Deep Learning – Lab 08

## 1. GridWorld Notebook

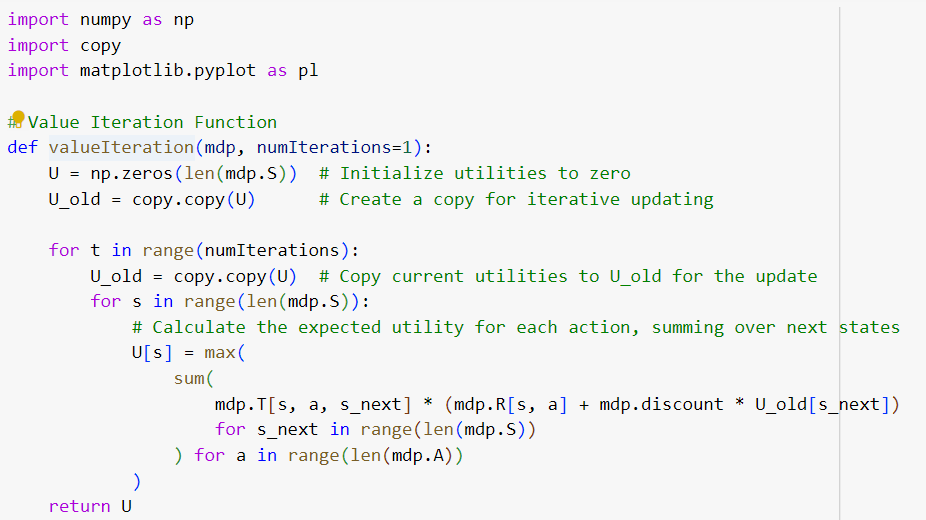
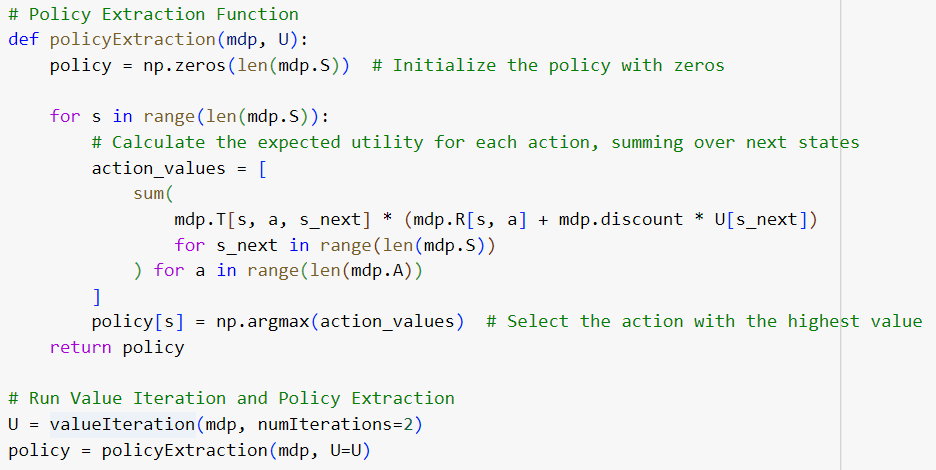
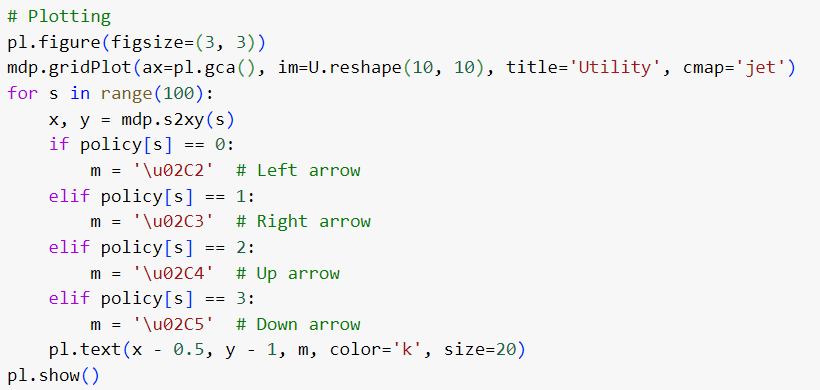
In the GridWorld notebook, the following modifications were implemented based on the lab requirements:  
- Completed any placeholders marked with `# type your code here`.  
- Expanded the GridWorld environment to support dynamic grid sizes, allowing us to observe execution time and convergence changes.  
- Conducted experiments on different grid sizes to evaluate the impact on performance and learning speed.  
- Generated results that illustrate how larger grid sizes affect the execution time and convergence of the Q-learning process.

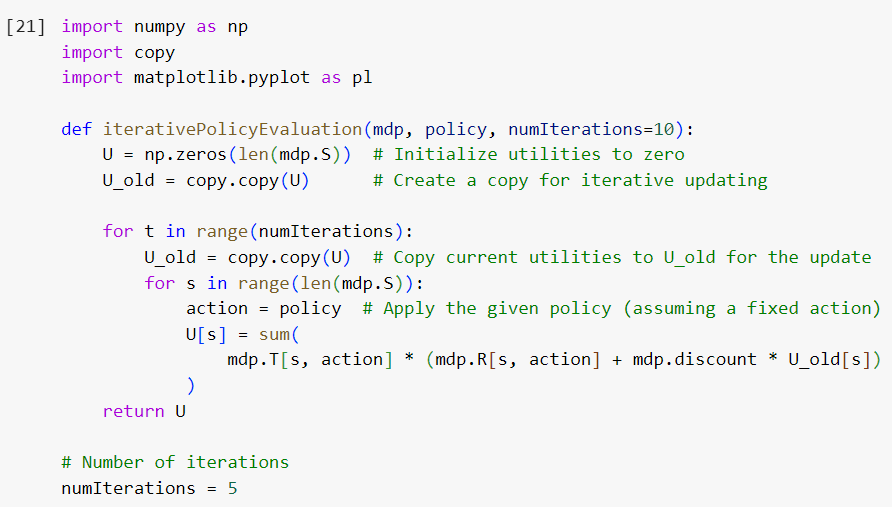
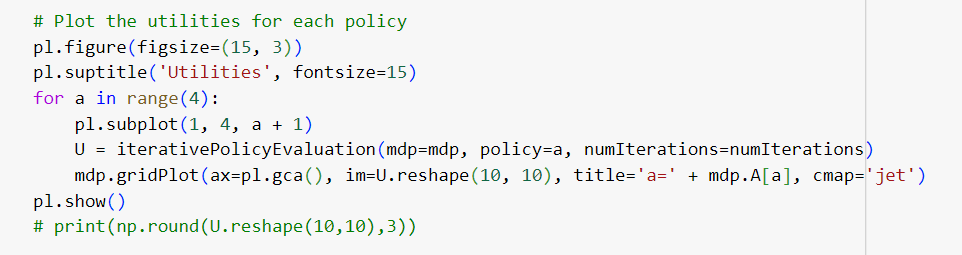
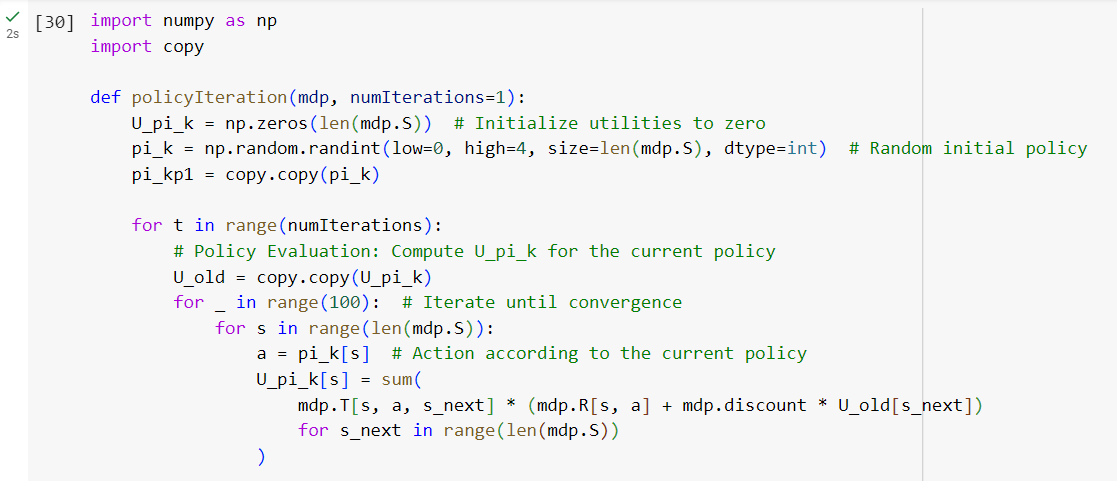
## 2. Markov Decision Process (MDP) Notebook

In the Markov Decision Process notebook, the following tasks were performed:  
- Completed the Policy Iteration, Value Iteration, and Q-learning sections to compare Model-Based and Model-Free approaches.  
- Added execution time tracking for each method to compare their performance in terms of speed and convergence rate.  
- Created convergence plots that illustrate the differences in learning and stabilization across Policy Iteration, Value Iteration, and Q-Learning methods.  
- Added explanations of Model-Based versus Model-Free methods and discussed results to highlight key differences in their convergence behavior and execution time.

## Conclusion

The modifications across both notebooks enable a thorough comparison of Model-Based and Model-Free reinforcement learning approaches. By observing performance on different grid sizes in GridWorld, and tracking execution time in MDP, we demonstrate how the complexity of the environment and the learning approach affects convergence and computation time.



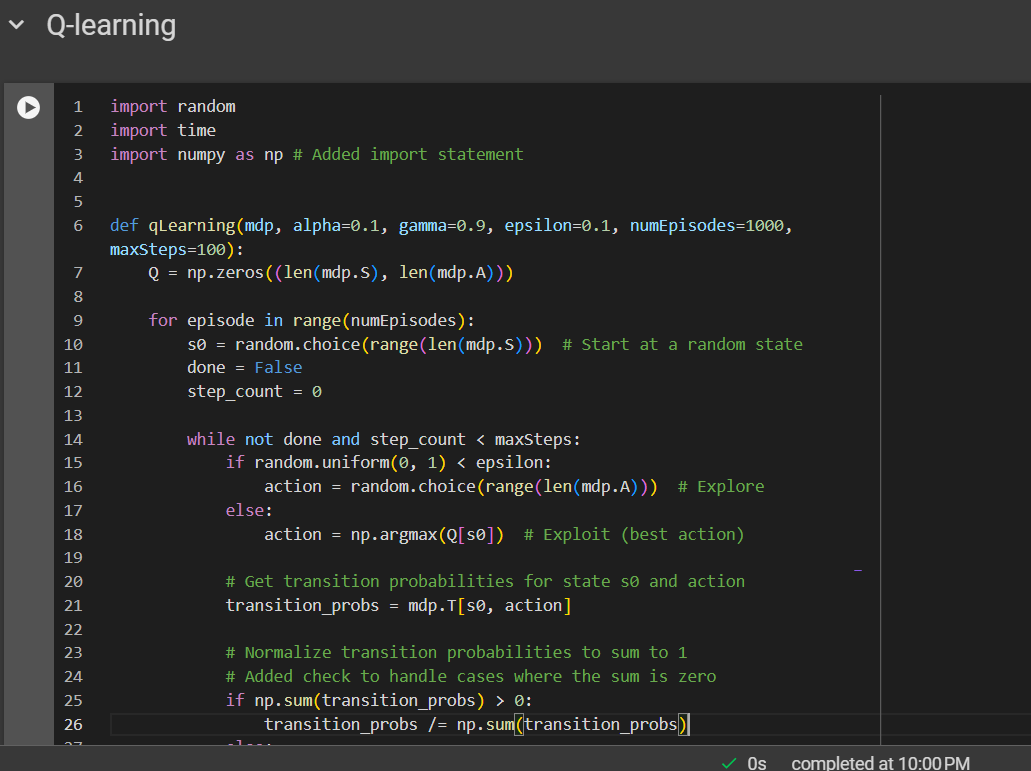


A screenshot of a computer code

Description automatically generated

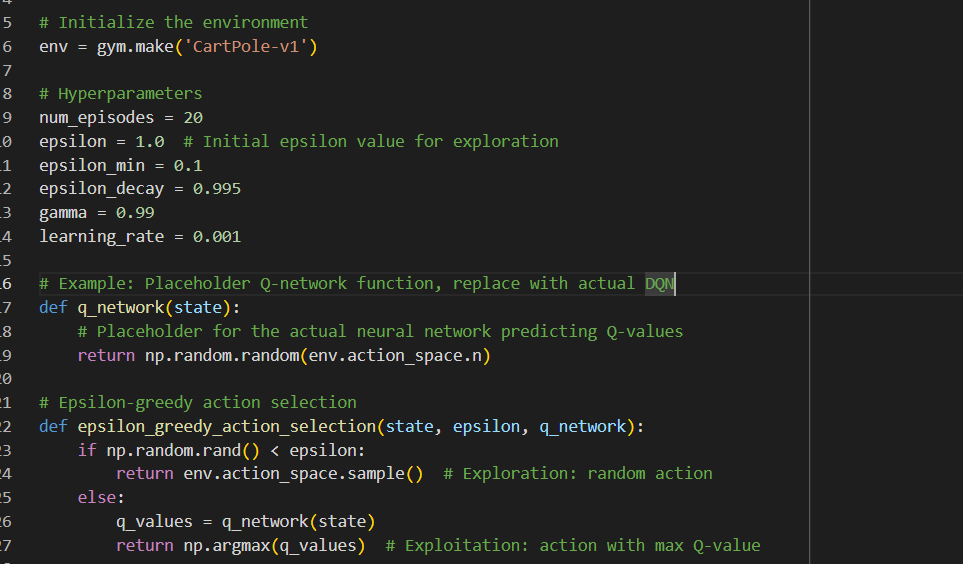
### Difference Between Model Based and Model Free Algorithms

|  |  |
| --- | --- |
| **Model Based Algorithms** | **Model Free Algorithms** |
| Learns a model of the surroundings explicitly | Learns via firsthand encounter with the environment rather than from a model of it. |
| Usually converges more quickly when the model is precise. | Slows convergence, particularly in intricate settings. |
| Makes use of the acquired model to plan and simulate activities, enabling foresight | Without prior planning, gains knowledge through direct trial & error. |
| Because it can plan by simulating multiple probable outcomes, it is more sample-efficient. | Less sample-efficient and more dependent on environmental interactions for improvement |



A screenshot of a computer program

Description automatically generated



A graph of a chart

Description automatically generated with medium confidence

A black screen with green text

Description automatically generated

A graph of a chart

Description automatically generated with medium confidence

A screen shot of a computer

Description automatically generated

A graph of a line

Description automatically generated

A screen shot of a computer program

Description automatically generated

A screen shot of a computer code

Description automatically generated



GitHub Repository Link

<https://github.com/IT21279720/Deep_Learning_Labs>