

# **ANL303A Fundamentals of Data Mining**

**Level:** 3

**Credit Units:** 5 Credit Units

**Language:** ENGLISH

**Presentation Pattern:** EVERY SEMESTER

## **Synopsis:**

This course introduces students to the concepts and applications of data mining. Students will be introduced to the methodology of data mining, the data preparation and data exploration process as well as data mining techniques for association, clustering and predictive modelling. The course gives students an introduction on how these techniques can be applied to business analytics problems.

## **Topics:**

- Overview of Data Mining
- Process of Data Mining
- Data Exploration
- Data Visualisation
- Data Cleaning
- Data transformation and reduction
- Association analysis
- Clustering analysis
- Decision Trees
- Model evaluation
- Business and data phases
- Model evaluation and deployment

## **Textbooks:**

by A.B.M Shawkat Ali, Saleh A. Wasimi: Data mining: methods and techniques 2007 Edition. Cengage  
ISBN-13: 9789814510875

by A.B.M Shawkat Ali, Saleh A. Wasimi: Data mining: methods and techniques 2007 Edition. Cengage  
ISBN-13: 9789814510875-AA

**Learning Outcome:**

- Differentiate the various aspects of data mining.
- Recommend data mining tools for association analysis, clustering and predictive modelling.
- Assess the following applications: association analysis with Apriori, clustering with K-means, classification with classification and regression tree.
- Categorise different aspects of data preparation.
- Discuss the use of data exploration techniques using summary statistics, dimension reduction and visualisation.
- Plan the process of data mining, i.e. CRISP-DM framework.
- Prepare data for mining and analysis.
- Execute techniques such as association analysis with Apriori, clustering with K-means, and classification with CART (classification and regression tree).
- Analyse data using exploration techniques such as summary statistics, dimension reduction and visualisation.
- Defend the use of appropriate data mining techniques for different business problems.
- Interpret the results of a data mining analysis.
- Evaluate the validity of different data mining models.
- Apply the above-mentioned data mining tasks using the software package specified in this course, examine the output produced by the software, and infer their implications for the problem(s) under consideration.

**Assessment Strategies:**

<b>Continuous Assessment Component</b>	<b>Weightage (%)</b>
PRE-COURSE QUIZ	5
GROUP BASED ASSIGNMENT	10
LAB REPORT	15
INDIVIDUAL CASE STUDY REPORT	20
<b>Sub-Total</b>	<b>50</b>

<b>Examinable Component</b>	<b>Weightage (%)</b>
Written Exam	50
<b>Sub-Total</b>	<b>50</b>

**Weightage Total** **100**