Computer Science Department

Course Syllabus

Course Information

- CS 422/522 Data Mining – Fall 2025
- M 6:15-9:00pm in HT319

Faculty Contact Information

- Dr. Jiang Li
- HT232; tentative office hours: Mon./Wed. 10:00-11:00am, or by appointment or drop-in. Either meet in person or through the zoom: https://zoom.us/j/97748140648?pwd=zy3fAevEHvUgUmbaGKcIBHa9qOqojU.1
- lij@hood.edu (preferred); 301-696-3564

1. Course Description

Introduces basic principles and methods for data analysis and knowledge discovery to computer science students. Topics include preprocessing, association, classification, and anomaly detection. Students develop basic skills for modeling and performance evaluation.

Prerequisite: MATH 213 or MATH 112 or PSY 211 or ECMG 212, Level III Mathematics Placement, and (CS 200 or CS 219); or Permission of Instructor

2. Course Objectives

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

- Explore and understand complex datasets, including techniques to preprocess and clean data, ensuring it is suitable for analysis.
- Learn various data mining techniques such as clustering, classification, association rule mining, and anomaly detection, and understand when and how to apply them to real-world problems.
- Evaluate model performance and explain model output, providing valuable insights for decision-making and problem-solving.
- Learn to visualize data mining results effectively, using charts, graphs, and other visualization methods to communicate findings to stakeholders clearly.
- Understand the ethical implications of data mining, including privacy concerns,

data security, model bias, and the responsible use of data, ensuring they approach data mining projects with integrity.
- Use available data analysis methods and tools.
- Effectively use professional literature.
- Effectively communicate with technical and non-technical audiences.

The following student outcomes are assessed in this course:
- 1. Analyze a complex computing problem and apply principles of computing and other relevant disciplines to identify solutions.
- 2. Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of computer science.
- 3. Communicate effectively in a variety of professional contexts.
- 6. Apply computer science theory and software development fundamentals to produce computing-based solutions.

3. Required Text And Tools

- *Introduction to Data Mining*, 2nd Edition, by Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, and Vipin Kumar.

4. Course Outline

Week	Topics
1	Orientation and Introduction (Chapter 1, Colab introduction)
2	Data (Chapter 2, Python review)
3	Labor Day, no class
4	Data Exploration (Introduction to numpy and pandas)
5	Classification: Basic Concepts and Decision Tree (Chapter 3)
6	Classification I (Chapter 4)
7	Classification II (Chapter 4)
8	Association Rule Mining (Chapter 5)
9	Fall Break
10	TBD
11	Clustering I (Chapter 7)
12	Clustering II (Chapter 8)
13	Anomaly Detection (Chapter 9)
14	Validity and reproducibility of results (Chapter 10)
15	Special Topics in Data Mining
16	Final Project presentation

5. Course Evaluation

Grades will consist of the following components:

Component	Percentage
Class participation	10%
Exercise/Homework	55%
Midterm project	15%
Final Project	20%

Point distribution guideline will be as follows:

Point Range	Letter Grade
95.00-100	A
90.00-94.99	A-
85.00-89.99	B+
80.00-84.99	B
75.00-79.99	B-
70.00-74.99	C+
65.00-69.99	C
60.00-64.99	C-
55.00-59.99	D
0-54.99	F

Note: I reserve the right to alter component weighting or provide a "curve" on an assignment as warranted.

All assignments will be graded and returned in a timely manner. When an assignment is returned, you will have a period of one week to contest any portion of the grade. The instructor who graded your assignment will be the person to resolve a grading conflict. The judgment of the instructor will be final in all such cases. When contesting a grade, you must be able to demonstrate how your particular solution is correct. Also, when contesting a grade, the instructor reserves the right to re-evaluate the entire project or exam, not just the portion in dispute.

6. Homework Guidelines

Software and Platforms
We will utilize several open-source data analysis packages and web-based tools in this course. Basic concepts in statistics, linear algebra, and machine learning are discussed along with course contents. For example, we do not perform eigenvalue decomposition, support vector machine (SVM) or Fourier transform by hand. Instead, we discuss where, when, and how to effectively apply these methods to extract interesting results.

In-session Exercises
For each major topic, there will be a couple of in-class activities planned. With these, you put what is covered in the lecture into action directly.

Homework

- There will be several out-of-class homework assignments. These are comprehensive, out-of-class activities with intermediate complexity that you complete on your own.
- Homework will be assigned at the end of lectures. As a general rule, these assignments will be due on Mondays. Please submit a paper copy and/or an electronic copy through the assignment link on Blackboard.
- You will need to read, discuss, write a one-page summary for, and present an important technical paper in data mining from the reading list provided or of your own choice with the instructor's approval. Details will be given when it is assigned. (Graduate students only.)
-

Midterm-project

- A comprehensive project covers the topics covered during the first half of the class.
- Details will be posted.
-

Final Project

Each student will work on a term project to be presented at the last class meeting.

- The term paper should be four to six pages double column, following IEEE/Computer Society transaction format, in PDF format and its original editable format (LaTex, Word, Writer, etc.)
- All source code, entire input data set, computation output, and documentation in a public GitHub account (when applicable).

7. Policies, Guidelines, Academic Honesty & Tips

- Adhering to the Academic Honesty Policy/Honor Code is student responsibility. Deviation from the policy will not be tolerated.
- Discussions with classmates are permitted but deliverables must be your own, individual work. This means you are free (and encouraged) to discuss assignments with other students outside of class; just don't share answers or code.
- **Conditional Generate AI Usage** In this course, the use of Generative AI is permissible only under specific circumstances, as outlined in the assignment instructions. When AI tool usage is allowed, exercise caution and ensure that you appropriately cite and attribute any content generated through their use. It is essential to strike a balance between leveraging the advantages of AI resources and preserving your individual creativity and problem-solving abilities. Failure to adhere to the specified Generative AI usage guidelines may result in academic consequences. Pay careful attention to the assignment requirements to determine whether AI tools are permitted and follow the instructions accordingly.

- Late assignment, up to 5 days, will be penalized by 20% each day; assignment more than 5 days late will receive a score of zero. No assignments will be

accepted after the last day of class. Assignments turned in after they have been collected in class are considered late. No exceptions unless prior-arrangement.

- You are responsible for the content of reading assignments, lectures, handouts, announcements and schedule changes made in class whether you are present. If you must miss a class, be sure to check with me to cover the missed topics.
- Attendance is expected at each class and lab meeting. It affects the class participation grading and lab grading, it is in your own best interest to attend class and labs, as your grade will almost certainly suffer indirectly if you choose not to attend.
- Finally, material in the course is, inherently, cumulative. Be aware, if you fall behind, it may be difficult to catch up. If you fall behind, ask for assistance as quickly as possible { I am here to help you).

Hood Policy Statements:

Attendance Policy

Regular class attendance is an essential part of the educational program of the College, and it is expected that no student will absent herself/himself without adequate reason. The College recognizes that there are other justifiable reasons for class absence: observance of religious holidays, participation as a representative of the College in athletic contests or cultural performances. Such absences cannot be justifiable if the student has previous excessive absences and has not made arrangements to make up the work missed prior to the absence.

Classes are held up to the date and hour preceding vacation and they resume promptly after the recess in accordance with the academic calendar.

The Academic Honor System (Student Handbook)

As a place of honor and respect, all members of Hood College assume the obligation to maintain the principles of honesty, responsibility, and intellectual integrity in all activities related to their Hood College experience. Students are expected to adhere to the highest standards of academic honesty and integrity in all coursework and related matters. It is the responsibility of each student to support these values through maturity of thought, expression, and action. Members of the faculty and staff are available to assist students in this process.

Communication Standards (Faculty Handbook)

The student's ability to write and speak standard English is an integral part of the grade evaluation for the course. Therefore, the quality of written and spoken work will be considered as a student's final grade is being calculated.

Accommodations for Students with Disabilities (Faculty Handbook)

In accordance with Section 504 of the Rehabilitation Act of 1973, and the Americans with Disabilities Act of 1990, Hood makes every reasonable effort to accommodate the needs of handicapped students. The Coordinator for Disability Services arranges for academic accommodations, such as note-takers or interpreters; the Office of the Dean of Students coordinates residential room accommodations for these students.

Hood College acknowledges and supports the rights of students with disabilities to have access to higher education and upholds Section 504 of the Rehabilitation Act of 1973, and the Americans with Disabilities Act of 1990. The College works diligently to eliminate barriers to the educational pursuits of our students wherever possible. Because each disability is unique, services are individually tailored to the needs of each student. If you are a student with a disability in need of classroom accommodations and have not already registered with Kate Gmuer, Coordinator of Disability Support Services, please contact her at 301-696-3421 or gmuer@hood.edu to discuss policies and procedures related to disability services and to establish the accommodations for which you are eligible.

Title IX Compliance and Reporting (Student Handbook)

Sexual misconduct is one of the most serious violations of Hood College's values of honor and respect and no form of sexual misconduct is acceptable within our community. Consistent with state and federal laws, including Title IX of the Education Amendments of 1972 as amended ("Title IX") and Title VII of the Civil Rights Act of 1964, the College has implemented measures to ensure that all allegations of sexual misconduct are investigated and resolved in a timely, confidential, fair, and impartial manner. The College has designated certain officials as "responsible officials". A report to these individuals is an official report to the College which necessitates a response by the College. Faculty members are considered responsible officials and, as such, are not considered confidential resources. Therefore, incidents shared with the Faculty must be reported to the Title IX Coordinator. Please be advised that if you wish to speak confidentially about an incident and/or to obtain an academic accommodation, you may contact the Wellness Center to speak with a counselor, Health Services Staff and/or the Dean of the Chapel.