



# CAP 4770 Data Mining Course Syllabus

## Fall 2023

**Semester:** Fall 2023

**Course Prefix/Number:** CAP 4770

**Course Title:** Data Mining

**Course Credit Hours:** 3.0

**Course Meeting Times/Places:**

Main (Pensacola) Campus, Bldg/Rm:78/103  
Tuesdays and Thursdays 9:30 AM - 10:45 AM

### Instructor and Contact Information:

Dr. Dallas Snider  
Office: Building 4, Room 326  
E-mail: [dsnider@uwf.edu](mailto:dsnider@uwf.edu)  
Tel: 850-473-7348

**Online Office Hours:** <https://uwf.zoom.us/j/84162379826>

- Mondays and Wednesdays (8:30AM - 9:30AM Central)
- Other days and times by appointment

**Course Web Site:** <https://uwf.instructure.com/courses/47927>

### Course Description:

This course will expose students to data mining concepts, techniques, and software utilized in the overall process of discovering knowledge within data.

### Required Textbook(s):

*Data Mining: Concepts and Techniques, Third Edition* by Jiawei Han, Micheline Kamber and Jian Pei, 2012.  
ISBN: 978-0-12-381479-1.

### Student Learning Outcomes:

Upon completion of this course students will be able to:

- Identify data mining functionalities
- Identify data warehousing functionalities
- Create processes to perform the following tasks on data sets: data cleaning, data integration and transformation, data reduction, discretization, and concept hierarchy generation
- Describe data mining primitives, languages, and system architectures
- Apply analytical characterization techniques
- Perform attribute relevance analysis
- Mine descriptive statistical measures in databases
- Mine association rules in databases
- Apply classification algorithms on databases
- Apply clustering algorithms on databases
- Analyze data with data mining software
- Distinguish outliers in databases



# CAP 4770 Data Mining Course Syllabus

## Fall 2023

### Topics:

- What is data mining?
- What is data warehousing?
- Data mining functionalities - what kinds of patterns can be mined?
- Data preprocessing, cleaning, integration, reduction and transformation, data reduction
- Discretization and concept hierarchy generation
- Data mining primitives, languages, and system architectures
- Attribute analysis
- Association rule mining
- Classification by decision tree induction
- Clustering
- Outlier analysis
- Issues in contemporary data mining applications.

### Technology Requirements:

- Students will need a good Internet connection, with an up-to-date operating system and web browser.
- Students will use R and R Studio by either
  1. Accessing the Hal Marcus College of Science and Engineering R Studio Server at <https://rstudio.hmcse.uwf.edu/>
  2. Installing R and R Studio on your personal computer if you like.

### Accessibility:

RStudio accessibility information can be found by going to the Help menu option and then choosing Accessibility or by going to the following website: <https://support.posit.co/hc/en-us/articles/360044226673-RStudio-Accessibility-Features>

### Grading/Evaluation:

#### Weights:

Grading criteria will be broken down as follows:

- |                |     |
|----------------|-----|
| • Quizzes      | 10% |
| • Midterm Exam | 15% |
| • Final exam   | 25% |
| • Assignments  | 50% |

### Rubrics:

All Homework assignments and the Project have unique grading rubrics associated with them. Each assignment will have differing numbers of points possible. These assignments will be manually graded by the instructor.

### Re-grading Assignments:

It is the student's responsibility to check graded assignments/tests when they are graded. I will gladly re-grade an assignment/test when a question or mistake is brought to my attention. To ensure fairness, I reserve the right to re-grade the entire assignment/test. As a result, your grade may increase, decrease, or remain the same. Grades will not be changed after a week from the date graded assignments/tests is graded.



# CAP 4770 Data Mining Course Syllabus

## Fall 2023

### Communication:

You are responsible for checking your UWF email **DAILY** to keep up with important announcements, assignments, etc. E-mail will be the primary form of communication between students and instructor. **Please do not use the Canvas email feature as the email I receive from the system is not mobile device friendly and it will delay my ability to respond to you.**

### Participation and Feedback:

Active participation and regular feedback is encouraged. Discussion is an important part of learning. Participation in discussions is an integral part of this course and comprises a significant part of your course grade. Effective communication between the instructor and students, and the students and students, will make the course more useful, interesting, and productive. Be sure to contact the instructor as soon as practicable should anything threaten to prevent you from keeping up with the graded assignments in the course.

### Overall Grading Scale:

A :	89.5 - 100
B+:	86.5 - 89.499
B :	79.5 - 86.499
C+:	76.5 - 79.499
C :	69.5 - 76.499
D+:	66.5 - 69.499
D :	59.5 - 66.499
F :	0 - 59.499

### Important Notes:

- Any changes to the syllabus or schedule made during the semester take precedence over this version. Check email regularly for up-to-date information.
- September 17<sup>th</sup> - Withdrawal deadline for ALL courses with partial refund; automatic grade of "W"
- November 13<sup>th</sup> - Withdrawal deadline; automatic grade of "W"

### Additional Syllabus Statements:

The following link provides policy statements that apply to all syllabi on all UWF campuses and are in effect even if not reproduced on the official course syllabus:

<https://confluence.uwf.edu/display/public/Additional+Syllabus+Statements>

The topics include but are not limited to

- Course Modality
- Recording in Class
- Academic Conduct
- TurnItIn
- Sexual Discrimination or Harassment Reporting
- Civil Discourse
- Health and Safety Protocols



# CAP 4770 Data Mining Course Syllabus

## Fall 2023

### Tentative Course Schedule:

Module	Week of	Topic/Activity
1	Aug 22	Chapter 1 – Introduction, What is data mining?
2	Aug 29	Chapter 2 – Data, Measurements, and Data Preprocessing Part 1
3	Sep 05	Chapter 2 – Data, Measurements, and Data Preprocessing Part 2
4	Sep 12	Chapter 3 – Data Warehousing and Online Analytical Processing
5	Sep 19	Chapter 4 – Pattern Mining: Basic Concepts and Methods
6	Sep 26	Chapter 6 – Classification: Basic Concepts and Methods
7	Oct 03	Chapter 7 – Classification: Advanced Methods
8	Oct 10	Midterm Exam
9	Oct 17	Chapter 8 – Cluster Analysis: Basic Concepts and Methods
10	Oct 24	Chapter 10 – Deep Learning
11	Oct 31	Chapter 11 – Outlier detection
12	Nov 07	Chapter 12 – Data mining trends and research frontiers
13	Nov 14	Ethics in Data Mining /Work on Project Part 1
14	Nov 21	Thanksgiving no class Work on Project Part 2
15	Nov 28	Finish Final Assignment, Final Exam Opens
16	Dec 04	<b>Final Exam Due on Monday December 4th at 11:59 PM Central</b>