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

   CMP 466 – Machine Learning and Data Mining

2. **Credit Hours:**

   3 – 0 – 3

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

   Prerequisites: CMP 305 and MTH 221 and (NGN 111 or NGN 211 or QBA 201 or STA 201 or STA 202)

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

   Name: Dr. Salam Dhou
   Email: sdhou@aus.edu
   Office: ESB 2179
   Phone: 06 515 2943
   Office Hours: Posted on office door and iLearn; also by appointment

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

   Introduces principles of machine learning and data mining methods for the discovery of knowledge from datasets. Covers key topics in machine learning and data mining including data preparation, dimensionality reduction, visualization, supervised and unsupervised learning, and association mining. Focuses on practical applications using state-of-the-art software tools.

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

   Textbook:

   - P.-N. Tan, M. Steinbach, A. Karpatne, and V. Kumar, *Introduction to Data Mining*, 2nd ed., Pearson, 2018.

   Other supplemental material:

   - A.C. Müller, S. Guido, *Introduction to Machine Learning with Python: A Guide for Data Scientists*, 1st ed., O'Reilly Media, 2016.
   - E. Alpaydin, *Introduction to Machine Learning*, 3rd ed., MIT Press, 2014.

7. **Course Learning Outcomes:**

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

   1. Perform machine learning steps including data preparation, task identification, model selection, and evaluation.
   2. Employ mathematical methods to explain the theoretical aspect of machine learning and data mining techniques.
   3. Select appropriate supervised learning methods including classification and regression for given problems and datasets.
   4. Use unsupervised learning methods such as clustering and association rule mining to discover patterns and relationships in datasets.
   5. Apply feature selection and dimensionality reduction methods.
   6. Use state-of-the-art software to explore and solve practical machine learning problems.

8. Teaching and Learning Methodologies:

Methods include lectures, class discussions, problem-based learning, and group work. Students learning is assessed via quizzes, exams, homework, and team project.

9. Course Topics and Schedule:

Topic/Activity	Weeks
Introduction to Artificial Intelligence, Matching Learning, Data Mining, and Python Primer.	Week 1
Data: Types, Quality, Preprocessing, Visualization and application using Python SciKit-Learn.	Week 2
Supervised learning: Basic Concepts, Parametric and Non-Parametric Methods, K-Nearest Neighbors. **Team Project Proposal Submission Deadline (September 11th)**	Week 3
Supervised learning: Decision Tree and application using Python SciKit-Learn.	Week 4
Supervised learning: Ensemble Methods, and application using Python SciKit-Learn.	Week 5
Supervised learning: Model Overfitting and Model Selection.	Week 6
Supervised learning: Imbalanced Classes and Evaluation Metrics.	Week 7
Supervised learning: Naïve Bayes and application using Python SciKit-Learn.	Week 8
Feature Selection and Dimensionality Reduction with Principal Component Analysis (PCA), and application using Python SciKit-Learn. **Midterm Exam (October 26th)- Common Exam- To be confirmed**	Week 9
Supervised learning: Linear Algebra Review, Simple Linear Regression, Multiple Linear Regression, Non-linear Regression.	Week 10
Supervised learning: Logistic Regression and application using Python SciKit-Learn. **Team Project Submission, Presentations & Q&A (November 7th)**	Week 11
Supervised learning: Support Vector Machines and application using Python SciKit-Learn toolkit.	Week 12
Unsupervised Learning (Clustering): Basic Concepts, k-means Clustering and application using Python SciKit-Learn toolkit.	Week 13
Unsupervised Learning (Clustering): Hierarchical Clustering and Application using Python SciKit-Learn toolkit	Week 14
Unsupervised Learning (Clustering): Density-based Clustering and Evaluation, and application using Python SciKit-Learn toolkit.	Week 15
Final Exam (Common Exam)	Week 16

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

Laboratory: none in this course.
Project: Team project assigned to teams of three students. The project requires students to choose a topic, select data accordingly, apply data preprocessing, select models, and apply different machine learning algorithms. Student are asked to evaluate their machine learning models, analyze the results, and state their implications through several submissions. Teams are expected to submit a final report and give an oral presentation. Students are evaluated individually based on their contribution in the project.

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

Assignment	Due Date (tentative)
Homework Assignment 1 – Introduction and data	Week #2
Homework Assignment 2 – Classification using decision trees	Week #5
Homework Assignment 3 – Classification using other methods	Week #8
Homework Assignment 4 – Clustering	Week #13
Team Project	Week #11

12. Student Evaluation:

Assessment	Weight	Due Date (tentative)
Quizzes	10 %	Pop quizzes
Homework	5 %	cf. section 11
Team Project - Proposal 4% - Project Report 10% - Presentation and Q&A 3%	17 %	cf. section 11
Midterm Exam	33 %	Week #9
Final Exam	35 %	Final Exams week

13. Assessment Instruments:

Assessment	Course Learning Outcomes
Quizzes	O1 – O5
Homework	O1 – O5
Project	O1, O3, O5, O6
Midterm Exam	O1– O3
Final Exam	O1 – O5

14. Contribution of Course to Program Outcomes:

BSCS Program Outcomes	Emphasis in this course	Course Learning Outcomes
(1) Analyze a complex computing problem and apply principles of computing and other relevant disciplines to identify solutions.	◑	O1, O2
(2) Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.	◑	O1– O6
(3) Communicate effectively in a variety of professional contexts.		
(4) Recognize professional responsibilities and make informed judgments in computing practice based on legal and ethical principles.		
(5) Function effectively as a member or leader of a team engaged in activities appropriate to the program's discipline.	◑	O6
(6) Apply computer science theory and software development fundamentals to produce computing-based solutions.	◑	O1– O6

Emphasis: ● High; ◑ Medium; ○ Low; Blank – Nothing Specific Expected

BSCoE Program Outcomes	Emphasis in this course	Course Learning Outcomes
(1) Identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics	◑	O1, O2
(2) Apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors	◑	O1– O6
(3) Communicate effectively with a range of audiences		
(4) Recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts		
(5) Function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives	◑	O6
(6) Develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions	◑	O1– O6
(7) Acquire and apply new knowledge as needed, using appropriate learning strategies		

Emphasis: ● High; ◑ Medium; ○ Low; Blank – Nothing Specific Expected

15. Letter Grade Policy:

Total (T)	Letter Grade
$94 \leq T$	A
$90 \leq T < 94$	A-
$85 \leq T < 90$	B+
$80 \leq T < 85$	B
$75 \leq T < 80$	B-
$70 \leq T < 75$	C+
$65 \leq T < 70$	C
$60 \leq T < 65$	C-
$50 \leq T < 60$	D
$T < 50$	F

Honor Code: AUS academic integrity code will be strictly enforced. Cheating will not be tolerated. If caught cheating, you will get an "XF" in the course. No electronics (phones, smart watches, laptops, etc.) are allowed during exams or quizzes. In addition, unsanctioned use of any Generative Artificial Intelligence model (e.g., ChatGPT or similar) in attempting quizzes, labs, exams, or homework

assignments will be considered a violation of the AUS academic integrity code.

Exam/ Quiz Info

Quizzes can be unannounced.

All exams and quizzes are closed-books, closed-notes, no internet is allowed, unless stated otherwise. A student will normally receive a score of 0 for any missed exam or quiz. At the instructor's discretion, this rule may be relaxed, only:

- in the case of extreme emergencies (e.g. death in family or severe illness) and
- an acceptable and verifiable excuse is provided, and/or
- permission was obtained from the instructor before the date of the exam/quiz

If the student has a valid reason for missing an exam or quiz (as mentioned above), no make ups will be given during the semester. The instructor will add the percentage of the missed exam to next exam, if any, or the final exam. If it is a quiz, the average quiz grade will be assigned.

If a student misses an assignment, exam, or class and has a valid written excuse, it has to be submitted within a maximum of 4 days of the missed activity. After this period, no further adjustments will be made. Sick Leaves have to be stamped by the AUS Health Clinic before submitting them to the instructor.

Homework/ Team Project Info:

Homework and team project submissions must be submitted by the deadline. Late submissions are penalized at the rate of 10% per calendar day, including weekends, that they are late. No submissions will be accepted after 4 days of the deadline. The student may be exempted from these penalties under the same conditions that an exam may be made up.

Grading Policy:

If a student disagrees with the grade of any course work, he/she has to contact the instructor within 4 calendar days of receiving the result. After this period, no revisions will be made.

Attendance/ Leave Policy

Attendance is mandatory. The AUS attendance policy of 15% maximum absences allowed will be enforced. Upon breaching the 15% absence rule the student will be automatically dropped from the course with a grade of W (regardless of whether or not the absence is justified). Attendance will be taken using iLearn Attendance tool any time during the lecture.
There is no difference between excused and unexcused absences. An absence is an absence.

Absences are counted from the first day of the term, not from the first day the student enrolls.
A late arrival will count as one-half of an absence.
Students leaving the class before it has been dismissed will be penalized one-half absence.
Students who show irresponsible behavior or weak attendance (if not dismissed), may not qualify for privileges that might be applied at the end of the semester.

Disruptive Activities All communication devices (e.g. mobile phones, etc.) must be turned off during class. Students causing disruption (e.g. making noise) during class will be asked to leave the classroom, thereby incurring one-half absence. Students causing disruption during an exam/quiz will be asked to leave and will not be given make up exam/quiz.

Generative AI tools: The use of generative AI tools, including ChatGPT and other similar tools, to complete or support the completion of any form of exams, assignments or project in this course is not allowed and would be considered academic misconduct.