Machine Learning and the Physical Sciences

Spring 2024, GP148/248 CME215

Course information

Course Faculty: Prof. Ching-Yao Lai
Office location: Mitchell room 357
Email: cyaolai@stanford.edu

Class: Tuesdays and Thursdays 9-10:20am
Classroom: Building 300 room 300

Office hours: Upon request by email

Course Teaching Assistant:
Josh Rines (jrines@stanford.edu)
Niall Coffey (nbcoffey@stanford.edu)

Summary:
This course provides a survey of the rapidly growing field of machine learning in the physical sciences. It covers various areas such as inverse problems, emulating physical processes, model discovery given data, and solution discovery given equations. It both introduces the background knowledge required to implement physics-informed deep learning and provides practical in-class coding exercises. Students have the opportunity to apply this emerging methodology to their own research interests across all fields of the physical sciences, including geophysics, climate, fluids, or other systems where the same technique applies. Students develop individual projects throughout the semester.

Recommended prerequisite:
Calculus (e.g. MATH 51), Differential Equations (e.g. MATH 131P), or equivalents. No background in machine learning is necessary, as the class will cover its fundamentals.

Grading scheme: Letter & CR/NC
Participation – 10 %
Oral presentations – 15 %
Final Project Paper - 75%

Credit: 3 Units

Ways of thinking/Ways of doing breadth requirement:

This course satisfies the following WAYS requirements:

- Scientific Method and Analysis (SMA)
- Applied Quantitative Reasoning (AQR)

Course learning goals/outcomes

- Learn useful concepts about machine learning
- Learn how to incorporate our physics knowledge into machine learning training
- See examples of how these can be applied to the physical sciences, e.g. climate science and geophysics. We welcome you to contribute to new applications!
- Gain practical experiences to use/implement the codes covered in class
- Utilize softwares like Google Colab and the Stanford cluster for quantitatively analyzing and visualizing dataset.
- Learn how to ask a new scientific question based on available data and knowledge
- Conceive and develop a course project that applies these novel and exciting tools to your research!

Coursework and assessments

Coding exercises will be distributed throughout the class that will not be graded. Students are evaluated based on their **course project**, **final paper**, and a **final presentation** of the projects**:**

1. For the **course project** the students should apply the method taught in class to their research fields or areas of interests and develop individual projects throughout the semester.
2. Students will have **final presentations** of their course project near the end of the class. This will be graded and count towards 15% of your course grade.
3. Students should submit a **final paper** (due on June 13 5pm) that summarizes the course project. The final paper should not be longer than 10 pages. Note that the project along with the final paper is a large part of your course grade (75%). Start your project and get familiar with codes early! We expect that you will be working intensively on the projects during weeks 5-8.

Basic criteria for the evaluation written final paper are:
- Sophistication of material presented
- Quality of exposition
- Amount of effort

Schedule

Course schedule is outlined below, subject to changes.

<table>
<tr><th>Week</th><th>Days</th><th>Lecture #</th><th>Topic</th></tr>
<tr><td rowspan="2">1</td><td>Apr 2</td><td>1</td><td>Introduction</td></tr>
<tr><td>Apr 4</td><td rowspan="3">2-4</td><td rowspan="3">Basics of neural networks
- Universal function approximation
- Back propagation
- 2 lectures of concepts and 1 lecture of in-class coding exercise</td></tr>
<tr><td rowspan="2">2</td><td>Apr 9</td></tr>
<tr><td>Apr 11</td></tr>
<tr><td rowspan="2">3</td><td>Apr 16</td><td rowspan="4">5-8</td><td rowspan="4">Physics-informed machine learning
- Concepts
- Fourier series analogy
- Examples of the Schrodinger, Burgers, and the Navier Stokes equations
- Overfitting</td></tr>
<tr><td>Apr 18</td></tr>
<tr><td rowspan="2">4</td><td>Apr 23</td></tr>
<tr><td>Apr 25</td></tr>
<tr><td rowspan="2">5</td><td>Apr 30</td><td>9</td><td>Inferring hidden parameters in fluid dynamics</td></tr>
<tr><td>May 2</td><td>10</td><td>Example of PINN physics-informed machine learning applied in Climate science and Geophysics</td></tr>
<tr><td rowspan="2">6</td><td>May 7</td><td>11</td><td>Discovering governing equations from data I</td></tr>
<tr><td>May 9</td><td>12</td><td>Demonstration of running codes on Colab (TA)</td></tr>
<tr><td rowspan="2">7</td><td>May 14</td><td>13</td><td>Discovering governing equations from data II</td></tr>
<tr><td>May 16</td><td rowspan="3">14-16</td><td rowspan="3">Basics of convolutional neural networks (CNN) and application to climate modeling</td></tr>
<tr><td rowspan="2">8</td><td>May 21</td></tr>
<tr><td>May 23</td></tr>
<tr><td rowspan="2">9</td><td>May 28</td><td rowspan="3">17-19</td><td rowspan="3">Student presentations on course projects
- Students should submit a final paper (due on June 13 5pm) that summarizes the course project. The final paper should not be longer than 10 pages.</td></tr>
<tr><td>May 30</td></tr>
<tr><td>10</td><td>June 4</td></tr>
</table>

Course materials

Course slides and notes will be provided on Canvas.

Technology: You will need to have access to a device that connects to the internet so that you can access email, Canvas, and the codes. We will also use your laptops to run Google Colabs in the cloud.

Github: We will provide sample codes in our class Github page:
https://github.com/GP248CME215
In the beginning of the class you will be asked to create a Github account to gain access to the codes. Note that some codes were written for educational purposes and not open source. If you intend to use these codes for your research towards a publishable paper, please discuss with Prof. Lai.

Tools (subject to changes): For your course projects we recommend starting with smaller datasets and shorter jobs that can be run on Google Colab, or locally on your personal computer with Python. The codes we provide are written in tensorflow; you are free to rewrite it in Pytorch.

Sample reading list:

1. Raissi, M., Perdikaris, P. and Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, pp.686-707.
2. Raissi, M., Yazdani, A. and Karniadakis, G.E., 2020. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science*, *367*(6481), pp.1026-1030.
3. Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S. and Yang, L., 2021. Physics-informed machine learning. Nature Reviews Physics, 3(6), pp.422-440.
4. Brunton, S.L., Proctor, J.L. and Kutz, J.N., 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, *113*(15), pp.3932-3937.
5. Rudy, S.H., Brunton, S.L., Proctor, J.L. and Kutz, J.N., 2017. Data-driven discovery of partial differential equations. Science Advances, 3(4), p.e1602614.
6. Kashinath, K., Mustafa, M., Albert, A., Wu, J.L., Jiang, C., Esmaeilzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A. and Manepalli, A., 2021. Physics-informed machine learning: case studies for weather and climate modelling. Philosophical Transactions of the Royal Society A, 379(2194), p.20200093.

Course policies

Presence and Participation:
Thinking through difficult issues is most productive and most fun when done out loud in the company of others! This is why attendance and participation is a large part of your course grade (10%).

The Honor Code:
You will be evaluated in this course as an individual and are expected to cite sources and

individuals from whom you have learned and borrowed as a display of academic, intellectual, and creative integrity. Please review Stanford's Honor Code, and documentation and citation resources from the Hume Center for Writing and Speaking. When in doubt, feel free to chat with Prof. Lai.

Academic Accommodations:
Stanford is committed to providing equal educational opportunities for disabled students. Disabled students are a valued and essential part of the Stanford community. We welcome you to our class.

If you experience disability, please register with the Office of Accessible Education (OAE). Professional staff will evaluate your needs, support appropriate and reasonable accommodations, and prepare an Academic Accommodation Letter for faculty. To get started, or to re-initiate services, please visit oae.stanford.edu.

If you already have an Academic Accommodation Letter, we invite you to share your letter with us. Academic Accommodation Letters should be shared at the earliest possible opportunity so we may partner with you and OAE to identify any barriers to access and inclusion that might be encountered in your experience of this course.

Students who are immunocompromised should register with the OAE as soon as possible.

Student athletes who anticipate challenges in being able to participate in class or submit assignments on time should speak to a course instructor or teaching assistant as soon as possible about available alternatives or allowances.