MA601: Topics in Mathematics

(Mathematical Foundation of Data-Driven Science and Machine Learning)

Spring 2022 (Jan 13 - April 29) - 3 Credit Hours

Instructor Contact Information

Name: Dr. Abd-AlRahman AlMomani (***Dr. AlMomani***)
Office Address: SC 383
Email Address: aaalmoma@clarkson.edu
Office Hours: MW 11:00 - 12:00 PM, 1:00 - 2:00 PM, and by appointment.

Masking

With regard to wearing masks, Clarkson is following the guidance and requirements of the State of New York. As of the time of writing this syllabus, mask wearing in classrooms will be required for the first few weeks of class. You will be notified when the university moves to its next phase and the guidance from CDC and St. Lawrence County Public Health indicates vaccinated individuals do not have to wear masks inside. You should have a mask with you at all times.

Course Description

The goal of this course is to develop a "Mathematical Foundation of Data-Driven Science and Machine Learning". Modern state-of-the-art methods in data driven modeling will be explored, focusing on the challenges and research directions/opportunities in these methods. The mathematical foundation of machine learning (ML) will be explored, focusing on the classification and clustering problems, with connections to the applications on dynamical systems (Physical, Biological and Chemical systems).

Prerequisites

I will assume the student has undergraduate-level knowledge of Linear Algebra, Matrix Computations, and Numerical Methods.

Instructor Participation

During the course, you can expect me to respond to your emails within 48 hours (excluding weekends). The submitted work will be graded within 7 calendar days of submission, and to be an active participant on the discussion board.

Delivery Method

Lectures will be delivered by in-person meetings. **Attendance is mandatory**. If not able to meet in-person (occasionally with permission, or by SAS approval for the semester), students will join the class via video conferencing and participate in activities.

Instructional Materials

We will source the material from different books. Excerpts from different books and published papers will be provided through the closed Moodle webpage for the class.

However, a very good reference that covers most of the material is:

- Data-Driven Science and Engineering, by Steven L Brunton, J Nathan Kutz.

and it could be helpful to have this book.

Technology

- Computer System & Software Requirements
- Software Accessibility Policies in General
- Software Privacy Policies in General

Minimum Technology Skills

- Use a learning management system (Moodle).
- Basic familiarity with LaTeX.
- Basic familiarity with Matlab OR Python programming.

Course Outcomes (CO)

This Class serves three main purposes:

1. Provide the students with the mathematical foundations that are necessary to carry on research in complex dynamical systems using the modern state-of-the-art methods in data-driven science and machine learning.
2. Provide the students with hands-on experience on implementation and using of these methods.
3. Sharpen the students' skills in documenting their research progress and communicating their progress and findings with their research advisors and other scientific community.

CO1: Explain similarities and differences between White, grey, and black box modeling.

CO2: Rigorously perform dimensionality reduction using PCA, with in-depth understanding of the geomatic view of SVD and PCA.

CO3: Explain similarities and differences between different spaces of basis functions in terms of complexity, and physical interpretation of the results.

CO4: Rigorously setup and perform data-driven modeling using sparse identifications of complex dynamics.

CO5: Explain similarities and differences between different regularization methods.

CO6: Explain similarities and differences between different classification and clustering methods.

CO7: Rigorously set up and perform classification and clustering tasks using different ML methods including Neural Networks) with applications focused on dynamical systems and complex fluid flows.

CO8: Understand the fundamentals of DMD, POD, and their applications, characteristics, and shortcomings.

CO9: Understand the fundamentals of complex networks of dynamical systems, and their application in some Biological and Chemical systems.

Course Schedule & Graded Activities

Topical Contents

1. **Basic concepts of inverse and ill-posed problems**
   1.1. Overview.
   1.2. Forward Problem.
   1.3. Numerical methods in solving the forward problem.
   1.4. N-Body problem and Chaos
   1.5. Inverse problem
   1.6. White, Grey, and Black box modeling.
   1.7. The Method of Least Squares
   1.8. Rank Deficiency and Ill-Conditioning
   1.9. Standard systems as a source of data:
      1.9.1. Lotka–Volterra equations.
      1.9.2. Duffing Oscillator.
      1.9.3. Lorenz System.
      1.9.4. Rossler System.
      1.9.5. Kuramoto Model.
      1.9.6. Kuramoto–Sivashinsky equation.
2. **SVD and Dimensionality Reduction**
   2.1. Singular Value Decomposition (SVD)
   2.2. Pseudo Inverse
   2.3. Revisiting the method of least squares.
   2.4. Principal Component Analysis (PCA)
   2.5. Eigenfaces example.
3. **Basis Functions**
   3.1. What is a Mathematical Model? In-depth view.
   3.2. Weierstrass Approximation Theory.
   3.3. Basis Functions.
   3.4. Power Polynomial (Challenges and advantages)
   3.5. Radial Basis Functions (RBF)
   3.6. Fourier Series (Discrete and Fast Fourier transforms)
   3.7. Wavelets
   3.8. Setup of Data-Driven Modeling.
   3.9. Challenges and Misconceptions.
4. **Machine Learning and Model Selection**
   4.1. On Ax = b (Linear Algebra, Dynamical Systems, and Sampling)
   4.2. Regression as an optimization problem.
   4.3. Gradient descent.
   4.4. Regression and sparse identifications.
   4.5. Hard-Thresholding
   4.6. L0 minimization and NP-Hardness.
   4.7. Tikhonov and the big picture of Regularization.
   4.8. L1 Regularization. (LASSO)
   4.9. L1-Magic (Compressed Sensing)

4.10. Subset Selection Problem and Greedy Algorithms.
4.11. Pareto Front
4.12. Model Selection: CrossValidation and Information Criteria

5. Clustering and Classification

5.1. Feature Selection and Data Mining
5.2. Supervised versus Unsupervised Learning
5.3. Unsupervised Learning: K-means Clustering
5.4. Unsupervised Hierarchical Clustering: Dendrogram
5.5. Mixture Models and the Expectation-Maximization Algorithm
5.6. Supervised Learning and Linear Discriminants Analysis (LDA)
5.7. Support Vector Machines (SVM)
5.8. Classification Trees and Random Forest
5.9. On Graph Theory and Spectral Clustering

6. Neural Networks and Deep Learning

6.1. Neural Networks: 1-Layer Networks
6.2. Multi-Layer Networks and Activation Functions
6.3. The Backpropagation Algorithm
6.4. The Stochastic Gradient Descent Algorithm
6.5. Deep Convolutional Neural Networks
6.6. Neural Networks for Dynamical Systems
6.7. The Diversity of Neural Networks

7. Data-Driven Dynamical Systems

7.1. Overview, Motivations, and Challenges
7.2. Dynamic Mode Decomposition (DMD)
7.3. Koopman Operator Theory
7.4. Data-Driven Koopman Analysis
7.5. Proper Orthogonal Decomposition (POD)
7.6. POD for Partial Differential Equations

8. Complex Networks Science

8.1. Networks of Dynamical Systems.
8.2. Erdös-Rényi random graph model.
8.3. Watts-Strogatz small-world model.
8.4. Barabási-Albert (BA) preferential attachment model.
8.5. Biological and Chemical network analysis.
8.6. Gene Regulatory Networks (GRN) and switching systems.

Dates	Module Title	Graded Activities (Course Outcome the Graded Activity Fulfills)
Jan 13- Jan 24	Ch1: Basic concepts of inverse and ill-posed problems	HW1 (CO1, CO2)
Jan 26- Jan 31	Ch2: SVD and Dimensionality Reduction	HW1 (CO1, CO2)
Feb 2 - Feb 9	Ch3: Basis Functions	HW2 (CO3)
Feb 11 - Mar 4	Ch4: Machine Learning and Model Selection	HW3 (CO4, CO5)
Mar 7 - Mar 23	Ch5: Clustering and Classification	HW4 (CO6, CO7)
Mar 25 - April 1	Ch6: Neural Networks and Deep Learning	HW5 (CO6, CO7)
April 4 - April 15	Ch7: Data-Driven Dynamical Systems	HW6 (CO8, CO9)
April 18 - April 29	Ch8: Complex Networks Science	HW6 (CO8, CO9)

Grading

Grade Ranges

Course Average	Grade	Quality Points
97+	A+	4.0
93-96	A	4.0
90-92	A-	3.667
87-89	B+	3.334
84-86	B	3.0
80-83	B-	2.667
76-79	C+	2.334
70-75	C	2.0
<70	F	0

Breakdown

Activity	Percent of Final Grade
Participation and Efforts	5%
Quiz	5%
HW	40%
Exam	20%
Project	30%
Total	

Activities and grading Criteria

Participation and Efforts: TBD

Quiz: (Weekly, on Monday)

HW: Each HW will be divided into three parts

A: Theoretical problems, proves, analysis.

B: Implementation and programming for specific tasks.

C: Will be assigned for some HW (not all of HW will have this part). It mainly requires you to review a paper related to our course and reproduce the paper results.

No Hand-written HW will be accepted. I will provide you with further instructions regarding documentation and submitting the code.

Exam: One Midterm exam (**February 23, 2022**).

Project: divided into three phases:

Phase 1: Proposal

Phase 2: Progress Report.

Phase 3: Final Report.

See the project instructions document for more information on the project topic and objectives.

Course Success

(Explain the estimated workload to your students. How can students be successful in this course?)

Please see tips for being a successful student and other helpful information from the Student Success Center.

Course Policies

Etiquette Expectations & Learner Interaction

Educational institutions promote the advance of knowledge through positive and constructive debate--both inside and outside the classroom. Please visit and follow: Netiquette and Electronic Learner Interaction Guidelines.

Late Work

TBD

Attendance

Lectures will be delivered by in-person meetings. **Attendance is mandatory.**

If you are unable to attend the class in person due to illness or exposure to infection, please notify me ahead of time. If you are in self-isolation, quarantine, etc. but can keep up with the material, please do so. Communicate all suspicion of illness or illness itself to the Student Health and Dean of Students, who will then inform the instructor in a manner that preserves your privacy as a student and as a patient.

Instructional Continuity

In case of an emergency, I may have to cancel the class on a short notice. I will communicate with you via email about this. If I have to stay away from campus for a longer period of time, e.g., to quarantine, then we will continue instruction online. In case that I am incapable of teaching, e.g., too ill for a prolonged period, then a substitute will be found.

Academic Unit Information

The Department of Mathematics is located on the 3rd floor of the Science Center (SC355). Please contact Tyler Hayes (Administrative Assistant, thayes@clarkson.edu or math@clarkson.edu) if you need any assistance that goes beyond this course.

Institutional Policies

Institutional Policies & Regulations

Academic Integrity

Academic Integrity, based on the values of honesty, trust, fairness, respect, and responsibility, is a fundamental principle of scholarship in higher education. Clarkson's Academic Integrity Policy prohibits: plagiarism (using another person's writing or copying any work without proper citation), falsification, unauthorized collaboration during a test or on an assignment, or substitution for another student to take an exam, course or test, and other forms of academic dishonesty.

If you are to benefit from this class and be properly evaluated for your contributions, it is important for you to be familiar with and follow Clarkson University's Academic Integrity policy. Please review this policy online (Undergraduate section IV – Academic Integrity, Graduate section IV – Academic Integrity). ***Work that violates this policy will not be tolerated***. Students who are found responsible for a violation of the Academic Integrity Policy will have both a university process sanction and an academic outcome, that could include a failing grade on the assignment or exam, or a failing grade for the course.

Please refer to **Clarkson Library's Guide to Plagiarism** and the **guide to Citing Sources** for assistance on avoiding plagiarism and properly citing sources.

Students with Disabilities Requesting Accommodation(s)

The University strives to make all facilities and programs accessible to students with permanent, ongoing, and temporary disabilities by providing appropriate and reasonable academic accommodations, as necessary. Disabilities that may benefit from reasonable accommodations include, but are not limited to, broken wrist, ADHD, surgery recovery, Learning Disability, concussion, visual impairment, etc. For more information and/or to request accommodations, contact the Office of Accessibility Services at oas@clarkson.edu or 315-268-7643.

Students with Disabilities Policy
Office of Accessibility Services Website

Other Policies of Note:

Student Regulation Requirements for Excused and Extended Absence

Undergraduate: III-F. Attendance
Graduate – II-F. Attendance

Grading System

Discrimination & Harassment

Religious Accommodations