Recently, there has been a surge of interest in exploiting geometric structure in data and models in Machine Learning. This course will ------give an overview of this emerging research area and its mathematical foundation, with a focus on recent literature and open problems. We will cover a range of topics at the intersection of Geometry and Machine Learning including Basic Differential Geometry, Graph Representation Learning, Manifold Learning, Graph Neural Networks, Machine Learning on Manifolds, and Geometric Deep Learning. Lectures will be complemented by student-led discussions of relevant papers.

Course Information

Class times: Mo/ Wed, 3-4.15pm, **Maxwell Dworkin G115**

Instructor: Melanie Weber (mweber@seas.harvard.edu)

Teaching Fellows: Henry Bae (henrybae@college.harvard.edu), Binita Gupta (bgupta@mde.harvard.edu), Victoria Tang (xutang@g.harvard.edu)

Office hours:

Melanie Weber: Wed, 4.30-6.30pm, Pierce 311 (15min time slots, please book here)

Teaching Fellows: Mon, 1:00-3:00pm, and Tue, 6:00-7.00pm in Maxwell-Dworkin 323

Course Description

We will discuss methods for Machine Learning in Non-Euclidean spaces, with a focus on understanding the key mathematical and algorithmic ideas. The course covers three areas of Geometric Machine Learning:

1. Geometric Representation Learning (~Weeks 1-4);
    - Learning in High Dimensions
    - Metric Embeddings
    - Manifold Learning
    - Shallow Graph Embeddings
2. Graph Neural Networks (~Weeks 5-8);
    - Graphs, Groups, Symmetries
    - Graph Learning Tasks
    - Message-Passing Paradigm, Architectures and Expressivity
    - E(3)-Equivariant Graph Neural Networks
    - Discrete Curvature for Graph Machine Learning
3. Machine Learning on Manifolds (~Weeks 9-13).
    - Manifolds, Tangent Spaces, Geodesics
    - Optimization on Manifolds
    - Geometric Deep Learning
    - Challenges and Progress in Geometric Machine Learning

The course mainly consists of lectures, which are complemented by student-led discussions of relevant papers (sign up here). A list of suggested papers can be found on the reference list for the course (marked with *). In addition, each student will conduct a course project, which may consist of reproducing or extending a paper in one of the three areas above or of applying geometric methods in the student's own area of research.

Goals

By the end of this course, students will be able to:

- Recognize geometric structure in data and in machine learning models.
- Utilize algorithms that exploit such geometric structure.

Prerequisites

This course targets PhD students with research interests in Machine Learning and Data Science. Masterâ€™s and advanced undergraduate students are welcome, especially those pursuing or interested in pursuing research in Machine Learning and Data Science. Prerequisites are APMTH 120 and COMPSCI 181 or equivalent. Introductory-level knowledge of Differential Geometry is helpful, but not required.

Grading

Students are required to lead at least one in-class discussion on a reading assignment, submit solutions for three problems sets (one for each of the main course topics) and to conduct a course project. The course grade is based on class participation and attendance (5%), a one-page summary for a reading assignment whose discussion they lead (10%), problem sets (30%) and the course project (55%); please see below for the detailed breakdown.

Final Grade Breakdown:

Participation	5% of final grade
Problem sets (3)	30% of final grade
Paper Presentation/ Reading summary (1 page)	10 % of final grade
Project Proposal (1 page)	10 % of final grade
Project Presentation (poster session)	10 % of final grade
Project Report (8 pages)	35 % of final grade

Submission deadlines as announced at the beginning of class are final. For problem sets only, late submissions within 5 days will receive partial credit. The late penalty is applied as a multiplier to your problem set score, which decreases linearly from 1.0 (full credit) for an assignment submitted in time, to 0.5 (half credit) for an assignment that is 5 days late. For example, if a student receives a score of 80 out of 100 on their first problem set, which was submitted one day late, the student receives a score of 80*0.9=72. Scores are rounded up to the next integer. Each student receives (2) â€œlate daysâ€ for the semester, which will waive up to two days of late penalties. They will be automatically applied to your scores in a way that maximizes your overall grade.

Academic Integrity

Students must submit their own work and original code (where applicable) but may discuss their work with their peers. Students are expected to list the names of students with whom they have collaborated on the problem sets. All written submissions and presentations must adhere to standard citation practices and acknowledge quotations, summaries, paraphrases, and arguments from any sources.

Attendance, Participation, and Classroom Climate

Students are expected to attend lectures. There will be weekly reading assignments in preparation for the lectures and in-class discussions; students are expected to come to class prepared and to participate actively.

Accommodations for students with disabilities

Harvard University values inclusive excellence and providing equal educational opportunities for all students. Our goal is to remove barriers for disabled students related to inaccessible elements of instruction or design in this course. If reasonable accommodations are necessary to provide access, please contact the Disability Access Office (DAO). Accommodations do not alter fundamental requirements of the course and are not retroactive. Students should request accommodations as early as possible, since they may take time to implement. Students should notify DAO at any time during the semester if adjustments to their communicated accommodation plan are needed.

Class Schedule

		Suggested		

Date	Topic	Readings	Materials	Assignments
01/22	Lecture 0: Introduction		slides	
	Part I: Geometric Representation Learning			
01/24	Lecture 1: Learning in High Dimensions	Bronstein et al. [ch.2]	slides	Homework 1 (due 02/21)
01/29	Lecture 2: Metric Embeddings	Magen, 2006 (lecture notes)	slides	
01/31	Lecture 3: Manifold Learning I	Tennenbaum et al., 2000	slides	
02/05	Lecture 4: Manifold Learning II	Belkin & Niyogi, 2003 Saul, Roweis (survey)	slides	
02/07	Lecture 5: Shallow Graph Embeddings	Hamilton, 2020 [ch. 3]	slides	
02/12	Student-led paper discussions I	Van der Maaten, Hinton (2008) Qiu et al. (2018) Bordes et al. (2013) Moon et al. (2019) Klimovskaia et al. (2020)		
02/14	Student-led paper discussions II	Ribeiro et al. (2017) Nickel, Kiela (2020) Chami et al. (2020) Gu et al. (2019) Bronstein et al. (2006)		
02/19	No class (President's Day)			
	Part II: Graph Neural Networks			
02/21	Lecture 6: Introduction to Graph Neural Networks I	Hamilton, 2020 [ch. 1.2] Hamilton, 2020 [ch. 5.1]	slides	Homework 1 due
	Lecture 7:			

02/26	Introduction to Graph Neural Networks II	Hamilton, 2020 [ch. 7.3]	slides	Homework 2 (due 03/20)
02/28	Lecture 8: Groups, Symmetries, Graphs	Bronstein et al., 2022 [ch. 3.1, 3.2]	slides	
03/04	Lecture 9: Equivariant Graph Neural Networks	Bronstein et al., 2022 [ch. 5.3, 5.4, 5.5]	slides	
03/06	Lecture 10: Discrete Curvature and Graph Machine Learning	see references in slides	slides	Course project proposals due
	Spring Recess			
03/18	Student-led paper discussions III	Bruna et al. (2014) Nguyen et al. (2023) Jing et al. (2021) You et al. (2019) Zitnik et al. (2018)		
03/20	Student-led paper discussions IV	BrasÃ³ et al. (2022) Huang et al. (2023) Velickovic et al. (2018) Ying et al. (2018) Zaheer et al. (2017)		Homework 2 due
	Part III: Machine Learning on Manifolds			
03/25	Lecture 11: Manifolds, Tangent Spaces, Geodesics	see references in slides	slides	
03/27	Lecture 12: Optimization on Manifolds I	Boumal, 2022 [ch. 1]	slides	Homework 3 (due 04/24)
04/01	Lecture 13: Optimization on Manifolds II	Boumal, 2022 [ch. 4]	slides	
04/03	Lecture 14: Geodesically Convex Optimization	Boumal, 2022 [ch. 11]	slides	
04/08	Lecture 15: Learning in High Dimensions, revisited	see references in slides	slides	
04/10	Lecture 16: Geometric Deep Learning I	Bronstein et al., 2021 [ch. 3.5, ch. 5]	slides	
04/15	Lecture 17: Geometric Deep		slides	

	Learning II			
04/17	Student-led paper discussions V	Finzi et al. (2021) Cohen et al. (2018) Batzner et al. (2022) Safaie et al. (2023)		
04/21	Course project posters due for print (optional). Please use the template (size 36 x 24).			
04/22	Student-led paper discussions VI	Talwalkar et al. (2008) Qi et al. (2017) Zhang et al. (2016) Wang et al. (2022) Gao et al. (2022)		
04/24	Hardness of Learning with Symmetry (Guest Lecture by Bobak Kiani (Harvard))	Reyzin (2020) [p. 1-6]	slides	Homework 3 due
04/29	Course Project Presentations (Poster Session)			
05/03	Course project report due			