Spring 2021

PHYS 7332: Network Data Science 2

Machine Learning with Graphs

Objective of the course is to …

familiarize the students with state-of-the-art descriptive and predictive algorithms on graphs.

This course covers a number of advanced topics in machine learning and data mining on graphs

- Vertex classification
- Graph clustering
- Link prediction and analysis
- Graph distance functions
- Graph embedding and representation learning
- Deep learning for graphs
- Anomaly detection
- Graph summarization
- Network inference
- Adversarial learning on networks
- Notions of fairness in social networks

Course work involves …

- Class presentations
- A semester-long project
  - You will be asked to replicate the findings of a recent paper (< 3 years old) on graph mining that was published in a top conference such as ACM KDD, SIAM SDM, IEEE ICDM, WWW (a.k.a. The Web Conference), WSDM, ICML, ICLR, NeurIPS, AAAI, and IJCAI.

Prerequisites

- Students are expected to have
  - a basic knowledge of calculus and linear algebra,
  - be familiar with probability theory and statistics, and
  - have good programming skills (e.g., Python, Julia, JAVA, C, Matlab, R, or any programming language of their preference).
- Basic knowledge of machine learning or data mining is helpful.

If you don't have the prerequisites, I suggest the following options

Option 1: Coursera Courses

- David Dye's on Mathematics for Machine Learning (audit for free)
- Karl Schmedders's An Intuitive Introduction to Probability (audit for free)
- Andrew Ng's on Machine Learning (free)

If you only have time to take one course, then take the Mathematics for Machine Learning course.

Option 2: Khan Academy Courses (free)

- Linear Algebra
- Calculus 1
- Statistics and Probability

Side Note: If you want to learn linear algebra from a world-renowned professor, I recommend taking the OCW MIT course on linear algebra with Gilbert Strang. The course is free and uses Matlab for its homework assignments. You can download Matlab for free since Northeastern has a site license for it.

Other resources

- Linear Algebra Review and Reference (Zico Kolter and Chuong Do, Stanford)
- Probability, Linear Algebra, and Differentiation (Iain Murray)
- Linear Algebra Tutorial (C.T. Abdallah, Penn)
- Probability Review (David Blei, Princeton)
- Probability Theory Review (Arian Maleki and Tom Do, Stanford)
- Theoretical CS Cheat Sheet (Princeton)
- Encyclopedia of Distances (Michel Marie Deza and Elena Deza)
- Mathworks Matlab Tutorials
- Ben Taskar's Matlab Tutorial