CSC4025Z (2023): Artificial Intelligence

Assignment 1: Application of Bayesian Networks

Lecturer: Dr. Jan Buys (jbuys@cs.uct.ac.za)
TA: Sello Ralethe (RLTSEL002@myuct.ac.za)

Deadline: Friday 5 May 2023 at 22:00.

You should work in groups of **2 to 3** students. For any questions related to the project please e-mail **BOTH** the lecturer and the TA. If you need any clarifications please contact us as early as possible.

The goal of the project is to develop a Bayesian Network-based AI system and to evaluate its performance. The system should be implemented in Python using pyAgrum. First complete tutorial on Bayesian networks to become familiar with the software.

You may select any scenario for decision making. For example, selling or buying some share on the stock market; making some decision based on the predicted weather; determining whether to perform a medical intervention for some condition; or modelling the risk of someone getting Covid-19. The model must be useful not just to yourself but to some wider target user community.

As part of the problem analysis, you should identify and incorporate knowledge from different sources. You should incorporate common knowledge (knowledge that is commonly accepted, but may not be written down anywhere) as well as formal (from published scientific sources) and/or expert knowledge (e.g. consultation with suitable medical experts). Bear in mind that the knowledge that you choose to incorporate in your model has a major influence on the potential usefulness and usability of the tool.

The network you design should have at a minimum five chance nodes. First design a Bayesian network with random variables (observed and unobserved). Then extend it to be a decision network with action and utility nodes.

Your submission should include the following deliverables (submit as a single zip file):
- **Project report**: The suggested length is 5 pages (this is a guideline, not a hard constraint). See the information about the structure below.
- **Code and data**: Submit the code for the project, as well as the Bayesian network model saved in BIF or BIFXML format. Also include any datasets used (if you use a training set >10MB only include a sample of the training data).

There will be a late penalty of 5% a day. The group leader should submit the assignment on Vula. Include the students numbers of all group members in the file name.

Use the following structure as a guide for writing your report:

1. Introduction
   - Background and objective: The objective should be phrased like: "The tool/network will be used to determine whether to ...[specify action/decision options] to bring about [some change] in [some condition]"
   - Potential user community

2. Problem analysis
   - Describe the potential factors to considered
   - Describe the data sets and expert knowledge used (support by references)
3. Decision Network model
   - Provide a diagram of your final model (you can generate a diagram with pyAgrum)
   - Explain how and why you arrived at this particular structure and weights
   - For complex models comprehensive probability tables are not necessary for the report
4. Model testing and evaluation
   - First show how the Bayesian network can be used to make inferences about events represented by unobserved random variables.
   - Then show how it can be applied as a decision network to achieve the objective defined in section 1
   - Show the efficacy and usefulness of your model in one or more use case scenarios
5. Conclusion
6. References

Marking rubric

Criteria	Mark
Problem formulation: Novelty, difficulty, and usefulness of problem	10
Model design: Knowledge acquisition for model development	10
Model design: Complexity and validity of the network	5
Evaluation: Appropriate test setup and evaluation	5
Analysis of model application for inference and decision making	10
Software: Code correctly implements model, appropriate choice of tools	5
Overall report quality	5
Total	50

Based on an assignment from Deshen Moodley.