UNIVERSITY OF
SOUTHERN MAINE

COS 472 & 572: Artificial Intelligence & Data Mining

Department of Computer Science

Fall 2024

Instructor Info

Dr. Yuqi Song

Phone: (207) 780-4794

Email: yuqi.song@maine.edu

Office: 286 Science Building, Portland

Student Hours:

Tuesday 10:45 am -11:45 am

Or by Appointment

Course Meetings

Payson Smith 204, Portland

Tu/Th 9:30-10:45 AM

Student Services and Policies Hub[1].

Contents

---

[1] https://mycampus.maine.edu/group/usm/student-services-and-policies-hub

1. Course Information

1A. Course Description

This course introduces the concepts, techniques, and applications of data mining and artificial intelligence (AI), and it mainly focuses on some methods and models that are useful in analyzing and mining real-world data. It will cover frequent regression classification, clustering, and classic deep learning framework. This course is designed to provide a multifaceted learning experience through a diverse of instructional formats: lectures, case studies, seminars and group projects. The course includes both theoretical foundations and practical exercises, preparing students for advanced studies or careers in AI and data mining.

1B. Course Materials & Books

Recommend (all optional thus not required)

- Data Mining: Concepts and Techniques (3rd Edition, 2012), Jiawei Han, Micheline Kamber, Jian.
- Introduction to Data Mining (2nd Edition, 2019), Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar.
- Python for Data Analysis (3rd Edition),(Open Edition), Wes McKinney, published by O'Reilly Media.
- Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems (3rd Edition, 2022) Aurélien Géron.
- Dive into Deep Learning, Aston Zhang, Alexander J. Smola, Zachary Lipton, Mu Li.

Supplemental

- Artificial Intelligence: A Modern Approach (4th edition), Stuart Russell and Peter Norvig.
- Artificial Intelligence: A Textbook, 2021, Springer.
- Deep Learning. Ian Goodfellow and Yoshua Bengio and Aaron Courville, 2016.
- Intro to Python for Computer Science and Data Science: Learning to Program with AI, Big Data and the Cloud, Paul J. Deitel & Harvey Deitel.
- Python for Data Analysis, Scott McCoy.

1C. Course Learning Outcomes

The main learning objectives of the course are to:

- Identify problems where artificial intelligence techniques are applicable.
- Apply selected basic AI techniques; judge applicability of more advanced techniques.
- Participate in the design of systems that act intelligently and learn from experience.

2. Coursework & Grading

2A. Course Assessment

Assessment Name	Value
Class attendance	10%
Quiz	10%
4 Assignments	20%
Individual competition	15%
Midterm	10%
Group project proposal	10%
Group project presentation, report, and source code	25%
Total:	**100%**

2B. Attendance

Regular attendance and fully engaged participation is expected of all students. You will be required to sign your name by the end of every class.

2C. Quizzes

There will be seven quizzes given throughout the semester. The quiz is scheduled in class and you will have a 20-minute window to complete it. A student's five best scores will be used to compute the quiz average. Since two quizzes will be dropped for all students, makeup quizzes will not be given. Each quiz takes 2%.

2D. Assignment

There are 4 assignments in total, they should be completed individually, and due on the day by 11:59 PM. For example, the first homework due is Sep 19, which means you should submit

your work before Sep 19 11:59PM. All assignments will be submitted and graded through Brightspace. Each assignment should be submitted with the assignment report (Microsoft Word or PDF file) and Python source code.

2E. Individual competition

It is an individual-based data mining competition on Kaggle. You have a long time to complete this competition and submit your results.

2F. Midterm

The midterm will be taken online through Brightspace and must be completed within the specified time.

2F. Group project

Hands-on project experience is crucial for a comprehensive understanding of data mining and AI techniques. Consequently, the project component will constitute 30% of the overall grade.

Project Timeline and Requirements:
1) Early semester: team member and topic selection: 1 to 3 members per team, more members, higher expectations. You will form project teams and select a topic that interests you. The instructor will provide three topics, you also can choose the topic of your own interest.

2) After midterm: each team will make their project proposal presentation during class. This presentation should introduce the specific problem being addressed, the methods chosen for the project (at least two different algorithms), and team members' division of labor.

3) In the last week, there will be a final project presentation, each group should make a 20-30 minute presentation and followed by the Q and A. This presentation should include a detailed explanation of the methods used; experimental results that highlight the effectiveness of your chosen methods; and demonstration of the project's code, etc. You also need to submit a final project report (4 to 8 pages research-paper-like report) and all source code.

3. Class Schedule

3A. Course Schedule

Week	Date	Topics	Note	Due
1	Sep 3	Introduction of Data mining and AI		
	Sep 5	Prepare for Data Mining: Python, Jupyter, Pandas, Scikit-learn		
2	Sep 10	Getting to Know Your Data		
	Sep 12	Data Preprocessing		
3	Sep 17	Data Visualization	Q1	
	Sep 19	Case study: data analysis		HW1
4	Sep 24	Logistic Regression		
	Sep 26	Classification, Decision Tree	Q2	
5	Oct 1	Naive Bayes		
	Oct 3	Seminar: how to do research		
6	Oct 8	Support Vector Machine		HW2
	Oct 10	Case study: classification task	Q3	
7	Oct 15	Fall break, No class		
	Oct 17	Midterm		
8	Oct 22	Ensemble Learning: Bagging, Boosting	Proposal	
	Oct 24	Cluster Analysis, K-Means Clustering	Proposal	
9	Oct 29	Density-based Clustering		
	Oct 31	Principle Component Analysis	Q4	
10	Nov 5	Case study: clustering task		
	Nov 7	Recommender System: Collaborative Filtering		
11	Nov 12	Seminar: technical writing		
	Nov 14	Recommender System: Latent Factor Models	Q5	
12	Nov 19	Material informatics		
	Nov 21	Deep learning introduction		HW3
13	Nov 26	DNN components I	Q6	
	Nov 28	Thanksgiving day, No class		
14	Dec 3	DNN components II		

	Dec 5	Case study: deep learning task	Q7	
15	Dec 10	Group project final presentation		
	Dec 12	Group project final presentations		HW4

3B. Schedule of seminars

- The course will feature a series of seminars. The exact scheduling details for these seminars will be posted here and announced well in advance to ensure ample notice.

4. Course-Specific Policies

4A. Handing in Assignments

All assignments and reports will be submitted and graded through Brightspace. Each assignment should include the assignment report (Microsoft Word or PDF file) and Python source code. You should complete all assignments individually, except for the Project.

4B. Late Assignments

Late assignments will be marked down 10% per day that they are late, and assignments submitted after seven days will not be accepted (except under special circumstances such as illness or other unanticipated impediments).

4C. Group project

For the group project, I encouraged you to work in teams of one to three. You can work together face to face or remotely using Zoom and screen sharing. Your roles and division of labor should be indicated in the report. Although it is a team project, I hope that you can actively participate in it and achieve real gains. Many students feel they learn more effectively by working with a partner. Teams should turn in only one report with both student names on it.

4D. Plagiarism

Plagiarism is turning in work that is not your own. Searching the internet for answers or using answers created by others is plagiarism and may result in failing the course as well as appropriate disciplinary action. It is your responsibility to not leave your work where others might copy it.

4E. Getting help

I want everyone to succeed. Do not put off getting help when you need it.

- Use the discussion board in Brightspace.
- Join student hours (or by appointment).
- Use the tutoring available through the Learning Commons which is available both on-campus and by Zoom. Learn more and schedule an appointment at: usm.maine.edu/learningcommons.

5. Academic Services & Policies

Below you'll find a brief list highlighting some of the most crucial student services and supports.

- **Request disability accommodations** | (207) 780-4706 | dsc-usm@maine.edu
- **Report Interpersonal violence** | (207) 780-5767 | usm.titleix@maine.edu
- **Report on-campus emergencies and safety concerns** | (207) 780-5211 or your local police agency.
- **Get academic help** | mycampus.maine.edu/group/usm/learning-commons
- **Get technology help** | usm.maine.edu/computing/helpdesk
- **Meet with an academic advisor |** usm.maine.edu/advising

For USM's most complete and current information on services available to students, as well as academic policies, use the QR Code to go to the [Student Services and Policies Hub webpage](2).