| CM 313 | Data and Visual Analytics in AI | L | T | P | Int. | Ext. | C |
|---|---|---|---|---|---|---|---|
| | | 3 | - | - | 30 | 70 | 3 |

**Course Objectives:**

The main objectives of this course are to:

1. This course introduces the visualization techniques of data.
2. To enable students to make more effective use of data.
3. To utilize various levels and types of summarization of data

**Course Outcomes:**

After successful completion of the course, students will be able to:

CO 1. Use basic data types and preprocessing techniques of data according to needs.
CO 2. Apply the data visualization through various graphs to analyse the data.
CO 3. Apply the visual distribution of data.
CO 4. Understand the multiple visual distribution of data.

**Course Content:**

| **UNIT – I** | | **12 Periods** |
|---|---|---|
| **Data Warehouse:** Data Warehouse: Basic Concepts, Data Warehouse Modelling: Data Cube and OLAP, Data Warehouse Design and Usage, Data Warehouse Implementation.<br>**Getting to know Data:** Data Objects and Attribute Types, Basic Statistical Descriptions of Data, Measuring Data Similarity and Dissimilarity. | | |
| **UNIT – II** | | **12 Periods** |
| **Data Mining:** What is Data Mining, Kinds of Data, Kinds of Patterns, Technologies Used, Major Issues in Data Mining.<br>**Data Pre-processing:** Data cleaning, Data Integration, Data Reduction, Data Transformation and Data Discretization. | | |
| **UNIT – III** | | **12 Periods** |
| **Visualizing Data:** Mapping data onto aesthetics, aesthetics and types of data, scales map data values onto aesthetics, visualizing amounts: bar plots, grouped and stacked bars, dot plots and heat maps, exploration of visualization tools.<br>**Visualizing Distributions:** Histograms and density plots - visualizing a single distribution, visualizing multiple distributions at the same time. Empirical cumulative distribution functions and q-q plots - empirical cumulative distribution functions, highly skewed distributions, quantile-quantile plots. | | |
| **UNIT – IV** | | **12 Periods** |
| **Visualizing Multiple Distributions:** Visualizing distributions along the vertical axis, visualizing distributions along the horizontal axis. Visualizing associations among two or more quantitative variables - scatter plots, scatter plot matrix, ggplots, correlograms, dimension reduction, paired data. | | |

**Learning Resources:**

**Text Book:**

1. Jiawei Han and Micheline Kamber, Data Mining- Concepts and Techniques, Morgan Kaufmann Publishers, Elsevier, 3rd Edition.
2. Claus O. Wilke, Fundamentals of Data Visualization, Oreilly publication, 1st Edition .

**erence Books:**

1. Arun K Pujari, Data Mining Techniques, 3rdEdition, Universities Press.
2. Kieran Healy, Data Visualization: A Practical Introduction 1stEdition, Princeton university press