M.Sc. (DATA SCIENCE) COURSE CURRICULUM

(W.E.F. THE ACADEMIC YEAR 2025-26)

Approved by

BOARD OF STUDIES

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

TELANGANA UNIVERSITY,

NIZAMABAD

M.SC. (DATA SCIENCE) COURSE CURRICULUM

(W.E.F. ACADEMIC YEAR 2025-2026)

I-SEMESTER

Paper Code	Title of the paper	Credits	No of Hours per week	Total Marks
Theory Papers				
MDS-101	Mathematical Foundations for Data Science.	3	3	100
MDS-102	Design and Analysis of Algorithms.	3	3	100
MDS-103	Java Programming.	3	3	100
MDS-104	Statistical Inference.	3	3	100
Practical Papers				
MDS-105	Statistical Process for Data Science using R-programming.	2	4	50
MDS-106	Design and Analysis of Algorithms Lab using Python.	2	4	50
MDS-107	Java Programming Lab.	2	4	50
MDS-108	Statistical Inference using Python.	2	4	50
	Total	**20**	**12+16**	**600**

II-SEMESTER

Paper Code	Title of the paper	Credits	No of Hours per week	Total Marks
Theory Papers				
MDS-201	Advanced Machine Learning Techniques	3	3	100
MDS-202	Artificial Intelligence	3	3	100
MDS-203	Optimization Techniques	3	3	100
MDS-204	Software Engineering	3	3	100
Practical Papers				
MDS-206	Advanced Machine Learning Techniques using Python Lab	2	4	50
MDS-207	Artificial Intelligence Lab	2	4	50
MDS-208	Optimization techniques using R	2	4	50
MDS-209	Data Handling using SPSS	2	4	50
	Total	**20**	**24**	**600**

III-SEMESTER

Paper Code	Title of the paper	Credits	No of Hours per week	Total Marks
Theory Papers				
MDS-301	Deep Learning Techniques	3	3	100
MDS-302	Computer Networks	3	3	100
MDS-303 (E-I)	a) Text Data Analytics b) Statistical Pattern Recognition c) Cloud Computing	3	3	100
MDS-304 (E-II)	a) Big Data Analytics b) Image Processing & Analysis c) Web technologies	3	3	100
Practical Papers				
MDS-305	Deep Learning Techniques & Computer Networks using Python	2	4	50
MDS-306	E-I & E-II using Python	2	4	50
MDS-307	Mini Project	4	8	100
	No of Credits	**20**	**12+16**	**600**

IV-SEMESTER

Paper Code	Title of the paper	Credits	No of Hours per week	Total Marks
Theory Papers				
MDS-401	Cryptography	3	3	50
MDS-402	Data Mining	3	3	50
MDS-403 E-III	a) Sentimental Analysis b) Computer Vision c) Machine Learning Operations	3	3	50
MDS-404 E-IV	a) Data Stream mining b) Web Mining c) Scalable Architecture	3	3	50
Practical Papers				
MDS-405	Cryptography & DM	2	4	50
MDS-406	E-I & E-II	2	4	50
MDS-407	Capstone Project	4	8	100
	No of Credits	20	12+16	**600**

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, I-SEMESTER

MDS-101-T: PAPER- I: MATHEMATICAL FOUNDATIONS FOR DATA SCIENCE

UNIT – I

Linear Algebra: Vector spaces, Subspaces, Basis and dimension of a vector space, linear dependence and independence, spanning set. Linear transformation, kernel, range, Matrix Representation of a linear transformation, Matrices: Trace and Rank of a Matrix and their properties, Determinants, Inverse, symmetric, orthogonal and idempotent matrices and their properties, Gauss elimination, row canonical form, diagonal form, triangular form and its Applications, Characteristic roots and vectors, Statement of Caley-Hamilton theorem and its applications, Orthogonal and Spectral decomposition of a real symmetric matrix, Singular value Decomposition.

UNIT – II

Combinatorics: Basic counting principle, inclusion-exclusion for two-sets, pigeonhole principle, permutations and combinations, Binomial coefficient and identities, generalized permutations and combinations. principle of inclusion-exclusion, applications of inclusion-exclusion. Recurrence Relations: introduction, solving linear recurrence relations, generating functions. Real time applications of combinatorial concepts.

UNIT-III

Graph Theory: Basic concepts to Graphs; Isomorphic graphs and simple problems; Trees: definitions, properties, and simple problems, tree traversals, spanning tree constructions: breadth and depth first search, minimal spanning tree constructions: Kruskal's algorithm. Planar and Hamiltonian graphs and simple problems and applications; Graph coloring and its applications. Real time applications of graphical concepts.

REFERENCE BOOKS

1. Gilbert Strang (2016): Introduction to linear algebra, 5/e., Wellesley-Cambridge.
2. David C. Lay (2019): Linear Algebra and Its Applications, Pearson, 5/e.
3. Joe L. Mott, Abraham Kandel, Theoder P. Baker, Discrete Mathematics for Computer Scientists and Mathematicians

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, I-SEMESTER

MDS-102-T: PAPER- II: DESIGN AND ANALYSIS OF ALGORITHMS

UNIT I

Introduction to Algorithms: Algorithm Specification, Performance Analysis, Randomized Algorithms. **Elementary Data Structures**: Stacks and Queues, Trees, Dictionaries, Priority Queues, Sets and Disjoint Set Union, Graphs.

Divide and Conquer: Binary Search, Finding the Maximum and Minimum, Merge Sort; Quick Sort, Selection sort, Strassen's Matrix Multiplication, Convex Hull.

Greedy Method: Knapsack Problem, Job Sequencing with Deadlines, Minimum-Cost Spanning Trees (Kruskal's & Prim's), Single Source Shortest Paths (Dijkstra's).

UNIT-II

Dynamic Programming: General Method, Multistage Graphs, All-Pairs Shortest Paths, Single-Source Shortest Paths, Optimal Binary Search Trees, 0/1 Knapsack, Traveling Salesperson Problem.

Back Tracking technique: General Method, 8-Queens Problem, Sum of Subsets, Graph Colouring, Hamiltonian Cycles, Knapsack Problem.

UNIT -III

Branch-Bound technique: General Method, 0/1 Knapsack Problem, Traveling Sales Person problem.

NP-Hard and NP-Complete Problems: Basic Concepts, Cook's Theorem, NP-Hard and NP-Complete problems. Graph Problems, NP-Hard Scheduling Problems, NP-Hard Code Generation, Some Simplified NP-Hard Problems.

REFERENCE BOOKS

1. E Horowitz, S Sahni, S Rajasekaran, (2007): Fundamentals of Computer Algorithms, 2/e, Universities Press.
2. T.H. Cormen, C.E. Leiserson, R.L Rivert, C Stein, (2010): Introduction to Algorithms, 3/e, PHI.
3. R. Pannerselvam (2007): Design and Analysis of Algorithms, PHI.
4. Hari Mohan Pandey, (2009): Design, Analysis and Algorithm, University Science Press.

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, I-SEMESTER

MDS-103-T: PAPER- III: JAVA PROGRAMMING

UNIT – I

Java Programming Fundamentals: Introduction, Overview of Java, Data Types, Variables and Arrays, Operators, Control statements, Classes, Methods, Inheritance, Packages and Interfaces. I/O basics: Byte & Character Streams, Reading Console input and output, Scanner Class, Console Class, Print Writer Class, String Handling, Exception Handling, Multithreaded Programming. Overview of Networking: Working with URL, connecting to a Server, Implementing Servers, serving multiple Clients, Sending E-Mail, Socket Programming, Internet Addresses, URL Connections. AWT: Introduction, AWT Class Hierarchy, Creating Container, Adding Components, Layout, Using Panel, TextField, TextArea, List, Checkbox, CheckBoxGroup, Choice, EventHandling, DialogBoxes, ScrollBar, Menu.

UNIT – II

Swing: Containment Hierarchy, Adding Components, JTextField, JPasswordField, JTable, JComboBox, JProgressBar, JList, JTree, JColorChooser, Dialogs. Java Database Connectivity (JDBC): Introduction, JDBC Drivers, JDBC Architecture, JDBC Classes and Interfaces, Loading a Driver, Making a Connection, Execute SQL Statement, SQL Statements, Retrieving Result, Getting Database Information, Scrollable and Updatable Resultset, ResultSet Metadata.

UNIT – III

Servlet: Introduction to Servlet, Servlet Life Cycle, advantages, Sharing Information, Initializing a Servlet, Writing Service Methods, Filtering Requests and Responses, Invoking Other Web Resources, Accessing the Web Context, Maintaining Client State, Finalizing a Servlet. Java Server Pages: Introduction to JSP, JSP Engine, Anatomy of a JSP Page, JSP Syntax, Life Cycle of a JSP Page, Creating Static Content, Creating Dynamic Content. .

REFERENCE BOOKS

1. Uttam K. Roy, Advanced Java programming
2. Herbertt Schildt, Java Complete Reference
3. Cay S. Horstmans, Gray Coronell, Core Java Vol. II – Advanced Features
4. Sharanam Shah, Vaishali Shah, Java EE 7 for Beginners

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, I-SEMESTER

MDS-104-T: PAPER- IV: STATISTICAL INFERENCE

UNIT-I

Estimation Theory: Basic concepts to estimation; Concepts, examples, applications and simple problems on criteria for good estimator: Unbiasedness, consistency, efficiency and sufficiency, Cramer-Rao inequality, Rao-Blackwell theorem, Fisher Information, Lehmann-Scheffe theorem. Simple Problems on UMVUE.

Methods of Estimation: Method of Moments, Least squares and Maximum Likelihood, Properties and Simple problems. Resampling methods: Jackknife, Bootstrap, Estimation of bias and standard deviation of point estimation by the Jackknife & Bootstrap methods with examples, U-statistic, Kernal and examples. Interval estimation, confidence level CI using pivots and shortest length CI and example problems.

UNIT-II

Testing of Hypotheses: Neyman-Pearson Lemma, Most Powerful tests, Uniformly Most Powerful tests, Likelihood ratio tests, Sequential Probability Ratio Tests.

Non parametric tests: One and two sample tests (Kolmogorov Smirnov, Kruskal Wallis & Friedman test, Kendal's tau, Ansari broadly tests)

UNIT-III

Non-parametric Density Estimation: Rosenblatt's naïve density estimator, its bias and variance. Consistency of Kernel density estimators and its MSE.

Simulation: Introduction, generation of random numbers for Uniform, Normal, Exponential, Cauchy and Poisson Distributions. Estimating the reliability of the random numbers. Priori and Posteriori distributions, conjugate families, Bayesian estimation of parameters, MCMC algorithms: Metropolis Hasting and Gibbs Sampler.

REFERENCES

1. Rohatgi, V.K.: An Introduction to Probability Theory and Mathematical Statistics (Wiley)
2. Gibbons: Non-Parametric Statistical Inference, (TMH)
3. Lehman, E. L.: Testing of hypothesis, John Wiley
4. Goon, Gupta and Das Gupta: Outlines of Statistics, Vol. II, World Press.
5. C.R. Rao – Linear Statistical Inference (John Wiley)

Instruction to Practical's of M.Sc. Data Science

1. The semesters I & II, each has **Four practical papers,** each has **two credits** with **4 hours of lab** with weightage of **50 marks each**.
2. **Each student has to spend minimum 60 hours** in lab for each practical paper and has to practice on various data sets available in various web sources.
3. Each practical record should contain all practical's mentioned in the syllabus and maintaining of **Practical records is mandatory** (submission at the time examination) irrespective of existence or non-existence of practical record marks allocation.
4. The **statistical analysis report should follow the steps of Data Analysis** depends on the practical (Problem, Data set considered (Data source), Data description, Data objectives, hypothesis framed on population, statistical techniques applicable (as per syllabus of paper), writing of base program code to familiar with computational procedure rather than usage of packages), outputs and results, data interpretation, conclusions on the data set).
5. Each practical record should be written with own hand writing (not computer printouts) and should be taken the signature of the concerned faculty with date of the practical done.
6. The Semester end practical exam question paper contains **answer any two out of the three** questions given in **2 hours duration** (including its implementation)
7. **The practical examination question paper is also common to all the students** of all colleges and conducted and scheduled by Head, Department of Statistics, O.U., Hyderabad, with the appointment of external examiners.
8. **All answers to the questions should be written in the practical answer booklets at the time of practical examination** (including executed program, output, interpretation and conclusions and incase of data analysis all the steps should be written as mentioned above.

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, I-SEMESTER

MDS-105-P: PAPER- V (PRACTICAL-I)

STATISTICAL PROCESS FOR DATA SCIENCE USING R

List of Practical's:

1. Basics of R-programming and R-studio for data handling.
2. Data understanding, data description, Measurement of scales, Data objectives, Formation of Hypothesis's. Sequential steps for writing data analysis report.
3. Evaluation of Data pre-processing steps. Data transformations (Standardize, Normalize, converting data from one scale to other scales etc.).
4. Evaluation of Descriptive Statistics to various measurement of scales data sets.
5. Data Visualization using R: Drawing One dimensional diagram (Pictogram, Pie Chart, Bar Chart), two-dimensional diagrams (Histogram, Line plot, frequency curves & polygons, ogive curves, Scatter Plot), other diagrammatical / graphical representations like, Gantt Chart, Heat Map, Box-Whisker Plot, Area Chart, Correlation Matrices.
6. Correlation Analysis (parametric and nonparametric), Simple and Multiple linear Regression model fitting and its analysis.
7. Testing of Hypothesis-I: Parametric tests (z-, χ^2, t-, F-tests, ANOVA).
8. Testing of Hypothesis-I: Non-Parametric tests (Sign test, Median, Wilcoxon sign rank, Mann-Whitney U, Run test).
9. Statistical analysis for qualitative data.
10. Applying the modelling process, Model evolution, over fitting, under fitting, cross validation concepts, Model Performance (train/test, K fold and leave out one approaches) for qualitative and Quantitative data.

Note: The implementation of the above list of practical's to be applied on the sample data sets available in various web sources and should be practiced by each student. For example, www.kaggle.com . contains thousands of data sets with different measurement of scales; few are: Fishers Iris Dataset; Online food dataset, Wine quality data set, water portability dataset, Heart data set, Protease data set, Mortagaze data set, flights dataset; Sustainable Development Data; Credit Card Fraud Detection; Employee dataset; Heart Attack Analysis & Prediction Dataset; Dataset for Facial recognition; Covid_w/wo_Pneumonia Chest Xray Dataset; Groceries dataset; Financial Fraud and Non-Fraud News Classification; IBM Transactions for Anti Money Laundering.

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, I-SEMESTER

MDS-106-P: PAPER- VI (PRACTICAL-II):

DESIGN & ANALYSIS OF ALGORITHMS LAB USING PYTHON

List of Practical's using python programming (not usage of packages)

1. Divide and Conquer method of implementation of sorting/ searching data / data set using:
   (i) Selection Sort
   (ii) Merge Sort
   (iii) Quick Sort.
   (iv) Construction of Heap, Maintain heap and Heap sort.
   (v) Binary Search.
   (vi) Strassen's Matrix multiplication
2. Greedy method implementation
   (i) Fractional Knapsack problem.
   (ii) Job Sequencing with Deadlines,
   (iii) Minimum-Cost Spanning Trees (Kruskal's & Prim's),
   (iv) Single Source Shortest Paths (Dijkstra's).
3. Dynamic programming technique implementation for
   (i) Travelling sales person problem.
   (ii) Multistage Graph problem,
   (iii) All-Pairs Shortest Paths (Warshal),
   (iv) Single-Source Shortest Paths (Bellman ford),
   (v) Optimal Binary Search Trees.
4. Back tracking technique implementation for
   (i) 0-1 Knapsack problem.
   (ii) 8-Queens Problem,
   (iii) Hamiltonian Graph problem
5. Branch and bound Implementation for
   (i) 0-1 Knapsack problem.
   (ii) Travelling salesman problem.

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, I-SEMESTER

MDS-107-P: PAPER- VII (PRACTICAL-III): JAVA PROGRAMMING LAB

List of Practical's using Java programming

1. Java programs to the Mathematical / Statistical applications to demonstrate core Java and OOPs concepts. (working with functions, classes, abstract class, interfaces, string handling and string buffer classes, user defined packages, event handling, exception handling, inheritance, polymorphism, multithreading, etc, Computation central tendencies, dispersions, moments, skewness, kurtosis, distributions, correlation and regression, Matrices)
2. Developing servlet applications on data bases. (like to accept H.T.No of a student from client and display the memorandum of marks from the server, Question-Answer Application using HttpServletRequest and HttpServletRequest interfaces etc).
3. Create JSP pages that prints (a) temperature conversion (from Celsius to Fahrenheit) chart; (b) current date and time (c). number of times page is referred after the page is loaded. simple JSP application to demonstrate the use of implicit object (at least 5). JSP Application to accept Registration Details from the user and store database table; Accept Registration Details from the user and store database table; Authenticate User Login as per the Registration Details. If Login Success then forward User to Index Page otherwise show Login failure Message; web Application to add items in the inventory using JSP.
4. Create GUI to present a set of choices for a user to select stationary products and display the price of Product after selection from the list; typical Editable Table which describing Employee for a software company; swing components using student registration form.
5. Create a Remote Object for simple arithmetic operators. Use AWT / SWING to create user interface.
6. Develop a Hibernate application to Store Feedback of Website Visitors in MySQL Database.
7. Write EJB applications using stateless session beans and state-full session beans.
8. Develop a Room Reservation System Application using Entity Beans.
9. Create Three-tire application using Servlets, JSP, EJB.

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, I-SEMESTER

MDS-108-P: PAPER- VIII (PRACTICAL-IV): STATISTICAL INFERENCE USING PYTHON

List of Practical's using Python programming

1. Data visualization: Diagrammatical / Graphical representation of the data in the form of dataset with different measurement of scales (Pictorial representation, Bar (simple, multiple, component, percent) and Pie Charts, Histogram, Line plot, frequency curves & polygons, ogive curves, Scatter Plot, Gantt Chart, Heat Map, Box - Whisker Plot, Waterfall Chart, Area Chart, Density Plot, Bullet Graph, Choropleth Map, Tree map, Path diagram, Network Diagram, Correlation Matrices).
2. Correlation and Regression Analysis (including simple (Pearsons and Spearman's), partial and multiple correlations, Simple, Multiple linear regression and logistic regression).
3. Parametric tests (z-, χ^2, t-, F-tests, ANOVA).
4. Non-Parametric tests (Sign test, Median, Wilcoxon sign rank, Mann-Whitney U, Run test, U-test, K-S test, Kruskal Wallis and Friedman test, Independence, goodness of fit, Kendal's tau, Ansari broadly tests).
5. Generation of Jackknife and Bootstrap samples and estimation of parameters and computation of bias.
6. Confidence Interval estimation for Binomial, Poisson, Normal and Exponential parameters.
7. Simulation: Generation of random numbers from various probability distributions (Uniform, Binomial, Poisson, Normal, Exponential, Gamma, Cauchy, Lognormal, and Weibull Distributions).
8. Bayesian estimation of parameters (using Metropolis Hasting and Gibbs Sampler).

M.SC. (DATA SCIENCE) COURSE CURRICULUM

(W.E.F. ACADEMIC YEAR 2024-2025)

M.SC. (DATA SCIENCE) I-YEAR, II-SEMESTER

Paper Code	Title of the paper	Credits	No of Hours per week	Total Marks
Theory Papers				
MDS-201	Advanced Machine Learning Techniques	3	3	100
MDS-202	Artificial Intelligence	3	3	100
MDS-203	Optimization Techniques	3	3	100
MDS-204	Software Engineering	3	3	100
Practical Papers				
MDS-206	Advanced Machine Learning Techniques using Python Lab	2	4	50
MDS-207	Artificial Intelligence Lab	2	4	50
MDS-208	Text Analytics Lab using Python	2	4	50
MDS-209	Data Handling using SPSS	2	4	50
	Total	**20**	**24**	**600**

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, II-SEMESTER

MDS-201: PAPER- I: ADVANCED MACHINE LEARNING TECHNIQUES

UNIT – I

Classification Techniques: Definitions, derivations, methods, properties, and applications of classifier algorithms: Logistic regression, Linear Discriminant Function (for binary outputs) with minimum mean squared error and using likelihood ratios (MVN populations), Bayes Mis-classification; Naïve Bayes classifier, Support Vector Machines, Decision Tree algorithms, Random Forest algorithm, Bootstrap algorithms (Bagging, Gradient, Ada- and XG-Boosting), KNN algorithm, Market-Basket Analysis.

UNIT – II

Multivariate Analysis Techniques: Definitions, goals, derivations, methods, properties, and applications of: Principal component analysis, Factor analysis, Multidimensional Scaling, Canonical Correlations and Canonical Variables, Conjoint Analysis, Path analysis, Correspondence analysis; Feature extraction and Feature selection techniques, Inter and intra class distance measures, Probabilistic distance measures.

UNIT - III

Cluster Analysis: Introduction, similarities and dissimilarities, Hierarchical and non-hierarchical clustering methods, Single, complete and average linkages, K-means, DBSCAN, CART, C4.5 methods.

Association Analysis: Problem definition, Frequent item set generation, Rule generation, Compact representation of frequent item sets, methods for generating frequent item sets, Aprori and FP-tree growth Algorithm.

REFERENCE BOOKS

1. Johnson, R.A, and Dean W. Wichern: Applied Multivariate Statistical Analysis.
2. Morrison, D: An Introduction to Multivariate Analysis.
3. Seber: Multivariate Observations
4. Anderson: An Introduction to Multivariate Analysis.
5. Bishop: Analysis of Categorical data.

SYLLABUS
M.SC. (DATASCIENCE) I-YEAR, II-SEMESTER

MDS-202-T: PAPER- II: ARTIFICIAL INTELLIGENCE

UNIT - 1

Introduction: History Intelligent Systems, Foundations of Artificial Intelligence, Sub areas of Al, Applications. Problem Solving - State - Space Search and Control Strategies: Introduction, General Problem-Solving Characteristics of problem, Exhaustive Searches, Heuristic Search Techniques, Iterative - Deepening A*, Constraint Satisfaction. Game Playing, Bounded Look - ahead Strategy and use of Evaluation Functions, Alpha Beta Pruning.

UNIT – II

Logic Concepts and Logic Programming: Introduction, Propositional Calculus Propositional Logic, Natural Deduction System, Axiomatic System, Semantic Table, A System in Propositional Logic, Resolution, Refutation in Propositional Logic, Predicate Logic, Logic Programming. Knowledge Representation: Introduction, Approaches to knowledge Representation, Knowledge Representation using Semantic Network, Extended Semantic Networks for KR, Knowledge Representation using Frames.

UNIT - III

Expert System and Applications: Introduction, Phases in Building Expert Systems Expert System Architecture, Expert Systems Vs Traditional Systems, Truth Maintenance Systems, Application of Expert Systems, List of Shells and tools. Uncertainty Measure - Probability Theory: Introduction, Probability Theory, Bayesian Belief Networks, Certainty Factor Theory, Dempster - Shafer Theory.

REFERENCE BOOKS

1. Saroj Kaushik, Artificial Intelligence, Cengage Learning India, First Edition, 2011.
2. Russell, Norvig, Artificial Intelligence: A Modern Approach, Pearson Education, 2nd Edition, 2004.
3. Rich, Knight, Nair, Artificial Intelligence, Tata McGraw Hill, 3rd Edition 2009.

SYLLABUS
M.SC. (DATASCIENCE) I-YEAR, II-SEMESTER
MDS-203: PAPER-III: OPTIMIZATION TECHNIQUES

UNIT–I

Meaning and scope of optimization techniques. Convex sets and their properties. General linear Programming Problem (LPP). Formulation of LPP. Statements of Fundamental theorem of LPP and other related theorems. Optimal Solution of LPP by graphical, Simplex and Charner's & two-phase methods. Concept of degeneracy and resolving it Concept of duality of LPP. Dual Primal relationship, Complementary slackness theorems, Fundamental Theorem of Duality. Dual simplex Algorithm.

UNIT-II

Transportation problem: Concept of transportation problem, TPP as a special case of LPP, Initial basic feasible solutions by North-West Corner Rule, Matrix minimum method and VAM. Optimal solution of TPP through MODI method for balanced and unbalanced transportation problems.

Assignment problem: Concept of Assignment problem, Assignment problem as special case of TP and LPP. Unbalanced assignment problem, optimal solution using Hungarian method and traveling salesman problem and its solution.

Problem of Sequencing: Optimal sequence of N jobs on two and three machines without passing.

UNIT-III

Integer Programming Problem: Gomory's cutting plane Algorithm for pure and mixed IPP Branch and bound Technique.

Networks: Basic concepts of Networks constraints; Construction of Network and critical path; PERT and CPM; Network flow problems. Time Cost Analysis.

Queuing Theory: Introduction, essential features of Queuing system, Operating characteristics of Queuing system (transient and steady states). Queue length, General relationships among characteristics. Probability distribution in queuing systems, distribution of Arrival and interarrival. Distribution of death (departure) process, service time. Classification of Queuing models and solution of Queuing models; M/M/1:∞/FIFO and M/M/1:N/FIFO

REFERENCES

1. Kantiswarup; Gupta P.K. and Singh, M.N.(1985) : Operations Research; Sultan Chand
2. Taha, H.A.(1982): Operations Research: An Introduction; MacMillan
3. Sharma,S.D.: Operations Research.
4. Hillier F.S. and Leiberman,G.J.(1962): Introduction to Operations Research; Holdon Day

SYLLABUS

M.SC. (DATASCIENCE) I-YEAR, II-SEMESTER

MDS-204-T: PAPER- IV: SOFTWARE ENGINEERING

UNIT – I

Software Engineering: The Nature of Software, Changing Nature of Software, Defining the Discipline, Software Process, Software Engineering Practice. The Software Process: A Generic Process Model, Defining a Framework Activity, Process Assessment and Improvement, Prescriptive Process Models, Specialized Process Models, Unified Process, Personal and Team Process Models. Defining Agility, Agile Process, Extreme Programming, Psychology of Software Engineering, Software Team Structures, Software Engineering Using the Cloud, Global Teams.

UNIT – II

Requirements: Core Principles of Modeling, Requirements Engineering, Establishing the Groundwork, Eliciting Requirements, Developing Use Cases, Building the Analysis Model, Requirements Analysis, UML Models That Supplement the Use Case, Identifying Analysis Classes, Specifying Attributes, Defining Operations, Class Responsibility-Collaborator Modeling, Associations and Dependencies, Analysis Packages. Design Concepts: Design within the Context of SE, Design Process, Design Concepts, Design Model, Software Architecture, Architectural Styles, Architectural Considerations, Architectural Design, Component, Designing Class-Based Components, Conducting Component-Level Design, Component-Based Development, User Interface Design Rules.

UNIT – III

Quality Management: Quality, Software Quality, Software Quality Dilemma, Achieving Software Quality, Defect Amplification and Removal, Reviews, Informal Reviews, Formal Technical Reviews, Elements of Software Quality Assurance, SQA Tasks, Goals, and Metrics, Software Reliability, A Strategic Approach to Software Testing, Test Validation Testing, System Testing, Debugging, Software Testing Fundamentals, White-Box Testing, Black-Box Testing, Path Testing, Control Structure Testing, Object-Oriented Testing Strategies & Methods, Security Engineering Analysis, Security Assurance, Security Risk Analysis.

REFERENCE BOOKS

1. Roger S Pressman, B R Maxim, Software Engineering – A Practitioner's Approach (8e)
2. Ian Sommerville: "Software Engineering".
3. Hans Van Vliet, Software Engineering.
4. D. Bell, Software Engineering for Students.
5. K.K. Aggarwal, Y. Singh, Software Engineering.
6. R. Mall, Fundamentals of Software Engineering

SYLLABUS

M.SC. (DATASCIENCE) I-YEAR, II-SEMESTER

MDS-205-P: PAPER- V (PRACTICAL-1):

ADVANCED MACHINE LEARNING USING PYTHON LAB

List of Practical's using Python

Note: preferably writing python code based on the procedure rather than usage of packages.

1. Implementation of classification techniques for the data sets and evaluation of its analysis using
    - (i) Logistic regression
    - (ii) Naïve Bayes classifier
    - (iii) Support Vector Machines,
    - (iv) Decision tree (ID-3) algorithms
    - (v) Random forest algorithm
    - (vi) Bagging and Boosting algorithms
    - (vii) KNN
2. Implementation of
    - (i) Principal component analysis
    - (ii) Factor analysis,
    - (iii) Multidimensional Scaling,
    - (iv) Path analysis.
3. Implementation of clustering using
    - (i) K-means,
    - (ii) DBSCAN,
    - (iii) CART,
    - (iv) C4.5 methods
4. Implementation of Association algorithms
    - (i) Aprori algorithm
    - (ii) FP-tree growth algorithm.

SYLLABUS

M.SC. (DATA SCIENCE) I-YEAR, II-SEMESTER

MDS-206-P: PAPER-VI (PRACTICAL-2):

ARTIFICIAL INTELLIGENCE USING R / PYTHON LAB

Programs List using R and Python

1. Implementation of A* and AO* algorithms.
2. Implementation of Alpha-beta pruning.
3. Implementation of search algorithms (BFS & DFS).
4. Implementation of Hill Climbing algorithm
5. Implementation of Gaming problems:
   (i) Tower of Hanoi problem
   (ii) Tic-Tac-Toe problem
   (iii) Water-Jug problem.
   (iv) 4-Queens problem.
   (v) 8 Puzzle problems.
   (vi) Monkey banana problem

SYLLABUS

M.SC. (DATASCIENCE) I-YEAR, II-SEMESTER

MDS-207-P: PAPER- VII (PRACTICAL-3):

OPTIMIZATION TECHNIQUES USING TORA & R

Programs List using R and TORA

1. Optimal Solution to a L.P.P. by
   (i) Graphical method
   (ii) Simplex method
   (iii) Charners Method
   (iv) Two-phase simplex method
   (v) Dual Simplex Method
   (vi) Duality
2. Transportation problem by MODI & stepping stone methods. (Balanced and unbalanced)
3. Assignment problem by Hungarian method (Balanced and unbalanced)
4. Traveling salesman problem by Hungarian method
5. Job sequencing problem for N jobs on 2 and 3 machines
6. Integer Programming Problem (Gomery's cutting plane and Branch & Bound methods)
7. Construction of Network diagram and finding Critical path using CPM and PERT.
8. Evaluation of Time cost analysis through CPM and PERT

SYLLABUS

M.SC. (DATASCIENCE) I-YEAR, II-SEMESTER

MDS-208-P: PAPER- VIII (PRACTICAL-4):

DATA HANDLING USING SPSS

List of Practical's

1. Basic operations of Data entry, Data import and export, I/O files handling etc.
2. **Data Visualization**: Pie diagram, Bar diagram, Histogram, Line plot, frequency curves & polygons, Scatter Plot, Gantt Chart, Box Plot.
3. **Descriptive Statistics:** Measures of Central Tendencies, Dispersions, Relative measures of Dispersions, Moments, Skewness, Kurtosis.
4. **Parametric Tests**: Testing for Mean(s), Variance(s), Proportion(s), ANOVA for one-way two-way and two-way with one and m-observations per cell and with & without interactions,
5. **Non–Parametric tests:** Sign test, Wilcoxon Sign Rank test, Mann-Whitney U-test, Run test, Kolmogorov Smirnov test, Chi-square test for goodness of fit and Chi-square test independence.
6. **Design and Analysis of Experiments:** Analysis of Variances for Completely randomized, randomized block and Latin Square Designs.
7. **Regression Analysis:** Analysis of Simple and Multiple Linear Regression models, Selection Best Linear Regression Model (All possible, forward, backward, stepwise and stage wise methods). Binary and multinomial Logistic regression models, Probit analysis.
8. **Multivariate Data Analysis:** Linear Discriminant Analysis, Principal Component analysis, Factor analysis, multi-dimensional scaling, Cluster analysis.