Exercise 1: Imitation Learning and Reinforcement Learning

Student: Erlis Lushtaku
Matriculation number: 5772476

Introduction

This report presents the results of experiments conducted as part of the Deep Learning Lab Course on Imitation Learning and Reinforcement Learning. The objective was to explore behavioral cloning and reinforcement learning techniques using the CarRacing environment from the OpenAI Gym benchmark suite. The experiments consisted of two main parts: imitation learning and reinforcement learning.

1.1. Imitation Learning

Hyperparameters and Training

lr=1e-4
batch_size=64
history_length=0
n_minibatches=1000
dropout_prob=0.2
optimizer=Adam
weight_decay=0.001

The training and validation accuracy can be seen in Fig.1 in the appendix.
The model achieved the highest score for mean: 885.6333333333162 and standard deviation std 9.726367370312698. The results for 15 episodes of testing can be seen in Table 1.

The modified_id_to_action function was also modified to make the model more accurate during testing. The amplitudes of doing a specific action like turning, accelerating, braking were made variable depending on the velocity of the car. For example, if the car is going fast we would like to turn the wheels less, we would like to brake during turns and accelerate less and the opposite goes for when the car is going slow. To find good parameters to multiply the velocity for each case a parameter search space was defined and the model was tested with different combinations. The best ones were put in the default values of the actions afterwards.

1.2. Reinforcement Learning: Deep Q-Networks (DQN)

Hyperparameters and Training

gamma=0.95
batch_size=64
epsilon=0.1
tau=0.01
lr=1e-4
history_length=0
replay_capacity=100000
num_episodes=1000
skip_frames=3
optimizer=Adam

The training and validation accuracy can be seen in Fig.2 in the appendix.
The model achieved the highest score for mean: 885.6333333333162 and standard deviation std 9.726367370312698. The results for 15 episodes of testing can be seen in Table 2.

Appendix:

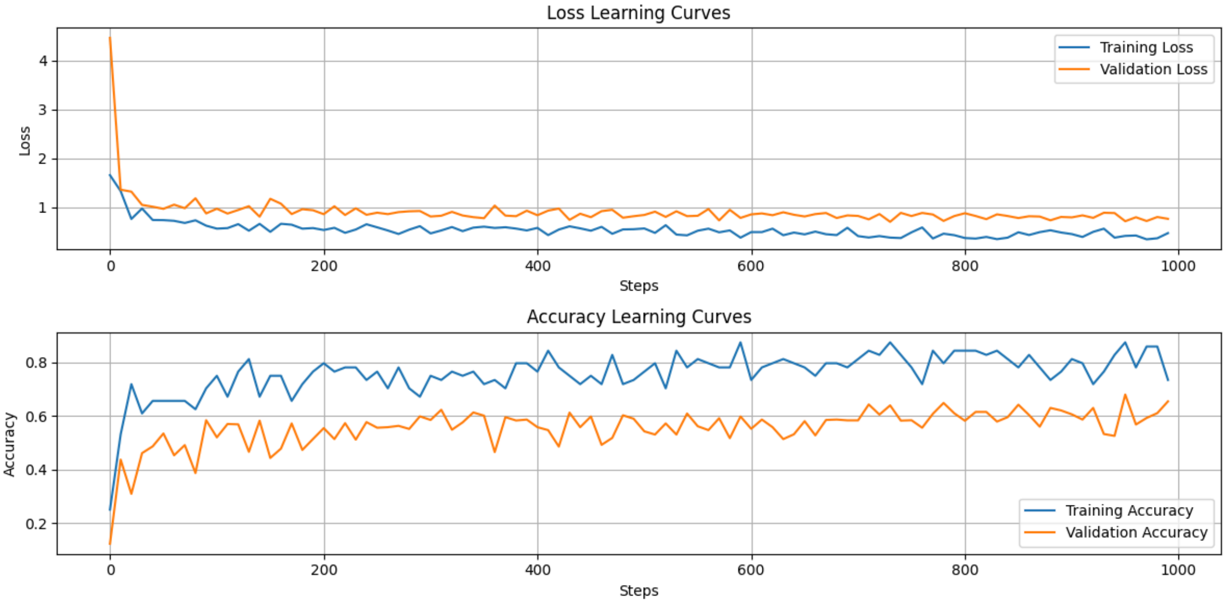

Fig.1: Training and validation performance for Imitation Learning agent

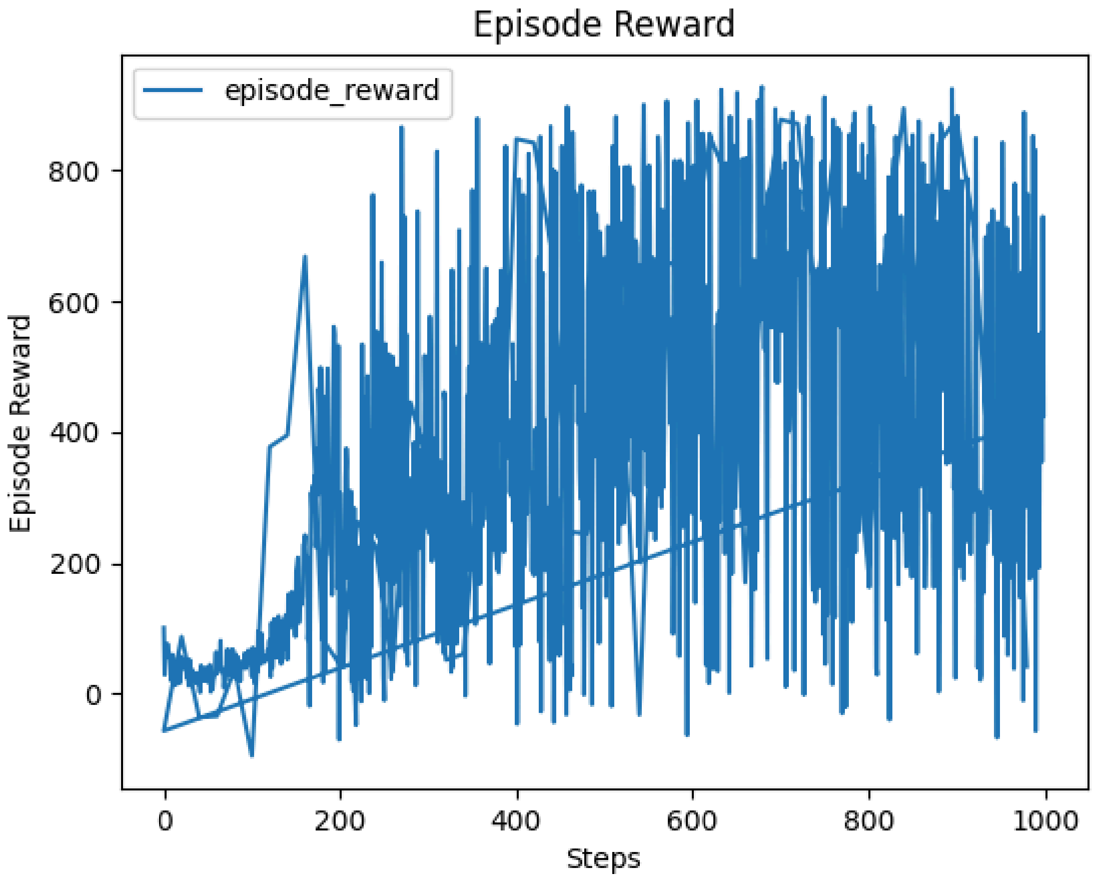

Fig.2: Training and validation performance for Reinforcement Learning agent

Episode	Reward
1	888.9999999999881
2	886.7999999999793
3	886.0999999999779
4	871.7999999999747
5	887.8999999999814
6	878.5999999999855
7	886.8999999999847
8	899.9999999999876

9	890.099999999984
10	890.5999999999868
11	896.7999999999877
12	879.3999999999867
13	859.5999999999735
14	887.2999999999811
15	893.599999999982
Mean and standard deviation	"mean": 885.6333333333162, "std": 9.726367370312698

Table 1: Results for 15 episodes of running the Imitation Learning agent

{"episode_rewards": [,,,,,,,,,,,,,,], "mean": 725.7682216556781, "std": 78.56869246889391}

Episode	Reward
1	737.4175226585887
2	789.9665505226246
3	748.4805755395432
4	727.1661016948939
5	796.5684210526061
6	791.7287356321644
7	788.3891891891619
8	512.6920634920482
9	728.7180722891354
10	764.454237288115
11	612.0205128204968
12	758.9744360901991
13	734.5046594981851

14	620.4609970674284
Mean and standard deviation	"mean": 725.7682216556781, "std": 78.56869246889391

Table 2: Results for 14 episodes of running the Reinforcement Learning agent