

Devi Ahilya University, Indore, India Institute of Engineering & Technology				IVYear B.E. (Computer Engg.) (Full Time)			
Subject Code & Name	Instructions Hours per Week			Credits			
8CERC3 Data Sciences	L	T	P	L	T	P	Total
Duration of Theory Paper: 3 Hours	3	1	2	3	1	1	5

Learning Objectives:

- This course will introduce students to data science basic principles and tools as well as its general mindset.
- Students will learn concepts, techniques and tools they need to deal with various facets of data science practice, including data collection and integration, exploratory data analysis, predictive modeling, descriptive modeling, data product creation, evaluation, and effective communication.

Prerequisites: Basic knowledge of algorithms and sufficient programming experience and familiarity with basic linear algebra, probability and statistics.

Course Outcome:

Students earned credits will develop ability to

CO. No.	Course Outcome	Program Outcomes
CO1	Students will be able to understand the fundamental concepts of data science, including data management, data warehousing, and data wrangling techniques.	PO1, PO2,PO3,PO4,PO5,PO9 PO12
CO2.	Apply statistical methods and exploratory data analysis techniques to summarize and interpret datasets.	PO1, PO2,PO3,PO4,PO5,PO9, PO10, PO12
CO3	Develop and analyze machine learning models using regression, classification, clustering, and other algorithms.	PO1, PO2,PO3,PO4,PO5, PO10,PO11
CO4	Demonstrate the ability to manage and process large-scale data using Big Data technologies such as Hadoop and Spark.	PO1, PO2,PO3,PO4,PO5, PO9
CO5	Evaluate advanced data science applications including social network analysis, recommender systems, and ethical data practices.	PO1, PO2,PO3,PO4,PO5, PO9,PO11

CO-PO Relationship

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	2	1	1	2				1			2
CO2	3	3	2	2	3				1	1		3
CO3	3	3	3	3	3					2	2	
CO4	3	3	2	2	3				1			
CO5	3	2	2	2	2				2		2	

1. * CO (rows) mention nil/very small/insignificant contribution to the PO(column)
2. 1 → relevant and small significance 2 → medium or moderate and 3 → strong

COURSE OF CONTENTS

Unit 1 : Introduction to Data Management

What is Data Science?, Data Science Languages, Data Warehousing & OLAP, Data Preparation, Data Wrangling etc.

Unit 2 : Statistics and EDA

Descriptive Statistics, Inferential Statistics, Exploratory Data Analysis, Hypothesis Testing etc.

Unit 3 : Machine Learning

Linear Regression, Supervised Classification – K-Nearest Neighbors, Clustering – K-Means etc., Decision Trees, Support Vector Machine and Neural Networks etc.

Unit 4 : Big Data Analytics

Introduction to Big Data and Hadoop, Managing Big Data, Introduction to SPARK, Big Data Analysis etc.

Unit 5 : Advance Topics

Mining Social Network graphs; Privacy, Security and Ethical Issues in Data Science; Data Visualization; Recommended Systems etc.

RECOMMENDED BOOKS

- [1] Cathy O'Neil and Rachel Schutt. Doing Data Science, Straight Talk From The Frontline. O'Reilly. 2014.
- [2] Jiawei Han, Micheline Kamber and Jian Pei. Data Mining: Concepts and Techniques, Third Edition. ISBN 0123814790. 2011.

- [3] Trevor Hastie, Robert Tibshirani and Jerome Friedman. Elements of Statistical Learning, Second Edition. ISBN 0387952845. 2009.
- [4] Foster Provost and Tom Fawcett. Data Science for Business: What You Need to Know about Data Mining and Data-analytic Thinking. ISBN 1449361323. 2013.
- [5] Allen Downey. Think Python. Oreilly. First Edition. ISBN 144933072X. 2012.
- [6] Joel Grus. Data Science from Scratch. Shroff. ISBN 9352130960. 2015.

Learning Outcomes:

At the conclusion of the course, students should be able to:

- Describe what Data Science is and the skill sets needed to be a data scientist.
- Use Python to carry out basic statistical modeling and analysis.
- Apply EDA and the Data Science process in a case study.
- Apply basic machine learning algorithms (Linear Regression, k-Nearest Neighbors (k-NN), K-Means) for predictive modeling.
- Create effective visualization of given data, reason around ethical and privacy issues in data science conduct and apply ethical practices. e