# PhD-course 02901 - 2010 edition

The course consists of four days (Mon-Thursday) of morning lecture+exercises and afternoon exercises. The last day (Friday) will have exercises in the morning and student presentation in the afternoon. The course (2.5 ects point) is passed with either a presentation or a report. The running theme of course will be probabilistic multivariate modeling and Bayesian inference. The exercises covers both theoretical, technical programming and application aspects. It will be up to the students to decide on what aspects to focus on. Specific machine learning application examples are used throughout the entire week.

## Introduction to Bayesian inference

Teacher: Mikkel N. Schmidt

Bayesian inference is the process of setting up a probabilistic model, fitting it to a set of data, and drawing conclusions by summarizing the posterior distribution of quantities of interest or making predictions for new observations. After introducing the Bayesian inference framework, we will discuss inference procedures including variational approximation and Markov chain Monte Carlo simulation. As a running example we will consider a Bayesian linear regression model. The aim of the day is to gain a strong understanding of the steps involved in Bayesian inference for a simple model, to serve as a foundation for the advanced models discussed in the rest of the course.

Reading material:

Chris Bishop, Pattern Recognition and Machine Learning, Cambridge University Press 2006, section 1.1-1.4, 3.3, 10.1, 10.3, and 11.1-11.3.

## Day 2: Factor Modeling

Teacher: Ole Winther

Factor modeling is a multivariate data analysis techniques that assumes that covariation in data can be described by a set latent unobserved factors. In its simplest form it can be viewed as a low rank approximation to the covariance of the data and it thus closely related to principal component analysis. The latent factor representation can often be given an interpretation. This has made the model useful in a very diverse range of disciplines ranging from social science to machine learning. In the lecture and exercises we will see different application examples such as collaborative filtering and gene expression analysis. The lecture will cover a gentle introduction to the basic model and learning in it with maximum likelihood and Bayesian methods. Then we will go on cover advanced topics such as Gibbs sampling, sparsity modeling, relation to independent component analysis and Bayesian networks and extensions to non-parametric factor modeling.

Reading material:

Chris Bishop, Pattern Recognition and Machine Learning, Cambridge University Press 2006, chapter 2, 11-11.3 and 12.

## Day 3: Mining Graphs by Relational Modeling

Teacher: Morten Mørup

Many types of data can naturally be represented as graphs/networks formed by pairwise relationships. In social sciences networks are commonly used to represent social interactions such as friendships, acquaintance and collaboration. In biology graphs naturally emerge in the representations of protein interaction and the wiring of neural systems the so-called connectome. In economics graphs are frequently used to represent relationships between entities such as cities, countries, politicians and companies. The infinite relational model (IRM) also denoted the stochastic block model is a Bayesian model for complex networks. The model has proven powerful in order to comprehend the dynamics of the many types of networks that naturally occur in practically all fields of science. We will derive the IRM model, understand the types of structure the model is able to account for, investigate methods for inferring the parameters of the model and finally apply the model on social networks.

Reading material:

Charles Kemp, Joshua B. Tenenbaum, Thomas L. Griffiths, Takeshi Yamada and Naonori Ueda "Learning Systems of Concepts with an Infinite Relational Model", Proceedings of the 21st National Conference on Artificial Intelligence (AAAI) 2006.

## Day 4: Manifold learning and visualization with kernels

Teacher: Lars Kai Hansen

The kernel trick is a conceptually simple yet powerful mechanism for non-linear generalization of familiar linear machine learning models such as principal component analysis, linear regression, and the linear discriminant. In this part of the course we will introduce kernel based models for non-linear manifold learning, their sparse approximations, some related probabilistic schemes, and finally present the NPAIRS workflow for production and validation of model visualizations.

Reading material:

C.M. Bishop, Pattern Recognition and Machine Learning, Springer (2006). Excerpts from chapters 6, 7, and 12.

S.C. Strother, et. al., The Quantitative Evaluation of Functional Neuroimaging Experiments, The NPAIRS Data Analysis Framework, NeuroImage 15(4):747-771 (2002).

L.S. LaConte et. al., The Evaluation of Preprocessing Choices in Single-Subject BOLD fMRI Using NPAIRS Performance Metrics. NeuroImage 18:10-27, (2003).

D.J. Jacobsen, L.K. Hansen and K.H. Madsen, Bayesian model comparison in non-linear BOLD fMRI hemodynamics Neural Computation 20:738-755 (2008).

## Day 5

Further exercise work in the morning and student exercises in the afternoon.