Prof. Dr. Frank Deinzer
Technical University of Applied Sciences
Würzburg-Schweinfurt

Portfolio Exam 2

Reinforcement Learning – BlackJack Player

Due date: 22.01.2025, 24:00

This portfolio exam is all about a self-learning BackJack player based on Reinforcement Learning methods. The basis of Reinforcement Learning are the methods from [2]. Basic concepts of the BackJack card game, various rule variations, and card counting methods can be found in [3].

Please note that this time it is necessary to pass both tasks (Kalman/Particle Filter and BlackJack Player), as we only have two portfolio components. If you fail one, you will fail the entire course.

Send me an email with the deliverables in a zip file attached. I must receive the email by the specified time.

Task P2.1

Realize a Reinforcement Learning implementation of a self-learning BlackJack player in a programming language of your choice. This implementation provides the basis for the paper from Task P2.2. It shall learn optimal policies for at least the following scenarios:

1. The *"Basic Strategy"* from [3].
2. The *"Complete Point-Count System"* from [3].
3. In addition to the basic rules, two rule variations of your choice shall be examined for their influence on the strategies from (1.) and (2.).
4. Consider improving the system from (2.) to be able to achieve higher profits on average. Note: Your system does not have to be suitable for humans. It may therefore be relatively complicated, e.g. with respect to card counting.

What profit can be expected for the different scenarios in a greedy evaluation of your approaches?

The deliverable for this task is the commented source code of your implementation and all logfiles that contributed to the results in Task P2.2. Do not use an external Reinforcement Learning framework.

Task P2.2

Prepare a research paper using the official IEEE conference template from [1] (format A4). Use either the LaTeX or the Overleaf template. Do not use Microsoft Word with the IEEE Word template for the paper. The absolute maximum length of the paper is 6 pages.

Make sure your paper has an appropriate structure and outline. If you are unsure how good papers are structured, read good papers and analyze their structure. An appropriate structure might look like this:

- *Abstract:* One-paragraph summary of the paper. The Abstract provides a short overview of the paper.
- *Introduction:* What is the topic and why is it worth studying? The Introduction commonly describes the topic under investigation, summarizes or discusses relevant prior research, identifies open questions and problems and provides an overview of the research that is to be described in greater detail in the sections to follow.
- *Description of Reinforcement Learning player:* What did you do? This is a section which details how the work was performed. It typically features a description of the methods that were involved. A rule of thumb is that this section should be sufficiently detailed for another

researcher to duplicate your research. You should also address the theoretical background of your specific problem solutions for the task at hand, taking into account the issues raised above.

- *Experiments and Evaluation:* How well does it work? The evaluation must show that your implmentation works correctly. Use appropriate datasets to represent specific properties of the algorithm and discuss them.
- *Conclusion:* A brief summary and outlook of questions to be explored in the future.
- *References:* List of articles, books, etc. cited. A list of the sources that are cited in the paper.

The practical part from Task P2.1 focuses on different aspects from the application point of view. The paper should integrate these findings into the paper. Among other things, the following scientific questions arise:

- How do the different learning algorithms behave for the task at hand? Which learning methods have specific advantages and disadvantages here?
- Can you roughly estimate the size of the state-action space for your implementation? Can one expect to achieve stable estimates $Q(\cdot, \cdot)$? If not, how do you deal with this?
- Can you explain why the rule changes you decided to make led to the policy changes you observed?

The deliverable for this task is a pdf version of your paper.

References

[1] IEEE Paper Template. `https://www.ieee.org/conferences/publishing/templates.html`.

[2] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2 edition, 2018.

[3] Edward O. Thorp. *Beat the Dealer*. Vintage, New York, 1966.