**HW11**

**Michelle Hui Sun. Monday 6:30pm Section 1**

1. **What parameters did you change?**

I changed first layer, second layer, epochs, batch size, epsilon\_min and the results are in the below table.

1. **What values did you try?**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Run number** | **First layer** | **Second layer** | **epochs** | **Batch size** | **Epsilon\_min** | **iteration** | **Average Rewards** | **Test Rewards** |
| 1 | 16 | 8 | 1 | 64 | 0.01 |  |  |  |
| 2 | 128 | 128 | 1 | 64 | 0.01 | 485 | 199.56280 | 225.221958 |
| 3 | 128 | 64 | 1 | 64 | 0.001 | 432 | 194.91354 | 251.71328056 |
| 4 | 128 | 64 | 1 | 128 | 0.01 | 419 | 199.9161498 | 238.38759 |
| 5 | 128 | 64 | 2 | 64 | 0.01 | 513 | 197.996149 | 248.342889 |

1. **Did you try any other changes that made things better or worse?**

Most of the changes i tired made the model become better.

1. **Did they improve or degrade the model? Did you have a test run with 100% of the scores above 200?**

The parameters I tried improved the model

I have two runs have all test run with 100% of the score above 200, which are shown below.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Run number** | **First layer** | **Second layer** | **epochs** | **Batch size** | **Epsilon\_min** | **iteration** | **Average Rewards** | **Test Rewards** |
| 3 | 128 | 64 | 1 | 64 | 0.001 | 432 | 194.91354 | 251.71328056 |
| 5 | 128 | 64 | 2 | 64 | 0.01 | 513 | 197.996149 | 248.342889 |

1. **Based on what you observed, what conclusions can you draw about the different parameters and their values?**

5.1 if the first layer and second layer becomes larger, the model performance become better.

5.2 increase batch size make the model process slightly faster. It improves the model performance (test rewards) marginally.

5.3 I also change the batch size from 1 to 2. It also slightly improve the model performance (test rewards).

5.4 I changed the epsilon, the larger the epsilon, the robot will have more random action, hence the model performance is better if we reduce the epsilon value (0.01 to 0.001).

1. **What is the purpose of the epsilon value?**

The bigger the epsilon, the movement/motion is more random. Hence the model might need a smaller epsilon to perform better.

1. **Describe "Q-Learning"**

Based on Wikipedia “ Q-learning is a model-free reinforcement learning algorithm to learn a policy telling an agent what action to take under what circumstances. It does not require a model of the environment, and it can handle problems with stochastic transitions and rewards, without requiring adaptations”.