

Meeting

05/27/2020

Shuo Zhang

Mujoco/Gym Experiments

Progress

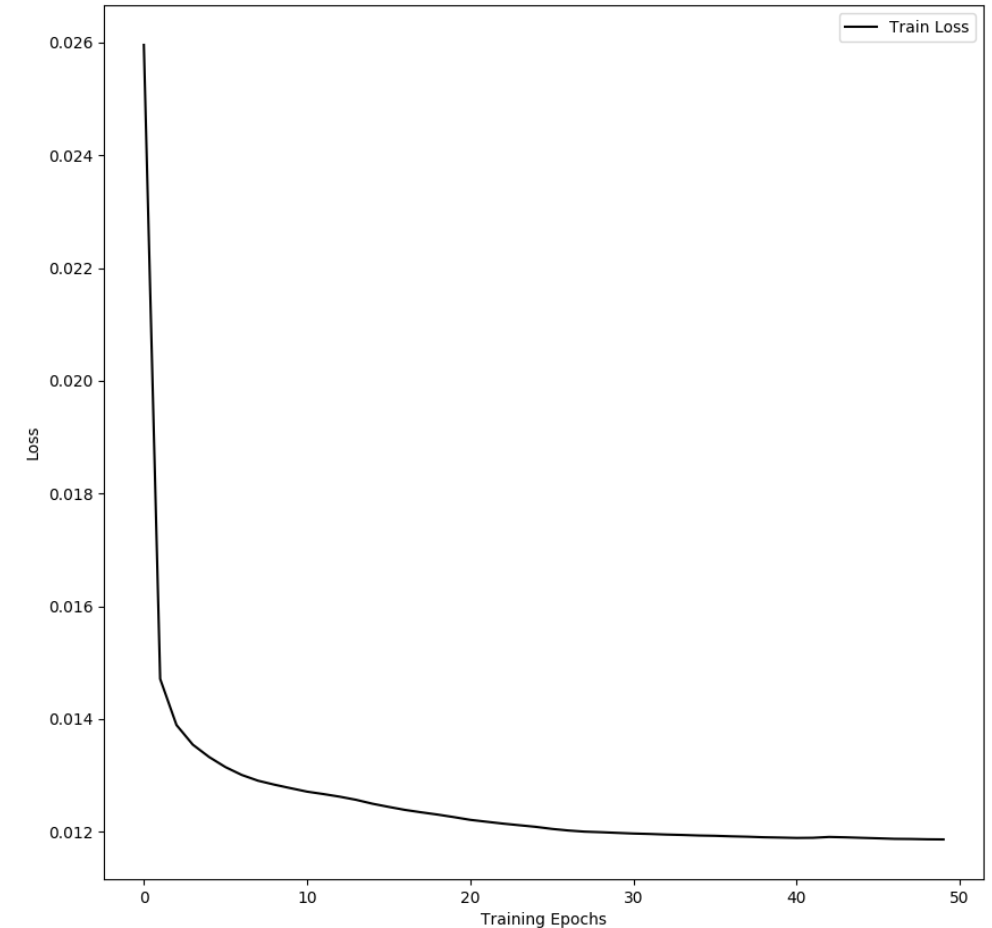
- Reacher-v2:
 - Collect Data: Done
 - Learn Dynamic Model: Done
 - Model Evaluation: Done
- Acrobot-v1;
 - Collect Data: Done
 - Currently Training: Should be over in 1 hour
 - Model Evaluation: Later

Collect Data

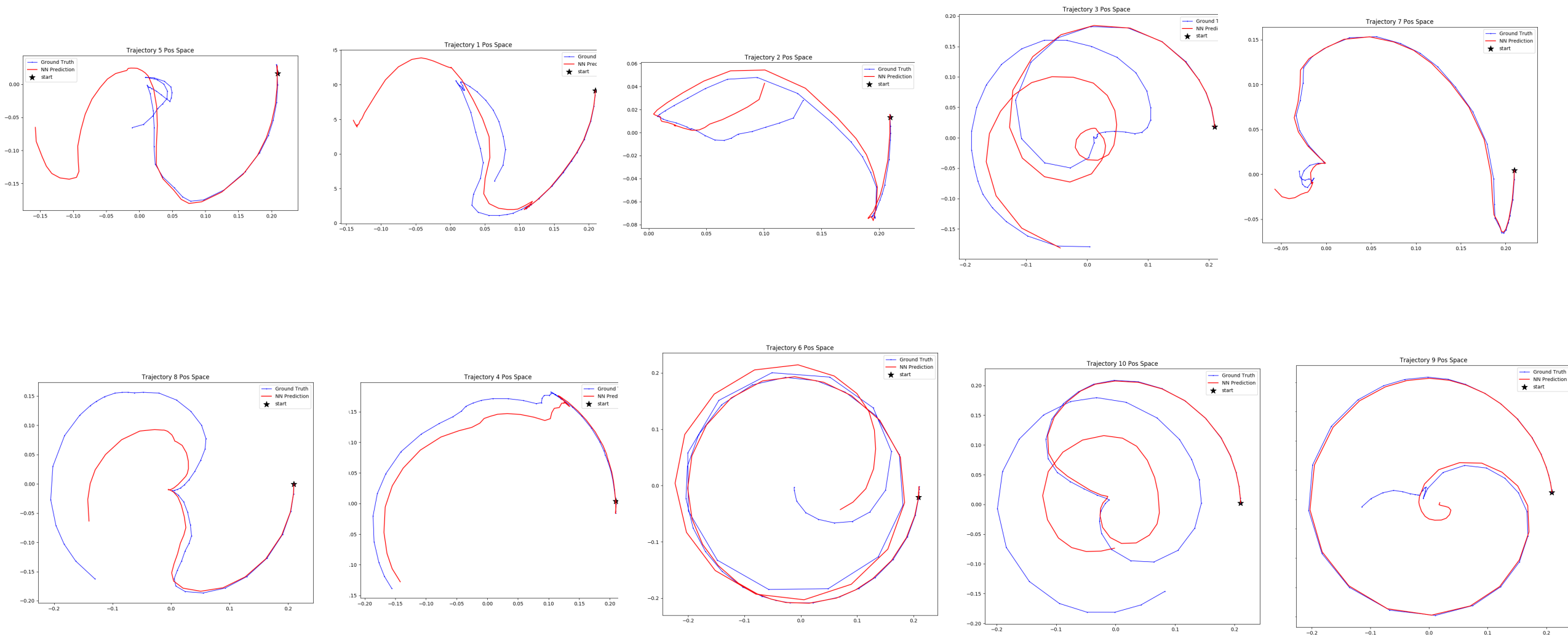
- 1 Million Steps for each task

Learn Transition Model

- Experimented different configs:
 - Learning Rate: [0.001, 0.0001]
 - Number of Hidden Nodes: [200, 512]
 - Number of Training Episodes: [20,50]
- Generally, 0.0001 is better than 0.001. 20 iterations of training are good enough. Little progress from iteration 21 to iteration 50, or even from iteration 6 to iteration 50.

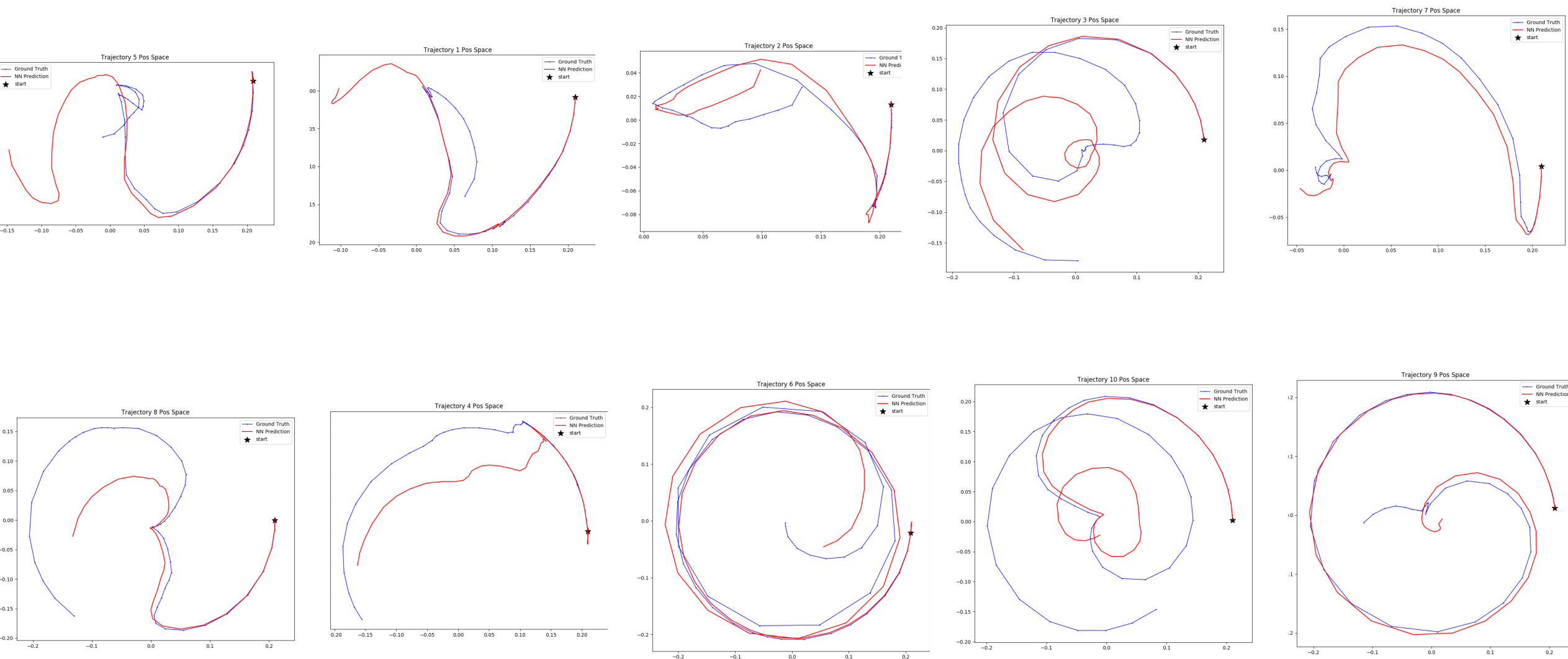


Model Evaluation: Reacher-v2



After 5 iterations of training

Model Evaluation: Reacher-v2



After 20 iterations of training