Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Scheme and detailed syllabus (DJS22)

Second Year B.Tech

in

Computer Science and Engineering (Data Science)

(Semester IV)

Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

Proposed scheme for Second Year B.Tech Program for Department of Computer Science and Engineering (Data Science)
Semester IV
(Academic Year 2023-2024)

Sr	Course Code	Course	Teaching Scheme (hrs.)				Continuous Assessment (A) (marks)			Semester End Assessment (B) (marks)						(A+B)	Total Credits	
			Th	P	T	Credits	Th	T/W	Total CA (A)	Duration	Th/Cb	O	P	O&P	Total SEA (B)			
1	DJS22DSC401	Statistics for Data Science	3	--	--	3	35	--	35	2	65	--	--	--	65	100	3	4
	DJS22DSL401	Statistics for Data Science Laboratory	--	2	--	1	--	25	25	--	--	--	--	--	--	25	1	
2	DJS22DSC402	Machine Learning - I	3	--	--	3	35	--	35	2	65	--	--	--	65	100	3	4
	DJS22DSL402	Machine Learning - I Laboratory	--	2	--	1	--	25	25	2	--	--	--	25	25	50	1	
3	DJS22DSC403	Design and Analysis of Algorithms	3	--	--	3	35	--	35	2	65	--	--	--	65	100	3	4
	DJS22DSL403	Design and Analysis of Algorithms Laboratory	--	2	--	1	--	25	25	2	--	--	--	25	25	50	1	
4	DJS22DSC404	Computer Communication and Networks	3	--	--	3	35	--	35	2	65	--	--	--	65	100	3	4
	DJS22DSL404	Computer Communication and Networks Laboratory	--	2	--	1	--	25	25	--	--	--	--	25	25	50	1	
5	DJS22DSL405	Data Engineering and Visualization Laboratory	1	2	--	2	--	50	50	2	--	--	--	50	50	100	2	2
6	DJS22IHC1	Universal Human Values	2	--	--	2	35	--	35	2	65	--	--	--	65	100	2	2
	DJS22IHT1	Universal Human Values Tutorial	--	--	1	1		25	25	--	--	--	--	--	--	25	1	1
7	DJS22ILLA2	Innovative Product Development - II	--	2	--	--	--	--	--	--	--	--	--	--	--	--	--	--
8	DJS22A3	Environmental Studies (Audit)	1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
		Total	**16**	**12**	**1**	**21**	**175**	**175**	**350**	**16**	**325**	**0**	**0**	**125**	**450**	**800**	**21**	**21**

Th	Theory	**T/W**	Termwork
P	Practical	**O**	Oral
T	Tuturial	**Cb**	Computer based

Prepared by Checked by Head of the Department Principal

Continuous Assessment (A):

Course	Assessment Tools	Marks	Time (hrs.)
Theory	a. One Term test (based on 40 % syllabus)	20	1
	b. Second Term test (next 40 % syllabus) / presentation / assignment / course project / group discussion / any other.	15	1
	Total marks (a + b)	35	--
Audit course	Performance in the assignments / qui / power point presentation / poster presentation / group project / any other tool.	--	As applicable
Laboratory	Performance in the laboratory and documentation.	25	
Tutorial	Performance in each tutorial & / assignment.	25	
Laboratory &Tutorial	Performance in the laboratory and tutorial.	50	

The final certification and acceptance of term work will be subject to satisfactory performance upon fulfilling minimum passing criteria in the term work / completion of audit course.

Semester End Assessment (B):

Course	Assessment Tools	Marks	Time (hrs.)
Theory / * Computer based	Written paper based on the entire syllabus.	65	2
	* Computer based assessment in the college premises.		
Oral	Questions based on the entire syllabus.	25	as applicable
Practical	Performance of the practical assigned during the examination and the output / results obtained.	25	2
Oral & Practical	Project based courses - Performance of the practical assigned during the examination and the output / results obtained. Based on the practical performed during the examination and on the entire syllabus.	As per the scheme	2

Prepared by Checked by Head of the Department Principal

Program: B.Tech in Computer Science and Engineering (Data Science)	S.Y. B.Tech	Semester: IV
Course: Statistics for Data Science (DJS22DSC401)		
Course: Statistics for Data Science Laboratory (DJS22DSL401)		

Prerequisite: Calculus.

Objectives:
To build the strong foundation in statistics which can be applied to analyze data and make predictions.

Outcomes: On completion of the course, learner will be able to:

1. Interpret data using descriptive statistics.
2. Demonstrate sampling distributions and estimate statistical parameters.
3. Develop hypothesis based on data and perform testing using various statistical techniques.
4. Perform analysis of variance on data.
5. Examine relations between data.

Statistics for Data Science (DJS22DSC401)		
Unit	**Description**	**Duration**
1	**Introduction to Statistics** Types of statistics, population vs sample, Measures of Central Tendency, Measures of position. Measures of dispersion, Moments, Pearson's β and γ coefficients, skewness, kurtosis, population parameters and sample statistics, histogram, frequency polygon. **Expected Values** The expected value of a Random Variable, Variance and Standard Deviation, Covariance and Correlation, Conditional Expectation and Prediction, Approximation Methods.	**06**
2	**Sampling Distribution** Exponential Family of Distributions, Population and Random Sampling, Sample mean, variance and standard deviation, Sampling from Normal distribution, Chi-Square, Student's t-distribution, F-distributions.	**06**
3	**Survey Sampling** The expectation and Variance of the Sample mean, Estimation of population variance, Estimation Ratio, Stratified Random Sampling: Defining sample size for proportionate, disproportionate random sample.	**05**
4	**Parameter Estimation** The method of moments, The method of Maximum Likelihood, The Bayesian Approach to Parameter Estimation.	**04**
5	**Testing Hypothesis and Goodness of Fit:** The Neyman-Pearson Paradigm, The Duality of Confidence Intervals and Hypothesis Tests, Generalized Likelihood Ratio Tests, Likelihood Ratio Test (LRT), Type-I and Type-II errors, Method of Evaluating Tests. Students's t-Test and Chi-Square Test.	**08**
6	**Analysis of Variance (ANOVA) for data analysis:** The One-way Layout: the *F* test, Non parametric method- The Kruskal Wallis Test. The Two-way Layout: Additive Parametrization, Randomized Block Design, Non parametric method-Friedman's Test.	**10**

	Analysis of Categorical Data: Fisher's Exact Test, The Chi Square Test of Homogeneity, The Chi Square Test of Independence	
	Total	**39**

Statistics for Data Science Laboratory (DJS22DSL401)	
Exp.	**Suggested Experiments**
1	To perform descriptive statistics on data.
2	To visualize descriptive statistics on data.
3	To prove central limit theorem.
4	To study sampling distributions and their parameters.
5	To perform Stratified Sampling.
6	To perform statistical estimation tests on data.
7	To calculate confidence interval for different parameters.
8	To perform correlation on given data.
9	To perform hypothesis test using t statistics.
10	To perform hypothesis test using Chi square.
11	To perform Kruskal-Wallis H Test.
12	To perform Friedman Test

Minimum 10 experiments from the above-suggested list or any other experiment based on syllabus will be included, which would help the learner to apply the concept learnt.
*The Term Work will be calculated based on Laboratory Performance (15m) and Quizzes (10m).

Books Recommended:

Textbooks:

1. S. P. Gupta, 'Statistical Methods', Sultan Chand, 46th Edition, 2021.
2. Thomas Hasalwanter, 'An Introduction to Statistics with Python', Springer, 1st Edition, 2016.
3. John A. Rice, 'Mathematical Statistics and Data Analysis', 3rd Edition, Thomson Learning, 1994.
4. Douglas C. Montgomery, Larry Faris Thomas and George C. Runger, 'Engineering Statistics', 3rd edition, John Wiley & Sons, 2003.

Reference Books:

1. Peter Bruce, Andrew Bruce, Peter Gedeck, 'Practical Statistics for data scientists 50+ Essential Concepts Using R and Python', Orelly, 2nd Edition, 2020.
2. Freedman, David, Robert Pisani, Roger Pervis, W. W. Norton, 'Statistics', 2007.
3. S. C. Gupta, V. K. Kapoor, Sultan Chand, 'Fundamentals of mathematical statistics', 10th Edition, 2002.

Web Links:

1. Essentials of Data Science With R Software _ 1: Probability and Statistical Inference, IIT Kanpur: https://nptel.ac.in/courses/111104146
2. Probability and Statistics: https://onlinecourses.nptel.ac.in/noc21_ma74/preview

Prepared by Checked by Head of the Department Principal

Program: B.Tech in Computer Science and Engineering (Data Science)	S.Y. B.Tech	Semester: IV
Course: Machine Learning - I (DJS22DSC402)		
Course: Machine Learning - I Laboratory (DJS22DSL402)		

Prerequisite:

1. Data Structures
2. Basic Probability
3. Statistics

Objectives:

1. To introduce the concepts of computation learning theory and techniques of Machine Learning.
2. To become familiar with regression, classification and clustering tasks.

Outcomes: On completion of the course, learner will be able to:

1. Classify given problems into classification, clustering and regression problems.
2. Apply machine learning techniques for a given problem.
3. Examine the dataset, choose appropriate algorithm and evaluate the results.
4. Design applications using machine learning algorithms.

Machine Learning - I (DJS22DSC402)		
Unit	**Description**	**Duration**
1	**Introduction to Machine Learning:** Types of Machine Learning, Issues in Machine Learning, Application of Machine Learning, Steps involved in developing a Machine Learning Application, Hypothesis and Inductive Bias, Bias-Variance Trade-off, Performance measures, Data Validation.	**04**
2	**Regression:** Linear Regression, Least Minimum Slope (LMS) algorithm, Gradient Descent, Lasso and Ridge Regression. Polynomial Regression. Logistic Regression, Maximum Likelihood Function.	**08**
3	**Classification:** Introduction to decision tree, Learning Decision tree using ID3 and Gini index; CART, Overfitting. Ensemble methods: Bagging (Random Forest) and Boosting (XG Boost).	**08**
4	**Bayesian Learning:** Introduction to Bayesian Learning, Naïve Bayes, Bayesian Network: Representation in Bayesian Belief Network, Inference in Bayesian Network, Applications of Bayesian Network	**06**
5	**Support Vector Machine:** Support Vectors, Functional Margin, Geometric Margin, Optimization problem, Lagrange Duality, KKT condition, Maximum margin with noise, Non-linear SVM and Kernel Function.	**06**
6	**Introduction to Clustering:** K-means, Adaptive hierarchal Clustering, Gaussian Mixture Models, Expectation Maximization.	**08**
	Total	**39**

Students should be encouraged to write these programs from scratch to develop better understanding of the algorithms. Last 30 minutes of the laboratory should be utilized as a discussion on available python libraries and hyperparameters.

Machine Learning - I Laboratory (DJS22DSL402)	
Exp.	**Suggested Experiments**
1	Perform Linear Regression.
2	Perform Logistic Regression.
3	Perform Decision Tree using GINI.
4	Perform CART decision tree algorithm.
5	Perform Ensemble methods.
6	Perform Bayesian Classification.
7	Perform Support Vector Machine.
8	Perform K-means clustering.
9	Perform Expectation –Maximization.
10	Mini project based on any machine learning application.

Above experiments or any other experiment based on syllabus will be included, which would help the learner to apply the concept learnt.

*The Term Work will be calculated based on Laboratory Performance (10m), Mini Project (5m) and Quizzes (10m).

Books Recommended:

Textbooks:

1. Ethem Alpaydın, 'Introduction to Machine Learning', MIT Press, 4th Edition, 2020.
2. Tom M.Mitchell, 'Machine Learning', McGraw Hill, 1st Edition, 2017.
3. Peter Harrington, 'Machine Learning In Action', DreamTech Press, 1st Edition, 2012.

Reference Books:

1. Andreas C. Müller and Sarah Guido, 'Introduction to Machine Learning with Python: A Guide for Data Scientists', O'reilly, 1st Edition, 2016.
2. Stephen Marsland, 'Machine Learning An Algorithmic Perspective', CRC Press, 2nd Edition, 2014.
3. Kevin P. Murphy, 'Machine Learning — A Probabilistic Perspective', MIT Press, Illustrated Edition, 2012.
4. Han Kamber, 'Data Mining Concepts and Techniques', Morgann Kaufmann Publishers, 3rd Edition, 2011.

Web Links:

1. Towards Data Science: https://towardsdatascience.com D
2. Machine Learning — Andrew Ng, Stanford University: https://youtube.com/playlist?list=PLLssT5z_DsKh9vYZkQkYNWcItqhlRJLN
3. Commonly used Machine Learning Algorithms: https://www.analyticsvidhya.com/blog/2017/09/common-machinelearning-algorithms/
4. A Tour to Machine Learning Algorithms: https://machinelearningmastery.com/a-tour-of-machine-learningalgorithms/

Prepared by Checked by Head of the Department Principal

Program: B.Tech in Computer Science and Engineering (Data Science)	S.Y. B.Tech	Semester: IV
Course: Design and Analysis of Algorithms (DJS22DSC403)		
Course: Design and Analysis of Algorithms Laboratory (DJS22DSL403)		

Pre-requisite:

1. Computer Programming
2. Data structures

Course Objectives: The objective of the course is to introduce important algorithmic design paradigms and approaches for effective problem solving. To analyze the algorithm for its efficiency to show its effectiveness over the others. In addition, the concepts of tractable and intractable problems and the classes P, NP and NP-complete problems will be introduced.

Outcomes: On completion of the course, learner will be able to:

1. Analyze the performance of algorithms using asymptotic analysis.
2. Solve the problem using appropriate algorithmic design techniques.
3. Able to prove that certain problems are NP-Complete.

Design and Analysis of Algorithms (DJS22DSC403)		
Unit	**Description**	**Duration**
1	**Introduction:** Introduction to Asymptotic Analysis, Analysis of control statements and loops, solving recurrence relations using tree, substitution, master method, analysis of quick sort and merge sort. Problem Solving using divide and conquer algorithm - Max-Min problem, Strassen's Matrix Multiplication.	**08**
2	**Greedy Method:** Introduction, control abstraction, Problem solving using - fractional knapsack problem, activity selection problem, job sequencing with deadline, find and union, Minimum Spanning trees (Kruskal's algorithm, Prim's algorithm), Single source shortest path (Dijkstra's algorithm), coin change problem.	**07**
3	**Dynamic Programming:** Introduction, principle of optimality, Components of dynamic programming, characteristics of dynamic programming, Fibonacci problem, Coin Changing problem, 0/1 knapsack (table and set method), Multistage graphs, All pairs shortest paths (Floyd Warshall Algorithm), Single source shortest path (Bellman-Ford Algorithm), Matrix Chain Multiplication, Optimal binary search tree (OBST-successful and unsuccessful search), Travelling salesperson problem, Johnson' algorithm for Flow shop scheduling, Longest Common Subsequence (LCS).	**10**
4	**Backtracking:** Introduction, Basics of backtracking, N–queen problem, Sum of subsets, Graph coloring, Hamiltonian cycles Generating permutation. **Branch-and-Bound:** Introduction, Control abstraction-LC BB, FIFO BB, LIFO BB, Properties, FIFO BB, LIFO BB, LC BB, Fifteen Puzzle problem, 0/1 Knapsack problem, Travelling Salesman problem, Job Sequencing with Deadline.	**07**

5	**String Matching Algorithms:** Introduction, The naive string-matching algorithm, The Rabin Karp algorithm, String matching with finite automata, The Knuth Morris Pratt algorithm	**03**
6	**Basics of Computational Complexity:** Complexity classes: The class P and NP, Polynomial reduction, NP Completeness Problem, NP-Hard Problems, NP Completeness problem using Travelling Salesman problem (TSP), Approximation algorithm using TSP.	**04**
	Total	**39**

Design and Analysis of Algorithms Laboratory (DJS22DSL403)	
Exp.	**Suggested Experiments**
1	Implementation of Min Max algorithm
2	Implementation of Strassen's Matrix Multiplication.
3	Implementation of Karatsuba algorithm for long integer multiplication
4	Fractional Knapsack implementation using greedy approach
5	Implementation of Activity selection using greedy approach
6	Implementation of Kruskal's/ Prim's algorithm using greedy approach
7	Implementation of job sequencing with deadline using greedy approach
8	Implementation of other greedy algorithms for example: tree vertex split, subset cover, container loading, coin changing, optimal; merge patterns (Huffman tree)
9	Implementation of Single source shortest path (Dijkstra's algorithm)
10	Implementation of Bellman Ford algorithm using Dynamic programming
11	Implementation of Longest Common Subsequence algorithm using Dynamic programming
12	Implementation of Travelling Salesperson problem using Dynamic programming
13	Implementation of multistage graphs/ all pair shortest path using dynamic programming.
14	Implementation of N-queen problem using Backtracking
15	Implementation of 15 Puzzle problem using Backtracking
16	Implementation of Knuth Morris Pratt string matching algorithm

Minimum 10 experiments from the above-suggested list or any other experiment based on syllabus will be included, which would help the learner to apply the concept learnt.
*The Term Work will be calculated based on Laboratory Performance (15m) and Assignments (10m).

Books Recommended:

Textbooks:

1. S. Sridhar, 'Design and Analysis of Algorithms', Oxford Education, 1st Edition, 2018.
2. Ellis Horowitz, Sartaj Sahni, S. Rajsekaran. 'Fundamentals of computer algorithms', University Press, 1st Edition, 2018.

Reference Books:

1. Sanjoy Dasgupta, Christos Papadimitriou, Umesh Vazirani, 'Algorithms', Tata McGraw- Hill, 1st Edition, 2023.
2. S. K. Basu, 'Design Methods and Analysis of Algorithm', PHI, 2nd Edition, 2013.
3. John Kleinberg, Eva Tardos, "Algorithm Design", Pearson, 1st Edition, 2013.
4. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, 'Introduction to Algorithms', 3rd Edition, The MIT Press, 2009.

5. Michael T. Goodrich, Roberto Tamassia, "Algorithm Design", Wiley Publication, 1st Edition, 2006.

Web Links:

1. NPTEL Course: https://onlinecourses.nptel.ac.in/noc19_cs47/preview
2. CodeChef: https://www.codechef.com/certification/data-structures-and-algorithms/prepare
3. Hackerrank: https://www.hackerrank.com/domains/algorithms

Program: B.Tech in Computer Science and Engineering (Data Science)	S.Y. B.Tech	Semester: IV
Course: Computer Communication and Networks (DJS22DSC404)		
Course: Computer Communication and Networks Laboratory (DJS22DSL404)		

Pre-requisite: System Fundamentals

Course Objectives: This course aims to provide students with a comprehensive understanding of computer networks, from the fundamental concepts of networking to advanced topics such as Internet of Things (IoT) architecture and wireless networks. Students will explore the key components of networking, including the network layer, transport and application layer protocols, and the interconnection of smart objects using IP. By the end of the course, students will be equipped with the knowledge and skills necessary to design, manage, and troubleshoot computer networks, with a focus on emerging technologies and applications in the field.

Outcomes: On completion of the course, the learner will be able to:

1. Develop a strong foundational understanding of computer networks, including the principles and concepts of network communication, protocols, and architectures.
2. Acquire knowledge and practical experience in the design and implementation of IoT architectures.
3. Explore the diverse applications of networking in the modern world, with a specific focus on interconnecting smart objects using IP.

Computer Communication and Networks (DJS22DSC404)		
Unit	**Description**	**Duration**
1	**Introduction to Computer Networks:** Basics of Computer Networks, Network Topologies and Protocols, OSI and TCP/IP Models, Network Devices and Components.	**05**
2	**Network Layer:** Services, Packet switching, ARP, RARP, Unicast Routing Algorithms-(DVR, LSR), IPv4 Addressing (Classfull and Classless), Subnetting, Supernetting design problems, IPv4 Protocol, IPV6 protocol.	**08**
3	**Transport & Application Layer:** Services, sockets, Transport Layer Protocols - User Datagram Protocol (UDP), Transmission Control Protocol (TCP), ARQ, Sliding Window Protocol Application layer protocols-HTTP, SMTP, DNS.	**08**
4	**Data link Layer:** Introduction, transmission medium, physical addressing, Error control (Hamming code, CRC), Flow control, Data-Link Layer Protocols: HDLC, Media Access Control: ALOHA, CSMA, Wired LANs: Ethernet, Wireless LANs.	**08**
5	**IoT Architecture and Technologies:** Introduction to the Internet of Things (IoT), IoT Architecture and Components, Communication Protocols for IoT (MQTT, CoAP, etc.).	**05**

6	**Interconnecting Smart Objects with IP** Architecture, IP Protocol Architecture, IPv6 for Smart Object Networks and the Internet of Things, Connectivity Models for Smart Object Networks The applications: Smart Cities and Urban Networks, Home Automation, Structural Health Monitoring.	**05**
	Total	**39**

Computer Communication and Networks Laboratory (DJS22DSL404)	
Exp.	**Suggested Experiments**
1	Learn to use commands like tcpdump, netstat, ifconfig, nslookup and traceroute. Capture ping and trace route PDUs using a network protocol analyzer.
2	Installation & Configuration of Network Simulator (NS2) in Linux environment. -Study of different topologies and create duplex link in NS2
3	Implementation of Distance Vector/ Link State Routing algorithm
4	Performance evaluation of Routing protocols using Simulation tool
5	Applications using TCP sockets like: a) Echo client and echo server b) Chat c) File Transfer
6	To design and configure Virtual Local Area Network and check the communication privacy among different sub networks.
7	Implement IoE based on IPv6 using packet tracer.
8	Simulate the home automation using Packet Tracer
9	Implementation of the Wireshark for Traffic analysis.
10	Case Study: On Smart Automation

Above experiments or any other experiment based on syllabus will be included, which would help the learner to apply the concept learnt.
*The Term Work will be calculated based on Laboratory Performance (15m) and Assignments (10m).

Books Recommended:

Textbooks:

1. James Kurose, 'Computer Networking:A Top-Down Approach', Pearson Education, 8th Edition, 2022.
2. Behrouz A. Forouzan, 'TCP/IP Protocol Suite', McGraw Hill Education, 4th edition, 2017.

Reference Books:

1. Maciej Kranz , "Building the Internet of Things" by Maciej Kranz, Wiley, 1st edition, 2016.
2. Rajkumar Buyya, Amir Vahid Dastjerdi, and Sriram Venugopal, "Internet of Things: Principles and Paradigms", Morgan Kaufmann Publishers, 2016.
3. Jean-Philippe Vasseur, 'Interconnecting Smart Objects with IP The Next Internet', Morgan Kaufmann Publishers, 2010.

Web Links:

1. Routing Protocol Information: https://www.cisco.com/c/en/us/support/docs/ip/routing-information-protocol-rip/13788-3.html.
2. Transmission of IPv6 Packets over IEEE 802.15.4 Networks: https://datatracker.ietf.org/doc/html/rfc4944
3. IPv6 in IoT: https://pianalytix.com/advantages-of-ipv6-in-iot/

Program: B.Tech in Computer Science and Engineering (Data Science)	S.Y. B.Tech	Semester: IV
Course: Data Engineering and Visualization Laboratory (DJS22DSL405)		

Prerequisite:

1. Mathematics for Intelligent Systems
2. Programming
3. Database Management System

Objectives: To develop skills of data analysis techniques for data modelling and visualization.

Outcomes: On completion of the course, learner will be able:

1. Apply visualization techniques to understand Data.
2. Apply ETL and perform OLAP operation.
3. Apply appropriate techniques to enhance data quality
4. Perform feature engineering to get data ready for modelling.

Foundation of Data Engineering Laboratory (DJS22DSL405)		
Unit	**Description**	**Duration**
1	**Data Warehousing, ETL Process and OLAP:** Introduction to Data Warehousing, Major steps in ETL process, Data extraction: Techniques, Data transformation: Basic tasks, Major transformation types, Data Loading: Applying Data, OLTP Vs OLAP, OLAP definition, Dimensional Modelling, Star Schema and Snowflake schema, Hypercubus, OLAP operations: Drill down, Roll up, Slice, Dice and Rotation, OLAP models: MOLAP, ROLAP.	**04**
2	**Data Preprocessing:** Data Quality: measurement error, noise, bias, accuracy, outliers, missing values, inconsistent values, duplicate values Data Cleaning: handling missing values and noisy data Data Transformation: smoothing, attribute construction, aggregation, normalization; Data Discretization: binning, histogram analysis, cluster Outlier Detection: types of outliers, challenges, statistical method, Distance-based method, Density-based method	**04**
3	**Feature Engineering:** Curse of Dimensionality, Feature selection: Univariate methods (Pearson correlation, F-score, Chi- square, Signal to noise ratio) and Multivariate methods (Forward selection, backward selection and stepwise selection), Feature extraction: principal component analysis	**05**

Visualization experiments can be performed using Tableau and Data Preprocessing experiments can be performed using Python/R.

Foundation of Data Engineering Laboratory (DJS22DSL405)	
Exp.	**Suggested Experiments**
Data Visualization Using Tableau	
1	Create new measures on a given dataset and visualize them using a bar graph.
2	Perform time series aggregation, apply filters on a given dataset, create line and area charts.
3	Apply maps, scatter plots on a given dataset and create a dashboard.
4	Perform joins, blends and create dual axis chart.
5	Perform table calculations, bins, distributions and create Heat maps.
6	Create an interactive data story.
Data Preprocessing:	
7	Perform Exploratory Data Analysis on a given dataset.
8	Perform Data cleaning on a given dataset.
9	Perform necessary Data Transformation on a given dataset.
Feature Engineering:	
10	Perform correlation analysis on a given dataset.
11	Perform dimensionality reduction using PCA.

A minimum of 10 experiments or any other experiment based on syllabus will be included, which would help the learner to apply the concept learnt.
*The Term Work will be calculated based on Laboratory Performance (15m), Mini Project (10m) and Quizzes (25m).

Books Recommended:

Textbooks:

1. Max Kuhn and Keijell Johnson, 'Feature Engineering and Selection: A practical Approach for Predictive Models', CRC Press, 1st Edition, 2020.
2. Roy Jafari, 'Hands-On Data Preprocessing in Python', Packt Publishing, 1st Edition, 2022.
3. Jason Browniee, 'Data Preparation for Machine Learning', Machine Learning Mastry 1st Edition, 1st Edition, 2021.
4. Bad Data Handbook: Cleaning Up the Data so you can get back to work, Ethan McCallum,O'Reilly, 1st Edition, 2012.

Reference Books:

1. S.C.Gupta and V.K.Kapoor, 'Fundamentals of mathematical statistics', Sultan Chand Publisher, 1st Edition, 2020.
2. Wes McKinney, 'Python for Data Analysis', O'Reilly, 2nd Edition, 2018.
3. Rayan Sleeper, 'Practical Tableau', O'Reilly, 1st Edition, 2018.
4. Jeffrey Shaffer, Steve Wexier, Andy Cotgreave, 'The Big Book of Dashboards: Visualizing your Data using Real-World Business Scenarios', Wiley, 1st Edition, 2017.
5. Paulraj Ponniah, 'Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals', Wiley, 2md Edition, 2011.
6. Han, Kamber, Morgan Kaufmann, 'Data Mining Concepts and Techniques', Elsevier 3rd Edition, 2013.

Web Links:

1. NPTEL Course: https://onlinecourses.nptel.ac.in/noc19_ge20/preview
2. Tableau: https://intellipaat.com/blog/what-is-tableau/
3. Data Warehouse: https://onlinecourses.nptel.ac.in/noc19_ge20/preview
4. Feature Engineering: https://towardsdatascience.com/what-is-feature-engineering-importance-tools-and-techniques-for-machine-learning-2080b0269f10

Program: B.Tech in Computer Science and Engineering (Data Science)	S.Y. B.Tech	Semester: IV
Course: Universal Human Values (DJS22IHC1)		
Course: Universal Human Values Tutorial (DJS22IHT1)		

Objectives:

1. Development of a holistic perspective based on self-exploration about themselves (human being), family, society, and nature/existence.
2. Understanding (or developing clarity) of the harmony in the human being, family, society, and nature/existence.
3. Strengthening of self-reflection.
4. Development of commitment and courage to act.

Outcomes: On completion of the course, learner will be able to:

1. Become more aware of themselves, and their surroundings (family, society, nature); they would become more responsible in life, and in handling problems with sustainable solutions, while keeping human relationships and human nature in mind. They would have better critical ability.
2. Become sensitive to their commitment towards what they have understood (human values, human relationships, and human society).
3. Apply what they have learnt to their own self in different day-to-day settings in real life, at least a beginning would be made in this direction.

Universal Human Values (DJS22IHC1)		
Unit	**Description**	**Duration**
1	**Introduction:** Need, Basic Guidelines, Content and Process for Value Education Purpose and motivation for the course. Self-Exploration–what is it? - Its content and process; 'Natural Acceptance' and Experiential Validation- as the process for self-exploration. Continuous Happiness and Prosperity- A look at basic Human Aspirations. Right understanding, Relationship and Physical Facility- the basic requirements for fulfilment of aspirations of every human being with their correct priority. Understanding Happiness and Prosperity correctly- A critical appraisal of the current scenario. Method to fulfil the above human aspirations: understanding and living in harmony at various levels.	**05**
2	**Understanding Harmony in the Human:** Harmony in Myself! Understanding human being as a co-existence of the sentient 'I' and the material 'Body'. Understanding the needs of Self ('I') and 'Body' - happiness and physical facility. Understanding the Body as an instrument of 'I' (I am being the doer, seer and enjoyer). Understanding the characteristics and activities of 'I' and harmony in 'I'. Understanding the harmony of I with the Body: Sanyam and Health; correct appraisal of Physical needs, meaning of Prosperity in detail. Programs to ensure Sanyam and Health.	**05**
3	**Understanding Harmony in the Family and Society:** Harmony in Human-Human Relationship. Understanding values in human-human relationship; meaning of Justice (nine universal values in relationships) and program for its fulfilment to ensure mutual happiness; Trust and Respect as the foundational values of relationship. Understanding the meaning of	**03**

	Trust; Difference between intention and competence. Understanding the meaning of Respect, Difference between respect and differentiation; the other salient values in relationship.	
4	**Understanding the harmony in the society (society being an extension of family):** Resolution, Prosperity, fearlessness (trust) and co-existence as comprehensive Human Goals. Visualizing a universal harmonious order in society- Undivided Society, Universal Order- from family to world family.	**03**
5	**Understanding Harmony in the Nature and Existence:** Whole existence as Coexistence Understanding the harmony in the Nature. Interconnectedness and mutual fulfilment among the four orders of nature recyclability and self-regulation in nature. Understanding Existence as Co-existence of mutually interacting units in all pervasive space. Holistic perception of harmony at all levels of existence.	**05**
6	**Implications of the above Holistic Understanding of Harmony on Professional Ethics:** Natural acceptance of human values. Definitiveness of Ethical Human Conduct. Basis for Humanistic Education, Humanistic Constitution and Humanistic Universal Order. Competence in professional ethics: a. Ability to utilize the professional competence for augmenting universal human order, b. Ability to identify the scope and characteristics of people friendly and eco-friendly production systems, c. Ability to identify and develop appropriate technologies and management patterns for above production systems. Case studies of typical holistic technologies, management models and production systems. Strategy for transition from the present state to Universal Human Order: a. At the level of individual: as socially and ecologically responsible engineers, technologists, and managers, b. At the level of society: as mutually enriching institutions and organizations.	**05**
	Total	**26**

Universal Human Values Tutorial (DJS22IHT1)	
Activity	**Suggested Tutorials**
1	Practice sessions to discuss natural acceptance in human being as the innate acceptance for living with responsibility (living in relationship, harmony, and co-existence) rather than as arbitrariness in choice based on liking-disliking.
2	Practice sessions to discuss the role others have played in making material goods available to me. Identifying from one's own life. Differentiate between prosperity and accumulation. Discuss program for ensuring health vs dealing with disease.
3	Practice sessions to reflect on relationships in family, hostel and institute as extended family, real life examples, teacher-student relationship, goal of education etc. Gratitude as a universal value in relationships. Discuss with scenarios. Elicit examples from students' lives.
4	Practice sessions to discuss human being as cause of imbalance in nature (film "Home" can be used), pollution, depletion of resources and role of technology etc.
5	Practice Exercises and Case Studies will be taken up in Practice (tutorial) Sessions e.g. To discuss the conduct as an engineer or scientist etc.

1. Term work shall consist of minimum 5 activities based on activities conducted.
2. Term work shall carry total 25 marks based on the performance in the tutorials.

The final certification and acceptance of term work will be subject to satisfactory performance of activities and upon fulfilling minimum passing criteria in the term work.

Books Recommended:

Textbooks:

1. Human Values and Professional Ethics by R R Gaur, R Sangal, G P Bagaria, Excel Books, New Delhi, 2010.

Reference books:

1. Jeevan Vidya: Ek Parichaya, A Nagaraj, Jeevan Vidya Prakashan, Amarkantak, 1999.
2. Human Values, A.N. Tripathi, New Age Intl. Publishers, New Delhi, 2004.
3. The Story of Stuff (Book).
4. The Story of My Experiments with Truth - by Mohandas Karamchand Gandhi
5. Small is Beautiful - E. F Schumacher.
6. Slow is Beautiful - Cecile Andrews
7. Economy of Permanence - J C Kumarappa
8. Bharat Mein Angreji Raj - PanditSunderlal
9. Rediscovering India - by Dharampal
10. Hind Swaraj or Indian Home Rule - by Mohandas K. Gandhi
11. India Wins Freedom - Maulana Abdul Kalam Azad
12. Vivekananda - Romain Rolland (English)
13. Gandhi - Romain Rolland (English)

Program: B.Tech in Computer Science and Engineering (Data Science)	S.Y. B.Tech	Semester: IV
Course: Innovative Product Development - II (DJS22ILLA2)		

Objectives:

1. To acquaint the students with the process of identifying the need (considering a societal requirement) and ensuring that a solution is found to address the same by designing and developing an innovative product.
2. To familiarize the students with the process of designing and developing a product, while they work as part of a team.
3. To acquaint the students with the process of applying basic engineering fundamentals, so as to attempt at the design and development of a successful value-added product.
4. To inculcate the basic concepts of entrepreneurship and the process of self-learning and research required to conceptualize and create a successful product.

Outcome:

Learner will be able to:

1. Identify the requirement for a product based on societal/research needs.
2. Apply knowledge and skills required to solve a societal need by conceptualizing a product, especially while working in a team.
3. Use standard norms of engineering concepts/practices in the design and development of an innovative product.
4. Draw proper inferences through theoretical/ experimental/simulations and analyse the impact of the proposed method of design and development of the product.
5. Develop interpersonal skills, while working as a member of the team or as the leader.
6. Demonstrate capabilities of self-learning as part of the team, leading to life-long learning, which could eventually prepare themselves to be successful entrepreneurs.
7. Demonstrate product/project management principles during the design and development work and excel in written (Technical paper preparation) as well as oral communication.

Guidelines for the proposed product design and development:

- Students shall form a team of 3 to 4 students (max allowed: 5-6 in extraordinary cases, subject to the approval of the department review committee and the Head of the department).
- Students should carry out a survey and identify the need, which shall be converted into conceptualization of a product, in consultation with the faculty supervisor/head of department/internal committee of faculty members.
- Students in the team shall understand the effective need for product development and accordingly select the best possible design in consultation with the faculty supervisor.
- Students shall convert the best design solution into a working model, using various components drawn from their domain as well as related interdisciplinary areas.
- Faculty supervisor may provide input to students during the entire span of the activity, spread over 2 semesters, wherein the focus shall be on self-learning.

- A record in the form of an activity logbook is to be prepared by each team, wherein the team can record weekly progress of work. The guide/supervisor should verify the recorded notes/comments and approve the same on a weekly basis.
- The design solution is to be validated with proper justification and the report is to be compiled in a standard format and submitted to the department. Efforts are to be made by the students to try and publish a technical paper, either in the institute journal, "Techno Focus: Journal for Budding Engineers" or at a suitable publication, approved by the department research committee/ Head of the department.
- The focus should be on self-learning, capability to design and innovate new products as well as on developing the ability to address societal problems. Advancement of entrepreneurial capabilities and quality development of the students through the year long course should ensure that the design and development of a product of appropriate level and quality is carried out, spread over two semesters, i.e. during the semesters III and IV.

Guidelines for Assessment of the work:

- The review/ progress monitoring committee shall be constituted by the Head of the Department. The progress of design and development of the product is to be evaluated on a continuous basis, holding a minimum of two reviews in each semester.
- In the continuous assessment, focus shall also be on each individual student's contribution to the team activity, their understanding and involvement as well as responses to the questions being raised at all points in time.
- Distribution of marks individually for the both reviews as well as for the first review during the subsequent semester shall be as given below:
    - Marks awarded by the supervisor based on log-book: 20
    - Marks awarded by review committee: 20
    - Quality of the write-up: 10
- In the last review of the semester IV, the marks will be awarded as follows.
    - Marks awarded by the supervisor (Considering technical paper writing) : 30
    - Marks awarded by the review committee: 20

Note- A Candidate needs to secure a minimum of 50% marks to be declared to have completed the audit course.

Review/progress monitoring committee may consider the following points during the assessment.

- In the semester III, the entire design proposal shall be ready, including components/system selection as well as the cost analysis. Two reviews will be conducted based on the presentation given by the student's team.
    - First shall be for finalisation of the product selected.
    - Second shall be on finalisation of the proposed design of the product.
- In the semester IV, the expected work shall be procurement of components/systems, building of the working prototype, testing and validation of the results based on work completed in semester III.
    - First review is based on readiness of building the working prototype.

- Second review shall be based on a presentation as well as the demonstration of the working model, during the last month of semester IV. This review will also look at the readiness of the proposed technical paper presentation of the team.

The overall work done by the team shall be assessed based on the following criteria;

1. Quality of survey/ need identification of the product.
2. Clarity of Problem definition (design and development) based on need.
3. Innovativeness in the proposed design.
4. Feasibility of the proposed design and selection of the best solution.
5. Cost effectiveness of the product.
6. Societal impact of the product.
7. Functioning of the working model as per stated requirements.
8. Effective use of standard engineering norms.
9. Contribution of each individual as a member or the team leader.
10. Clarity on the write-up and the technical paper prepared.

- The semester reviews (III and IV) may be based on relevant points listed above, as applicable

Guidelines for Assessment of Semester Reviews:

- The write-up should be prepared as per the guidelines given by the department.
- The design and the development of the product shall be assessed through a presentation and demonstration of the working model by the student team to a panel of Internal and External Examiners, preferably from industry or any research organizations having an experience of more than five years, approved by the Head of the Institution. The presence of the external examiner is desirable only for the 2nd presentation in semester IV. Students are compulsorily required to present the outline of the technical paper prepared by them during the final review in semester IV.

Program: B.Tech in Computer Science and Engineering (Data Science)	S.Y. B.Tech	Semester: IV
Course: Environmental Studies (DJS22A3)		

Pre-requisite: Interest in Environment and its impact on Human

Objectives:

1. Understand environmental issues such as depleting resources, pollution, ecological problems and the renewable energy scenario.
2. Familiarize environment related legislation

Outcomes: On completion of the course, learner will be able to:

1. Understand how human activities affect environment
2. Understand the various technology options that can make a difference
3. Apply what they have learnt to their own self in different day-to-day settings in real life, at least a beginning would be made in this direction.

Environmental Studies (DJS22A3)		
Unit	**Description**	**Duration**
1	**Social Issues and Environment:** Ecological footprint and Carrying Capacity, Depleting nature of Environmental resources such as soil, water minerals and forests, Carbon emissions and Global Warming.	**04**
2	**Technological Growth for Sustainable Development:** Social, Economical and Environmental aspects of Sustainable Development, Renewable Energy Harvesting, Concept of Carbon credit, Green Building, Power and functions of Central Pollution Control Board and State Pollution Control Board.	**04**
3	**Green Technology:** History, Agenda, and Challenges Ahead. Sustainable Cloud Computing, and Risk Management, Sustainable Software Design, Data Center Energy Efficiency, Thin-Client and Energy Efficiency.	**05**
	Total	**13**

Books Recommended:

Textbooks:

1. Erach Bharucha, 'Textbook for Environmental Studies For Undergraduate Courses of all Branches of Higher Education', Universities Press, 1st Edition, 2019.
2. Mohammad Dastbaz, Colin Pattinson, Babak Akhgar, Morgan and Kaufman, 'Green Information Technology A Sustainable Approach', Elsevier, 1st Edition, 2015.
3. R. Rajagopalan, 'Environmental Studies From Crisis to Cure', 2012.

Reference Books:

1. Paulina Golinska, Marek Fortsch, Jorge Marx-Gómez, 'Information Technologies in Environmental Engineering: New Trends and Challenges', Springer, 2011.

Prepared by Checked by Head of the Department Principal