SAMRAT VIKRAMADITYA VISHWAVIDYALAYA, UJJAIN MP

(Session -2022-2023 onwards)

B.Sc. (Hon's) Computer Science FOUR YEARS (EIGHT SEMESTERS), CBCS SCHEME (NEP)

B.Sc. (Hon's) Computer Science VIII Sem

S.No.	Paper code	Course Component and Name of Course	Credits			Marks		Total
			T	P	Total	Max Marks CCE Internals	Max Marks Theory Externals	
1.	BSCH 801 Major-1	Big Data Technologies	3	0	3	40	60	100
2.	BSCH 802 Major-2	Machine Learning	3	0	3	40	60	100
3.	BSCH 803	Dissertation	4	0	4	40	60	100
4.	BSCH-804	Field Project/Research Project/Internship/ Apprenticeship	10	0	10	100	300	100
Total Credits and marks			20		20	220	380	600

PART A: Introduction				
Program: Honors with Research		Class: VIII Sem	Year: IV Year	Session: 2024-25
Subject: Data science				
1.	Course Code	**BSCH-801**		
2.	Course Title	**Big data Technology**		
3.	Course Type (Core Course/Elective/Generic Elective/ Vocational	**Major-I**		
4.	Pre-Requisite (if any)	To study this course, a student must have the basic knowledge of DBMS and Data Mining.		
5.	Course Learning Outcomes(CLO)	• Student must be able to understand the building blocks of Big data. • Student must be able to articulate the programming aspects of cloud computing. • Student must be able to understand the specialized aspects of big data with the help of different big data applications. • Student must be able to represent the analytical of big data. • Student must be know the recent research trends related to Hadoop file system, Map redeuce and Google file system etc.		
6.	Credit Value	3 Credit		
7.	Total Marks	Max. Marks : 100	Min. Passing Marks: 35	

PART B: Content of the Course		
Total No. of Lectures (in hours per week): 01 Hours per day		
Total Lectures: 60 Hours		
Unit	Topics	No. of Lectures
I	Introduction to Big Data Platform – Challenges of Conventional Systems - Intelligent data analysis – Nature of Data - Analytic Processes and Tools - Analysis vs Reporting – Modern Data Analytic Tools. Big Data Analytics Process, Big Data Analytics for Business. Identifying problem and solving problem in Big Data environment. Analyzing Unstructured vs. Structured Data, Databases.	12
II	Hadoop and MapReduce Introduction to Hadoop, Hadoop architecture, A Brief History of Hadoop, Apache Hadoop and the Hadoop Ecosystem, Hadoop Releases; Hadoop Distributed File system: Design of HDFS, HDFS Concepts.	12
III	Introduction to MapReduce: MapReduce Basic Concepts, Understanding the Map Reduce architecture, Writing MapReduce Programs.understanding Map phase, shuffling, sorting, and reducing phase.	12

IV	Spark Introduction to Spark, Resilient Distributed Dataset (RDD), RDD Operations: actions and transformation functions. Spark Data frames, operations on Data frames: Join, group by, aggregate, handling missing data.	12
V	:Sparks and MLLib Sparks and its basic operations. MLLib: Data types, Basic statistics, Classification(Logistic regression, Decision tree classifier)and linear regression model generation, Model Evaluation, Collaborative filtering, and Clustering.	12

PART C: Learning Resources

Textbooks, Reference Books, Other Resources

Suggested Readings:

1. Arvind Sathi, "Big Data Analytics: Disruptive Technologies for Changing theGame", 1st Edition, IBM Corporation, 2012.

2. Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", 1st Edition, Wiley and SAS Business Series, 2012.

3. Tom White,"Hadoop: The Definitive Guide", 3rd Edition, O'Reilly Media, 2012

4. Donald Miner, Adam Shook, Eric Sammer, "Hadoop Operation", O'Reilly 2012.

5. Donald Miner, Adam Shook "MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems", O'Reilly 2012.

6. Chuck Lam, "Hadoop in Action", Manning Publications, 2010.

Suggested equivalent online courses:

https://onlinecourses.nptel.ac.in/noc20_cs92/preview
https://archive.nptel.ac.in/courses/106/104/106104189/
https://nptel.ac.in/courses/106104189
https://archive.nptel.ac.in/noc/courses/noc21/SEM2/noc21-cs86/

Part D: Assessment and Evaluation

Suggested Continuous Evaluation Methods:

Maximum Marks: **100**
Continuous Comprehensive Evaluation (CCE): **40** Marks
University Exam (UE): **60** Marks

Internal Assessment: Continuous Comprehensive Evaluation (CCE)	Class Test Assignment/Presentation	20 20 **Total Marks: 40**
External Assessment: University Exam (UE) Time: 03.00 Hours	Section (A): Objective type Section (B): Short Questions Section (C): Short Questions	04 × 05 = 20 08 × 05 = 40 **Total Marks: 60**

PART A:				
Program: Honors with Research		Class: VIII Sem	Year: IV Year	Session: 2024-25
Subject: Computer Application				
1.	Course Code	**BSCH -802**		
2.	Course Title	**Machine Learning**		
3.	Course Type (Core Course/Elective/Generic Elective/ Vocational	**Major -II**		
4.	Pre-Requisite (if any)	To study this course, a student must have the basic knowledge of Artificial Intelligence.		
5.	Course Learning Outcomes(CLO)	• Students will be able to explain the need of machine learning and model building. • Understand the concept to apply the supervised algorithms. • Develop a skill to implement unsupervised algorithms for problem solving. • Understand the concept of reassurance learning algorithms. • Able to apply the learning algorithms in real world problem solving.		
6.	Credit Value	3		
7.	Total Marks	Max. Marks : 100	Min. Passing Marks:35	

PART B: Content of the Course		
Total No. of Lectures (in hours per week): 03 Hours per week		
Total Lectures: Hours		
Unit	Topics	No. of Lectures
I	Introduction to Machine learning: Type of Learning and Examples, basic concepts in machine learning, Computational Learning theory, Introduction to Parametric Models – Non-Parametric Models – Probability Basics.	12
II	Supervised Learning Algorithms Supervised Machine Learning Algorithms, working of supervised machine learning algorithm, Naive Bayes algorithm, decision tree, Support Vector Machines, KNN, Random Forest algorithm.	12
III	Clustering- K-means -EM Algorithm- Mixtures of Gaussians - Dimensionality Reduction - Factor analysis – Feature Selection - Principal Component Analysis - Probabilistic PCA - Independent components analysis - Singular Value Decomposition.	12
IV	Reinforcement Learning Algorithms Reinforcement Machine Learning Algorithms, working of reinforcement machine learning algorithm, Finite Markov Decision Processes, Dynamic Programming, Monte Carlo Methods	12
	Ranking: Priority Inbox - Ordering Email Messages by Priority - Writing a Priority	

V	Inbox - Spam Filtering - Analyzing Social Graphs - Social Network Analysis - Hacking Twitter Social Graph Data - Analyzing Twitter Networks – Case Study.	12

PART C: Learning Resources

Textbooks, Reference Books, Other Resources

Suggested Readings:

- Ethem Alpaydın(2020), Introduction to Machine Learning, The MIT Press Cambridge, Fourth Edition, MIT Press Hardcover.
- Shai Shalev-Shwartz, Shai Ben-David(2014), Understanding Machine Learning: From Theory to Algorithms, Cambridge University Press.
- V Kishore Ayyadevara(2018), Pro Machine Learning Algorithms A HandsOn Approach to Implementing Algorithms in Python and R, Apress
- Kevin P. Murphy(2022) , Probabilistic Machine Learning an Introduction. The MIT Press

Suggestive digital platform web links:

1. https://alex.smola.org/drafts/thebook.pdf
2. https://onlinecourses.nptel.ac.in/noc21_cs85/preview
3. https://onlinecourses.nptel.ac.in/noc21_cs70

Part D: Assessment and Evaluation

Suggested Continuous Evaluation Methods:

Maximum Marks: **100**

Continuous Comprehensive Evaluation (CCE): **40** Marks

University Exam (UE): **60** Marks

Internal Assessment: Continuous Comprehensive Evaluation (CCE)	Class Test Assignment/Presentation	20 20 **Total Marks: 40**
External Assessment: University Exam (UE) Time: 03.00 Hours	Section (A): Objective type Section (B): Short Questions Section (C): Short Questions	$04 \times 05 = 20$ $08 \times 05 = 40$ **Total Marks: 60**