



**Mehran University of Engineering and Technology, Jamshoro**  
**Department of Software Engineering**

**ORIGINAL SUBMITTED SYLLABUS**

<b>Title of Subject</b>	:	<b>DATA MINING CONCEPTS &amp; TECHNIQUES</b>
<b>Code</b>	:	SE611
<b>Discipline</b>	:	Software Engineering (1 <sup>st</sup> Semester)
<b>Effective</b>	:	14 ME SE and onwards
<b>Pre-requisite</b>	:	None
<b>Assessment</b>	:	10% Sessional 30% Mid Semester 60% Final Semester Examination
<b>Credit Hours</b>	:	3+0
<b>Minimum Contact Hours:</b>		<b>Marks : 100 + 00</b>

---

**Specific Objectives of course:**

- This course on data mining will cover methodology, major software tools and applications in this field. By introducing principal ideas in statistical learning, the course will help students to understand conceptual underpinnings of methods in data mining. The focus will be on the breadth of data mining and its applications in information technology.

**Course outline:**

**Introduction:** Relational Databases, Data Warehouses, Transactional Databases, Introduction to Data Mining, Knowledge Discovery Process, Major Applications of Data Mining, Major Issues in Data Mining.

**Data Preprocessing:** Why Preprocess the Data, Descriptive Data Summarization, Major Tasks in Data Preprocessing (Data cleaning, integration and transformation), How to Handle Missing Data, How to Handle Noisy Data, Data reduction Methods, Discretization and Concept hierarchy for numeric data.

**Mining Frequent Patterns, Association & correlations:** Relationships & Patterns, Market basket analysis, Frequent Itemsets, Closed Itemsets and Association Rules, Frequent Pattern Mining,

**Classification & Prediction:** Data Mining Functionalities, Classification Schemes, Descriptive & Predictive Data Mining, Issues Regarding Classification and Prediction, Classification Methods.

**Association Mining:** Association Mining, Support & Confidence, Apriori Algorithm, Sampling Algorithm, Frequent-Pattern Tree Algorithm, Partition Algorithm.

**Cluster Analysis:** Introduction to clustering, Major clustering methods, Outlier Analysis

**BOOKS RECOMMENDED**

1. Jiawei Han, Micheline Kamber, and Jian Pei. Data Mining: Concepts and Techniques (Latest Edition). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
2. Thomas Connolly and Carolyn E. Begg. Database Systems: A Practical Approach to Design, Implementation and Management. (Latest Edition) Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
3. Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. Introduction to Data Mining, (Latest Edition). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

---

**Approval:**

Board of Studies:  
AS&RB:  
Academic Council:

Res.No.02  
Res.No. 127.85  
Res.No. 83.20

Dated:13-02-2014  
Dated: 10-03-2014  
Dated: 30-05-2014