

**Prasad V. Potluri Siddhartha Institute of Technology:: Vijayawada.
Department of Computer Science and Engineering**

I/II M.Tech. (CSE) - (First Semester)

17CSCS1T5C

**ADVANCED DATA MINING
Elective - I**

Credits: 4

Lecture: 4 Periods/week

**Internal Assessment: 40 Marks
Semester end examination: 60 Marks**

Course Description

Data Mining studies algorithms and computational paradigms that allow computers to find patterns and regularities in databases, perform prediction and forecasting, and generally improve their performance through interaction with data. The knowledge discovery process includes data selection, cleaning, coding, using different statistical and machine learning techniques, and visualization of the generated structures. The course will cover advanced classification techniques, pattern mining, cluster analysis and various methodologies for mining complex data types.

Course Outcomes:

At the end of the course the student will be able:

CO1: To understand the Data mining Concepts and the methods for pattern mining.

CO2: To demonstrate the Classification Methods.

CO3: To explain Cluster Analysis in high-dimensional data.

CO4: To understand the various data mining methodologies for complex data types.

Unit-1:

Data mining Overview: Why Data Mining? , What is data mining?, What kind of Data can be Mined?, What kind of patterns can be minded? Major Issues in Data Mining.

Advanced pattern mining: Pattern Mining, Pattern Mining in multilevel, multidimensional space, Constraint-Based Frequent Pattern Mining, Mining High Dimensional Data and Colossal Patterns

Unit-2:

Classification: Advanced Methods - Classification by back propagation, support vector machines, lazy learners, other classification methods – genetic algorithms, rough set approach, fuzzy set approach;

Unit-3:

Cluster Analysis : Density based methods –DBSCAN, OPTICS, DENCLUE; Grid-Based methods – STING, CLIQUE.

Advanced Cluster Analysis: Exception – maximization algorithm; clustering High-Dimensional Data; Clustering Graph and Network Data.

Unit-4:

Temporal and Spatial Data Mining: Introduction; Temporal Data Mining – Temporal Association Rules, Sequence Mining, GSP algorithm, SPADE, SPIRIT Episode Discovery, Time Series Analysis, Spatial Mining – Spatial Mining Tasks, Spatial Clustering. Data Mining Applications.

TEXT BOOKS:

1. Data Mining Concepts and Techniques, Jiawei Han Micheline Kamber, Jian pei, Morgan Kaufmann 3editon (Unit 1,2,3)
2. Data Mining Techniques – Arun K pujari, Universities Press. (Unit-4)

REFERENCE BOOKS:

1. Introduction to Data Mining – Pang-Ning Tan, Vipin kumar, Michael Steinbach, Pearson.
2. Data Mining Principles & Applications – T.V Sveresh Kumar, B.Esware Reddy, Jagadish S Kalimani, Elsevier.