Research Diary

Advanced Deep Network Development

Daniel Kuknyo

ELTE IK

2022/23 Fall Semester

Our GitHub repo

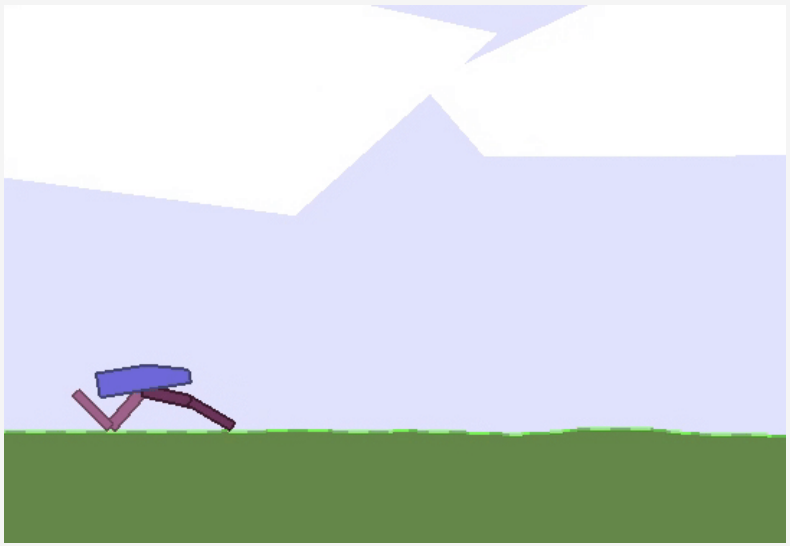
- Please visit the project file for more:
- `https://github.com/basictask/SpikingNeuralNetworks/tree/daniel`
- For any more questions: daniel.kuknyo@mailbox.org

Week 1.

- We've received the code base from our supervisor.
- We've spent most of our work to debug the code.
- Got it up and running on everyone's environment.

Week 2.

- Everyone chooses their own environment: discrete, continuous, convolutional.
- In my case the problem is the continuous control.
- The environment is BipedalWalker. Goal: to predict the inclination angles of the joints at each time frame.

Figure: 

Week 3.

- The chosen environment is tested with the base code.
- Q-learning is implemented. It's not performing well.

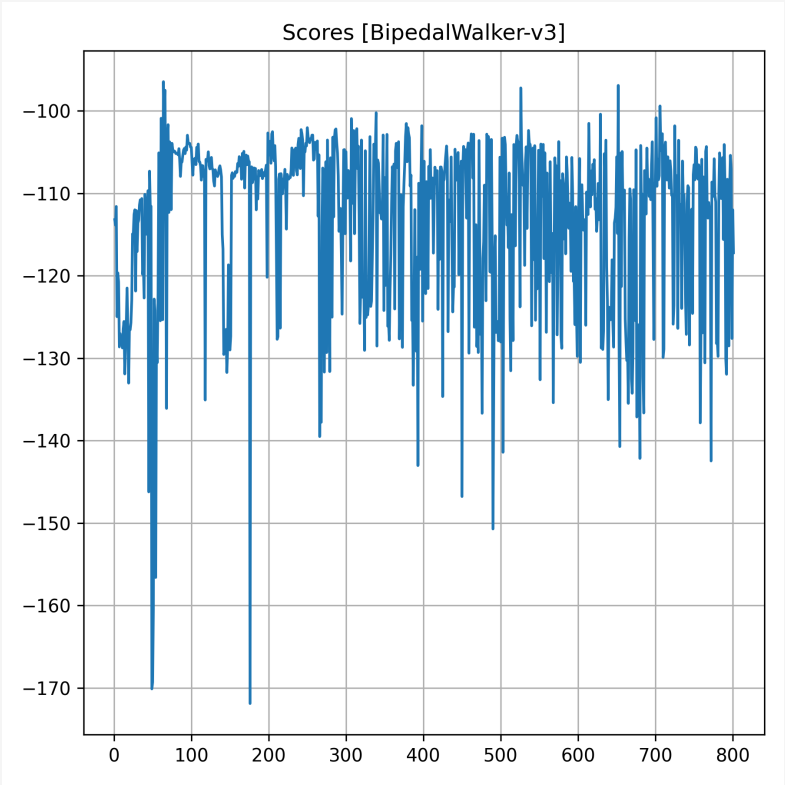


Figure:

Week 4.

- Action space discretization is implemented.
- I have developed an approach that subdivides a large action vector into separate angles.

Week 3.

- Soft actor-critic model is implemented.
- The performance is increasing.
- It's two of the same architectures of Q-learning models with different input/output sizes.

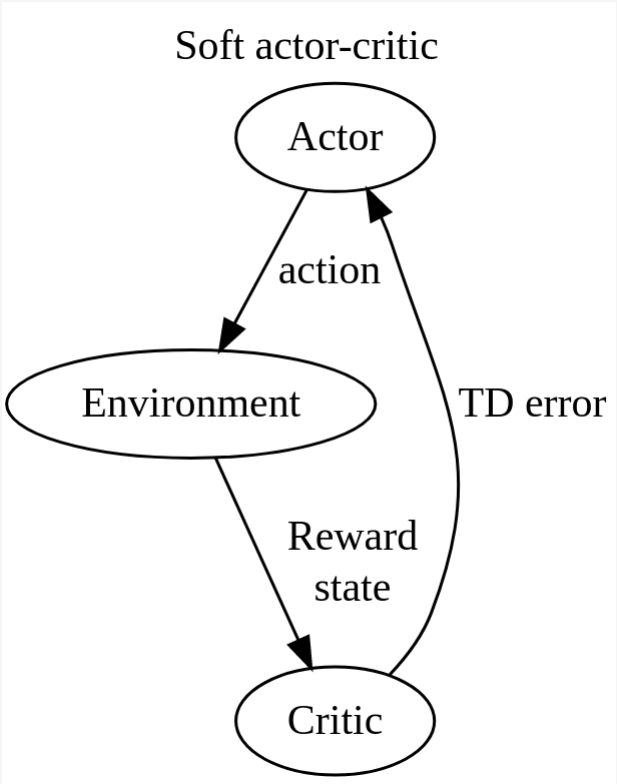


Figure:

Week 5.

- The soft actor-critic with Q-learning was not a big success.
- The next step was to implement continuous control for soft actor-critic models.
- In this case the model is directly predicting the inclination angles for the 4 joints of the walker.
- From now on the networks have different architectures, usually the actor has more neurons.

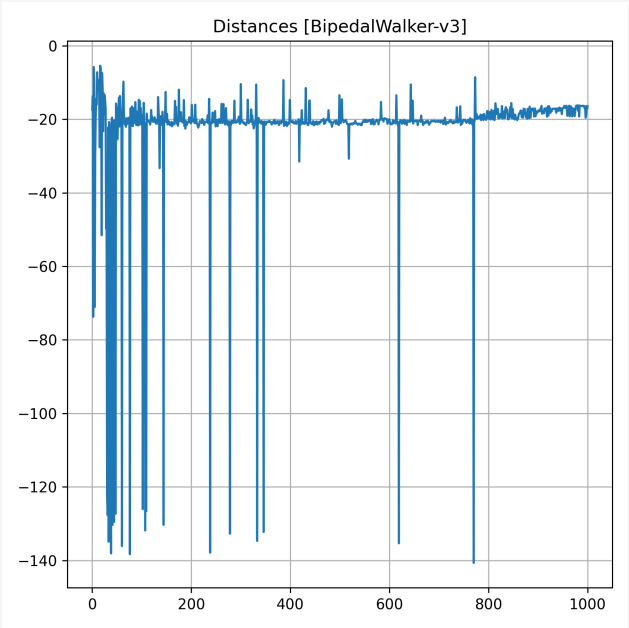
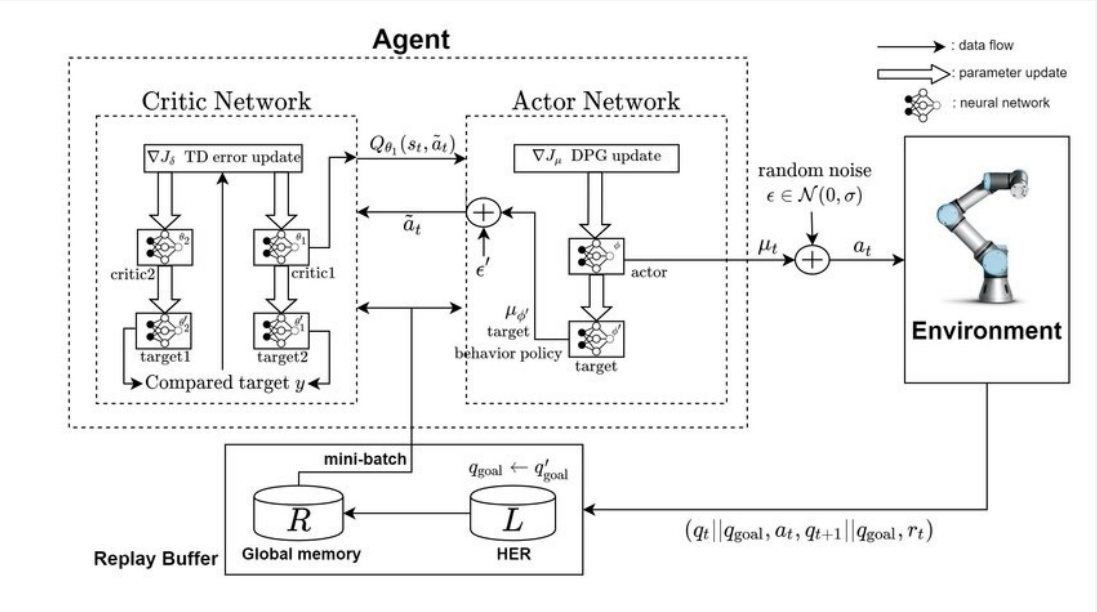


Figure:

Week 6.

- The soft actor critic was the best model so far, but the environment still wasn't solved very well.
- This gave reason for further experimentation. The next architecture to be implemented was to add another critic to the prediction process.
- This is called the TD3 architecture. The first model is based on spiking neural networks for both critics and the actor as well.

Figure:



Week 7.

- I had a lot of trust in the TD3 architecture.
- It takes a lot of time to teach however.
- The next week was only based on teaching and hyperparameter fine-tuning.

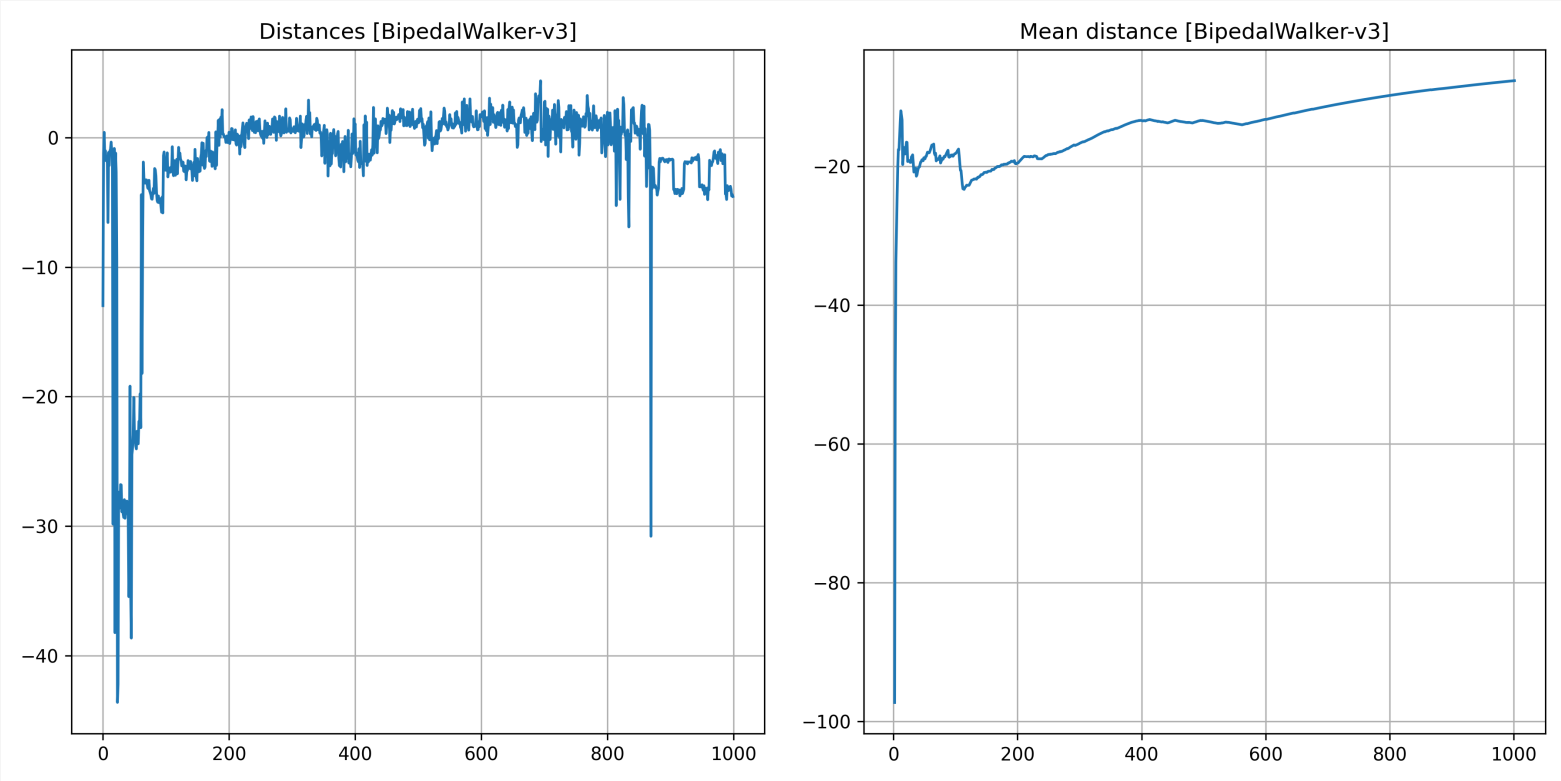


Figure:

Week 8.

- The last model to be implemented is an experiment that's based on involving a regular feedforward network as the actor while keeping the 2 spiking critic networks.
- This model was by far the best performing. This can be attributed to the fact that sometimes the spiking networks have outliers as outputs.
- The BipedalWalker environment is not robust to noise. It's very easy to tip the walker because it's unbalanced by default.
- Below two separate runs can be seen with different parameters.

Figure: