CSE-772: Probabilistic Machine Learning

Instructor: Piyush Rai

1 Prerequisites

Familiarity with the basic concepts in linear algebra, statistics and probability. Having already done some coursework on these would be ideal but we will nevertheless cover some of the essentials concepts during the beginning of the course. Some prior familiarity with machine learning (e.g., via a course like CSE-771) will be helpful (but not required). Some assignments may have a programming component. The students should be comfortable with basic programming in MATLAB or Octave (other programming languages such as Python may be allowed with instructor's permission).

2 Grading

Grading will be roughly based on: 30% assignments (4 or 5 assignments; some may involve a programming component), 15% mid-sem, 15% end-sem, 40% class-project. Note: A solid/publishable piece of work from the class project may automatically lead to an A grade.

3 Course Description

Estimating the parameters of the underlying model that is assumed to have generated the data is central to any machine learning problem. Probabilistic modeling offers principled and rigorous ways to model data of diverse types, characteristics, and peculiarities, and offers algorithms to uncover the model parameters and make inferences/predictions about the data. This course will expose the students to the basic concepts/algorithms used in probabilistic modeling of data and we will gradually work our way up to use these as building blocks for solving more complex machine learning problems. We will also, at various points during this course, look at how the probabilistic modeling paradigm naturally connects to other dominant paradigm which is about treating machine learning problems as optimization problems, and understand the strengths/weaknesses of both these paradigms, and how they also complement each other in many ways.

Tentative list of topics: A rough outline of the course is given below.

- Introduction to probabilistic modeling of data

- Basic methods for parameter estimation in probabilistic models: MLE and MAP estimation

- Common probability distributions, conjugate priors and exponential family

- Introduction to Bayesian learning

- Case studies: Bayesian linear regression and classification, sparse linear models

- Latent variable models for clustering: mixture models

- Latent variable models for dimensionality reduction: factor analysis, probabilistic PCA and matrix factorization

- Latent variable models for modeling sequence and time-series data: hidden Markov models, linear dynamical systems

- Latent variable models for structured prediction

- Learning and inference in probabilistic graphical models

- Approximate Bayesian inference (MCMC, Variational Bayes, Expectation Propagation)

- Online approximate Bayesian inference

- Bayesian learning with kernels: Gaussian Processes (or "Bayesian SVM")

- Topic models

- Deep learning

4   Reference materials

There will not be any dedicated textbook for this course. In lieu of that, we will have lecture slides, online notes, tutorials, and papers for the topics that will be covered in this course. Some recommended (although not required) books are:

- Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007.

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

- Simon J. D. Prince, Computer Vision: Models, Learning, and Inference, Cambridge Univ.Press, 2012