1. **Course Number and Course Title:**
   CMP 49409 Machine Learning and Data Mining

2. **Credits Hours:**
   3 – 0 – 3

3. **Prerequisites and/or Co-Requisites:**
   STA 201, or STA 202, or QBA 111, or NGN 111, and MTH 221, and CMP 120.

4. **Name and Contact Information of Instructor:**
   Dr. Salam Dhou
   Office: M-329
   Email: sdhou@aus.edu
   Phone: (06) 515-2943
   Office Hours: as posted on *iLearn*; also by appointment

5. **Course Description (Catalog Description):**
   Introduces the principles of data mining and machine learning tools and techniques for the automatic discovery of knowledge from datasets. The course covers key topics in data mining and machine learning including data cleaning and preparation, visualization, modeling, clustering, classification, and association rules mining. The course studies examples of practical applications using state-of-the-art software tools.

6. **Textbook and other Supplemental Material:**
   Textbook:
   - Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Addison Wesley, 2nd edition, 2013.

7. **Learning Outcomes:**
   Upon completion of the course, students will be able to:
   1. Explain the importance of data mining to IT and society.
   2. Define data mining process and steps including data preparation, cleansing, and task identification.
   3. Design tasks and select accordingly data, knowledge representations, and methods.
   4. Apply data mining techniques and tools to solve real world problems.
   5. Use parametric and non-parametric classification techniques.
   6. Apply supervised and Un-supervised clustering.
   7. Use state-of-the-art software to explore and apply data mining theories.

8. **Teaching and Learning Methodologies:**
   Methods include lectures; problem and project based learning methods (assignments, projects, and exams) and class discussions.

9. **Course Topics and Schedule:**

Topic	Weeks
Introduction: data mining and machine learning	2
Data: types, quality, preprocessing, , similarity measures, and visualization	2

Classification: Basic Concepts, Decision Trees, and Model Evaluation	2
Classification: Nearest-Neighbor classifiers, Artificial Neural Networks.	2
Association Analysis: Basic Concepts and Algorithms	2
Association Analysis: Advanced Concepts	2
Cluster Analysis: Supervised and Unsupervised Clustering	2
Review and evaluation, class presentations	1
Total:	**15**

10. Schedule of Laboratory and other Non-Lecture Sessions:

None

11. Student Evaluation:

Assessment	**Weight**	**Due Date (tentative)**
Quizzes	10%	TBA
Homeworks, projects	20%	TBA
Midterm I exam	20%	Week 7
Midterm II exam	20%	Week 11
Final Exam	30%	Week 16

12. Contribution of Course to Program Outcomes

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

Program outcome	**Emphasis in this course**
a) An ability to apply knowledge of computing and mathematics appropriate to the discipline	●
b) An ability to analyze a problem, and identify and define requirements	◑
c) An ability to design, implement, and evaluate a computer-based system	◑
d) An ability to function effectively on teams to accomplish a common goal	○
e) An understanding of professional, ethical, legal, security and social issues	
f) An ability to communicate effectively with a range of audiences	
g) An ability to analyze the local and global impact of computing	
h) Recognition of the need for and an ability to engage in continuing professional level.	○
i) An ability to use current techniques, skills, and tools necessary for computing practice.	◑
(j) An ability to apply mathematical foundations, algorithmic principles, and computer science	◑
(k) An ability to apply design and development principles in the construction of software	

Emphasis: ● high; ◑ medium; ○ low