

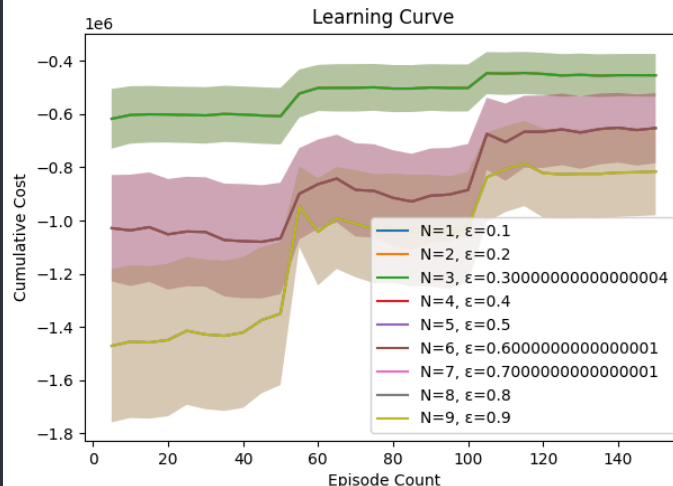
Assignment 1: Q-Learning for GridWorld

Name: Vinicius Eller

Question 1-2)

```
random_generated = ""
simple = ""

random_generated = ""
simple = ""
```



code: <https://github.com/Vigrel/Q-Learning-for-GridWorld.git>

Refs: Géron, Aurélien. Hands-On Machine Learning with Ten... (chapter 18). O'Reilly Media

Question 3)

The selection of epsilon and step count significantly shapes the learning curve in N-Step Q-Learning. Epsilon governs the exploration-exploitation balance, impacting the agent's trade-off between discovering optimal actions and exploiting known ones. Higher epsilon values promote exploration but may impede convergence due to increased noise, while lower values favor exploitation and quicker convergence but risk suboptimal policies. The step count (N) dictates the temporal credit assignment length, influencing the speed of adaptation to changing dynamics. Larger N values enhance long-term planning but may slow down learning due to delayed rewards. The impact of epsilon and step count varies across environments, with complex ones benefiting from higher exploration and longer-term planning to navigate intricate state spaces. While environments with more immediate feedback benefit from a more exploitative strategy and shorter temporal credit assignment.