Term Project: Reinforcement Learning - Part III

Ongoing changes in the operational environment of critical infrastructure due to external factors like fluctuations in supply and demand call for robust cybersecurity solutions to enhance resilience of the operational technology that enables the continuous operation and delivery of services. Likewise, the rise of new generations of cyber threats demands more sophisticated and intelligent cyber defence solutions equipped with autonomous agents capable of learning to make decisions without the knowledge of human experts when a timely response is critical. A promising paradigm is reinforcement learning: "an approach to machine learning that is sufficiently different from supervised and unsupervised machine learning to be often considered the third leg of the machine learning stool. Although reinforcement learning can be used for many different tasks, its most common application is in learning to control the behaviours of autonomous systems."[1]

Several reinforcement learning methods for automated network intrusion detection and mitigation have been proposed in recent years. For the purpose of broadening your horizons of how advanced machine learning can be used for cybersecurity in practice, specifically also for automated intrusions detection, you are supposed to explore the very idea of reinforcement learning by studying a high-level (fun) tutorial introduction:

Reinforcement Q-Learning from Scratch in Python with OpenAI Gym[2]

From this intuitive tutorial you will learn:

1) What Reinforcement Learning is and how it works
2) How to work with OpenAI Gym
3) How to implement Q-Learning in Python

Problem Description

As the CEO of an investment company, you have been observing significant losses due to lack of experience among your employees. In response, you've decided to develop a model to assist your employees in making investment decisions. This model will not only aid your team but also help you identify employees who may intentionally make poor investment decisions, potentially harming your company. Utilizing this model, you can detect anomalous behaviour or poor investment decisions by employees in the company.

---

[1] Fundamentals of Machine Learning for Predictive Data Analytics, John D. Kelleher, Brian Mac Namee, and Aoife D'Arcy, The MIT Press, 2020.

[2] URL: https://www.learndatasci.com/tutorials/reinforcement-q-learning-scratch-python-openai-gym/

Your company typically invests in Stocks, Real Estate, Commodities, Cryptocurrencies, and Forex, with budgets ranging from 1 to 100 million dollars. The allocation of funds for each investment sector has already been determined by another department in the company, with investments in Stocks at $5M, Real Estate at $17M, Commodities at $11M, Cryptocurrencies at $9M, and Forex at $7M. Each investment also yields a profit for the company.

The investment department can work with budgets ranging from **1to 100 million**, and after selecting an investment action, they calculate their success as follows:

budget – investment amount + profit

Your task is to train an agent capable of making investment decisions regardless of the budget available.

Here are the key aspects of this problem:

- **State space** consists of budgets, expressed as integer values from 1 to 100.

- **Actions** involve investing in Stocks, Real Estate, Commodities, Cryptocurrencies, and Forex.

- **Rewards** are calculated as "*log*(*budget – investment amount + profit*)". Better decisions result in higher rewards, as they generate more profit for the company. In cases where investments exceed the available budget, leading to the company going into debt, negative rewards are computed as (*budget – investment amount + profit*). This means unaffordable investments result in more significant negative rewards compared to affordable investments.

<u>NOTE:</u> Given the challenge of predicting profits in the stock market, for the purpose of this question, profit is calculated as a random number with a maximum of double the investment amount and a minimum of zero.

The provided code will <u>**generate a sample training dataset**</u> for this problem. Your task is to train a Reinforcement Learning model and address the following aspects:

1. **Explain the hyperparameters** of this model (*alpha, gamma, epsilon*) and how they influence the training process.

2. **Obtain the Q-Table** of the trained model and provide your interpretation of this table based on the values of the hyperparameters you chose.

3. **Compute the policy** of the model and provide your interpretation.

4. **Apply this model to unseen data** and extract the optimal action table for each state.

Thank you!