University of Lincoln
School of Computer Science
Assessment Briefing 2023-2024

The use of AI tools to generate all or part of your assessment submission is **not** permitted unless specifically mentioned below.

Module Code and Title: CMP9794M Advanced Artificial Intelligence

Contribution to Final Module Mark:
50%

Description of Assessment Task and Purpose:

This individual assessment, referred to as **item 1**, is an assignment. Your task is to use materials covered during the lectures and workshops of this module to implement a software application to solve Medical Diagnosis problems by performing probabilistic reasoning from data. The sources of data and random variables of each of these datasets are:

- **Cardiovascular Disease**: age, hight, weight, gender, systolic blood pressure (ap_hi), diastolic blood pressure (ap_lo), cholesterol, glucose, smoking, alcohol intake, physical activity, target.
- **Diabetes**: pregnancies, glucose, blood pressure, skin thickness, insulin, bmi, diabetes pedigree function, age, outcome.

Whilst you are encouraged to use both datasets, you should use at least one dataset and justify your choice. But DO NOT use the original sources of data. The data will be pre-processed by the delivery team and will be made available via Blackboard in week 2 or 3.

Task 1: Bayesian Networks (50%)

This task consists of implementing *Bayesian Networks* to answer probabilistic queries such as:

```
P(target=0|height=164, weight=70, systolic_blood_pressure=130,
           diastolic_blood_pressure=90, glucose=1, smoke=0, alcohol=0)
P(target=1|height=164, weight=64, systolic_blood_pressure=180,
           diastolic_blood_pressure=90, glucose=1, smoke=1, alcohol=0)
P(outcome=0|glucose=109, bmi=25.4, age=25)
P(outcome=1|glucose=183, bmi=23.3, age=58)
```

Your choice of solution, to be implemented in Python, should make use of the appropriate feature values, which can be discrete or continuous. For the former, code for data discretisation will be provided to you. Your solution for this task can elaborate on (or compare between) discrete or Gaussian Bayesian Networks, which should aim for the following:

1. Read the datasets (in CSV format) available via Blackboard.
2. Read a structure of each of your Bayesian networks from a configuration file, which can be a predefined structure or a learnt one. Whilst baseline structures will be provided by the delivery team, you are encouraged to implement learnt structures using one of the provided algorithms during the lectures.
3. Learn the parameters of the Bayesian networks (one network per dataset) using Maximum Likelihood Estimation for example.
4. Answer probabilistic queries using one of the algorithms provided in the module using either exact or approximate inference (or both if you wish to do so).
5. Quantify the performance of your nets using training & test sets and relevant metrics.

Whilst the algorithms of choice can be those discussed/provided during the lectures and workshops, at least one of those algorithms should be an implementation programmed by yourself—not provided as part of the module and not part of a publicly available library. Please justify your choices of algorithms, metrics and/or libraries. You should compare the performance of these algorithms in terms of predictive power (disease classification accuracy), AUC (area under the curve) score, statistical distances (e.g., Kullback-Liebler Divergence, Brier score), training and test times (in seconds), among others.

Task 2: Gaussian Processes (50%)

This second task asks you to compare the performance of your solution in Task 1 against *Gaussian Processes* (GPs), which will be discussed during the module for classifying feature sets to make predictions. Whilst you are encouraged to implement GPs as much as possible and/or to combine them with Bayesian Networks (if you see an opportunity to do that), you are allowed to use publicly libraries if you wish to use them. In any case, your comparison should make use of metrics consistently in both tasks to report and discuss your results.

In this assessment you are allowed to make use of publicly available resources including libraries or chatbots such as ChatGPT. But note that your report should be written by yourself (even if parts of your solutions are derived from a chatbot). In other words, resources like ChatGPT should be used to increase your understanding instead of writing the assignment for you. So, please indicate and justify in your report the methods used for solving the tasks above and read the Criterion Reference Grid for details on how your work will be graded.

Learning Outcomes Assessed:

- [LO1] Critically appraise a range of AI techniques for knowledge representation, reasoning and decision-making under uncertainty, identifying their strengths and weaknesses, and selecting appropriate methods to serve particular roles.
- [LO2] Design and develop a software algorithm for solving complex AI problems in an application domain of interest.

Knowledge & Skills Assessed:

- *Subject Specific Knowledge, Skills and Understanding*: e.g., literature searching, referencing, project planning, techniques and skills subject-specific knowledge.
- *Professional Graduate Skills*: e.g., independence and personal responsibility, adaptability, verbal communication, written communication, creativity, critical thinking, IT skills, problem solving, effective time management, working under pressure to meet deadlines.
- *Emotional intelligence*: e.g., self-awareness, self-management, motivation, resilience, self-confidence.

Assessment Submission Instructions:

You must make an electronic submission of your work in PDF format, NOT MS Word, by using the assessment link on Blackboard for this component. You must attend the lectures and workshops for further details, guidance and clarifications regarding these instructions.

Your submission must also include **a video of up to 3-minutes** (in MP4 format or any other compressed format) explaining and/or highlighting key aspects of your solutions. Use the tools of your choice to create your video—an example tool among others is VLC media player.

DO NOT include this briefing document with your submission.

The deadline for submission of this work is included in the School Submission dates (Hand in Dates SPREADSHEET) on Blackboard.

Date for Return of Feedback:

Please see the School assessment dates spreadsheet.

Format for Assessment:

Your submission describing the solutions to the medical problems above should include a concise report of **3 pages including references using the IEEE template provided via Blackboard**, which must be submitted as a PDF file on Blackboard. Whenever possible, you should cite previous works from the related literature to justify your arguments or choices.

The **software implemented** (source code) to solve the targeted problems should be submitted as a ZIP file in the 'assignment supporting documentation' on Blackboard. Such a ZIP file should also include an MP4 **video of up to 3-minutes** explaining your solutions.

Feedback Format:

Written and numerical feedback will be provided via Blackboard, and additional feedback can be provided upon request in a meeting or via email.

Additional Information for Completion of Assessment:

This assessment is an individually assessed component. Your work must be presented according to the Lincoln School of Computer Science guidelines for the presentation of assessed written work. Please make sure you have a clear understanding of the grading principles for this component as detailed in the accompanying Criterion Reference Grid. You are expected to take the following into account:

- Your submitted report should be a PDF file generated by one of the provided templates in MS Word or Latex via Blackboard. It should be a concise report of maximum **3 pages** in total including references.
- Submissions failing to meet the length requirements above or omitting source code or video **will not be marked and will receive a mark of zero**.
- Please make sure that you submit **your own work** (writing, results) and not somebody else's. Failure to do so will incur plagiarism or collusion, which will the reported to the School for investigation of potential academic misconduct.

If you are unsure about any aspect of this assessment component, please seek the advice with a member of the delivery team.

Assessment Support Information:

Assignment support will be provided during workshop sessions and surgery hours.

Important Information on Dishonesty, Plagiarism and AI Tools:

University of Lincoln Regulations define plagiarism as 'the passing off of another person's thoughts, ideas, writings or images as one's own...Examples of plagiarism include the unacknowledged use of another person's material whether in original or summary form. Plagiarism also includes the copying of another student's work'. Plagiarism is a serious offence and is treated by the University as a form of academic dishonesty.

Please note, if you use AI tools in the production of assessment work **where it is not permitted**, then it will be classed as an academic offence and treated by the University as a form of academic dishonesty.

Students are directed to the University Regulations for details of the procedures and penalties involved. For further information, see www.plagiarism.org