

Name of the Faculty: Krishna Reddy Polepalli  
 Name of the Academic Program: CSE  
 Email: pkreddy@iiit.ac.in

Course Code: CS4.405

Title of the course: Data Analytics-I

L-T-P: **3-1-1**. Credits: **4** (L= Lecture hours, T=Tutorial hours, P=Practical hours)

TYPE-WHEN: Fifth semester and onwards

### 1.Prerequisite Course / Knowledge:

- (i) Data and Applications, *or equivalent courses that cover Data modelling, normalization, SQL*
- (ii) First courses on programming, data-structures and algorithms
- (iii) Basics of Python language, to be able to use relevant libraries and toolkits for data analytics

### 2.Course Outcomes (COs)

Objective: In a computerized and networked society, vast amount of data is being collected every day in multiple domains. We are drowning in data, but starving for knowledge or actionable insights. Data mining or data analytics constitute a collection of concepts and algorithms, which are being developed to answer “how” questions by extracting interesting and useful knowledge of from large data. Data analytics based platforms are being operated in multiple domains to extract valuable and actionable insights from the data to improve the business performance. The objective of this first level course is to learn the important concepts and algorithms related to data mining functionalities such as summarization, pattern mining, classification, clustering and outlier analysis.

The Course Outcomes (COs) are as follows:

- After completing the course successfully, the students are able to
  - CO-1. describe the concepts of data summarization, data warehousing, pattern mining, classification and clustering approaches
  - CO-2. perform the task of data summarization, pattern mining, classification and clustering based on the requirement.
  - CO-3. prescribe a single or a combination of data summarization, pattern mining, classification and clustering approaches for the problem scenario of a business/organization.
  - CO-4. construct the improved data analytics methods for existing services.
  - CO-5. formulate new data mining problems for creating new services and design the corresponding solutions

### 3. Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs) – Course Articulation Matrix

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO1 0	PO1 1	PO1 2	PSO 1	PSO 2	PSO 3	PSO 4
CO 1	3	2	3	3	3	3	2	2	3	2	2	3	2	2	2	3
CO 2	3	1	3	1	3	2	1	1	3	3	2	3	2	1	2	3
CO 3	1	3	1	3	1	1	1	1	1	1	1	3	3	1	2	3
CO 4	1	2	1	3	3	1	1	1	3	1	2	3	3	1	2	3
CO 5	2	3	3	1	1	1	2	1	3	3	2	3	2	2	2	3

‘3’ in the box denotes ‘High-level’ mapping, 2 for ‘Medium-level’ mapping, 1 for ‘Low’-level’ mapping.

#### 4.Detailed Syllabus

(please list the order in which they will be covered)

Unit 1: Introduction, data summarization through characterization, discrimination and data warehousing techniques (9 hours)

Unit 2: Concepts and algorithms for mining patterns and associations (9 hours)

Unit 3: Concepts and algorithms related to classification and regression (9 hours)

Unit 4: Concepts and algorithms for clustering the data (9 hours)

Unit 5: Outlier analysis and future trends. (3 hours)

- Five mini projects related to the above syllabus will be done by students in the laboratory

#### Reference Books and materials:

1. Book: Jiawei Han and Micheline Kamber, Data Mining: Concepts and Techniques, Third edition, 2012, Elseiver Inc.
2. Book: Pang-Nong Tan, Michael Steinbach and Vipin Kumar, Introduction to Data Mining, 2006, Pearson Education.
3. Research Papers: About 25 research papers from the proceeding of the conferences and journals related to data summarization, data warehousing, pattern mining, classification, clustering, outlier detection.

#### 5. Teaching-Learning Strategies in brief

Lectures by integrating ICT into classroom teaching, weekly tutorials involving problem solving and active learning by students and Project-based Learning by doing 5 mini-projects in laboratory by the students

#### 6.Assessment methods and weightages in brief

Two Class Room tests: 10 marks; Mid Semester Examination in theory: 20 marks , End Semester Examination in Theory: 40 marks, Assessment of 5 mini projects in Laboratory: 30 marks

\*\*\*\*\*