COMP7015 Artificial Intelligence **Semester 1, 2024–25**

Course Project Instructions

Instructor: Dr. Kejing Yin Oct. 7, 2024

1 Overall Requirements

1. You should form a group of **one to four** people to carry out the project.
2. Milestones (important dates):
   (a) **Oct. 22, 2024 (Tuesday)**: Submit a project proposal if you choose Topic 3 (open topic). This is not needed if you choose Topic 1 or 2.
   (b) **Oct. 25, 2024 (Friday)**: Deadline of group information registration (names and student IDs of your group member) and topic selection. You will be notified your group number afterwards. Please register as early as you can to allow more time to work on the projects.
   Link: `https://hkbu.questionpro.com/comp7015-2425-reg`. *(Only one registration is needed for each group.)*
   (c) **Nov. 22, 2024 (Friday)**: Submission of the **source codes**, a **project report**, and the prediction for test set (for option 1 only).
   (d) **Nov. 29 – Nov. 30, 2024**: **In-person presentation**. Detailed arrangements will be announced after the registration deadline.
3. The final submissions should include the followings:
   (a) The project report should be a pdf file with at most five A4 pages (single column), describing the work your group has done and the results you observe. At the end of the report, you must include a section to faithfully report the contribution of each group member.
   (b) The source codes should be compressed into a ".zip" file. We will evaluate your code on PCs located at FSC 8/F, so ensure that your code runs smoothly in that environment.
   (c) For option 1 only: The test set prediction should be in a ".csv" file. See details in option 1.
4. In the **In-person Presentation**, each group has around 8 minutes to do a presentation. Each group member needs to present his/her contribution. You are highly encouraged to use as much visualizations as possible (e.g., figures, tables, and graphs) to make your presentation clearer. There will be Q&A following the presentation.
5. We have zero tolerance to cheating and plagiarism. Your project report and source codes will be scanned by anti-plagiarism platforms, anyone who plagiarizes and whom were plagiarized will receive zero score and be reported to the departmental exam committee for further penalty.
6. **Resubmission of previous work**: Please be noted that you cannot submit your previous work (e.g., published papers and work submitted to other courses) or otherwise it will be considered self-plagiarism.
7. **Third-party library policy**: In option 2 (adversarial search), you need to implement the algorithms on your own. Using third-party libraries that contain the adversarial search functions are not allowed. In options 1 and 3, you can use third-party libraries but you need to explicitly acknowledge them in your report and presentation.
8. If you take a cue from other sources, *e.g.*, books, literature, and online resources, you should clearly cite them in your Final Project report. Note that it is explicitly prohibited to copy source codes from online resources or that generated by generative AI tools like ChatGPT.

2 Topics

You can choose one topic from the following:

Topic 1: ICU Mortality Prediction

Predicting mortality risk in ICU is an important application of machine learning in healthcare. With a more accurate and earlier prediction of risks, clinicians could better allocate resources and potentially improve patient outcome. In this topic, we extract a toy dataset from an open-source and real-world database. There are two files provided to you:

- `mortality_traindata.csv` contains the training data along with their corresponding labels and you can use it to train a model.
- `mortality_testdata.csv` contains a test dataset where the labels are not available.

In this project, you are required to:

1. Select 3–4 suitable machine learning models for this dataset.
2. Use appropriate regularization techniques and compare the performance before and after applying regularization.
3. Try to turn the hyperparameters in the models you selected and observe how they affect the performance.
4. Follow the machine learning workflow introduced in the lecture to select the best machine learning model and hyperparameters.
5. Generate the predictions for the test set and save the results to a csv file "`predictions.csv`". Use 1 for positive predictions and 0 for negative predictions, output one row for each sample and follow the same order as that in the testdata csv file).

We will measure the F1 score using the reserved test set labels.

Topic 2: Adversarial Search for Gomoku

In this topic, you are required to implement the adversarial search algorithms to develop an agent that plays the Gomoku game. In this game, two players will take turns to mark the spaces in a 9×9 grid with X or O. The player who succeeds in placing five of their marks in a horizontal, vertical, or diagonal row is the winner.

In this project, you are required to:

1. Implement the functions necessary for formulating this deterministic game.
2. The game developed should be able to accept input from human player (either terminal or GUI is acceptable) and take actions that maximize the agent's score.
3. Implement the minimax search algorithm to determine the next move.
4. Implement the alpha-beta pruning algorithm and analyze the improvement in terms of running time.
5. Define a heuristic evaluation function and implement the heuristic alpha-beta search (cut-off search).
6. Implement Monte Carlo tree search (MCTS) to determine the next move.

Topic 3: Open Topic

If you are not interested in the two problems introduced above, you can propose another problem/dataset to work on for your course project. If you decide to do so, there are a few more requirements:

1. Submit a project proposal (at most two A4 pages) by Oct. 22, 2024 via direct email to Dr. Kejing Yin. You should clearly describe the problem you would like to work on, the motivation of choosing the topic, your plans to carry out the project, and the datasets and algorithms to be used.
2. Depending on the scope and estimated workload, the course instructor team will approve or reject your proposal. If your proposal is rejected, you need to choose from topics 1 or 2 to proceed.
3. You do not need to limit yourself to the algorithms/models introduced in the course. Any AI-related methods can be used as long as they are suitable for the problems you proposed.
4. If needed, the course instructor and TAs will give brief feedback to your proposal to make sure that it is feasible.