Deep Learning LAB

Name: Tushar Dilip Maske
Matr. No.: 4967240

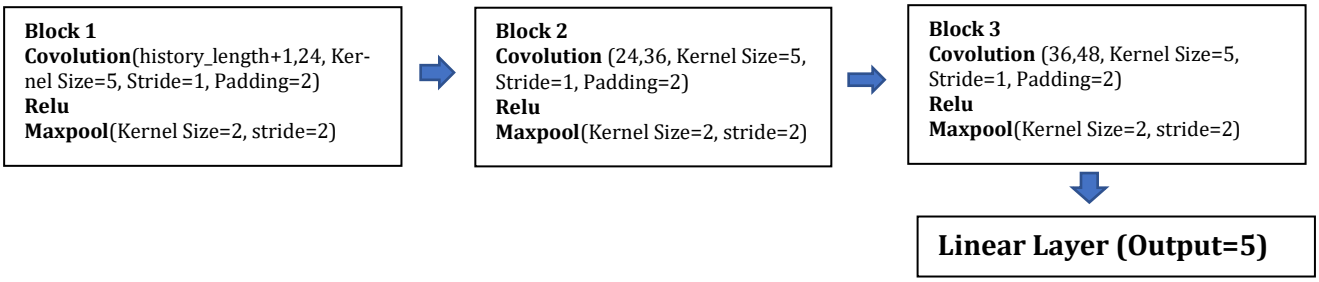
Assignment 3: Robotics

Task 2: Imitation Learning

Model Training and Validation

Fig. 1 consist of the results of training and validation accuracies for progressing minibatches.
The results are generated without history and with history value N = {1,3,5}.
Also, the training samples have been chosen randomly with uniformly distributed actions.

CNN Network:
Linear Layer (Output=5)



Training Parameters:
Number of Minibatches=10000, Batch Size=65, Learning Rate=1e-4

Results:

History Length	Resulting Values
0	426.65
1	587.26
3	102.83
5	63.67

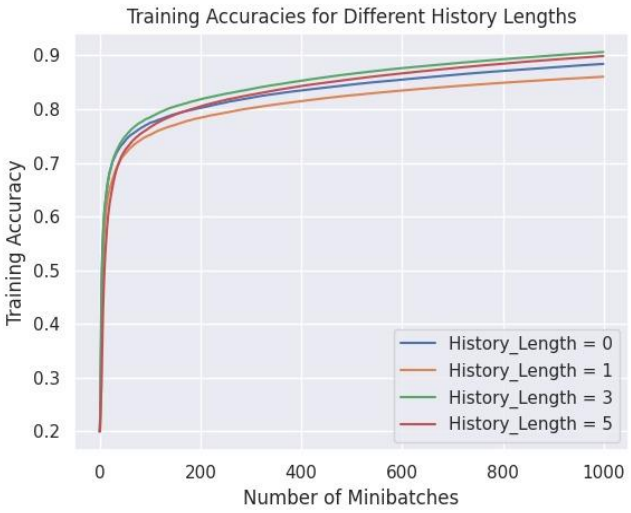


Fig.1: Imitation Learning Training Accuracies

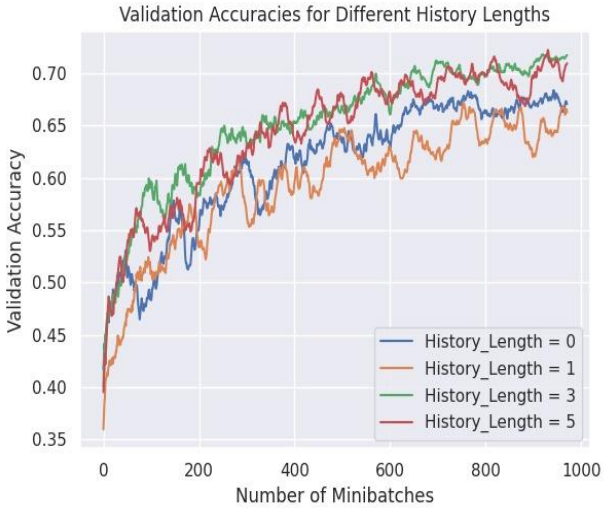


Fig.2: Imitation Learning Validation Accuracies

Task 3: Reinforcement Learning: Deep Q-Networks

3.1 CartPole

Network: Given MLP Network

Results:

Mean: 925.86, Standard Deviation: 123.95 (Mean of the readings from fig 4)

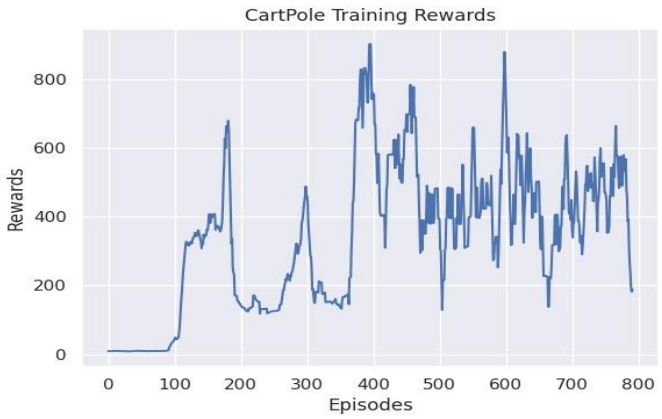


Fig.3: Cartpole Training Rewards

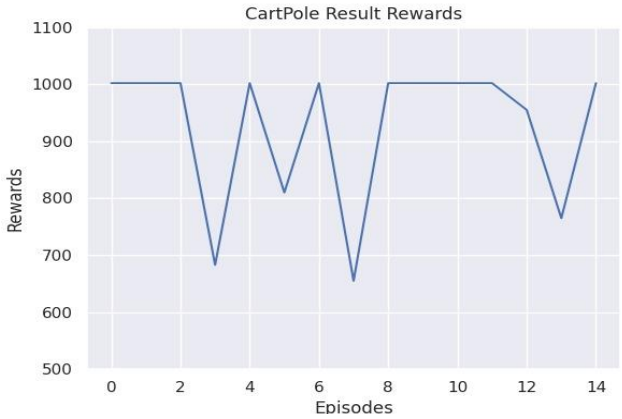


Fig.4: Cartpole Result Rewards

3.2 CarRacing

Network:

Block 1
Covolution(history_length+1,24, Kernel Size=5, Stride=1, Padding=2)
Relu
Maxpool(Kernel Size=2, stride=2)

→

Block 2
Covolution (24,48, Kernel Size=5, Stride=1, Padding=2)
Relu
Maxpool(Kernel Size=2, stride=2)

→

Linear Layer (Output=5)

Settings:

Initial timestep = (1/10)*Maximum_Timesteps, then after every 50 episodes, increase by 100
Action Selection= Random with probability -> {a=[0,1,2,3,4], p=[0.5, 0.2, 0.2, 0.05, 0.05]}

Results:

Mean: 674.57, Standard Deviation: 149.37 (Mean of the readings from fig 6)

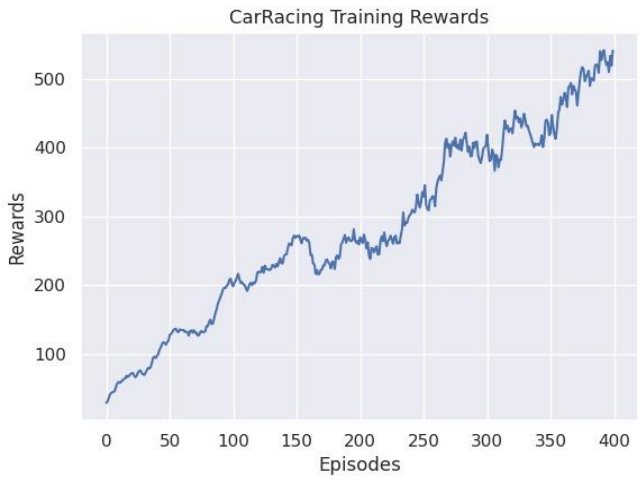


Fig.5: CarRacing Training Rewards

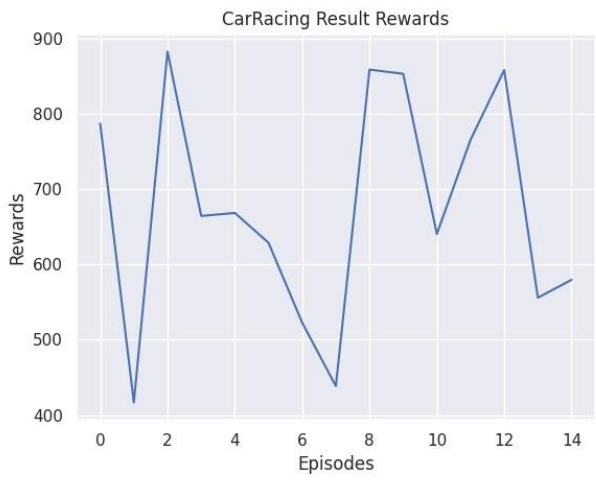


Fig.6: CarRacing Result Rewards