Ciências ULisboa | Informática

Aprendizagem Automática Avançada
(2º Ciclo de Informática)

Bayesian Learning Exercises

Problem 1 *

The following problem is adapted from `https://www.ida.liu.se/ext/caisor/TDDC65/dectree-exercise/page-100930.html`.

You may use pomegranate's (`https://pomegranate.readthedocs.io/en/latest/`) implementation of Conditional Probability Tables and Bayesian Networks.

Pomegranate has a tutorial on Bayesian Networks available `https://pomegranate.readthedocs.io/en/latest/BayesianNetwork.html`.

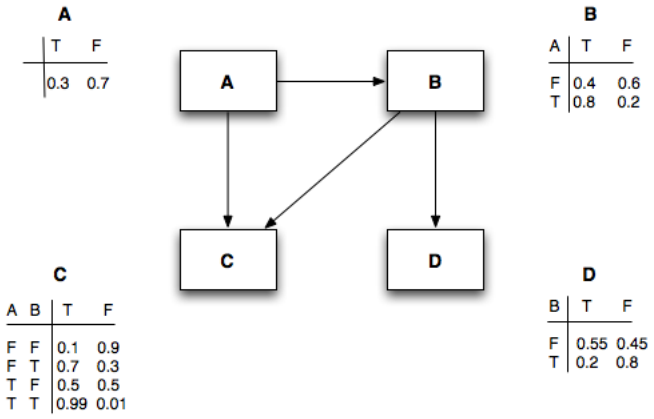

Figure 1: Bayesian Network / Causal Net. Variables A, B, C and D are represented as nodes, and a truth table is given alongside them.

Given the bayesian network provided in Figure 1, calculate the following queries:

a) $P(A = T \mid C = T, D = T)$

b) $P(A = T \mid D = F)$

c) $P(B = T \mid C = T)$

d) $P(B = T \mid A = T, C = T)$

e) $P(C = T \mid A = F, B = F, D = F)$

Problem 2 *

For problem 2, compare Naïve Bayes and Random Forest results in the cases indicated below. You may use scikit-learn for both.

You may use a dataset you are interested in exploring or any other dataset we have explored in class (e.g., iris, breast cancer) to solve the following exercises.

a) Use Naïve Bayes to classify your dataset, exploring different parameter values.

b) Use Random Forest to classify your dataset, exploring different parameter values.