ENEE759G: Advanced Topics in Computer Engineering - Unsupervised Learning

Fall 2022 (MW 2-3:15, AJC 2134)
Instructor: Joseph JaJa
Course Syllabus

Course Objectives: The course will cover core statistical machine learning techniques for unsupervised learning. Topics covered will include: density estimation, latent variable models, nonlinear dimensionality reduction strategies, mixture models, clustering, directed and undirected graphical models and inference, and generative deep models.

Course prerequisites: Graduate standing.

Prerequisite topics: A strong undergraduate background in probability and statistics, linear Algebra, advanced calculus, algorithms, and nonlinear optimization is required for this course.

Textbooks: No textbook is required for this course but reference materials will be posted on the ELMS web site of the class. The following textbooks can also serve as references.

T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, second edition, Springer, 2009.

K. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.

S. Theodoridis and K. Koutroumbas, Patter Recognition, fourth edition, Elsevier, 2009.

Core Topics:

1. **Introduction and Background**
    - Basic framework and concepts
    - Nonparametric Probability density estimation
    - Maximum Likelihood Estimation and Bayesian Estimation.
    - Matrix Factorization (including SVD) and Nonlinear Optimization
2. **Latent Variable Models**
    - Principal Component Analysis
    - Introduction to Factor Analysis
    - Independent Component Analysis
    - Nonnegative Matrix Factorization
    - The EM algorithm with application to Gaussian Mixture Models
3. **Nonlinear Dimensionality Reduction**
    - Kernel PCA
    - Laplacian Eigenmaps
    - Local Linear Embedding and Isometric Mapping

4. **Clustering**
    - Proximity measures and evaluation methodologies
    - The k-means algorithm and its variant the k-medoid algorithm

- Hierarchical clustering
- Spectral clustering
- Overview of other types of clustering algorithms such as Self Organizing Maps, density based clustering, and affinity propagation.

5. **Directed Graphical Models**
    - Basic definitions and properties
    - Naïve Bayesian networks
    - Inference: Exact and Approximate
6. **Undirected Graphical Models**
    - Basic definitions and concepts
    - Markov properties
    - Factor graphs
    - Inference algorithms
7. **Deep Generative Models**
    - RBMs and deep belief networks
    - Deep auto-encoders

Course Grade: Homeworks 15%; Lecture Scribing 15%; Scholarly Paper 35%; Exam: 35%

Homeworks (15%): Assignments will be given out throughout the semester, which will provide a complementary technical background to the material covered in class.

Lecture Scribing (15%): Each student is expected to scribe two topics selected from the scribing signup sheet. Students are responsible for providing comments for all scribed notes (see ELMS Collaborations). Each scribe should be uploaded as a pdf file (with the appropriate topic name) under the ELMS Discussions environment within a week after the topic is covered in class. Grading is based on the quality of the submitted pdf files (which may be revised after considering the comments provided by other students) and feedback provided for other topics.

Scholarly Paper (35%): Each student is required to write a scholarly paper focusing on a single topic from unsupervised learning. The paper should include a detailed technical background, the state of the art of the topic, and suggestions for possible research directions. It is expected that an extensive list of references be included in the paper. **Paper outlines have to be submitted by October 26, 2022.** The outline should be around 5 pages and must describe the general topic to be covered in the scholarly paper, a brief background, and a list of at least 10 references.

Exam (35%): The exam will be held in class on **November 30, 2022**, and will be closed books, closed notes. The questions on the exam will essentially be from the lecture material covered in class.

Contact Information

Instructor:	Joseph JaJa
Office:	3214 Iribe
Office Hours:	M,W 3:30-5
Email:	josephj@umd.edu
Phone:	405-1925