

CS 480: Data Mining/Machine Learning
Section 1
James Madison University
Fall 2019 (3 credits)

1 Basic Course Information

1.1 Course Description and Goals

Can computers learn? And if so, how can a computer learn? In this course, you will be introduced to machine learning, a technique which applies algorithms that enable systems to "learn by example". Google search utilizes it to complete the search criterion as you are typing, and it enables Netflix to make good movie recommendations. Machine learning is prevalent in many fields: autonomous driving, detecting credit card fraud and cyber attacks, and organizing/searching through your every growing set of photos on your phone. Using machine learning foregoes writing very complex functional programs and instead utilizes one of a set of core learning algorithms. For example, imagine writing a traditional program to decode a zip code from a picture of an envelope. Machine learning utilizes examples of hand written zip codes (digits) and "learns" from these examples to recognize/decode zip codes from millions of pieces of mail each day. This course focuses on a balance of theoretical and practical knowledge. Popular methods such as neural networks, deep learning, and support vector machines will be covered. Small projects will allow students to apply these techniques and showcase their results in a quantitative manner. Data science, which includes the field of machine learning, is the fastest growing subfield in computer science (with the highest predicted job growth over the next 10 years), so, I hope you will join me for this exciting course.

1.2 Meeting Times and Locations

Section	Days	Time	Location
01	T/R	11:00-12:15	ISAT/CS Building 246

1.3 Instructor

Name Dr. Kevin Molloy
Office ISAT/CS 216
Email molloykp@jmu.edu
Office Hours T 9:00 - 10:45
W 13:00 - 16:30

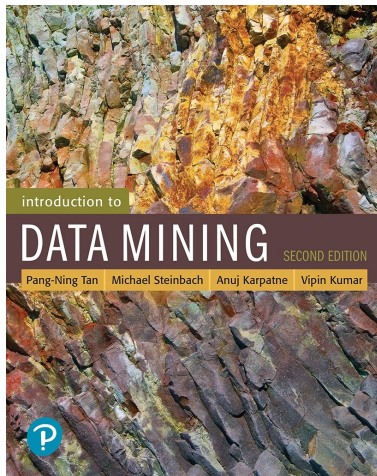
1.4 Website: <https://w3.cs.jmu.edu/molloykp/teaching/cs480ML>

Much of the information for this course will be disseminated via this website. You should check this website often (at least once a week) for announcements and updates.

1.5 Prerequisites

A grade of “C-” or better in **CS 327 (discrete structure II)**.

1.6 Required Texts



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

1.7 Computing Resources

You will require access to a computer for this class. A machine running Windows, Mac OS, or Ubuntu/Linux will work fine. My opinion is that Mac OS and Ubuntu/Linux machines are easier to use. Information is available on the website on how to utilize a virtual machine that will supply all the software you will need for this class (and provide a common environment to what you will experience in the labs).

1.8 Expectations/Keys to Success

Homework

In a three-hour course, you should expect **six hours** of homework per week. How you manage your schedule is up to you, but do spend some time each day on this course.

Preparing for Class

The material (reading and/or videos) for each class is detailed on the class calendar (available on the website). You will have a **quiz each week or every other week** based on this material, so, a key to success is reading this material to reinforce the material discussed in class.

Programming Assignments

Programming assignments (PA) will utilize the python programming language. Each assignment can take about between 4 and 15 hours (final project) to complete. Don't wait until the week before the assignment is due to get started.

Seeking Help

Piazza is a great platform to see help in this class. However, I generally do **not** answer Piazza questions over the weekend. If you choose to complete assignments at the last minute or after the deadline, please keep this in mind. I will make sure any questions posted over the weekend are answered on Monday. Please ask questions using Piazza first if at all possible. I have it set up so that I get an email when a question is posted to Piazza, so emailing me is not quicker and by posting to Piazza you will have a chance of being answered by a classmate, TA, or another faculty member. Email should be reserved for questions whose answers would only benefit you personally or only I would know the answer to.

1.9 Communication

We will use a number of different tools for communication in this course. These include:

- https://w3.cs.jmu.edu/molloykp/teaching/cs480ML/cs480ML_2019Fall/ is our central course web site. The announcements, discussion board, videos, and documents posted there are part of the required reading for the course.
 - Use public posts on Piazza to discuss the material related to this course.
 - **Canvas** will be used to submit assignments and disseminate grades

- **Mail the professor** if you have logistic or personal issues to discuss such as setting up an appointment outside of office hours, if a health problem arises, or if you have a personal emergency.
- **Office Hours** No appointments are required to attend office hours or you can make an appointment with me.

1.10 Attendance and Participation

Regular attendance and fully engaged participation is expected. Your grade will be partially based on in-class activities, so, attendance will affect your grade.

Please silence your cell phone while class is in session. If you have a laptop or tablet, you are encouraged to bring it to class and use it to work along with programming examples and exercises. Mute the volume to avoid unintended interruptions, and do not use any electronic devices for activities that may distract other students. Repeated violations of this policy may result in disciplinary action or a grade penalty in the course.

I strongly encourage you to check the main website and the Piazza web forum regularly for important announcements (usually regarding programming projects). You may also use the Piazza forum to ask general questions of interest to the class as a whole (e.g., administrative issues or project clarification questions) as well as to offer each other general advice on class assignments. However, do not post any information that would violate the university academic integrity policy. If you are unsure about this, please email me for approval before you post.

2 Methods of Evaluation and Grading Policies

You are responsible for all material discussed in lecture and discussion section and posted on the class web page, including announcements, deadlines, policies, etc.

Your final course grade will be determined according to the following percentages:

Component	Count	Weight
Labs/In-class work	6	10%
Homework/Programming	3	20%
Exams	4	50%
Final Programming Assignment	1	15%
Round Table Discussion	1	5%

Letter grades will be assigned on the scale A=90-100, B=80-89, C=70-79, D=60-69, F=0-59, with potential minor adjustments after considering the overall performance of the class and actual distribution of numeric scores. I will use + and - grades at my discretion.

2.1 Exams

We will have a total of four exams . Two exams will be take home and two exams will be in class. You must be present or give a valid doctors excuse or similar note confirming your absence if you miss any exam. For the in class exams, If you must be absent during an exam for a legitimate reason, you must contact me at least one week beforehand to make special arrangements. Failure to make prior arrangements for a missed exam will result in a zero grade. Except in extraordinary situations, you will not be excused from exams. Your instructor will be the sole arbiter of whether a situation qualifies as extraordinary. Hence, you should behave as if you will not be excused from exams for any reason.

2.2 Grading Disputes

If you believe I have made an error while grading your work or calculating your final score, please bring it to my attention after class or during office hours. If I determine that there has been a simple mistake, I will fix it immediately and no formal request is necessary.

If you believe an exam question or assignment has been graded unfairly, you must submit a written formal request for a regrade via email. Such requests must be submitted within one week of when the assignment in question is returned to you. **Any coursework submitted for reconsideration may be regraded in its entirety, which could result in a lower score if warranted.**

3 Course Policies

Important announcements will be made in class and/or on the class website. Please make it a habit to check the web page and or your email daily during the week.

Although every effort has been made to be complete and accurate, unforeseen circumstances arising during the semester could require the adjustment of any material given here. Consequently, given due notice to students, I reserve the right to change any information on this syllabus or in other course materials.

You are permitted to use course materials for your own personal use only. Course materials may not be distributed publicly or provided to others (excepting other students in the course), in any way or format unless explicitly allowed.

3.1 Programming Assignments(PA's)

PA's must be submitted electronically following the instructions given in class and on the website. Assignments may not be submitted by any other means (e.g., do not email your projects to me unless I request that). It is your responsibility to test your program and verify that it works properly before submitting it.

All assignments are due at 23:00 (11:00pm) on the day indicated on the assignment unless noted otherwise.

3.2 Final Programming Project

The final programming project will be from one of 4 or so activities, but will be an individual project. Part of the project will involve a peer review during the development process and peer grading/feedback for a scientific presentation.

If you are involved in independent research with others at JMU, please advise me, as it is possible that this research could be used for this project.

3.3 Missed and Late Assignment Policy

It will not be possible to receive credit for in-class work that is missed due to absence except in the case of prearranged absences for sports, academic conferences, or other CS related activities. Missing class for these reasons requires coordination with me prior to missing class (no consideration will be given for coordinating this after the missed class). In recognition of the fact that absences are occasionally unavoidable, I will drop the lowest score for in-class activities.

Assignments may be submitted up to 48 hours late for a 15% penalty per 24-hour period. For example, a submission that would have earned 90 points in an on-time submission will earn $90 \times 0.85 = 76.5$ points if submitted up to 24 hours late, or $90 \times 0.60 = 54$ points if submitted up to 48 hours late. If you make multiple submissions, I will typically grade the latest submission. If you wish me to grade a different submission, you must indicate this before the 48-hour late period is over.

Regardless of the above policy, I reserve the right to refuse to grade any programs submitted after the beginning of the second class period following the project deadline, because I may discuss the solution in class.

Project extensions will not necessarily be granted due to server congestion, system problems, network problems, power outages, etc., so do not wait to submit a program until the night it is due. No consideration in grading will be made for errors made in transferring files or submitting the wrong version of your project. Having a working, non-submitted version will not count; only submitted code will be counted.

You will be responsible for developing your own techniques for testing your projects before submitting it. I will grade your assignment based on test cases not provided to you in advance. Your code will be graded on a combination of correctness, completeness, documentation, and code style.

I will be exploiting electronic methods to detect plagiarism (this is a machine learning class, so, be aware that machine learning methods will be used). You should be able to explain your code to me. See Section [4.2](#) for more details.

3.4 Adding and Dropping the Course

Students are responsible for adding and dropping the course and verifying these actions in MyMadison. The last date to drop this class with a "W" grade is October 24th. Please consult the appropriate [registrar dates](#) for other deadlines. I will not give “WP” or “WF” grades to students requesting a drop after the deadline except in extraordinary circumstances.

3.5 Disability Accommodations

If you need an accommodation based on the impact of a disability, you must contact the [Office of Disability Services](#) if you have not previously done so. Disability Services will provide you with an Access Plan letter that will verify your need for services and make recommendations for accommodations to be used in the classroom. Once you have shown me this letter, we will sit down and review the course requirements, your disability characteristics, and your requested accommodations to develop an individualized plan appropriate for this course. I will not make any accommodations without the appropriate documentation, as I am not qualified to diagnose disabilities.

3.6 Excused Absences

All University’s policies apply during the semester. Some of these policies appear in the Undergraduate Catalog.

Missing an exam for reasons such as illness, religious observance, participation in required university activities, or family or personal emergency (such as a serious automobile accident or the funeral of a close relative) all are circumstances that *may* qualify as an excused absence. Where possible you should attempt by all means necessary to attend and take exams at their regularly scheduled class period.

If you must be absent during an exam for a legitimate reason, you must contact me at least one week beforehand to make special arrangements. Failure to make prior arrangements for a missed exam will result in a zero grade. Excused absences will be granted at my discretion and only with appropriate documentation. Please contact me as soon as possible if you wish to request an excused absence.

Observance of religious events will be accommodated for students of any faith.

3.7 Classroom Behavior

Students are expected to maintain a high level of civility for all participants in and out of class meetings. This includes respecting the beliefs of participants of all genders, ethnicities, and social backgrounds. Harassment of any type will not be tolerated and failure to behave in a respectful manner will result in referrals to University Counseling or the Office of Student Judicial Affairs. Any instances of sexual harassment will be reported to the Office of Equal Opportunity according to the following policy:

<https://www.jmu.edu/JMUpolicy/policies/1340.shtml>

3.8 Inclement Weather

This class will operate in accord with JMU's inclement weather policy available at <http://www.jmu.edu/JMUpolicy/1309.shtml>.

3.9 Religious Observation Accommodations

I will give reasonable and appropriate accommodations to students requesting them on grounds of religious observation. If you require such accommodations you must notify me at least two weeks in advance.

.

4 Academic Honesty and Collaboration

4.1 Academic Honesty

You are expected to comply with the JMU Honor Code as stated in the Student Handbook and available from the [Honor Council website](#) on all assignments, projects, and exams.

Consulting with other students about problems and solutions is not necessarily a violation of the honor code, depending on the particular assignment. All final work turned in for an assignment must be your own unless it is a group project. In particular, you may not share source or binary code on programming assignments unless the project specification explicitly allows it. If you are in doubt about whether something is an honor code violation, please contact me immediately.

If I find evidence of a violation of the honor code, I will bring the matter to the attention of the involved individuals via email and request a face-to-face meeting. As per section IV of the honor code, first time student offenders may agree that a violation has occurred and accept an appropriate penalty by submitting an "Informal Resolution Agreement Form" to the honor council. If the student is not a first-time offender or if there is disagreement about the violation or penalty, the matter will be referred to the honor council under section V of the honor code.

4.2 PRIME DIRECTIVE

PRIME DIRECTIVE: Be able to explain your own work including homework code and exam solutions.

Nearly all cheating in programming can be averted by adhering to the PRIME-DIRECTIVE. Students may be asked at any time to explain code or exam solutions they submit. Inability to do so

will be construed as evidence of misconduct. More specific guidelines are given below.

Thou Shalt Not

For the purposes of this course, the following actions constitute scholastic misconduct (cheating):

- Directly copying someone else's solution to a homework problem, including student solutions from a previous semester
- Directly copying an answer from some outside source such as the Internet or friend for a homework problem
- Making use of an Instructor Solution manual to complete homework problems
- Paying someone for a homework solution or submitting someone else's work as your own
- Posting solutions to any web site including our course web site
- Collaborating or copying someone else's answer during an exam
- Aiding or abetting any of the above
- Witnessing any of the above and failing to report it the instructor immediately

Fair Collaboration

The purpose of this course is to learn about programming and learning from one another is a great help. To that end, the following actions **will NOT be considered cheating in this course.**

- Talking to other students in the course about HW problems and informally describing how a problem may be solved.
- Getting or giving help fixing a bug or two: a second set of eyes is a great boon to finding that misplaced semicolon that is preventing your code from compiling.
- Searching the Internet for alternative presentations of a programming concept.
- When unsure whether collaboration is fair or not, stop the activity until it can be cleared with instructor.

4.3 Penalties

Any instance of misconduct will be referred to the honor board.

Acknowledgments

Many of the course materials we will be using this semester, including portions of this syllabus, were developed by other faculty members in the JMU Computer Science department and from the book's authors. Contributors include, David Berstein, John Bowers, Alvin Chao, Chris Mayfield, Nathan Sprague and Dee Weikle.