

Summer School on
“Modeling and Decision Making Using Bayesian Statistics”
4th – 8th June, 2012, Aalto University, Department of Applied
Mechanics, Marine Technology in Espoo, Finland

Take-home assignment

Assignment 1: Building Bayesian networks

Recall the continuous Bayesian network model that you have developed in Exercise 2, session “04 Exercises: Bayesian Networks.”: Extend this model to at least 10 variables. Specify discretization for each of the nodes and use it as a Bayesian network (i.e., conduct diagnostic reasoning). Describe three different test scenarios for the model, i.e., three real or reasonably realistic cases with which you have tested the model. Report your observations.

Assignment 2: Learning Bayesian networks

The file iris.txt (attached) contains a set of data from the Machine Learning Repository at the University of California, Irvine (<http://archive.ics.uci.edu/ml/>), describing properties of three classes of iris flowers. It is possible to learn a Bayesian network that will classify the flowers based on their properties. Create a collection of Bayesian networks for the purpose of this classification and verify their classification accuracies. What classification accuracy can you get? In order to standardize the results among all students in the class, please use the “leave-one-out” cross-validation approach in testing your models.

Hint: The class variable is discrete but the feature variables are continuous. In order to be able to use any of the **GeNIe** algorithms, you will need to discretize the feature variables.

Please send a report of the assignment, including your GeNIe model(s) to marek@sis.pitt.edu by Friday, 15 June 2012. You can work in groups of two. Please send just one report with the names and Email addresses of both team members.

In case of any problems with or any questions about the software, please get in touch with my team through the **GeNIe** and **SMILE**® Forum (<http://genie.sis.pitt.edu/forum/>).