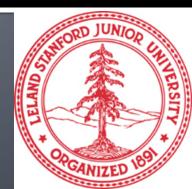
# Stanford CS224W: Machine Learning with Graphs

CS224W: Machine Learning with Graphs Jure Leskovec, Stanford University http://cs224w.stanford.edu

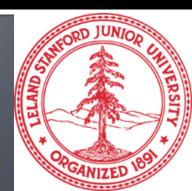


## CS224W Course Logistics

- The class meets Tue and Thu 10:30-11:50am Pacific Time on Zoom.
  - Videos of the lectures will be recorded and posted on Canvas.
- Structure of lectures:
  - 60 minutes of a prerecorded lecture.
  - During this time we will be using Piazza Live Q&A
  - 20 minutes of a live Q&A session

# Stanford CS224W: Course Logistics

CS224W: Machine Learning with Graphs Jure Leskovec, Stanford University http://cs224w.stanford.edu



## Logistics: Teaching Staff

#### Instructor



Jure Leskovec

#### Advisor



Michele Catasta

#### **Teaching Assistants**



Jiaxuan You Head TA



**Div Garg** 



Jonathan Gomes-Selman



Weihua Hu



Natasha Sharp Course Coordinator



Jingjing Tian



Zhitao (Rex) Ying



**Zecheng Zhang** 

## **Course Outline**

Date	Topic	Date	Topic
Tue, Jan 12	Introduction; Machine Learning for Graphs	Tue, Feb 16	Reasoning over Knowledge Graphs
Thu, Jan 14	Traditional Methods for ML on Graphs	Thu, Feb 18	Frequent Subgraph Mining with GNNs
Tue, Jan 19	Node Embeddings	Tue, Feb 23	Community Structure in Networks
Thu, Jan 21	Link Analysis: PageRank	Thu, Feb 25	Traditional Generative Models for Graphs
Tue, Jan 26	Label Propagation for Node Classification	Tue, Mar 2	Deep Generative Models for Graphs
Thu, Jan 28	Graph Neural Networks 1: GNN Model	Thu, Mar 4	Scaling Up GNNs
Tue, Feb 2	Graph Neural Networks 2: Design Space	Tue, Mar 9	Learning on Dynamic Graphs
Thu, Feb 4	Applications of Graph Neural Networks	Thu, Mar 11	GNNs for Computational Biology
Tue, Feb 9	Theory of Graph Neural Networks	Tue, Mar 16	GNNs for Science
Thu, Feb 11	Knowledge Graph Embeddings	Thu, Mar 18	Industrial Applications of GNNs

#### Logistics: Website

- http://cs224w.stanford.edu
  - Slides posted before the class
- Readings:
  - Graph Representation Learning Book by Will Hamilton
  - Research papers
- Optional readings:
  - Papers and pointers to additional literature
  - This will be very useful for course projects

#### Logistics: Communication

- Piazza Q&A website:
  - http://piazza.com/stanford/win2021/cs224w
    - Register with your @stanford.edu email
  - Please participate and help each other!
    - Don't post code, annotate your questions, search for answers before you ask
    - Given COVID/virtual class, this will be the main mode of communication
- To reach course staff (prof/TAs), <u>always</u> use:
  - cs224w-win2021-staff@lists.stanford.edu
- We will post course announcements to Piazza (make sure you check it regularly)

#### Work for the Course & Grading

- Final grade will be composed of:
  - Homework: 30%
    - Homework 1, 2, 3, each worth 10%
  - Coding assignment: 30%
    - 5 Coding assignments using Google Colab, each worth 6%
  - Course project: 40%
    - Proposal: 30%
    - Final report: 70%
  - Extra credit: Piazza participation, code contribution
    - Used if you are on the boundary between grades

#### Homework, Write-ups

- Assignments are long and take time (~10h) Start early!
  - A combination of data analysis, algorithm design, and math
  - Generally due on Thursdays 23:59 Pacific Time
- How to submit?
  - **Upload via Gradescope (http://gradescope.com)** 
    - You will be automatically registered to Gradescope once you officially enroll in CS224W
    - Each answer must start on a new page. Read carefully the course info page!
  - Both homework (including code) and project deliverables must be uploaded to Gradescope!
- Total of 2 Late Periods (LP) per student:
  - Max 1 late period per assignment (no LP for final report)

#### **Honor Code**

- We strictly enforce the <u>Stanford Honor Cøde</u>
  - Violations of the Honor Code include:
    - Copying or allowing another to copy from one's own paper
    - Unpermitted collaboration

    - Giving or receiving unpermitted aid on a take-home examination
    - Representing as one's own work the work of another
    - Giving or receiving aid on an assignment under circumstances in which a reasonable person should have known that such aid was not permitted
  - The standard sanction for a first offense includes a onequarter suspension and 40 hours of community service.

#### **Course Projects**

- Course project:
  - Make predictions on a network dataset
- Performed in groups of up to 3 students:
  - Fine to have groups of 1 or 2. The team size will be taken under consideration when evaluating the scope of the project in breadth and depth. But 3 person teams can be more efficient.
  - Project is the important work for the class
  - Teaching staff will help with problems and data
  - More details to follow.
- Read: <a href="http://cs224w.stanford.edu/info.html">http://cs224w.stanford.edu/info.html</a>

## Course Schedule

Week	Assignment	Due on (11:59pm PT)
1	Colab 0	
1	Colab 1	Thu, Jan 28
2	Homework 1	Thu, Feb 4
3	Colab 2	Thu, Feb 11
	<b>Project Proposal</b>	Thu, Feb 11
4	Homework 2	Thu, Feb 18
5	Colab 3	Thu, Feb 25
6	Homework 3	Thu, Mar 4
7	Colab 4	Thu, Mar 11
8	Colab 5	Thu, Mar 18
	Project Report	Sun, Mar 21 (No Late Periods!)

#### Prerequisites

- The course is self-contained.
- No single topic is too hard by itself.
- But we will cover and touch upon many topics and this is what makes the course hard.
  - Good background in:
    - Machine Learning
    - Algorithms and graph theory
    - Probability and statistics
  - Programming:
    - You should be able to write non-trivial programs (in Python)

### **Graph Machine Learning Tools**

- We use <u>PyTorch Geometric</u> (PyG)
- We further recommend:
  - DeepSNAP: Library that assists deep learning on graphs.
    - Flexible graph manipulation, standard data split pipeline, ...
  - GraphGym: Platform for designing Graph Neural Networks.
    - Modularized GNN implementation, simple hyperparameter tuning, flexible user customization
  - Both platforms are very helpful for the course project (save your time & provide advanced GNN functionalities)
- Other network analytics tools: SNAP.PY, NetworkX