Special Topics in Computing Science

CMPT 983

Graph Representation Learning

Simon Fraser University

Fall 2021

Instructor: Oliver Schulte

For course details such as scheduling, contact information, office hours etc., please see the course website at courses.cs.sfu.ca

Overview

Graph data represent relationships between entities in a domain. They are a common data type, which makes them important for many applications. Domains with major graph datasets include the following: enterprise data management through relational databases, social networks, bioinformatics (e.g. protein-protein interactions), information extraction in natural language processing, where knowledge graphs represent a large amount of information that can be extended through on-line sources. While graph data are powerful and widely available, they are a challenge for

standard machine learning methods that are designed for independent data points. The goal of this course is to introduce students to the special challenges of learning from graph data, and to the machine learning methods that have been developed to address them.

This course is an advanced seminar course. Readings will include current research.

Objectives

- Learn the main theoretical ideas behind graph representation learning.
- Ability to read research papers in this area.
- Become familiar with a set of computational tools for machine learning with graph data.
- The course project should evaluate one or graph learning tools on a realistic task.

The initial lectures will provide introduction to the necessary background. *The course is a seminar course, meaning that students are expected to give presentations.*

Prerequisites

- Having passed CMPT 726 or equivalent. Consult with me on what is “equivalent”.

Topics

- Types of Graph Data: Homogeneous, Heterogeneous, Multi-Relational
- Traditional Methods: Node Features, Graph Kernels, Spectral Analysis, Exponential Random Graph Model
- Node Embeddings
- Graph Neural Networks
- Graph Generative Models
- Advanced Topics from current research (methodology and applications)

Textbook

Graph Representation Learning, William Hamilton, 2020, Morgan and Claypool. A preprint is available https://www.cs.mcgill.ca/~wlh/grl_book/ .

Grading and Expectations

About half the course is devoted to student presentations and course discussions. Presentations and projects can be done individually or in groups. All group members receive the same grade. The grading scheme is as follows.

- Exercises/Quizzes: 10%.
- Textbook/Topic Presentation: 30%.
- Project Presentation 15%
- Final Project 45%

Accommodations. If a student misses a graded course component with an acceptable excuse, I will give them a final exam and move the weight of the missed class component to the exam.

Exercises and Quizzes

It is important for the quality of the class discussion that each student keeps up with the class readings. To give you feedback on your understanding, I will give regular quizzes on-line and conduct in-class surveys using mobile Canvas; **you must be equipped to take part in these quizzes.** I may also give small homework exercises if this helps to fill gaps or reinforce important points. *I will discard the worst 2 quizzes for each student*, to allow for technical difficulties, occasional absences etc.

Presentations

- A topic presentation introduces advanced or additional topics beyond what we cover in the main lectures. Examples include:
    - Parts of the textbook
    - A research/survey paper you are interested in
    - A topic from your own research
- The project presentation should outline an idea for a project. The purpose is to give you feedback early.

- The final presentation of the project will show what you created and results.

Projects

Every student should carry out a course project. I will suggest course projects, but you are free to design your own. A typical course project applies graph learning to a dataset that is of interest to the student, for example:

- A dataset from your thesis research
- A dataset from a competition (Kaggle, OpenGB)