

**1. Course Number and Course Title:**

COE 637 Data Mining and Knowledge Discovery

**2. Credits Hours:**

3 – 0 – 3

**3. Prerequisites and/or Co-Requisites:**

Prerequisite: Admission to the MSCoE Program Co-requisites:  
None

**4. Name and Contact Information of Instructor:**

Dr. Michel Pasquier

Office: EB2-213

Email: mpasquier@aus.edu

Phone: (06) 515-2883

Office Hours: as posted on *iLearn*; also by appointment

**5. Course Description (Catalog Description):**

Introduces the principles of data mining and knowledge discovery (KDD). Covers key topics including data preparation, visualization, pattern recognition and statistical machine learning techniques, experimental validation, and model interpretation. Studies examples of practical applications using state-of-the-art software in the field, such as R or Weka.

**6. Textbook and other Supplemental Material:**

Textbook:

- Witten I.H., Frank E., and Hall M.A., Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 4th edition, 2016.

Other supplemental material:

- Han J., Kamber M., and Pei J., Data Mining: Concepts and Techniques, Morgan Kaufmann, 3rd edition, 2012.
- Hastie T., Tibshirani R., and Friedman J., The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2nd edition, 2009.
- Weka 3 Data Mining software (free, open source), University of Waikato, 2014. □ Selected papers and data sets.

**7. Learning Outcomes:**

Upon completion of the course, students will be able to:

1. Explain the importance of data mining and knowledge discovery to IT and society.
2. Design tasks and select accordingly data, knowledge representations, and methods.
3. Understand the fundamentals of statistics and probability theory.
4. Apply and evaluate data mining techniques and tools to solve real world problems.
5. Use modern software, such as R or Weka, to explore and apply data mining theories.
6. Critically assess current research work in data mining and knowledge discovery.
7. Conduct and present independent research on data mining and knowledge discovery.

**8. Teaching and Learning Methodologies:**

Methods include lectures; problem and project based learning methods (assignments, class presentation, survey report, exams and research project) and class discussions.

#### 9. Course Topics and Schedule:

Topic	Weeks
Introduction: data mining and knowledge discovery	1
Input data: concepts, instances, attributes, preparation	1
Output knowledge: tables, trees, rules, visualization	1
Basic methods: inferring rules, decision trees, clustering	3
System evaluation: testing, validation, performance	2
Advanced methods: learning, classification, prediction	3
Data transformation: sampling, projections, cleansing	1
Ensemble learning: combining models, bagging, boosting	1
Review and evaluation, class presentations	3
<b>Total:</b>	<b>16</b>

#### 10. Schedule of Laboratory and other Non-Lecture Sessions: None

#### 11. Out-of-Class Assignments with Due Dates:

Assignment	Due Date (tentative)
Homework 1: Introduction to Weka data mining software	End of Week 3
Homework 2: Data preparation and mining basics	End of Week 5
Homework 3: Data cleaning and preparation for modeling	End of Week 7
Homework 4: Data selection and feature reduction	End of Week 9
Homework 5: Classification and prediction	End of Week 11
Homework 6: Data mining application case study	End of Week 13
Theme-based survey report	End of Week 8
Research project	End of Week 15

#### 12. Student Evaluation:

Assessment	Weight	Due Date (tentative)
Homework	10%	See Section 11

Class presentation	10%	Week 14
Survey report	15%	See Section 11
Research project	20%	Week 9
Midterm exam	20%	Week 8
Final Exam	25%	TBD

### 13. Contribution of Course to Program Outcomes

This course contributes to the accomplishment of the following program outcomes:

Program Outcome	Extent of Contribution
1. Perform research emphasizing creativity, independent learning and scientific methods in a chosen area of computer engineering.	●
2. Apply advanced mathematics and engineering knowledge in identifying, formulating and solving engineering problems.	●
3. Select and use techniques, skills and modern tools necessary for research or professional practice.	●
4. Communicate effectively.	○
5. Recognize the need for, and engage in, lifelong learning.	
6. Attend to professional and ethical responsibilities.	○

Extent of contribution: ● high; ● medium; ○ low