T-61.5140 Machine Learning: Advanced Probabilistic Methods
Hollmén, Ilin
Examination, 7th of January, 2011 from 9 to 12 o'clock.

In order to pass the course and earn 5 ECTS credit points, you must also pass the term project. Results of this examination are valid for one year after the examination date. Information for Finnish speakers: Voit vastata kysymyksiin myös suomeksi, kysymykset on ainoastaan englannin kielellä. Information for Swedish speakers: Du får också svara på svenska, frågorna finns dock endast på engelska.

1. How can you generate samples from a Laplace distribution. Write a sketch of a program generating the samples. You have access to a computer that is able to generate samples from the uniform distribution. Hint: The probability density function for an Laplacian random variable is here: $p(x \mid \mu, b) = \frac{1}{2b} exp^{(-\frac{x-\mu}{b})}$.

2. Write the algorithm for Gibbs sampling and write the distributions to sample from in the case of $p(x_1, x_2, x_3, x_4)$.

3. For the Bayesian network that decomposes the joint probability as in $p(x_1, \ldots, x_5) = p(x_1)p(x_2|x_1)p(x_3|x_1)p(x_4|x_2, x_3)p(x_5|x_4)$, draw the corresponding graphical representation. Assuming all the variables have discrete values $x_i \in \{0, 1, 2, 3\}$, give the sizes of the tables representing the probabilities for the conditional probability distributions. Moreover, derive the junction tree representation (and name the steps). Draw the resulting junction tree.

4. Write the probability $p(\mathbf{x})$ for the finite mixture model of exponential distributions, name the parts of the mixture model, and derive the E-step and the M-step of the Expectation-Maximization (EM) algorithm. Hint: The probability for an exponentially distributed random variable can be calculated with the following equation: $p(x \mid \lambda) = \lambda e^{-\lambda x}, x \geq 0$.

5. Swine influenza virus H1N1 is a cause of a respiratory disease and a public health risk. Currently, it is thought to be a risk for a pandemic, that is, a global outburst of the disease. Christopher W. Olsen states in his article "INFLUENZA: Pigs, people and Public Health" that the clinical signs or symptoms of influenza in pigs and people are remarkably similar, with fever, lethargy (=weakness, lack of energy), lack of appetite and coughing in both species. Additionally, assume the existence of a diagnostic test T1 that can detect the H1N1 virus (or the disease) with a probability 0.6 and gives a false positive test result with a probability 0.1. Furthermore, you have a diagnostic test T2 that detects the virus with probability 0.99 and gives a false positive test results with a probability 0.03. Also, you may use the knowledge that a visit to areas where the swine influenza is common increases the risk of the disease.

Build a Bayesian network for diagnosis of the swine influenza with the above information. Present the random variables in the model, their dependencies graphically, give a joint probability distribution for the model, give the probability tables (with probability values as you perceive them). Explain the use of the model in a diagnostic setting. Perform the diagnostic inference (using your favorite method) for one healthy individual and one flu patient once the test result T1 is available, and once the test results T1 and T2 are available. Comment the information value of the diagnostic tests T1 and T2 based on the posterior probabilities of the healthy individual and the sick patient.