validation based protocols, multiple granularities. Recovery System – failure classification, storage structure, recovery and atomicity, log- based recovery, Shadow paging, buffer management, failure with loss of non-volatile storage, advanced recovery techniques, remote backup systems.

Dr. Anil K. Singh

Dr. Anil K. Singh

[Handwritten signature]

[Handwritten signature]



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

OVERVIEW OF STORAGE AND INDEXING: Tree structured indexing – intuition for tree indexes, Indexed sequential access method (ISAM), B+ Trees – a dynamic tree structure.

Reference Books:

- Raghurama Krishnan, Johannes Gehrke, Database Management Systems, 3rd edition, Tata McGraw Hill, New Delhi, India.
- Elmasri Navate, Fundamentals of Database Systems, Pearson Education, India.
- Abraham Silberschatz, Henry F. Korth, S. Sudarshan (2005), Database System Concepts, 5th Edition, McGraw-Hill, New Delhi, India.
- Peter Rob, Carlos Coronel (2009), Database Systems Design, Implementation and Management, 7th edition

Course Code- B260902T

Course Name- Introduction to Machine Learning

(Credits: 04)

Course Objectives:

CO1: Explain the Bayesian concepts of and design/implement NBC & GDA as generative classifiers and Bayesian linear and logistic regression as a discriminative model.

CO2: Appreciate the importance of exponential family. Mixture models, sparse linear models, and kernels in supervised learning and apply EM, SVMS, kernels, and Gaussian processes for regression and classification tasks.

CO3: Discuss the theory of and formulate probabilistic graphical models: HMM, MRF, and CRF in order to solve real-life problem

CO4: Illustrate the mathematical theory of perception models and backpropagation algorithm and implement the models/BPA in Python

CO5: Explain the theory of and implement unsupervised machine learning algorithms: dimensionality reduction and clustering algorithms in Python.

Unit-1:

Introduction, generative models for discrete data (Bayesian concept learning, Naïve Bayes' classifier), Gaussian discriminant analysis, Inference in jointly Gaussian distributions, Bayesian statistics, Bayesian linear and logistic regression.

Unit-2:

General linear models and exponential family. Mixture models and EM algorithm, Sparse linear models, Review of SVM, Multiclass SVM, kernels for building generative models, Multiple kernels, kernels for strings, trees, and graphs, Gaussian Processes.

Unit-3:



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Graphical models-DIRECTED Graphical models (Bayesian networks), Markov and Hidden Markov Models, Markov Random fields, Conditional Random fields, Neural Networks- Perception, MLP, and back propagation, Methods of acceleration of convergence of BPA.

Unit-4:

Dimensionality Reduction and Cluster Analysis: Factor analysis, Kernel PCA. Independent Component Analysis, ISOMAP, LLE, feature Selection, DBSCAN Algorithm, DENCLUE, Spectral clustering.

Reference Books:

- Kevin P. Murphy, Machine learning a Probabilistic Perspective, MIT Press, 2012
- Jeeva Jose, Introduction to Machine Learning using Python.
- Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007
- Simon Hay kin, Neural Network a comprehensive foundation, Pearson Education- 1994
- Tom Mitchell, Machine Learning – McGraw Hill Science, 1997.

Course Code- B260808T

(Credit 04 each)

Course Name- DESIGN AND ANALYSIS OF ALGORITHMS

Course Objectives:

- CO1. Identify various Time and Space complexities of various algorithms
- CO2. Understand Tree Traversal method and Greedy Algorithms
- CO3. Apply Dynamic Programming concept to solve various problems
- CO4. Apply Backtracking, Branch and Bound concept to solve various problems
- CO5. Implement different performance analysis methods for non-deterministic algorithms

SYLLABUS

UNIT - I

INTRODUCTION: Algorithm, pseudo code for expressing algorithms, performance analysis-space complexity, time complexity, asymptotic notation- big (O) notation, omega notation, theta notation and little (o) notation, recurrences, probabilistic analysis, disjoint set operations, union and find algorithms.

UNIT - II

DIVIDE AND CONQUER: General method, applications-analysis of binary search, quick sort, merge sort, AND OR Graphs.

GREEDY METHOD: General method, Applications-job sequencing with deadlines, Fractional knapsack problem, minimum cost spanning trees, Single source shortest path problem.

UNIT - III

GRAPHS (Algorithm and Analysis): Breadth first search and traversal, Depth first search and

D. S. Chakraborty

Dr. R. S. Singh

[Handwritten signature]

[Handwritten signature]



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

traversal, Spanning trees, connected components and bi-connected components, Articulation points.

DYNAMIC PROGRAMMING: General method, applications - optimal binary search trees, 0/1 knapsack problem, All pairs shortest path problem, Travelling sales person problem, Reliability design.

UNIT - IV

BACKTRACKING: General method, Applications- n-queen problem, Sum of subsets problem, Graph coloring and Hamiltonian cycles.

BRANCH AND BOUND: General method, applications - travelling sales person problem, 0/1 knapsack problem- LC branch and bound solution, FIFO branch and bound solution.

UNIT - V

NP-HARD AND NP-COMPLETE PROBLEMS: Basic concepts, non-deterministic algorithms, NP-hard and NP-complete classes, Cook's theorem.

Reference Books:

1. Ellis Horowitz, Satraj Sahni, Rajasekharam (2007), Fundamentals of Computer Algorithms, 2nd edition, University Press, New Delhi.
2. R. C. T. Lee, S. S. Tseng, R.C. Chang and T. Tsai (2006), Introduction to Design and Analysis of Algorithms A strategic approach, McGraw Hill, India.
3. Allen Weiss (2009), Data structures and Algorithm Analysis in C++, 2nd edition, Pearson education, New Delhi.
4. Aho, Ullman, Hopcroft (2009), Design and Analysis of algorithms, 2nd edition, Pearson education, New Delhi.

B260904P Practical – 2 Lab based on the courses B260901T to B260903T

Credits:04

B260905R Master's Dissertation

Credits: 04

Deccath

[Handwritten signature]

[Handwritten signature]

[Handwritten signature]



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Semester IV

Course Code- B261001T

Course Name- Cryptography & Data Privacy

(Credits: 04)

Course Objectives:

CO1: Understand fundamental concepts of cryptography.

CO2: Describe the difference among symmetric, asymmetric and public key Cryptography.

CO3: Define basic requirements of cryptography.

CO4: Apply concepts of Encryption & Decryption.

CO5: Describe the process for implementing cryptographic systems.

UNIT 1:

Introduction to cryptology: Monoalphabetic and Polyalphabetic cipher, The Shift Cipher, The Substitution Cipher, The Affine Cipher, The Vigenere Cipher, The Hill Cipher, Cryptanalysis, Some Cryptanalytic Attacks, Stream & Block ciphers, Mode of operations in block cipher.

UNIT 2:

Shannon's Theory of Perfect Secrecy: Perfect Secrecy, Random Numbers, Pseudorandom Numbers. DES & AES: The Data Encryption Standard (DES), Feistel Ciphers, Description of DES, Security analysis of DES, Differential & Linear Cryptanalysis of DES, The Advanced Encryption Standard(AES), Description of AES, analysis of AES, Prime Number Generation: Trial Division, Fermat Test, Carmichael Numbers, Miller Rabin Test

UNIT 3:

Public Key Cryptography: Principle of Public Key Cryptography, *RSA Cryptosystem*, Factoring problem, Cryptanalysis of RSA, Quadratic Residue Problem, Diffie-Hellman (DH) Key Exchange Protocol, Discrete Logarithm Problem (DLP), *ElGamal Cryptosystem*, ElGamal & DH, Algorithms for DLP. Elliptic Curve, Elliptic Curve Cryptosystem (ECC), Elliptic Curve Discrete Logarithm Problem (ECDLP).

UNIT 4:

Cryptographic Hash Functions: Hash and Compression Functions, Security of Hash Functions, Message Authentication Codes.

UNIT 5:

Digital Signatures: Security Requirements for Signature Schemes, Signature and Hash Functions, RSA Signature, ElGamal Signature, Digital Signature Algorithm (DSA), ECDSA.

Reference Books:

- Wenbo Mao, Modern Cryptography: Theory and Practice. Pearson Education, 2004
- W Starling, Cryptography and Network security, Pearson Education, 2004.

Deoath

2
Modette
2019



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

- J Buchmann, Introduction to Cryptography, Springer (India) 2004
- D R Stinson, Cryptography: Theory and Practice. CRC Press, 2000.
- Bruce Schenier, Applied cryptography, John Wiley & Sons, 1996.
- B Forouzan, Cryptography and Network security, Tata McGraw Hill, 2011

Course Code- B261002T

Course Name- Deep Learning & Neural Networks

(Credits: 04)

Course Objectives:

- CO1. Understand the theoretical foundations, algorithms and methodologies of Neural Network.
- CO2. Recognize the characteristics of deep learning models that are useful to solve real-world problems.
- CO3. Design and develop an application using specific deep learning models.
- CO4. Implement different deep learning algorithms.
- CO5. Depict practical knowledge in handling and analysing real world applications.

UNIT 1:

Introduction to Deep Learning: Definition, Applications, Neural Networks, Machine Learning vs Deep Learning, Deep Learning Libraries (Tensor flow, Keras, PyTorch), Types (Supervised Learning, Unsupervised Learning, Reinforcement Learning) and their Comparison. Datasets: Numerical Data, Categorical Data, Data Quality, Data Remediation, Data Pre-processing (Dimensionality Reduction, Feature Transformation, Feature Subset Selection).

UNIT 2:

Neural Networks: Basics, Types, Intuitions, Neurons, Kernels, Biases, Weights, Initialization, Gradient Descent, Heuristics, Training (Holdout Method, K-Fold n Cross-Validation Method, Bootstrap Sampling, Lazy vs Eager Learner), Evaluation (Regression, Classification and Clustering), Perceptrons, Derivatives, Computation graph, Vectorization, Broadcasting, Propagation (Forward and Back), Parameters vs Hyper parameters.

UNIT 3:

Deep Feed Forward Network: Feed-forward Networks, Gradient-based Learning, Hidden Units, Architecture Design, Computational Graphs, Back Propagation, Regularization, Parameter Penalties, Data Augmentation, Multitask Learning, Bagging, Dropout and Adversarial Training and Optimization.

UNIT 4:

Deepest

[Handwritten signature]

[Handwritten signature]

[Handwritten signature]



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Convolution Networks: Convolution Operation, Pooling, Basic Convolution Function, Convolution Algorithm, Unsupervised Features and Neuroscientific for convolution Network. Sequence Modelling: Recurrent Neural Networks (RNNs), Bidirectional RNNs, Encoder Decoder Sequence-to-Sequence Architectures, Deep Recurrent Network, Recursive Neural Networks and Echo State networks.

Reference Books:

- Goodfellow L., Bengio Y. and Courville A., Deep Learning, MIT Press, 2016.
- Patterson J. and Gibson A., Deep Learning: A Practitioner's Approach, O'Reilly First Edition, 2017.
- Haykin S., Neural Network and Machine Learning, Prentice Hall Pearson Third Edition, 2009.

DSE Courses-II (Any One)

(Choose any one from following 1-4)

1. Course Code- B261004T

(Credit 04 each)

Course Name- Reinforcement Learning (Mathematical Approach)

Course Objectives: CO1. Explore a computational approach to learn from the environment.

CO2. Understand key concepts and application of Reinforcement learning

CO3. Understand the theoretical background of the subject.

CO4. Use it in some practical applications.

CO5. Have an idea about different kinds of learning methods.

SYLLABUS

UNIT 1:

Introduction: Problem formulation and motivating examples, Elements of Reinforcement Learning, History of Reinforcement learning, Probability, Eigen Value, Eigen Vector, Steady State.

UNIT 2:

Evaluative Feedback n-Armed Bandit Problem, Action Value Methods, Evaluation vs. Instruction, Incremental Implementation, Tracking non-stationary problems, Problem Definition Agent-Environment Interface, Goals and rewards, Returns and Utility, The Markov Property, Value functions, Optimal value functions, Optimality and approximation.

UNIT 3:

Dynamic Programming Policy Evaluation, Policy improvement, Policy Iteration, Value Iteration, Asynchronous Dynamic Programming. Monte Carlo Methods Monte Carlo



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Policy Evaluation, Estimation of Action Values, Monte Carlo Control, Incremental Implementations.

UNIT 4:

Temporal Difference Learning Classical Conditioning, TD Prediction, Optimality of TD (0), TD(Lambda), Q-Learning, Actor-Critic Method. Function Approximator Grandmother Cell Theory, Function Approximators, Perceptron-Linear Function approximator, Neural Network-Non- Linear Function Approximator, Brief of Genetic Algorithm.

Reference Books:

1. Reinforcement Learning: An Introduction, Richard S. Sutton & Andrew G. Barto.

2. Course Code- B261005T

(Credit 04 each)

Course Name- AI in Cybersecurity & Adversarial Machine Learning

Course Objectives:

- CO1. Understand the concept of security, types of attack experienced.
- CO2. Understand the concepts of encryption and authentication to deal with attacks.
- CO3. Understand the concept of Intrusion prevention detection and firewall
- CO4. Apply concepts of Encryption & Decryption.
- CO5. Understand the concept of attack detection and prevention.

SYLLABUS

UNIT 1:

Critical characteristics of Information - NSTISSC Security Model - Components of information System - SDLC - Information assurance - Security Threats and vulnerabilities - Overview of Security threats - Security Standards.

UNIT 2:

Classical Cryptography - Symmetric Cryptography- Asymmetric Cryptography - Modern Cryptography -Access Control - DRM - Steganography - Biometrics.

UNIT 3:

Network Security - Intrusion Prevention, detection and Management - Firewall - Ecommerce Security - Computer Forensics - Security for VPN and Next Generation Networks.

UNIT 4:











CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Host and Application security -Control hijacking, Software architecture and a Simple buffer overflow - Common exploitable application bugs, shellcode - Buffer Overflow - Side-channel attacks - Timing attacks, power analysis, Cold-boot attacks, defenses - Malware - Viruses and worms, spyware, key Loggers, and botnets; defenses auditing, policy - Defending weak applications - Isolation, sandboxing, virtual machines.

Reference Books:

1. William Stallings, Cryptography and Network Security: Principles and Practice, 6th Edition, PHI, 2014.
2. Michael E. Whitman and Herbert J Mattord, Principles of Information Security, 6th Edition, Vikas Publishing House, 2017.

3. Course Code- B261006T

(Credit 04 each)

Course Name- Natural Language Processing (NLP) & AI Applications

Course Objectives:

- CO1. Acquire basic understanding of linguistic concepts.
- CO2. Acquire basic understanding of natural language complexity, variability.
- CO3. Acquire basic understanding of machine learning techniques as applied to language.
- CO4. Use Naive Bayes Probabilistic language modeling and its Applications
- CO5. Implement N-grams Models

SYLLABUS

UNIT 1:

Introduction: Natural Language Processing, hands-on demonstrations. Ambiguity and uncertainty in language. The Turing test. Regular Expressions, Chomsky hierarchy, regular languages, and their limitations. Finite-state Automata, Practical regular expressions for finding and counting language Phenomena, exploring a large corpus with regex tools. An introduction to Programming in Python- Variables, numbers, strings, arrays, dictionaries, Conditionals, iteration. The NLTK (Natural Language Toolkit) String Edit Distance and Alignment Key algorithmic tool: dynamic programming, a simple Example, use in optimal alignment of sequences. String edit operations, edit Distance, and examples of use in spelling correction, and machine translation.

UNIT 2:

Context Free Grammars Constituency, CFG definition, use and limitations. Chomsky Normal Form, Top-down parsing, bottom-up parsing, Non Probabilistic Parsing Efficient CFG parsing with CYK, another dynamic Programming algorithms. Early parser. Designing a little grammar, and parsing with it on some test data-



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

Examples of applications in natural language. Information Theory The —Shannon game- motivated by language! Entropy, Cross entropy, Information gain. Its application to some language phenomena.

UNIT 3:

Language modelling and Naive Bayes Probabilistic language modeling and its Applications, Markov models, N-grams. Estimating the probability of a word, and smoothing. Generative models of language. Part of Speech Tagging and Hidden Markov Models, Viterbi Algorithm for Finding HMM Path Dynamic programming with Hidden Markov Models, and its use for part-of-speech tagging, Chinese word segmentation, prosody, information Extraction, etc.

UNIT 4:

Probabilistic Context Free Grammars, Weighted context free grammars, Weighted CYK. Pruning and beam search, Parsing with PCFGs, A tree bank, The probabilistic version of CYK, how do humans parse, Experiments with Eye tracking, Modern parsers, Maximum Entropy Classifiers, The maximum Entropy principle and its relation to maximum likelihood, Maximum entropy Classifiers and their application to document classification, sentence Segmentation, and other language tasks. Maximum Entropy Markov Models& Conditional Random Fields Part-of- speech tagging.

Reference Books:

1. Jurafsky and Martin, —Speech and Language Processing, 2nd Edition, PrenticeHall, 2008.
2. Manning and Schütze—Statistical Natural Language Processing, MIT Press, 2001.
3. James Allen, —Natural Language Understanding, The Benajmins /Cummings Publishing Company, 1998.

4. Course Code- B261007T

(Credit 04 each)

Course Name- Data Mining & Warehousing

Course Objectives:

CO1. Identify the scope and necessity of Data Mining& warehousing for the society

Deoathi

[Handwritten signature]

[Handwritten signature]

[Handwritten signature]



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data Science

CO2. Describe various Data models and design methodologies of Data Warehousing destined to solve the root problems

CO3. Understand various tools of Data mining and their techniques to solve the real time problems

CO4. Analyze the data, identify the problems, and choose the relevant algorithms to apply.

CO5. Understand to assess the Pros and cons of various algorithms and analyze their behavior on real datasets.

SYLLABUS

UNIT 1:

Introduction – Steps in KDD – System Architecture – Types of Data- Data Mining functionalities- Classification of Data mining systems –Integration of data mining system with a data warehouse – Issues- Data Processing –Data Mining Application.

UNIT 2:

Data warehousing Components-Building a data warehouse-Multi Dimensional Data Model –OLAP Operation in the Multi-Dimensional Model – Three Tier Data Warehouse Architecture – Schemas for Multidimensional Data Model – Online Analytical Processing (OLAP) –OLAP vs OLTP integrated OLAM and OLAP Architecture.

UNIT 3:

Mining frequent patterns – Associations and correlations – Mining methods – Finding Frequent item set using Candidate Generation – Generating Association rules from Frequent Item sets – Mining Frequent item set without candidate Generation- Mining various kinds of association Rules – Mining Multi level Association rule- Mining Multi-Dimensional Association Rule- Mining Correlation analysis- Constraint based association mining.

UNIT 4:

Mining Classification and prediction – Issues regarding Classification and Prediction – Classification by Decision Tree Induction –Bayesian classification Naïve Bayesian Classification - Linear Regression frequent patterns – Associations and correlations- Cluster analysis – Types of data in Cluster Analysis-Social Impacts of Data Mining-Recent trends in Data Mining.

Reference Books:

1. Sandy Ryza, Uri Laserson, Sean Owen, Josh Wills, Advanced Analytics with Spark: Patterns for Learning from Data at Scale, Oreilly, 2015.



CHHATRAPATI SHAHU JI MAHARAJ UNIVERSITY, KANPUR

STRUCTURE OF SYLLABUS FOR THE

**PROGRAM: M.Sc. in Mathematics with Artificial Intelligence & Data
Science**

B261008R Master's Dissertation / Research Thesis
Credits: 04

Deccan

[Handwritten signature]

Madhwa

[Handwritten signature]