COURSE DESCRIPTION

Course Code Number	CS-505	Course Title	Data Mining Concept and Technique
Semester hours	37.5 hours (1 period - 50 minutes)	**Credit Units**	3
Prerequisite	CST-204, CS-304 CS-404	**Course Coordinator**	Daw Khaing, Associate Professor, Faculty of Information Science

Course Description:

The course is designed to expand students' knowledge and skill gained in database management courses and look in depth at data warehousing and data mining methods. The course examines the database architectures and technologies required for solving complex problems of data and information management, information retrieval and knowledge discovery facing modern organizations. This course also provides hands-on experience with state-of-the-art data mining methods tools using WEKA software.

Course Contents:

1. Introduction to Data Mining
2. Getting to know your data
3. Data Preprocessing
4. Mining frequent patterns, Associations and Correlations, Basic Concept and Method
5. Classification Basic Concept
6. Cluster Analysis: Basis Concept and Method

Course Objectives:

This course is intended to

- Introduce the basic concepts and techniques of data mining.
- Develop skills of using recent data mining software for solving practical problems.
- Gain experience of doing independent study and research
- Study the methodology of engineering legacy database for data mining to derive business rules for decision support systems.
- Develop and apply critical thinking, problem solving and decision making skills.

Learning Outcomes:

Students who complete the course should be able to

- Know the latest development of knowledge discovery and data mining concepts and techniques.
- Utilize the theories and algorithms for data mining and knowledge discovery.
- Manipulate the possibilities and fundamental limitations that are included in data preprocessing steps of data mining.
- Analyze and compare the performance of different mining methods on a wide range of datasets using data mining methods such association, classification and clustering analysis.
- Apply the data mining method using WEKA tools for analysis.
- Implement the relevant applications in specific domains such as medicine and health care, market basket analysis, etc.

Software or Tools

- **WEKA software**

Reference

Jiawei Han, Micheline Kamber and Jian Pei, Data Mining – Concepts and Techniques. Morgan Kaufmann, Third Edition, 2011.

Reference Books

1) Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Addison Wesley, 2006.
2) Ian H. Witten and Eibe Frank, Data Mining – Practical Machine Learning Tools and Techniques (2nd Ed.), Morgan Kaufmann, 2005.
3) S.M. Weiss and N. Indurkhya, Predictive Data Mining, Morgan Kaufmann, 1998.
4) Margaret H. Dunham, Data Mining – Introductory and Advanced Topics, Prentice Hall, 2003.

Exam Assessment

Exam Paper   60%

Tutorials/Test (3 times)   10%

Project /Presentation   10%

Assignment/Attendance (5+5)   10%

Quiz/Discussion (6 times)   10%

<div align="center">

University of Computer Studies
2017-2018 Academic Year
B.C.Sc (Fifth Year)

</div>

CS-505   : Data Mining   (First Semester)

Text Book   : Data Mining Concepts and Techniques (Jiawei Han, Micheline Kamber, Jian Pei)

Period   - 45 periods for 15 Weeks including Practice (50 minutes for 1 period)

No.	Chapter	Page	Period	Remark
1	**1 Introduction**	1-34	5	
	1.1 Why Data Mining?			
	1.2 What is Data Mining?			
	1.3 What Kinds of Data Can Be Mined?			
	1.4 What Kinds of Patterns Can Be Mined?			
	1.5 Which Technologies Are Used?			
	1.6 Which Kinds of Applications Are Targeted?			
	1.7 Major Issues in Data Mining			
	1.8 Summary			
	1.9 Exercises			
2	**2 Getting to Know Your Data**	39-79	10	
	2.1 Data Objects and Attributes Types			
	2.2 Basic Statistical Description of Data			
	2.3 Data Visualization			
	2.4 Measuring Data Similarity and Dissimilarity			
	2.5 Summary			
	2.6 Exercises			
3	**3 Data Preprocessing**	83-121	10	
	3.1 Data Preprocessing: An Overview			
	3.2 Data Cleaning			
	3.3 Data Integration			
	3.4 Data Reduction			
	3.5 Data Transformation and Data Discretization			
	3.6 Summary			
	3.7 Exercises			
4	**6 Mining Frequent Patterns, Associations, and Correlation: Basic Concepts and Methods**	243-273	5	
	6.1 Basic Concepts			
	6.2 Frequent Itemset Mining Methods			
	6.3 Which Patterns Are Interesting? Pattern Evaluation Methods			
	6.4 Summary			