

DAIDS-1-02T: Introduction to Data Science

Total Marks: 100
 External Marks: 70
 Internal Marks: 30
 Credits: 6
 Pass Percentage: 40%

Course: Introduction to Data Science	
Course Code: DAIDS-1-02T	
Course Outcomes (COs)	
After the completion of this course, the students will be able to:	
CO1	Understand tools and techniques to analyze and extract insights from data received from different data sources such as social media, IoT devices, and sensors.
CO2	Understand the general techniques and frameworks that can be used to handle special types of data, such as acoustic, image, sensor, and network data
CO3	Apply mathematical or logical operations to the data to derive new insights.
CO4	Apply tools for understanding complex data structures and relationships.
CO5	Explore various applications of data science in the field of business, energy, health care, biotechnology, manufacturing, telecommunication, pharmaceuticals etc.

Detailed Contents:

Module	Module Name	Module Contents
Module 1	Data Science	Introduction to Data Science, Landscape-Data to Data science, Data Growth-issues and challenges, data science process. foundations of data science. Messy data, Anomalies and artefacts in datasets. Cleaning data.
Module II	Data Acquisition and Processing	Introduction, Structured Vs Unstructured data, data preprocessing techniques including data cleaning, selection, integration, transformation and reduction, data mining, interpretation.
Module III	Representation of data	Special types-acoustic, image, sensor and network data. Problems when handling large data – General techniques for handling large data, distributing data storage and processing with Frameworks
Module IV	Data Science Ethics	Doing good data science – Owners of the data - Valuing different aspects of privacy - Getting informed consent - The Five Cs – Diversity – Inclusion – Future Trends.
Module V	Data Modeling:	Basics of Generative modeling and Predictive modeling. Charts-histograms, scatter plots, time

		series plots etc. Graphs, 3D Visualization and Presentation.
Module VI	Applications of Data Science	Business, Insurance, Energy, Health care, Biotechnology, Manufacturing, Utilities, Telecommunication, Travel, Governance, Gaming, Pharmaceuticals, Geospatial analytics and modeling

Books

<ol style="list-style-type: none"> 1. Sinan Ozdemir, “Principles of Data Science”, Packt Publishing, 2016 2. Joel Grus, “Data Science from Scratch”, O’Reilly, 2016 3. Foster Provost & Tom Fawcett, “Data Science for Business”, O’Reilly, 2013 4. Roger D. Peng & Elizabeth Matsui, “The Art of Data Science”, Lean Publishing, 2015 5. Peter Bruce, Andrew Bruce, Peter Gedeck, “Practical Statistics for Data Scientists-2e: 50+ Essential Concepts Using R and Python”, O’Reilly
--