Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Shri Vile Parle Kelavani Mandal's

Scheme and detailed syllabus of DJS23 Honors Program in Data Analytics

With effect from the Academic Year: 2026-2027

Scheme for Honors in Data Analytics (Academic Year 2026-2027)

Sr. No.	Course Code	Course	Teaching Scheme			Semester End Examination (A)						Continuous Assessment (B)						Aggregate (A+B)	Credits
			Theory (hrs.)	Practical (hrs.)	Tutorial (hrs.)	Duration	Theory	Oral	Pract	Oral & Pract	SEE Total(A)	Term Test 1 (TT1)	Term Test 2 (TT2)	Term Test 3 (TT3)	Term Test Total (TT1+TT2 +TT3)	Term work	CA Total (B)		
SEM III																			
1	DJS23BCH1201	Fundamentals of Data Mining	3	--	--	2	60	--	--	--	60	15	15	10	40	--	40	100	3
SEM IV																			
2	DJS23BHI1251	Data Analytics and Visualization	3	--	--	2	60	--	--	--	60	15	15	10	40	--	40	100	3
	DJS23BHI1251L	Data Analytics and Visualization Laboratory	--	2	--	2	--	25	--	--	25	--	--	--	--	25	25	50	1
SEM V																			
3	DJS23BHI1301	Natural Language Processing	3	--	--	2	60	--	--	--	60	15	15	10	40	--	40	100	3
	DJS23BHI1301L	Natural Language Processing Laboratory	--	2	--	2	--	25	--	--	25	--	--	--	--	25	25	50	1
SEM VI																			
4	DJS23BHI1351	Time Series and Forecasting Analytics	3	--	--	2	60	--	--	--	60	15	15	10	40	--	40	100	3
	DJS23BHI1351L	Time Series and Forecasting Analytics Laboratory	--	2	--	2	--	25	--	--	25	--	--	--	--	25	25	50	1
SEM VII																			
5	DJS23BHI1401	Business Intelligence	3	--	--	2	60	--	--	--	60	15	15	10	40	--	40	100	3
		Total	**15**	**6**	**0**	**16**	**300**	**75**	**0**	**0**	**375**	**75**	**75**	**50**	**200**	**75**	**275**	**650**	**18**

Prepared by **Checked by** **Head of the Department** **Vice Principal** **Principal**

Semester End Examination (SEE) – A

Course	Assessment Tools	Marks	Time (hrs.)
Theory / * Computer based	Written paper based on the entire syllabus.	60	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

Continuous Assessment (B):

Course	Assessment Tools	Marks	Time (mins)
Theory	a. Term Test 1 (based on 40 % syllabus)	15	45
	b. Term Test 2 (on next 40 % syllabus)	15	45
	c. Assignment / course project / group discussion / presentation / quiz/ any other.	10	–
	Total marks (a + b + c)	40	–
Audit course	Performance in the assignments / quiz / 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	

Program: B.Tech in Computer Science and Engineering(IoT and Cybersecurity with Block chain Technology)	S.Y.B.Tech	Semester : III
Course : Fundamentals of Data Mining	Course Code: DJS23BCH1201	

Teaching Scheme (Hours / week)				Evaluation Scheme								
				Semester End Examination Marks (A)			Continuous Assessment Marks (B)					Total marks (A+B)
Lectures	Practical	Tutorial	Total Credits	Theory			Term Test 1	Term Test 2	Term Test 3	Total		
				60			15	15	10	40		100
				Laboratory Examination			Term work			Total Term work		
3	--	--	3	Oral	Practical	Oral & Practical	Laboratory Work	Tutorial / Mini project / presentation/ Assignment				--
				--	--	--	--	--		--		

Pre-requisite:

1. Basic programming concepts
2. Probability distributions

Objectives:

1. To understand data mining concepts.
2. To learn Data mining techniques and algorithms.
3. Comprehend the data mining environments

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

1. Identify concepts of Data Warehouse architecture, dimensional modelling, STAR, Snowflake, and Fact Constellation schemas used in strategic decision-making.
2. Apply data preprocessing techniques and perform Knowledge Discovery in Databases (KDD) for data mining applications.
3. Apply association rule mining techniques using Apriori, FP-Growth, Partition, and Pincer-Search algorithms.
4. Apply classification techniques including Decision Trees, Bayesian classifiers, ensemble methods, and evaluate classifier performance.
5. Perform clustering and outlier analysis using partitioning and hierarchical clustering techniques.
6. Apply Web Mining, Text Mining, and Spatial Data Mining techniques for analyzing structured and unstructured data.

Detailed Syllabus: (unit wise)		
Unit	**Description**	**Duration**
1	**Data Warehousing:** Introduction to Data Warehouse and Dimensional modelling: Introduction to Strategic Information, Need for Strategic Information, Features of Data Warehouse, Data warehouse versus Data Marts, Data warehouse versus Data Lake, Top-down versus Bottom-up approach. Data warehouse architecture, metadata, E R modelling versus Dimensional Modelling, Information Package Diagram, STAR schema, STAR schema keys, Snowflake Schema, Fact Constellation Schema, Factless Fact tables, Update to the dimension tables, Aggregate fact tables. Cloud Data warehousing: Introduction to cloud data warehousing, Key features, Traditional warehousing Vs Cloud data warehousing , Six Steps to Getting Started with Cloud Data Warehousing	**08**
2	**Data Mining:** Integration of Data mining system with a Data warehouse, Data Preprocessing: Major tasks in preprocessing, Data Cleaning: Missing values, Noisy data; Data Integration: Entity Identification Problem, Redundancy and Correlation Analysis, Tuple Duplication, Data Value Conflict Detection and Resolution; Concept hierarchy generation for Nominal data. KDD vs Data Mining, Major issues in Data Mining. DM Applications	**06**
3	**Association Rules:** Association Rule, Methods to Discover Association Rules, A Priori Algorithm, Partition Algorithm, Pincer-Search Algorithm, Dynamic Itemset Counting Algorithms, FP-Tree Growth Algorithm, Discussion on Different Algorithms, Incremental Algorithms, Border Algorithms, Generalized Association Rule, Association Rules with Item Constraints.	**06**
4	**Classification:** Classification and Prediction ,Basic concepts, Decision Tree Induction, Attribute Selection Measures using Information Gain, Tree pruning Bayes Classification Methods: Bayes' Theorem, Naïve Bayesian Classification Model Evaluation & Selection: Metrics for Evaluating Classifier Performance, Random Forests, Lazy learner.	**06**
5	**Clustering and Applications:** Cluster Analysis and Requirements of Cluster Analysis Partitioning Methods: k-Means, k-Medoids Hierarchical Methods: Agglomerative, Divisive Evaluation of Clustering: Assessing Clustering Tendency, Outlier Analysis.	**08**

6	**Web Mining:** Web Mining, Web Content Mining, Web Structure Mining, Web Usage Mining, Text Mining, Unstructured Text, Episode Rule Discovery for Texts, Hierarchy of Categories, Text Clustering. Introduction to Time series mining . Spatial Data mining: Spatial Data, Spatial Vs. Classical Data Mining, Spatial Data Structures, Mining Spatial Association and Co-location Patterns	08
	Total	**42**

Books Recommended:

Text Books

1. Data Mining – Concepts and Techniques – Jiawei Han & Micheline Kamber, 3rd Edition Elsevier,2011.

2. Data Mining Introductory and Advanced topics – Margaret H Dunham, PEA, 2006.

3. Data Mining Techniques, Arun K Pujari, University Press,2013.

4. Cloud Data Warehousing, Joe Kraynak and David Baum, John Wiley & Sons, Inc.,2020

Reference Books

1. Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques (Second Edition), Morgan Kaufmann, 2005.

Web resources:

1. https://www.javatpoint.com/data-mining
2. https://www.spiceworks.com/tech/big-data/articles/what-is-data-mining/

Online Courses: NPTEL / Swayam

1. Course on- Data Mining

- https://onlinecourses.nptel.ac.in/noc21_cs06/preview

Evaluation Scheme:

Semester End Examination (A):

Theory:

1. Question paper based on the entire syllabus total comprising of 60 marks.
2. Total duration allotted for writing the paper is 2 hrs.

Continuous Assessment (B):

Theory:

1. Term Test 1 (based on 40 % syllabus) of 15 marks for the duration of 45 min.
2. Term Test 2 (on next 40 % syllabus) of 15 marks for the duration of 45 min.
3. Term test 3 can be conducted as Assignment / course project / group discussion /presentation / quiz/ any other for 10 marks.

Prepared by	Checked by	Head of the Department	Vice Principal	Principal