Advanced Topics Of Deep Learning - Mid Semester Project

Ofek Naim (ID. 314728996), Hinoy Solomon (ID. 205417678)

Submitted as mid-semester project report for Advanced Topics Of Deep Learning course, Colman, 2024

1 Introduction

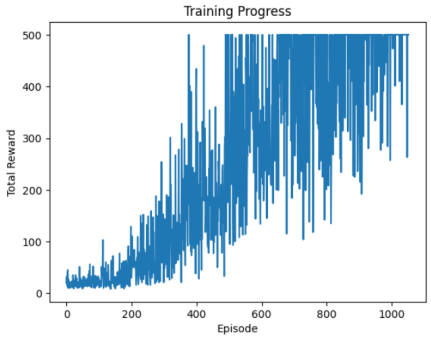
In this project, we implemented, trained, and evaluated policy-based reinforcement learning models using the REINFORCE algorithm on the CartPole-v1 environment. The objective was to balance a pole on a moving cart by learning an optimal policy through interaction with the environment. To achieve this, we designed three different neural network architectures—Small, Medium, and Large—and trained them separately to analyze their learning performance. Each model was trained using a policy gradient approach, where it learns by optimizing the probability of taking actions that lead to high rewards. After that, we implemented Imitation Learning to train an agent to play CartPole-v1 using demonstrations from an expert model. Unlike traditional reinforcement learning, where the agent learns through trial and error, imitation learning enables the agent to learn directly from expert actions.

2 Training RL Agent

We force the training model to stop when he reach average reward of 480 since the last 10 episodes, with max of 500 episodes. In all of our architecture we use the same activation function - RELU, the same activation function for output layer - Softmax, and the same optimizer - Adam. Considering them the best for our exercise. Here I will show what architecture I used for each train model:

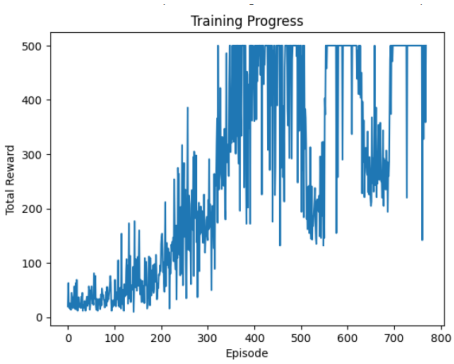
2.0.1 Architecture One:

in this architecture we used a neural network with two hidden layer of 24 neurons each. we achieve our goal after 1052 episodes



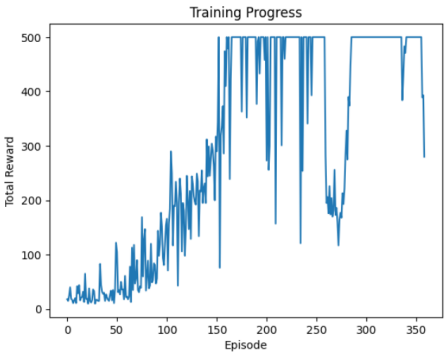
2.0.2 Architecture Two:

In this Architecture we used 2 hidden layer, first hidden layer with 64 neurons and the second with 32 neurons. we achieve our goal after 769 episodes an improvment



2.1 Architecture Three:

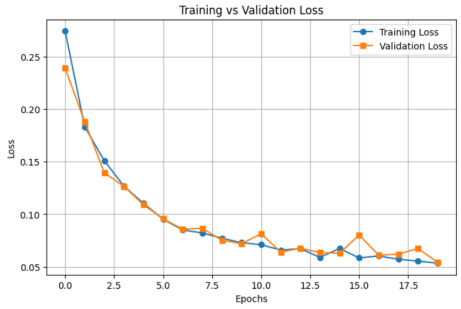
In this Architecture we used 3 hidden layer, first hidden layer with 256 neurons, the second with 64 neurons and third hidden layer with 32 neurons. we achieve our goal after 359 episodes



Therefor, we choose Architecture to be our best model to continue with for our Imitation Learning.

3 Imitation Learning

After getting our successful model, we create a dataset of state and action using our model. Using this dataset, we will train our Imitation model using supervised learning. We used 25,000 examples and the Imitation learning was doing great.



With high accuracy of 98% over the test data

```
157/157 ━━━━━━━━━━━━━━━━━━━━ 1s 4ms/step - accuracy: 0.9793 - loss: 0.0570
Test Accuracy: 98.02%
```

Testing our model: getting 500 reward average

```
Episode 1: Reward = 500.0
Episode 2: Reward = 500.0
Episode 3: Reward = 500.0
Episode 4: Reward = 500.0
Episode 5: Reward = 500.0
Episode 6: Reward = 500.0
Episode 7: Reward = 500.0
Episode 8: Reward = 500.0
Episode 9: Reward = 500.0
Episode 10: Reward = 500.0
✅ Imitation Model Average Reward over 10 episodes: 500.00
```

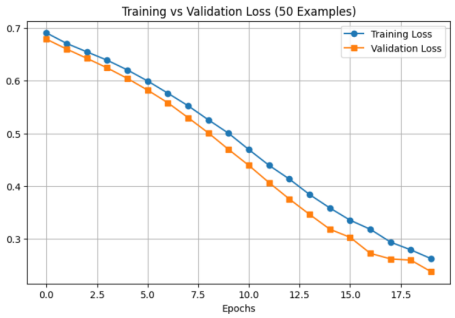
4 Testing different size of dataset

4.1 General approach

We were asked to evaluate our model with different numbers of examples: 50, 100, 500, 1000 we choose our examples randomally. And these are the results:

4.2 50 Examples

The results were not good, probably because 50 examples are not enough.

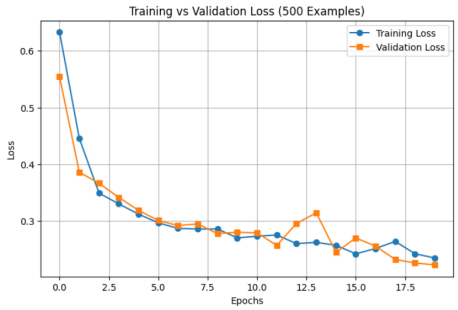


4.3 100 Examples

The results were good, but still not satisfactory.

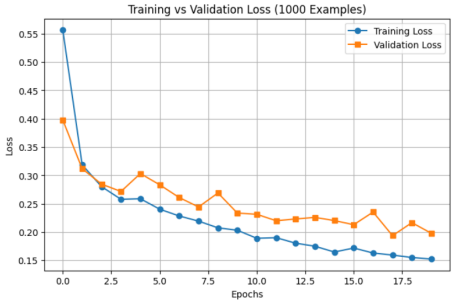
4.4 500 Examples

The results are really good, starting to meet our expectations.



4.5 1000 Examples

We're starting to see improvement and better results, but it seems we've slightly exceeded the optimal amount, and there is no significant improvement.



4.5.1 Videos

50 examples: https://drive.google.com/file/d/10KTT28YOUwTqV0tgV9rBquuG9atB8Ld3/view?usp=sharin

100 examples: https://drive.google.com/file/d/1o8y27wyF-XKVwKf-TF7lCDf5WIPDudou/view?usp=sh

500 examples: https://drive.google.com/file/d/1Ygxc5p4TfWqUYpF07QTxM9xgryMvC$_xf$/$view?usp$ = $sharing$

1000 examples: https://drive.google.com/file/d/1aBR2ZPvT434nyX604$_4VNeCzBQ8ejYJs$/$view?usp$ = $sharing$

5 Notebooks

5.1 Expert model training

https://colab.research.google.com/drive/16xsmNd7vHFIgWlyQHpnOdT2-zNyjCsP-?usp=sharing

5.2 Imitation model training

https://colab.research.google.com/drive/10s1YHJiviES0X6Hwoms0HmPL2hE-NK6m?usp=sharing

5.3 Inference

https://colab.research.google.com/drive/1kjeITXqE7tuVFxb5$_Ti74BN0Dca0VTj-?usp$ = $sharing$