Department of Computational and Data Sciences

4400 University Drive, MS 6A12, Fairfax, Virginia 22030
Phone: 703-993-9298 Fax: 703-993-9300
Email: cds@gmu.edu Web: https://cos.gmu.edu/cds

CDS 403
Machine Learning

1. General Information

Instructor:	Dr. Anamaria Berea
Where:	Virtual
When:	Alt. Synchronous (Wednesday 4.30pm)/ Asynchronous
Course website:	The GMU Blackboard Website
Code repository:	https://gitlab.com/aberea/cds403
Credits:	3
Prerequisites:	CDS 130, CDS 101, CDS303
Office Hours:	By appointment
Instructor email:	aberea@gmu.edu
TA:	NA

2. Course Description

This course teaches the fundamentals of machine learning and helps develop necessary skills for developing computational models of real-world systems, that we can use for pattern recognition, prediction and forecasting, and many more. This is a hands-on coding class that will teach you algorithms such as classification, supervised and unsupervised learning, NLP and image processing, and many more.

3. Learning Outcomes

By the end of the course, students will

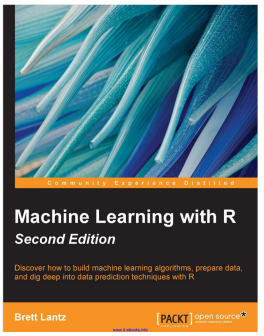
- have a fundamental knowledge on ML,
- be able to write intermediate level code,
- be able to create and run basic ML models of real-world systems,
- have experience with model development software.

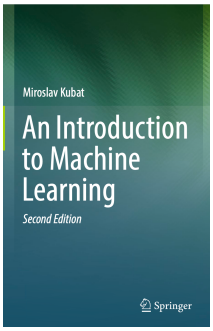
4. Format

The course will be taught virtually through coding exercises and supported with additional material disseminated through the course website and code repository.

5. Textbooks and Slides

We have no required textbooks for the class because there is not a comprehensive book that covers all the topics we will learn throughout the course. The following are free e-books that you may find helpful:





6. Technology Requirements

Hardware: You will need to access to a Windows, Mac, or Linux computer with at least 2 GB of RAM and a fast, reliable broadband Internet connection (e.g., cable, DSL). For the amount of computer hard disk space required for the course, consider and allow for the space needed to: (1) install the required and recommended software and (2) save your course assignments.

Software: This course will be using Python and/or R as a programming language. Prior knowledge of these languages is required.

7. Course outline (Tentative)

In this course we will cover the following topics (please note that the topics and their order are subject to change at the discretion of the instructor, any changes will be announced in class or via Blackboard):

Week Of	Topic	What is Due
August 24	Introduction to Machine Learning	
August 31	Understanding Data for ML	HW 1
September 7	Lazy Learning	HW2
September 14	Probabilistic Learning	HW2
September 21	Classification Algorithms	HW3
September 28	Forecasting and Predictive Algorithms	HW4
October 5	Neural Networks 1	HW5
Fall Break		
October 19	Neural Networks 2	HW6
October 26	Neural Networks 3	HW7
November 2	Support Vector Machines	HW8
November 9	Pattern Recognition	HW9
November 16	Clustering Methods	HW10
Thanksgiving Recess		
November 30	Evaluating and Improving Model Performance	Class Project
December 7	Student Presentations	
TBD	**Final Exam**	

8. Grades

Each homework and written exam will be given a numerical grade on a 0-100 scale. Some homework assignments may include bonus tasks which can increase the total score over 100. At the end of the term, the final mark will be totaled as a weighted average according to the following weights:

Average Homework Score	**40%**
Class Project	**30%**
Final Exam Score	**30%**

Please note that the average homework score is calculated by dividing the total scores of all homework by the total number of homework assignments. Final grades at the end of the course will be assigned based on the following table, independent of the relative standing in the class.

Final Mark	Corresponding Grade
>96.7	A+
93.3 – 96.6	A
90.0 – 93.2	A-
86.7 – 89.9	B+
83.3 – 86.6	B
80.0 – 83.2	B-
76.7 – 79.9	C+
73.3 – 76.6	C
70.0 – 73.2	C-
60.0 – 69.9	D
<60.0	F

9. Exams

The course includes mandatory class project and final exams. You are responsible for all the content covered before the exam. The exam will be conducted in the form of a take-home exam using your own personal electronic devices. There will be NO GROUP work in the exam, meaning it will be done individually. The honor code will be in effect in ensuring that all work turned in will be YOUR OWN.

A student who cannot write a course examination or complete a course homework because of an incapacitating illness, severe domestic affliction, or other compelling reasons can apply for an extension of time. Note that such extensions will be evaluated case by case. There is no guarantee that the instructor will grant the extension.

10. General guidelines for homework preparation and submission

- All homework assignments are mandatory. A single homework may involve a combination of questions requiring you to write code and textual answers.
- Grades of assignments will be based on conciseness and completeness of your answers. Please write to the point and explicitly address the question or task. Avoid using unnecessary graphics (figures, tables, graphs) unless they serve a specific purpose.
- Programming assignments:
  - should be submitted as a single file per homework named according to the format: homework_number.R. For instance, if you are turning in Homework 4, then your file name should be "homework_4.R".
  - should start with comment lines that show (1) student's full name, (2) assignment number, and (3) question number. See below for an example:

```
########################################
##                Name:  First  M.  Last
##                Assignment:  4
########################################
```

- Written assignments should be typed in a single word/PDF document. In other words, one word/PDF file will contain all written answers. Assignments that are hand-written in any form (including scanned ones) will NOT be accepted.
- In case of more than one file needs to be submitted, you should submit a single ZIP file containing all the assignment files.
- Please make sure you have a backup of all the materials you submit.
- Assignments should be submitted **through the Blackboard course website**. Every assignment should allow you to submit the work multiple times up until the deadline. The last version of the assignment will be graded. If you turn in a newer version of your homework, you must include ALL files associated with that homework.

Please note: Assignments should be submitted only through the Assignment submission section of the Blackboard system - DO NOT email assignments directly to the instructor.

11. Late submission

Assignments submitted **after the due date will not be accepted**. Exceptions to this policy may be made given serious circumstances at the discretion of the Instructor.

Please note: Deferral of a work is a privilege and not a right; there is no guarantee that a deferral will be granted. Please make sure you notify the instructor as soon as you know a deferral is required.

12. Attendance

Attendance is not part of the grade but for the synchronous classes it is highly recommended. It is worth noting that course slides may not convey all the information given during the lecture.

13. Course website

The course has a Blackboard website. This website will provide you a portal through which you may obtain lecture notes, retrieve assignment data and review links to additional materials, and receive special announcements. You are required to visit the course website regularly and follow all announcements. Please notify ITU (and, if necessary, the instructor) if you encounter any problems accessing the Blackboard website.

14. Electronic communication, office hours and support

All course-related submission of assignments should be made through the course Blackboard website. Please **DO NOT** email your assignment submissions to the instructor unless the Blackboard website is down for an extended period.

Students are encouraged to contact the instructor and the assistants for any questions regarding the course content. The ideal time to contact the instructor for such inquiries is the office hours which is provided under General Information. The instructor will notify the students via GMU email if there are any temporary or permanent changes in office hours. There is no need to make an appointment for the office hours; first come, first served. For other times, students MUST contact the instructor via email to make an appointment. Ideally, the instructor will respond to course-related student emails within 2 business days. If you include "[CDS 403]" in the email subject, it will help to accelerate the response time.

15. Students with disabilities

Disability Services at George Mason University is committed to providing equitable access to learning opportunities for all students by upholding the laws that ensure equal treatment of people with disabilities. If you are seeking accommodations for this class, please first visit http://ds.gmu.edu/ for detailed information about the Disability Services registration process. Then please discuss your approved accommodations with me. Disability Services is located in the Student Union Building I (SUB I), Suite 2500. Email: ods@gmu.edu | Phone: (703) 993-2474

16. Expectations from students

- *Academic Integrity:* Students must be responsible for their own work, and students and faculty must take on the responsibility of dealing explicitly with violations. The tenet must be a foundation of our university culture. [See http://academicintegrity.gmu.edu].
- *Honor code:* Students must adhere to the guidelines of the George Mason University Honor Code. [See https://oai.gmu.edu/mason-honor-code/].
- *MasonLive/Email (GMU Email):* Students are responsible for the content of university communications sent to their George Mason University email account and are required to activate their account and check it regularly. All communication from the university, college, school, and the program will be sent to students solely through their Mason email account.
- *University Policies:* Students must follow the university policies (See http://universitypolicy.gmu.edu) including the Responsible Use of Computing [See https://universitypolicy.gmu.edu/policies/responsible-use-of-computing/).

17. Frequently asked questions

- *What software do we need?*

R.

- *What are the required knowledge and skills to be successful in this course?*

This course has a prerequisite of CDS 130, CDS303 and CDS101, which touches upon some programming concepts and also intro to ML. So, you are expected to have computing skills like running a program and developing algorithms. Python to R is transferable knowledge.

- *Have another question?*

Ask the instructor via email. If your question applies to all students, it will be shared here as well.

Disclaimer: Any typographical errors in this Course Outline are subject to change and will be announced in class. The date of the final examination is set by the Registrar and takes precedence over the final examination date reported by the instructor.

Notes: (1) Recording is permitted only with the prior written consent of the professor or if recording is part of an approved accommodation plan. (2) The format and template of this syllabus is prepared based on the syllabus of Dr. Andreas Zufle's GGS 787 - Scientific Data Mining for Geo-informatics 2018 course.