University of North Carolina at Charlotte
College of Computing and Informatics

Course Number and Title: ITIS/DSBA/HCIP 6162: Data Mining, Fall 2025

Credits: 3 Graduate Credits

Time, Days, and Location:
1 pm-3:45 pm on Thursdays at the Dubois Center 601

Faculty Information: Xi (Sunshine) Niu, Ph.D., Associate Professor
Office: Woodward 310G
Email: xniu2@charlotte.edu
Office hours: 10am-12pm Wednesdays

Teaching Assistant: Mr. Xiangcheng Wu
Email: xwu20@charlotte.edu
Office hours:
1) 2pm-3pm Mondays; 2) 2pm-3pm Wednesdays
Zoom: through Canvas

Course Description:

This course is about data mining. It is an essential part of AI, which is one of the hottest topics in computer science today. Data mining is a fast-evolving field, especially for recent five years.

The availability of large amounts of data has created unprecedented opportunities to leverage computational and statistical approaches to turn data into actionable knowledge. This course covers general techniques for analyzing large amounts of **numeric** and **text** data. The entire data mining process is covered in this course: setting up a problem, data preprocessing, model constructions, model evaluations, and interpretations in decision making.

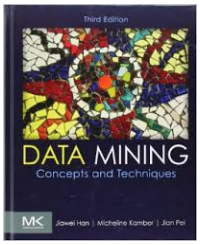
This course covers both classical data mining approach (e.g., Apriori, Random Forest, etc) as well as the recent deep learning models (e.g., RNN, CNN, BERT). In addition, the recent rise of large language models (LLMs), especially ChatGPT, has brought global excitement. We will have LLMs as one of our topics.

Required Textbooks and Papers:

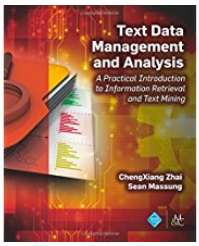
We will use **textbooks**, **online learning materials,** and **academic papers** as our learning materials. For all the papers, Dr. Niu will provide the full-text versions.

For textbooks, Dr. Niu will provide the electronic copies. The two textbooks are listed as below:

Title: Data Mining: Concepts and Techniques
Author(s): Jiawei Han, Micheline Kamber, and Jian Pei
Edition: 3rd Edition
Publisher: Morgan Kaufmann
Year: 2011



Title: Text Data Management and Analytics: A Practical Introduction to Information Retrieval and Text Mining
Author(s): ChengXiang Zhai and Sean Massung
Publisher: ACM and Morgan & Claypool Publishers
Year: 2016



Evaluation Methods:

Course grading will be based on these activities.

Activities	**Point**
In-Class Quizzes	2 points x 12 = 24 points
After-Class Homework	4 points x 13 = 52 points
Final Team Project	12 points
Class Attendance	12 points
Total	**100 points**

Grade Scale:
A = 90 points – 100points
B = 80 points – 89 points
C = 70 points – 79 points
U = Below 70 points

Weekly Lesson Schedule:

Date	Contents
Aug 21	Syllabus Lesson 1: Getting to Know Your Data
Aug 28	Lesson 2: Pattern Mining
Sept 4	Lesson 3: Machine Learning 1
Sept 11	Lesson 4: Machine Learning 2
Sept 18	Lesson 5: Cluster Analysis
Sept 25	Lesson 6: Word Association Mining
Oct 2	Lesson 7: Topic Modeling
Oct 9	Student Fall Recess – No Classes
Oct 16	Lesson 8: Introduction to Deep Learning
Oct 23	Lesson 9: RNN and CNN
Oct 30	Lesson 10: Word Embedding
Nov 6	Lesson 11: Transformers and BERT
Nov 13	Lesson 12: Large Language Models
Nov 20	Lesson 13: Deep Reinforcement Learning
Nov 27	Thanksgiving Break – No Classes

Course Policies:

Course Credit Workload:
This 3-credit course requires 9-12 hours effort (including the class time) for this course each week for 13 weeks. Efforts may include but is not limited to: required reading, homework assignments, and studying for quizzes.

Class Attendance Policy:
Attending every class is mandatory. Class attendance entails being prepared, present, and attentive for the entire class period. **Class attendance also means NOT being late and NOT leaving early.**

One or two absences (either excused or unexcused) is OK without losing the class attendance points, but you will lose the in-class quiz points for those absence(s). Three or four absences will lose half of your class attendance points (1/2 x 12 = 6 points). Five or more absences in total will lose all your class attendance points (12 points). For each absence, the student is responsible for catching up with all covered materials and assignments.

Late Submissions:
For assignments, unexcused late submission (according to the Canvas timestamp and the "late" flag) will receive a grade of 0. You should plan sufficiently for completing and submitting assignments. Should an emergency arise that greatly disrupts one's ability to complete an assignment, please send an email to Dr. Niu **before** the due

date with a plan for submission after the due date. You need to receive Dr. Niu's permission for late submission.

Policies on the Use of Generative Artificial Intelligence (AI):

In this course, students are permitted to use generative artificial intelligence (AI) tools like ChatGPT and Gemini to support their work. To maintain academic integrity, students must disclose any use of generative AI by including the following statement in their assignments: "I acknowledge the use of [generative AI tool Name] in this assignment. The [generative AI tool Name] was used in the following way(s) in this assignment: [e.g., writing the code in which portion of the assignment, correcting the code in which portion of the assignment, etc]."

Be aware that students are responsible for any errors or information that is misrepresented or inaccurate (i.e., hallucinations) that generative AI tools produce when submitting work that includes AI-generated material.

Important Note on Data Protection and Privacy: When using generative AI tools, it is important to be aware that the data you supply might be used for training AI models or other purposes. Consequently, there is no guarantee that the information you provide will remain confidential. You should exercise caution and avoid sharing any sensitive or private information when using these tools. Examples of such information include personally identifiable information, protected health information (PHI), financial data, intellectual property, original research, and any other data that might otherwise be legally protected.

Special Needs and Religious Accommodation:

If you have a documented disability and require accommodation in this course, contact the Office of Disability Services (https://ds.uncc.edu/students/academic) the first week of the semester. Accommodations for learning will be arranged by that office and communicated to Dr. Niu.

It is the obligation of students to provide faculty with reasonable notice of the dates of religious observances on which they will be absent by submitting a Request for Religious Accommodation Form to their instructor prior to the census date for enrollment for a given semester. The census date for each semester (typically the tenth day of instruction) can be found in UNC Charlotte's Academic Calendar (https://registrar.uncc.edu/printable-calendar).

Copyright and Permissions:

My lectures and course materials, including presentations, quizzes, code demonstrations, homework problems and answers are protected by copyright. I am the exclusive owner of copyright in those materials I create. You may make copies of course materials for your own educational use. However, you may not, nor may you knowingly allow others to reproduce or distribute these materials publicly without my written consent. This includes providing materials to commercial websites and other similar services. Students who publicly distribute or display or help others

publicly distribute or display copies or modified copies of my course materials are in violation of University Policy 406, The Code of Student Responsibility.

University Policies:

Code of Student Responsibility: https://legal.uncc.edu/policies/up-406
Code of Student Academic Integrity: https://legal.uncc.edu/policies/up-407
Diversity and Inclusion: https://diversity.uncc.edu/
Sexual Misconduct and Interpersonal Violence:
https://legal.uncc.edu/policies/up-502
Standard for Responsible Use: https://oneit.uncc.edu/iso/standard-responsible-use