XCS224W: Machine Learning with Graphs

Syllabus and Course Information

Welcome

Welcome to XCS224W: Machine Learning with Graphs! This professional course is based on graduate-level material from Stanford's on-campus course CS224W, adapted for a professional certificate format. In this course you will:

- Learn from Stanford graduate lecture videos (Winter 2021) that have been edited for easier navigation, reference, and review.
- Complete guided homework assignments implementing content covered in the course lectures.
- Receive support from Stanford-affiliated Course Facilitators.
- Connect to a cohort of peers from diverse locations and professional backgrounds.

Course Launch

All lecture videos will be available on the first day of the course (January 31) at 12:00pm Pacific Time. Course assignments will be made live on the days indicated in the calendar below, without exception. Maintaining the assignment schedule enables Course Facilitators to be most effective in providing support and answering questions on subject matter throughout the course.

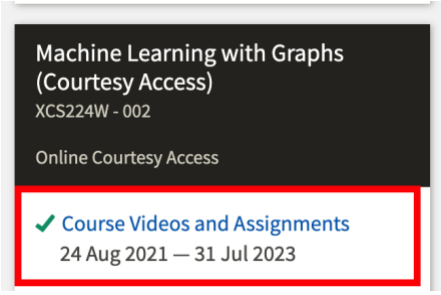
Getting Started

This course will use different tools to distribute content, run assignments, and deliver support. They are:

1) **SCPD Learning Management System** – accessed via the mystanfordconnection site which you used to apply to and enroll in this course.
2) **Slack** – for additional course support and class discussions.
3) **Google CoLab** – for assignment notebooks and data.

Accessing Your Course

1. On **January 31st after 12pm Pacific Time**, log in to the mystanfordconnection account you used when applying for the Artificial Intelligence Professional Program.
2. XCS224W: Machine Learning with Graphs will be visible as a live course. Click the link titled "Course Videos and Assignments" to enter our learning management system.



Joining Slack

In addition to individual support from Course Facilitators (more details and guidelines in Course Facilitators, Support, and Guidelines section below), the cohort will have a Slack workspace to ask additional questions and discuss course topics. An email invitation to http://xcs224w-scpd.slack.com/ will be sent to your email address on file with SCPD on **January 28th**.

If you have previously joined an SCPD or Stanford Slack Workspace for a different course in the AI Professional Program, we have found that Slack does not send a notification when our staff invites you to an additional workspace. Instead, you are automatically re-activated in the system. On **January 28th** you should proceed directly to http://xcs224w-scpd.slack.com/ à I have a guest account à Log in using your credentials.

Google CoLab

Course notebooks and homework will be posted in public Google CoLab links. In order to complete them, you'll need to make a copy within your own Google account. If you are unfamiliar with CoLab, you can explore some resources here and here to get started.

Course Calendar

Below is a ***potential*** pacing guide for moving through the course, however you are free to watch videos at a faster or slower pace depending on your availability. You are welcome to move through the videos at a different pace if you prefer:

WEEK	POTENTIAL VIDEO PACING	ASSIGNMENTS	
1	Module 1: Traditional Method for ML on Graphs	**Colab 0, CoLab 1 Open & CoLab 2 Open**	**January 31**
2	Module 2: Introduction to Graph Neural Networks	**CoLab 3 Open**	**February 11**
		CoLab 1 Due	**February 13**
3	Module 3: Theory of Graph Neural Networks		
4	Module 4: Knowledge Graphs	**CoLab 4 Open**	**February 25**
		CoLab 2 Due	**February 27**
5	Module 5: Network Motifs and Community Structure		
6	Module 6: Generative Models for Graphs	**CoLab 5 Open**	**March 11**
		CoLab 3 Due	**March 13**
7	Module 7: Advanced Topics in GNNs		
8	Module 8: Special Topics and Research on GNNs	**CoLab 4 Due**	**March 27**
9			
10			
		CoLab 5 Due	**April 10**

Assignments and Quizzes

Each CoLab will consist of a notebook that contains guided cells as well as homework prompts. You'll upload your notebooks to the SCPD Gradescope autograder, which will report back your current score on the assignment. You are allowed to submit any assignment unlimited times before the stated due date so you can continue to get feedback and de-bug.

CoLab 1

In this Colab you will write a pipeline for learning node embeddings over a classic graph in network science, the the Karate Club Network. First, you will explore multiple graph statistics for this graph. You will then transform the graph structure into a PyTorch tensor, so that you can perform machine learning over the graph. Finally, you will finish implementing your first graph learning algorithm: a node embedding model.

CoLab 2

In this Colab you will construct your first graph neural network using PyTorch Geometric (PyG) and apply the model on two Open Graph Benchmark (OGB) datasets. First, you will learn how PyTorch Geometric stores the graphs in PyTorch tensor. You will then load and take a quick look on one of the Open Graph Benchmark (OGB) datasets using the ogb package. Finally, you will build you own graph neural network using PyTorch Geometric.

CoLab 3

In this Colab you will first implement the well known GraphSAGE (Hamilton et al. (2017)) and GAT (Veličković et al. (2018)) GNN layers directly using Pytorch Geometric (PyG). You will test these custom layers on the CORA dataset, a standard graph benchmark dataset. Then you will work through a tutorial on DeepSNAP, a Python library assisting efficient deep learning on graphs, where you will learn to split graphs in different settings and apply dataset transformations. Finally, you will use DeepSNAP to train a simple GNN model to perform edge property prediction (link prediction).

CoLab 4

In this Colab, you will shift your focus from homogenous graphs to heterogeneous graphs. Heterogeneous graphs extend traditional homogenous graphs by incorporating different node and edge types. In this notebook, you will first learn how DeepSNAP stores and represents heterogeneous graphs as PyTorch Tensors. Then, you will build your own heterogenous graph neural network models, where you will use the knowledge from Colab 3 + Pytorch Geometric (PyG) to implement two different heterogeneous message passing layers. Lastly, using DeepSNAP, you will test your models on a node property prediction task: the heterogeneous ACM node prediction dataset.

CoLab 5

In this final Colab you will experiment with scaling up GNNs using PyTorch Geometric (PyG), DeepSNAP, and NetworkX. First, you will use PyG's NeighborSampler to scale up the training and testing of a GNN on the OGB arxiv dataset. Then, using DeepSNAP and NetworkX, you will implement a simplified version of the NeighborSampler and run experiments with different sampling ratios on the Cora graph. Finally, you will partition the Cora graph into clusters using different partition algorithms and then train a vanilla Cluster-GNN.

Late Assignments and One-time Penalty Waiver

All assignments can be turned in **up to five days late and are assessed a penalty of -1 point per day**. After five days, the submission link will close, and entries will no longer be accepted.

We understand that personal or professional events may cause you to miss a deadline on an assignment. Each student can use a **one-time penalty waiver to remove a late penalty.** The penalty waiver **cannot** be split into smaller parts (e.g., you cannot use two days on Assignment 3 and three days on Assignment 4.). In order to use your penalty waiver, contact your Course Facilitator and SCPD staff.

Passing the Course and Earning the Certificate

In order to earn the Certificate of Achievement for this course, you must achieve a final cumulative score of 70% or higher. Once you have successfully completed the course and the post-class survey, a digital Record of Completion will be emailed to you and the Certificate of Achievement will be mailed in approximately four weeks.

If you are interested in calculating your progress along the way, it may be helpful to know that there is a total of 300 base points in the course (meaning 210 to achieve 70%)

Deliverable	Regular Points
Colab 1	30
Colab 2	45
Colab 3	90
Colab 4	75
Colab 5	60
Total Available	**300**
Minimum Passing Total	**210 (70% of Regular Total)**

Honor Code

Students will be asked to review and maintain the standards set forth by the Stanford Honor Code when completing quizzes and assignments in this course. You can review the section labeled Violations of the Honor Code for representative examples relevant to this course.

Students are strongly encouraged to form study groups, discuss, and work on homework problems in groups and help each other; However, each student must write down the solutions independently and cannot refer to written notes from the joint session. **In other words, you must understand the solution well enough in order to reconstruct it independently**. Further, because we occasionally reuse problem set questions from previous years, you are expected not to copy, refer to, or look at the solutions when preparing your answers. It is an honor code violation to intentionally refer to previous year's solutions.

After completing this course, you are welcome to share your experience and credential with others; However, it is considered a violation of the honor code to share assignment solutions including on public platforms such as GitHub. Faculty in the computer science department have strongly encouraged us to refrain from posting solutions for assignments, thus we ask that you **DO NOT** share the exact code.

Course Facilitators, Support, and Guidelines

You have a wide range of support available to you throughout the course. You will be assigned and receive contact information for an individual Course Faciltiator (CF) who will act as your primary point of contact.

Below is a summary of the available resources and course support:

Office Hours

Your CF will be in touch with availability and scheduling logistics for video conference office hours. Office hours may be conducted using the Zoom conference service or via Slack video (more information below on the course Slack workspace). Your CF will provide further information on how they will schedule and run office hour sessions.

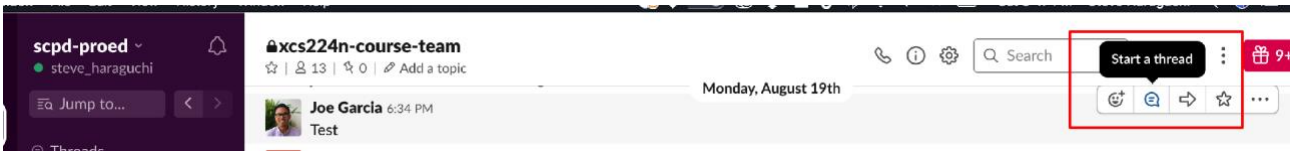
Email

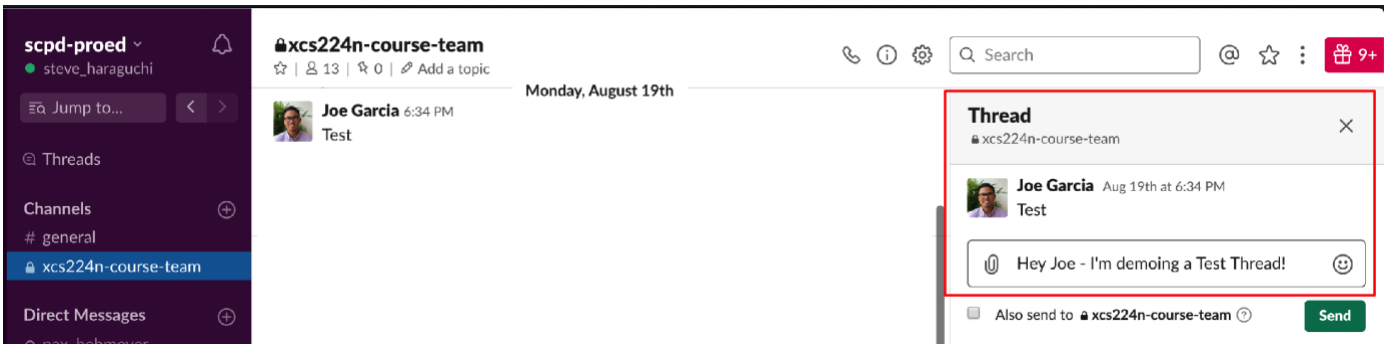
Your CF will also be available to answer questions via email – a Stanford contact address will be provided when you are first connected to your CF.

Slack Workspace – Usage and Guidelines

In addition to the individual and small group support provided by CFs, Slack will be a place where questions may be posed to the entire community (this is the fastest way to get an answer!). **In order to keep the Slack workspace readable, searchable, and useful to all, please follow the following guidelines:**

Reply in Threads to Keep Conversation Organized – When you are replying to a post or joining a conversation, respond by starting or joining a threaded conversation, rather than responding in the full flow of the standard timeline. See below for an example of how to respond in a threaded conversation to Joe's test message:





Use Multi-line Messages – Even if messages are threaded, you will soon see that Slack becomes unmanageable unless people use **single, multi-line messages instead of multiple, single-line messages**. Especially for mobile Slack users, it gets out of control!

Rather than the following:
"Hey all I have a question" [RETURN] <-- Creates new message
"I am a little confused about the quiz" [RETURN] <-- Creates new message
"I'm getting F for Question 40, but it seems like T is better" [RETURN] <-- Creates new message
---
Instead, try this!
"Hey all I have a question" [SHIFT+RETURN] <-- Creates new line in SAME message
"I am a little confused about the quiz" [SHIFT+RETURN] <-- Creates new line in SAME message
"I'm getting F for Question 40, but it seems like T is better" [SHIFT+RETURN] <-- Creates new line in SAME message
{RETURN} <-- Posts message

Note on Code Assignments and Debugging

While the course team is here to help and support your experience, it is ultimately your responsibility to write, test, and de-bug your own code. CFs may view and provide guidance on your work; however they will not send you exact answers on what to insert into your assignments. Additionally, before reaching out to a CF or Slack for help, it's expected that you have taken the reasonable steps of reading and performing an analysis yourself. This policy is meant to ensure that you leave the course having mastered the material and enable CFs to focus attention on questions where their guidance is most impactful.

Drop/Transfer Policy

You may drop this course for a full refund up until January 31, 2022 – the first day of the course. Once the course has begun, if you request to drop the course by Friday at 5:00pm PST on the third week of the cohort (**February 18, 2022**) you will be reimbursed 100% of your tuition minus a drop fee of $100. Beyond the third week of the course, tuition refunds are not granted. Up until **February 18, 2022,** you may also request to transfer your enrollment to a future cohort of XCS224W or another course in the AI Professional Program, also for a transfer fee of $100. To drop or transfer the course, send an email to xcs224w-staff@stanford.edu

Questions and Contacts

For course-specific questions or concerns (content, assignments, CF support), please contact your designated Course Facilitator.

For other course related questions, email xcs224w-staff@stanford.edu