Practical 2: Bayesian Networks

CS5010 AI Principles

Due date: Thursday 18th November 2021
30% of the total grade

Aims

The aims of this practical are two-fold:

- gain experience in applying Bayesian networks to solve real world problems;
- understand the advantages but also limitations of Bayesian networks.

Part 1: Application of Bayesian Networks

You are given the following description on COVID-19 and its available tests.

> *There are two COVID tests available in the UK: at-home rapid antigen test and PCR test. The first one can be taken at home and it takes around 30 mins to get a result while the PCR test needs to be carried out at a lab. Researches have found that the rapid antigen test's accuracy is affected by whether the subject has COVID symptoms or not. If the subject has COVID symptoms, the sensitivity and specificity are* 85% *and* 98% *respectively. However, without symptoms, the sensitivity and specificity drop to* 59% *and* 96%*.*
>
> *PCR test on the other hand is considered as a gold-standard test. It is sometimes prescribed as a second confirmation test. PCR test accuracy is affected by how the sample is collected. If the sample is collected by trained clinical staff, the sensitivity and specificity are* 99.2% *and* 99.98%*. On the other hand, the accuracies for self-administrated samples drop to* 95% *and* 99% *respectively.*
>
> *According to the latest statistics, the national COVID rate currently in the UK is roughly* 2% *(1.33 million active COVID case divided by the total population 68 million). It is estimated about one third of COVID cases are asymptomatic and about 9 in 10 PCR testing samples are collected by testee themselves.*

Read the description carefully and design a Bayesian network that is suitable to answer the following queries:

q1: Testee A has been tested positive using an at-home COVID rapid antigen test. He/She feels pretty normal and bears no obvious COVID symptom. How likely is A infected with COVID?

q2: Testee B, who has no connection to A, has been tested positive using an at-home COVID rapid antigen test. How likely is B infected with COVID?

q3: Should they take a second confirming PCR lab test and why ? If so, should they take the sample by themselves or by trained staff ?

Specifically, you should:

1. Present your network in a graphical representation and state all the assumptions you have made.
    - briefly describe the steps you have followed to reach the final Bayesian Network;
    - how many parameters are required for your Bayesian network ?
2. Answer the above three queries by making probabilistic inferences with your Bayesian network. You should provide intermediate steps for the inference results.

Part 2: Assessing the potential role of Bayesian Networks (AI) in judicial system

We have discussed the Sally Clark case during lectures. Sally Clark became the victim of a miscarriage of justice when she was found guilty of murdering her two infant children. A paediatrician professor applied statistics wrongly and concluded she was guilty. The conviction was largely based on the flawed inference. We have also showed in lectures that, by using a correct Bayesian Network, such a miscarriage of justice can be easily avoided.

In light of this, one may argue: **human experts should be replaced with AI, namely Bayesian Networks, when making probabilistic inferences at a judicial setting.**

Write a short paragraph (less than 1500 words; and format of bulleted lists of statements is fine) on the above argument. You should cover the following aspects and focus on the technical aspects of Bayesian Networks.

- What are the pros and cons of AI versus human in uncertainty reasoning ?
- Are current state-of-the art techniques in Bayesian Networks mature enough to fully replace humans? If not, what are the technical difficulties ?

Submission

Hand in via MMS: a report in the PDF format.

- the report must contain sections which correspond to the two parts described in this specification
- and it must address the questions associated with each part.

For Part 1, if you have used any software package to help you construct the network, you may include the code together with the PDF report in a compressed zip file. This is optional.

Marking and Extensions

This practical will be marked according to the guidelines at
https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/feedback.html
It will be based on the quality of your answers to both of the questions. The answers should be **concise**, **easy to follow**, and **straight to the point**. Good submissions should demonstrate good and **in-depth** understanding of the topic. The 1500 word limit for part two is a rough guide. You should definitely **NOT** aim to use all the 1500 word limit.

Also note that:

- Standard lateness penalties apply as outlined in the student handbook at https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/assessment.html
- You must reference any external sources used. Guidelines for good academic practice are outlined in the student handbook at https://info.cs.st-andrews.ac.uk/student-handbook/academic/gap.html

Lei Fang
Oct 2021