

Spring 2023

Seminar in Data Science and Information Technology: Advanced Learning Representations

Instructor: Yang Li
TAs: Guanlin Wu and Tao Shi

Description

This is a seminar course on machine learning geared towards first or second-year graduate students who are interested in the fundamental aspects and the state-of-art developments of learning techniques. In particular, we will cover a number of representation learning topics, including feature learning for structured data, metric learning, structured prediction and techniques for learning from large-scale data.

In this course, classical literatures in different topics of representation learning will be selected and discussed every week. In addition, each student will be asked to select a topic and present their reflection about the selected literatures. The presentation will be conducted in the manner of interactive discussions, so students can share their opinions with each other. This class also contains several in-class coding exercises for certain learning algorithms.

Prerequisites

Students need to have taken a graduate-level introductory machine learning class before taking this class. e.g. Learning from Data or its equivalent. Sufficient background in math, probability and optimization for understanding research papers is also expected.

Homework assignments

The homework assignments include two parts.

1. A short summary for each paper that we discuss in class, with answers to the discussion questions;
2. Coding exercises following the topic of each class.

Grading

Your overall grade will be determined roughly as follows:

ACTIVITIES	PERCENTAGES
Class participation	20 %
Homework	60 %
Final presentation and report	20%

Class participation is scored based on the attendance and participation in seminar discus-

sions.

Reference Material

The primary reading material of this class consists of selected book chapters and papers. See the course schedule for detail.

Course Schedule Note: reading materials will be selected from the following sources.

Class	Topic	Reading & Reference Material
1	Introduction	“Representation Learning: A Review and New Perspectives” by Yoshua Bengio, Aaron Courville, and Pascal Vincent
2	Manifold Learning: Laplacian Eigenmaps	“Laplacian Eigenmaps for Dimensionality Reduction and Data Representation” by Mikhail Belkin and Partha Niyogi
3	Non-linear Dimension Reduction	“A tutorial on Variational Autoencoders” by Carl Doersch
4	Deep Generative Models	“Generative Adversarial Nets” by Ian J. Goodfellow et al.
5	Deep Generative Models	“Variational Inference with Normalizing Flows” by Danilo Jimenez Rezende and Shakir Mohamed
6	Distance Metric Learning	“A Unifying Mutual Information View of Metric Learning: Cross-Entropy vs. Pairwise Losses” by Malik Boudiaf et al.
7	Causal Inference and Stable Learning	Guest Lecture
8	Attention and Transformer	TBA
9	Deep Graph Representation Learning	“Representation Learning on Graphs: Methods and Applications” by William L. Hamilton et al.
10	Student Presentations	
11	Student Presentations	