Machine Learning for Scientific Modeling: Data-Driven Discovery of Differential Equations is a special course taught in Spring 2019 focusing on applications of machine learning to the calibration and discovery of differential equations in applications. Below are a few highlights from the literature in this rapidly evolving field. During the course, we will aim to provide some of the background and present and discuss many of these research results in an informal environment.

Figure 1: (From article 1 below) Discovery of the fractional order $\alpha$ in an equation of the form $D^\alpha u = f$, where $D^\alpha$ is a Riemann-Liouville fractional derivative, by performing physics-informed Gaussian process regression on 5 data points for $u$ (left) and 4 data points for $f$. With this small dataset, the pictured regression recovers the $\alpha$ used to generate the data within 1%.

Figure 2: (From article 2 below) Discovery of conservation law for a double-pendulum by measuring dynamics in time with motion sensors and performing free-form symbolic regression (genetic algorithm) on conserved quantities.
Figure 3: (From article 3 below) Discovery of Navier-Stokes dynamics for fluid flow by performing sparse regression with a library of various derivative terms on simulated data.

Figure 4: (From article 4 below) Training a deep neural network to reproduce the dynamics of a Korteweg-de Vries equation. The neural network is trained on a dataset different from the one shown in the figure, but is able to correctly predict the dynamics of the pictured dataset.
Machine Learning for Scientific Modeling:
Data-Driven Discovery of Differential Equations

Instructors:
Mamikon Gulian: mamikon_gulian@brown.edu

Room:
Room 118 in 170 Hope St.

Time and Date:
9:00 am to 10:20 am on Tuesday/Thursday.

Office Hours:
At the first lecture, we will schedule two hours per week, most likely on Tuesday or Thursday afternoon. Additional office hours can always be scheduled by appointment (just ask me after lecture, or send me an email).

Course Description and Learning Goals:
This junior/senior level course will explore the use of Machine Learning to automate the discovery and calibration of models involving differential equations directly from data. After introducing the basic machine learning tools – Gaussian Processes and Neural Networks – we will see how they can be combined with ODE and PDE computational methods to generate models in physics, medicine, and finance. The course will progress to a survey of recent research works on the topic; a sample of possible articles is


We will also introduce fractional-order and nonlocal differential equations, and spend some time on the intersection of this area with machine learning. Students will complete final projects including a presentation on a relevant research work.

Prerequisites:
Required: Multivariable Calculus (Math 0180, 0200 or 0350), Linear Algebra (Math 0520 or 0540), ODEs (Math 1110 or APMA 0350). Recommended: Partial Differential Equations and a Probability or Statistics course. No prior knowledge of Machine Learning is required – the basics will taught in the course.
Course Outline:

- Part I (4 weeks): Review and survey of ordinary differential equations, partial differential equations. Review of conditional probability and Bayesian reasoning. Introduction of Gaussian models. This will be drawn from Chapters 1-4 of Murphy. There will be weekly problem sets (roughly 15 problems per week) and two quizzes on these topics.

- Part II (3 weeks): Introduction to Machine Learning. This will focus on Gaussian Processes and Neural Networks. References will be Chapter 15 and 16 of Murphy. There will reading and problem sets assigned on both topics, and a quiz on each of these two topics.

- Part III (4 weeks): Presentation of selected research articles on Machine Learning for Scientific Modeling/Computing. This part will also help inform students for making a decision of what to focus on for their final project. Students will be assigned articles to read, as well as small problem sets (roughly 5 questions) about the articles to guide their reading.

- Part IV (3 weeks): Student presentations and projects. The presentation must focus on a specific paper or cluster of papers on a subject, and describe the methodology, implementation, and impact of the paper.

Textbook:

*Machine Learning: A Probabilistic Perspective*, by Kevin P. Murphy (MIT Press, 1st edition). The Brown University Library has an electronic copy of this textbook that you can read freely. For the basics of Machine Learning (Probability, Bayesian methods, Gaussian Processes, and Neural Nets), we will read parts of Chapters 1-4, 15, and 16. Another textbook we may draw from is the online draft of *Bayesian Reasoning and Machine Learning* by David Barber (2016 Draft, 686 pages), that I will provide. This has similar content to the textbook of Murphy, in a more informal style. Eventually, the course will transition to focusing on research articles that I will provide prior to the lectures.

Web Page: I will make a course page to host course material on: [www.math.brown.edu/~mgulian](http://www.math.brown.edu/~mgulian)

Evaluation: The course grade is calculated as follows:

- Homework Problems 25%
- In-Class Quizzes 25%
- Final Presentation 50%

A tentative course grading for your total percentage $S$ is as follows: $S \geq 90\%$ is an A, $90\% > S \geq 80\%$ is a B, $80\% > S \geq 70\%$ is a C, and anything below is a failing grade. I may curve grades, but these percentages will always guarantee you a grade at least as high as indicated here. The course environment will be informal and if you are stuck on anything, just write to me with questions or to arrange some office hours.

Course-Related Work Expectations: Over 14 weeks, students are expected to spend roughly 3 hours per week in class and in office hours combined (roughly 40 hours total). Outside of class, students can expect to spend 10 hours per week on problem sets or article reading (140 hours). In addition, the final project/presentation is expected to take roughly 20 hours of preparation.

Collaboration Policy and the Academic Code: While students are allowed (and even encouraged) to work together and/or ask each other questions about homework problems, it is unacceptable to copy or submit another student’s work, calculations, or final answers without solving the problem yourself. Cheating on exams or any other form of academic dishonesty is prohibited by Brown’s Academic Code. All students should be aware of this code, and they should understand that violating the code can have serious consequences.