COMP9418 — Advanced Topics In Statistical Machine Learning
Tentative Outline

Lecturer: Edwin V. Bonilla
School of Computer Science and Engineering
UNSW Sydney

May 11, 2017

1 Course Description

This course provides an in-depth study of statistical machine learning approaches. The focus will be on methods for learning and inference in structured probabilistic models, with a healthy balance of theory and practice. This course is aimed at students who are willing to be go beyond basic understanding of machine learning. The course provides fundamental support for those willing to intensify their knowledge in the area of big data analytics.

It will cover topics on exact and approximate inference in probabilistic graphical models; learning in structured latent variable models; posterior inference in non-parametric models based on Gaussian processes; and relational learning.

2 Teaching Strategies and Rationale

The course involves lectures and practical work. Lectures aim to summarize the concepts and present case studies. The tutorial / lab exercises aim to reinforce the topics covered in lectures but they are not assessed. The quizzes and assignments aim to do the same but are assessed.

3 Learning Outcomes

1. Derive statistical independence assumptions from a given graphical representation of a probabilistic model
2. Understand and implement exact inference methods in graphical models including variable elimination and the junction tree algorithm
3. Understand and implement approximate inference algorithms in graphical models including sampling and variational inference
4. Derive and implement maximum likelihood learning approaches to latent variable probabilistic models
5. Understand and apply posterior inference and hyper-parameter learning in models based on Gaussian process priors
6. Understand and apply basic methods for structured prediction and relational learning

4 Resources

4.1 Prescribed Resources

Book Machine Learning: A Probabilistic Perspective. Kevin P. Murphy. MIT Press. 2012.

4.2 Recommended Resources

Book Bayesian Reasoning and Machine Learning. David Barber. Cambridge University Press. 2012.

Book Gaussian Processes for Machine Learning. Carl Edward Rasmussen and Christopher K. I. Williams. The MIT Press. 2006.

Book Probabilistic Graphical Models: Principles and Techniques. Daphne Koller and Nir Friedman. MIT Press. 2009.

5 Assumed Knowledge

- Knowledge of machine learning at the level of COMP9417. This is a pre-requisite but, since this is the first time the course is offered, this can be waived subject to LiC's approval. However, *students are not entitled to any consideration if they discover that they do not have sufficient background.*
- Solid mathematical background including linear algebra, basic probability theory, multivariate calculus.

These courses are a good indication of the knowledge required (note that they are not official pre-requisites):

- Linear Algebra - MATH2501
- Several Variable Calculus - MATH2011
- Theory of Statistics - MATH2801

6 Assessment

10% Quizzes: There will be a number of "take-home" quizzes during the semester to act as both a mechanism for students to check their understanding of the material and as a small assessment item.

15% Assignment 1

15% Assignment 2

60% Final Exam

7 Weekly Syllabus (very tentative)

W1 Intro to probabilistic graphical models

- Probability theory refresher
- Directed graphical models (Bayesian networks)
- Undirected graphical models (Markov random fields)
- Exact inference via variable elimination

W2–W3 Inference in graphical models

- The junction tree algorithm
- Sampling
- Variational inference

W4 Learning in graphical models

- Maximum likelihood
- Learning with latent variables and the EM algorithm

W5–W6 Some basic graphical models in machine learning

- Clustering and mixture models
- Dimensionality reduction and probabilistic principal component analysis
- Latent factor models
- Independent component analysis

W7–W8 Dynamic and structured graphical models

- Hidden Markov models
- Conditional random fields

W9–W11 Gaussian processes

- From linear regression to linear prediction and beyond
- Gaussian processes for regression
- Hyper-parameter learning
- Gaussian processes for classification
- Scalability to large datasets
- Variational learning of Gaussian process models
- Gaussian processes and deep Bayesian neural networks

W12 Relational learning (Guest lecture)