Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

Scheme and detailed syllabus of DJS23 Honors Program in Data Analytics

With effect from the Academic Year: 2024-2025

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 Honors in Data Analytics (Academic Year 2024-2025)

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)	Th	O	P	O & P	Total SEA (B)		
Sem III																
1	DJS23BCH1301	**Fundamentals of Data Mining**	3	--	--	3	40	--	40	60	--	--	--	60	100	**3**
Sem IV																
2	DJS23BCH1401	**Statistics for Data Science**	3		--	3	40	--	40	60	--	--	--	60	100	**3**
Sem V																
3	DJS23BCH1501	**Data Visualization**	3	--	--	3	40	--	40	60	--	--	--	60	100	**3**
4	DJS23BLH1501	**Data Visualization Laboratory**	--	2	--	1	--	25	25	--	25	--	--	25	50	**1**
Sem VI																
5	DJS23BCH1601	**Machine Learning Fundamentals**	4	--	--	4	40	--	40	60	--	--	--	60	100	**4**
6	DJS2 BLH1601	**Machine Learning Fundamentals Laboratory**	--	2	--	1	--	25	25	--	25	--	--	25	50	**1**
Sem VIII																
7	DJS23BCH1801	**Big data Analytics**	3	--	--	3	40	--	40	60	--	--	--	60	100	**3**
		Total	**16**	**4**	**--**	**18**	**200**	**50**	**250**	**300**	**50**	**--**	**--**	**350**	**600**	**18**

Continuous Assessment (A):

Course	Assessment Tools	Marks	Time (hrs.)
Theory	One Term test (based on 40 % syllabus)	15 each	1
	Second Term test (next 40 % syllabus) / presentation / assignment / course project / group discussion / any other.		As applicable
Audit course	Performance in the assignments / quiz / power point presentation / poster presentation / group project / any other tool.	10	
Laboratory	Performance in the laboratory and documentation.	--	
Tutorial	Performance in each tutorial & / assignment.	--	
Laboratory &Tutorial	Performance in the laboratory and tutorial.	--	

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.

Continuous Assessment (B):

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.	--	As applicable
Practical	Performance of the practical assigned during the examination and the output / results obtained.	--	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.	--	2

<table>
<tr><td colspan="8">Program: B.Tech. in CSE(IoT and Cyber Security with Blockchain Technology)</td><td colspan="3">Semester : III</td></tr>
<tr><td colspan="8">Course : Fundamentals of Data Mining</td><td colspan="3">Course Code:(DJS23BCH1301)</td></tr>
<tr><td colspan="4" rowspan="3">Teaching Scheme (Hours /week)</td><td colspan="7">Evaluation Scheme</td></tr>
<tr><td colspan="3">Semester End Examination Marks (A)</td><td colspan="3">Continuous Assessment Marks (B)</td><td>Total marks (A+B)</td></tr>
<tr><td rowspan="3">Lectures</td><td rowspan="3">Practical</td><td rowspan="3">Tutorial</td><td rowspan="3">Total Credits</td><td colspan="3">Theory</td><td>Theory</td><td>Term Test +Assignment</td><td>Total Term Test</td><td></td></tr>
<tr><td colspan="3">60</td><td>60</td><td>30 +10</td><td>40</td><td>100</td></tr>
<tr><td colspan="3">Laboratory Examination</td><td colspan="2">Term work</td><td rowspan="2">Total Term work</td><td rowspan="2">--</td></tr>
<tr><td>3</td><td>--</td><td>--</td><td>3</td><td>Oral</td><td>Practical</td><td>Oral & Practical</td><td>Laboratory Work</td><td>Tutorial / Mini project / presentation/ Journal</td></tr>
<tr><td></td><td></td><td></td><td></td><td>--</td><td>--</td><td>-</td><td>--</td><td>--</td><td>--</td><td></td></tr>
</table>

Pre-requisite:

1. Database Management Systems

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. Gain practical skills in collecting and preprocessing the raw data from various sources
2. Characterize the various kinds of patterns that can be discovered by association rule mining.
3. Develop a deep understanding of various classification algorithms.

4. Understand clustering and various clustering methods.
5. Learn various techniques that can be applied to extract valuable insights from web data.
6. Explore current research trends and emerging technologies in data mining.

Detailed Syllabus: (unit wise)		
Unit	Description	Duration
1	**Data Mining:** Data–Types of Data–, Data Mining Functionalities– Interestingness Patterns– Classification of Data Mining systems– Data mining Task primitives –Integration of Data mining system with a Data warehouse– Major issues in Data Mining–Data Preprocessing. KDD vs Data Mining, DBMS vs DM, Other Related Areas, DM Techniques, Other Mining Techniques, Issues and Challenges in DM, DM Applications- Case Studies	9
2	**Association Rules:** What is an 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	9
3	**Classification:** Classification and Prediction – Basic concepts–Decision tree induction– Bayesian classification, Rule–based classification, Lazy learner.	07
4	**Clustering and Applications:** Cluster analysis–Types of Data in Cluster Analysis–Categorization of Major Clustering Methods– Partitioning Methods, Hierarchical Methods– Density–Based Methods, Grid–Based Methods, Outlier Analysis.	06
5	**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	04
6	**Advanced Concepts:** Basic concepts in Mining data streams–Mining Time–series data—Mining sequence patterns in Transactional databases– Mining Object– Spatial– Multimedia–Text and Web data – Spatial Data mining– Multimedia Data mining–Text Mining– Mining the World Wide Web	04
	Total	**39**

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.

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 will be based on the entire syllabus summing up to 60 marks.
2. Total duration allotted for writing the paper is 2 hrs.

Continuous Assessment (B):

Theory:

1. Two term test of 15 marks each, Assignment / course project / group discussion /presentation / quiz/ any other 10 marks
2. Total duration allotted for writing the paper is 45 min.

Prepared by Checked by Head of the Department Vice Principal Principal