CPIT-440 Syllabus

Catalog Description

CPIT-440 Data Mining and Warehousing
Credit: 3 (Theory: 3, Lab: 0, Practical: 1)
Prerequisite: CPIT-340
Classification: Elective

The objective of this course is to explore the different knowledge extraction methods and its representation techniques as well as knowledge engineering. It also introduces the different basic artificial intelligence theories that qualify the students to understand the contents of the course. Topics include an introduction to data mining and warehousing, data warehousing and OLAP technology, classification and prediction, mining frequent patterns, associations and correlations, review of probability and statistics, and data preprocessing.

Class Schedule

Lab/Tutorial 90 minutes 1 times/week

Meet 50 minutes 3 times/week or 80 minutes 2 times/week

Jiawei Han, Micheline Kamber, , "Data Mining" 3 edition (2012)
ISBN-13 9789380931913 **ISBN-10** 9380931913

Grade Distribution

Week	Assessment	Grade %
7	Exam 1	15
9	Homework Assignments 1	5
10	Exam 2	15
12	Homework Assignments 2	5
13	Lab Exam	10
15	Group Project	10
16	Exam	40

Topics Coverage Durations

Topics	Weeks
Introduction to Data Mining and Warehousing	1
Data	2
Data Preprocessing	2
Data Warehousing and OLAP Technology	2
Mining Frequent Patterns, Associations and Correlations	2
Classification and Prediction	3
Cluster Analysis	3

Last Articulated

March 6, 2017

Relationship to Student Outcomes

a	b	c	d	e	f	g	h	i	j	k	l	m	n
x		x						x					

Course Learning Outcomes (CLO)

By completion of the course the students should be able to

1. **Design and implement data mining solution to a given problem (c)**
2. Use numerical and graphical methods to summarize data (i)
3. **Identify and remove missing values, noise and outliers in presented data (i)**
4. Identify relationships in presented data and Measuring Data Similarity and Dissimilarity (a)
5. Differentiate between data warehouse and operational database (a)
6. **Design a data warehouse schema using fact and dimension tables (c)**
7. Describe the OLAP operations to analyze data in a data cube (a)
8. **Use frequent pattern mining methods to extract association rules from given data (i)**
9. Characterize and rank the association rules using support, confidence and lift (a)
10. **Use classification to build prediction models (i)**
11. Validate a prediction model using split and cross validation methods (i)
12. Compare and contrast classifiers using various evaluation measures (a)
13. Differentiate between hierarchical and partitioning clustering methods (a)
14. Use clustering algorithms to form clusters in given data (i)

Coordinator(s)

Dr. Reem Alotaibi, Associate Professor