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 DJ19 Honors Program in Smart Computing

With effect from the Academic Year: 2024-2025

Scheme for Honors in Smart Computing (Academic Year 2024-2025)

Sr	Course Code	Course	Teaching Scheme(hrs)				Continuous Assessment (A)			Semester End Assessment (B) (marks)					Aggregate (A+B)	Total Credits
			Th	P	T	Credits	Th	T/W	Total CA (A)	Th	O	P	O &P	Total SEA (B)		
SEM V																
1	DJ19ICCHN1C1	Smart Technologies	4	--	--	4	25	--	25	75	--	--	--	75	100	4
SEMVI																
2	DJ19ICCHN1C2	Cognitive Computing	4	--	--	4	25	--	25	75	--	--	--	75	100	4
3	DJ19ICCHN1L2	Cognitive Computing Laboratory	--	2	--	1	--	25	25	--	--	--	25	25	50	1
SEM VII																
4	DJ19ICCHN1C3	IoT Data Analytics	4	--	--	4	25	--	25	75	--	--	--	75	100	4
5	DJ19ICCHN1L3	IoT Data Analytics Laboratory	--	2	--	1	--	25	25	--	--	--	25	25	50	1
SEM VIII																
7	DJ19ICCHN1C4	Social Cyber Security	4	--	--	4	25	--	25	75	--	--	--	75	100	4
		Total	16	4	0	18	100	50	150	300	0	0	50	350	500	18

<table>
<tr><td colspan="8">Program: B.Tech. in CSE(IoT and Cyber Security with Blockchain Technology)</td><td colspan="3">Semester : VII</td></tr>
<tr><td colspan="8">Course : IoT Data Analytics</td><td colspan="3">Course Code: DJ19ICCHN1C3</td></tr>
<tr><td colspan="8">Course: IoT Data Analytics Laboratory</td><td colspan="3">Course Code: DJ19ICCHN1L3</td></tr>
<tr><td colspan="4" rowspan="3">Teaching Scheme (Hours / week)</td><td colspan="7">Evaluation Scheme</td></tr>
<tr><td colspan="4">Semester End Examination Marks (A)</td><td colspan="3">Continuous Assessment Marks (B)</td><td rowspan="2">Total marks (A+ B)</td></tr>
<tr><td rowspan="2">Lectures</td><td rowspan="2">Practical</td><td rowspan="2">Tutorial</td><td rowspan="2">Total Credits</td><td colspan="4">Theory</td><td>Term Test 1</td><td>Term Test 2</td><td>Total</td></tr>
<tr><td colspan="4">75</td><td>25</td><td>25</td><td>25</td><td>100</td></tr>
<tr><td rowspan="3">4</td><td rowspan="3">2</td><td rowspan="3">--</td><td rowspan="3">4</td><td colspan="3">Laboratory Examination</td><td colspan="2">Term work</td><td rowspan="2">Total Term work</td><td rowspan="3">50</td></tr>
<tr><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>25</td><td>15</td><td>10</td><td>25</td></tr>
</table>

Pre-requisite:

1. Introduction to IoT
2. Machine Learning

Objectives:

1. Recognize the importance of data analytics in extracting meaningful insights from IoT-generated data.

2. Acquire hands-on experience with big data technologies and tools, such as Hadoop, Spark, and NoSQL databases.

3. Understand the fundamental concepts and principles of data stream processing and its significance in real-time analytics.

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

1. Recognize the importance of data analytics in extracting meaningful insights from IoT-generated data.
2. Understand the building blocks of Big Data Analytics.
3. Apply fundamental enabling techniques like Hadoop and Map Reduce in solving real world problems.
4. Analyze different NoSQL systems and how it handles big data.
5. Apply advanced techniques for emerging applications like stream analytics.
6. Apply recommendation systems in real-world scenarios across various domains

Detailed Syllabus: (unit wise)		
Unit	Description	Duration
1	Data Analytics 1.1 Introduction, Data Analysis, significance, Importance, Types, Need, Benefits, IoT analytics Tools, IoT analytics Use cases, 1.2 Machine learning, Types of Machine Learning Models, Model Building Process, Modelling Algorithms, Model Performance, 1.3 Big Data Platform, Big Data Pipeline, IoT analytics Use cases, IoT + Big Data Analytics + IIoT	06
2	Introduction to Big Data 2.1Classification of Digital Data, Structured and Unstructured Data – Introduction to Big Data: Characteristics – Evolution – Definition - Challenges with Big Data - Other Characteristics of Data - Why Big Data - Traditional Business Intelligence versus Big Data - Data Warehouse and 2.2Hadoop Environment Big Data Analytics: Classification of Analytics – Challenges - Big Data Analytics important - Data Science - Data Scientist - Terminologies used in Big Data Environments - Basically Available Soft State Eventual Consistency - Top Analytics Tools	12
3	Hadoop HDFS and Map Reduce 3.1 Distributed File Systems: Physical Organization of Compute Nodes, Large-scale File-System Organization. 3.2 Map Reduce: The Map Tasks, Grouping by Key, The Reduce Tasks, Combiners, Details of Map Reduce Execution, Coping With Node Failures. 3.3 Algorithms Using Map Reduce: Matrix-Vector Multiplication by Map Reduce, Relational-Algebra Operations, Computing Selections by Map Reduce, Computing Projections by Map Reduce, Union , Intersection, and Difference by Map Reduce	07

	3.4 Hadoop Limitations	
4	NoSQL 4.1 Introduction to NoSQL, NoSQL Business Drivers 4.2 NoSQL Data Architecture Patterns: Key-value stores, Graph stores, Column family (Bigtable)stores, Document stores, Variations of NoSQL architectural patterns, NoSQL Case Study 4.3 NoSQL solution for big data, Understanding the types of big data problems; Analyzing big data with a shared-nothing architecture; Choosing distribution models: master-slave versus peer-to-peer; NoSQL systems to handle big data problems.	06
5	Mining Data Streams 5.1 The Stream Data Model: A Data-Stream-Management System, Examples of Stream Sources, Stream Queries, Issues in Stream Processing. 5.2 Sampling Data techniques in a Stream 5.3 Filtering Streams: Bloom Filter with Analysis. 5.4 Counting Distinct Elements in a Stream, Count Distinct Problem, Flajolet-Martin Algorithm, Combining Estimates, Space Requirements 5.5 Counting Ones in a Window: The Cost of Exact Counts, The Datar-Gionis-Indyk-Motwani Algorithm, Query Answering in the DGIM Algorithm, Decaying Windows.	04
6	Real-Time Big Data Models 6.1 A Model for Recommendation Systems, Content-Based Recommendations, Collaborative Filtering 6.2 Case Study: Product Recommendation 6.3 Social Networks as Graphs, Clustering of Social-Network Graphs, Direct Discovery of Communities in a social graph	04
	Total	**39**

List of Laboratory Experiments:

Sr. No.	Title of Experiments (Minimum any eight)
1	Hadoop HDFS Practical: -HDFS Basics, Hadoop Ecosystem Tools Overview. -Installing Hadoop. -Copying File to Hadoop. -Copy from Hadoop File system and deleting file. -Moving and displaying files in HDFS. -Programming exercises on Hadoop
2	Use of Sqoop tool to transfer data between Hadoop and relational database servers. a. Sqoop - Installation. b. To execute basic commands of Hadoop eco system component Sqoop.
3	To install and configure MongoDB/ Cassandra/ HBase/ Hypertable to execute NoSQL commands
4	Experiment on Hadoop Map-Reduce: -Write a program to implement a word count program using MapReduce.
5	Experiment on Hadoop Map-Reduce: -Implementing simple algorithms in Map-Reduce: Matrix multiplication, Aggregates, Joins, Sorting, Searching, etc

6	Create HIVE Database and Descriptive analytics-basic statistics.
7	Data Stream Algorithms (any one): - Implementing DGIM algorithm using any Programming Language - Implement Bloom Filter using any programming language Implement Flajolet Martin algorithm using any programming language
8	Social Network Analysis using R (for example: Community Detection Algorithm)
9	Data Visualization using Hive/PIG/R/Tableau/.
10	Exploratory Data Analysis using Spark/ Pyspark.
11	Mini Project: One real life large data application to be implemented (Use standard Datasets available on the web). - Streaming data analysis – use flume for data capture, HIVE/PYSpark for analysis of twitter data, chat data, weblog analysis etc. - Recommendation System (for example: Health Care System, Stock Market Prediction, Movie Recommendation, etc.) Spatio Temporal Data Analytics

Any other experiment based on syllabus may be included, which would help the learner to understand topic/concept.

Books Recommended:

Text Books

1 Anand Rajaraman and Jeff Ullman ―Mining of Massive Datasets‖, Cambridge University Press,2016.

2 Alex Holmes ―Hadoop in Practice‖, Manning Press, Dream Tech Press, second edition 2015.

3 Dan Mcary and Ann Kelly ―Making Sense of NoSQL‖ – A guide for managers and the rest of us, Manning Press,2013.

4 DT Editorial Services, ―Big Data Black Book‖, Dream Tech Press,2016.

Reference Books

1 Bill Franks , ―Taming The Big Data Tidal Wave: Finding Opportunities In HugeData StreamsWithAdvancedAnalytics‖,Wiley,2012

2 Chuck Lam, ―Hadoop inAction‖, Dreamtech Press,2011.

3 Jared Dean, ―Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners‖,Wiley India Private Limited, 2014.

4 Jiawei Han and Micheline Kamber, ―Data Mining: Concepts and Techniques‖, Morgan Kaufmann Publishers, 3rd edition, 2010.

Web resources:

1. https://www.coursera.org/specializations/big-data#courses
2. https://www.digimat.in/nptel/courses/video/106106169/L01.html
3. https://www.coursera.org/learn/nosql-databases#syllabus
4. https://www.coursera.org/learn/basic-recommender-systems#syllabus

Online Courses: NPTEL / Swayam

1. Course on- Big Data Computing

https://onlinecourses.nptel.ac.in/noc20_cs92/preview

Evaluation Scheme:

Semester End Examination (A):

Theory:

1. Question paper based on the entire syllabus will comprise of 5 questions (All compulsory, but with internal choice as appropriate), each carrying 15 marks, total summing up to 75 marks.

2. Total duration allotted for writing the paper is 3 hrs.

Laboratory:

1. Oral and practical examinations will be based on the entire syllabus.

Continuous Assessment (B):

Theory:

1. Two term tests of 25 marks each will be conducted during the semester out of which one will be a compulsory term test (on minimum 02 Modules) and the other can either be a term test or an assignment on live problems or a course project.

2. Total duration allotted for writing each of the paper is 1 hr.

3. Average of the marks scored in the both the tests will be considered for final grading.

Laboratory: (Term work)

Laboratory work will be based on **DJ19ICCHN1C3** with a minimum of 08 experiments.

The distribution of marks for term work shall be as follows:

1. Laboratory work (Performance of Experiments): 15 Marks

2. Journal Documentation (Write-up and Assignments): 10 Marks

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

Prepared by Checked by Head of the Department Vice Principal Principal