**All Results and Figures**

**0. Review**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Environment(Data Amount) | Open Loop Control | | Closed Loop Control | |
| A\* + rollout | PPO + rollout | (A\* based)LQR | AIP |
| Real Hand(100% of 150K) | Not yet | Not yet | Not yet | Not yet |
| Gazebo Hand(0.1% of 1.6M) | Done | Done | Done | Done |
| Reacher-v2(0.1% of 1M) | Done | Done | Done | Not yet |
| Acrobot-v1(100% of 1M) | Done | Done | Not Available | NotAvail |

**1. Success Rate**

Table 1. Gazebo Hand (0.1% Model)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Goal Location | 0 | 2 | 7 | 8 | 15 |
| A\* | 0% | 100% | 100% | 0% | 0% |
| PPO | 0% | 100% | 100% | 100% | 0% |
| LQR | 100% | 100% | 100% | 100% | 100% |

Table 2. Reacher (0.1% Model)

|  |  |  |  |
| --- | --- | --- | --- |
| Goal Location | 1 | 2 | 5 |
| A\* | 0% | 100% | 0% |
| PPO | 0% | 0% | 0% |
| LQR | 100% | 0% | 100% |

Table 3. Acrobot (100% Model)

|  |  |
| --- | --- |
| Goal Height | 1.0 |
| A\* | 60% |
| PPO | 30% |
| LQR | - |

**2. Gazebo Hand (0.1% Model) Plots**

* A\*

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA screenshot of a cell phone

Description automatically generated

* PPO(Separate Model: Goal location not as a part of state; Sparse Reward)

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

* LQR

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

**3. Reacher (0.1% Model) Plots**

* A\*

A screenshot of a cell phone

Description automatically generatedA close up of a map

Description automatically generatedA screenshot of a cell phone

Description automatically generated

* PPO(General Model: Goal location as a part of state; Continuous Reward)

A screenshot of a cell phone

Description automatically generatedA close up of a logo

Description automatically generatedA close up of a logo

Description automatically generated

* LQR

A screenshot of a cell phone

Description automatically generatedA close up of a map

Description automatically generatedA screenshot of a cell phone

Description automatically generated

* AIP

Chart

Description automatically generatedDiagram

Description automatically generatedChart, bubble chart

Description automatically generated

Diagram

Description automatically generated

**4. Acrobot (100% Model) Plots**

* A\*

A screenshot of a computer

Description automatically generated

* PPO(Sparse Reward)

A screenshot of a cell phone

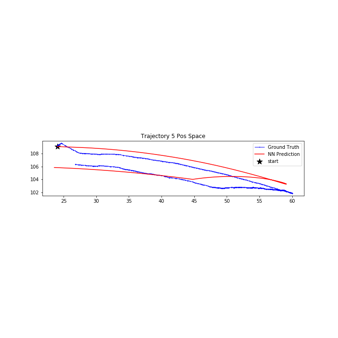
Description automatically generated

**Appendix**

**1. Transition Model:**

* Real Hand(100% Model):

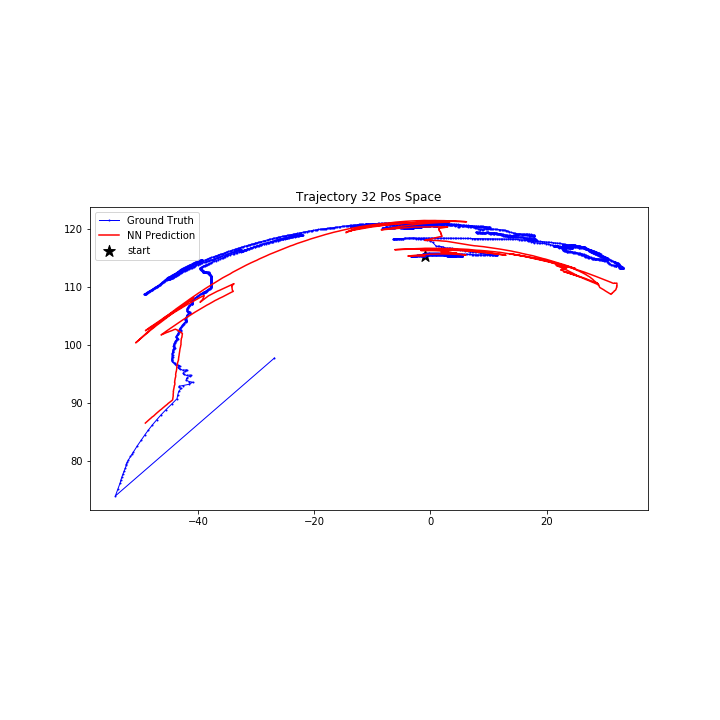
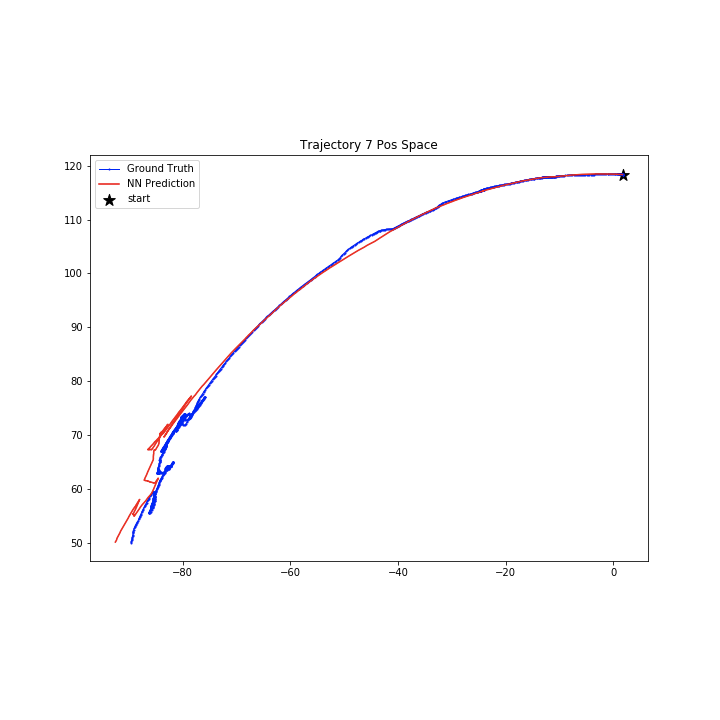
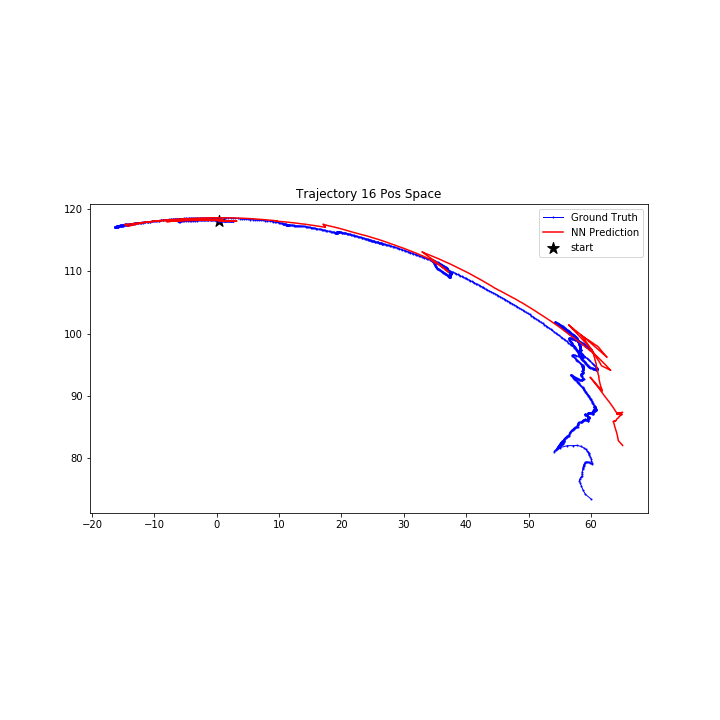
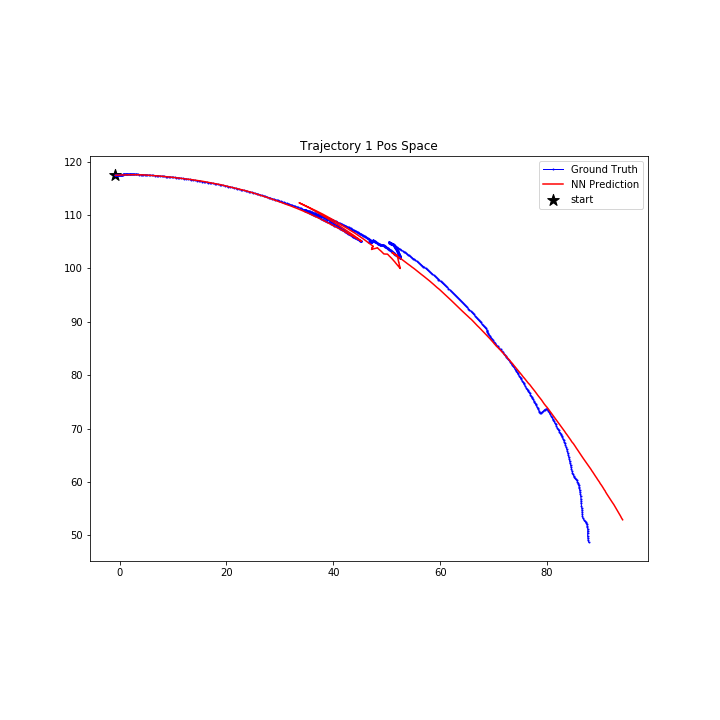
A close up of a device

Description automatically generatedA close up of a map

Description automatically generatedA close up of text on a black background

Description automatically generated

* Gazebo Hand(100% Model):



* Gazebo Hand(0.1% Model):

A close up of a map

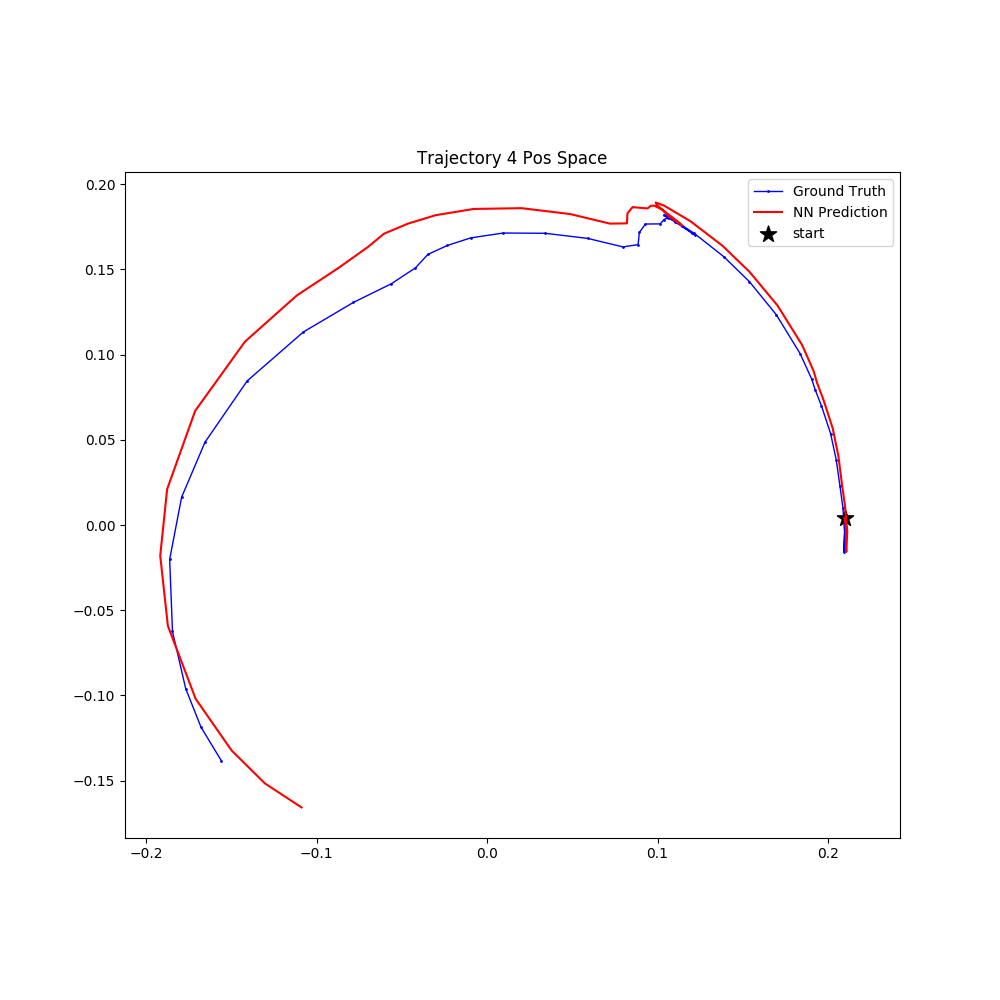
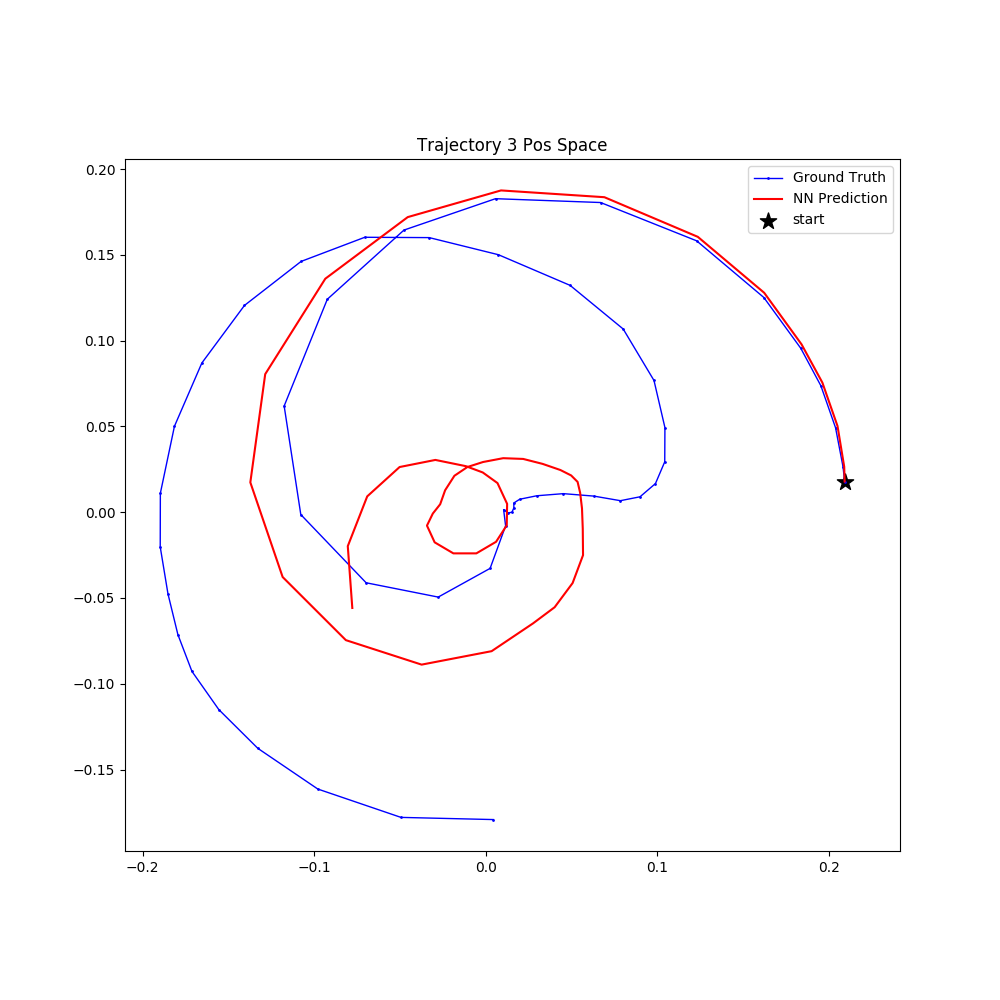
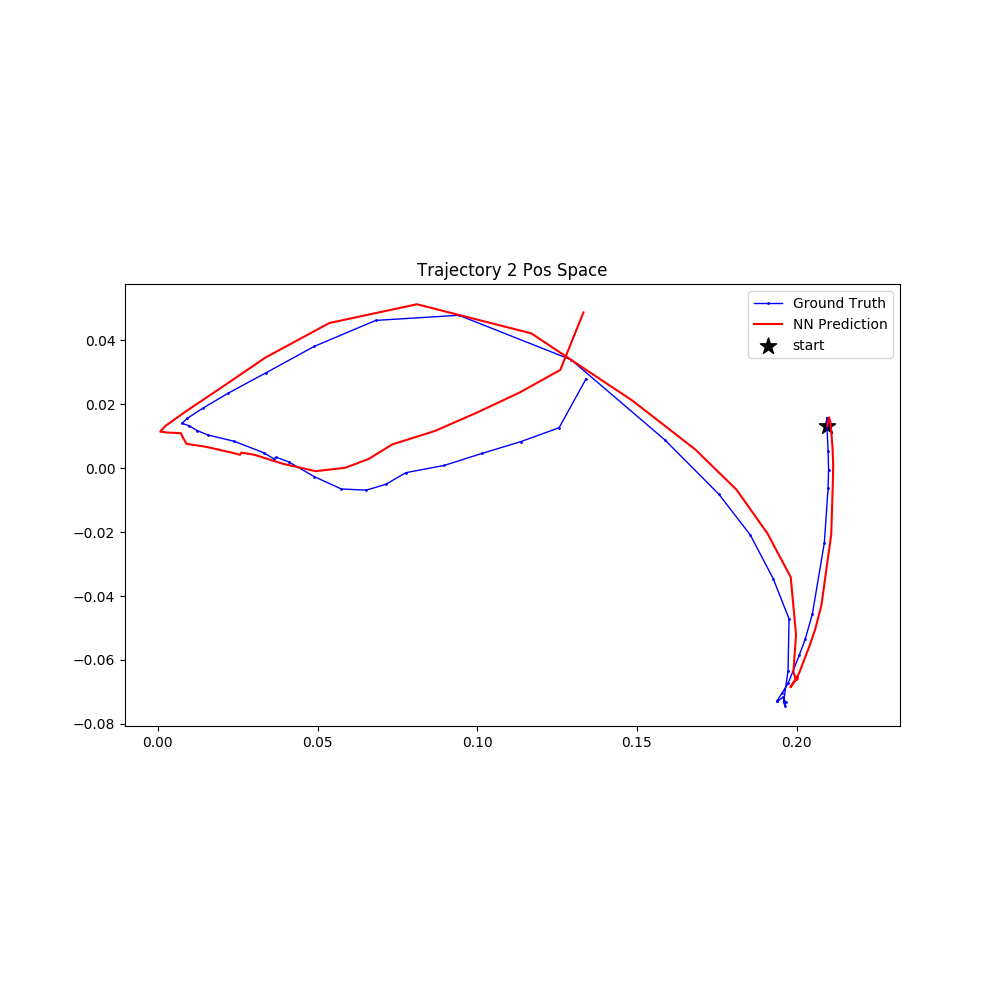
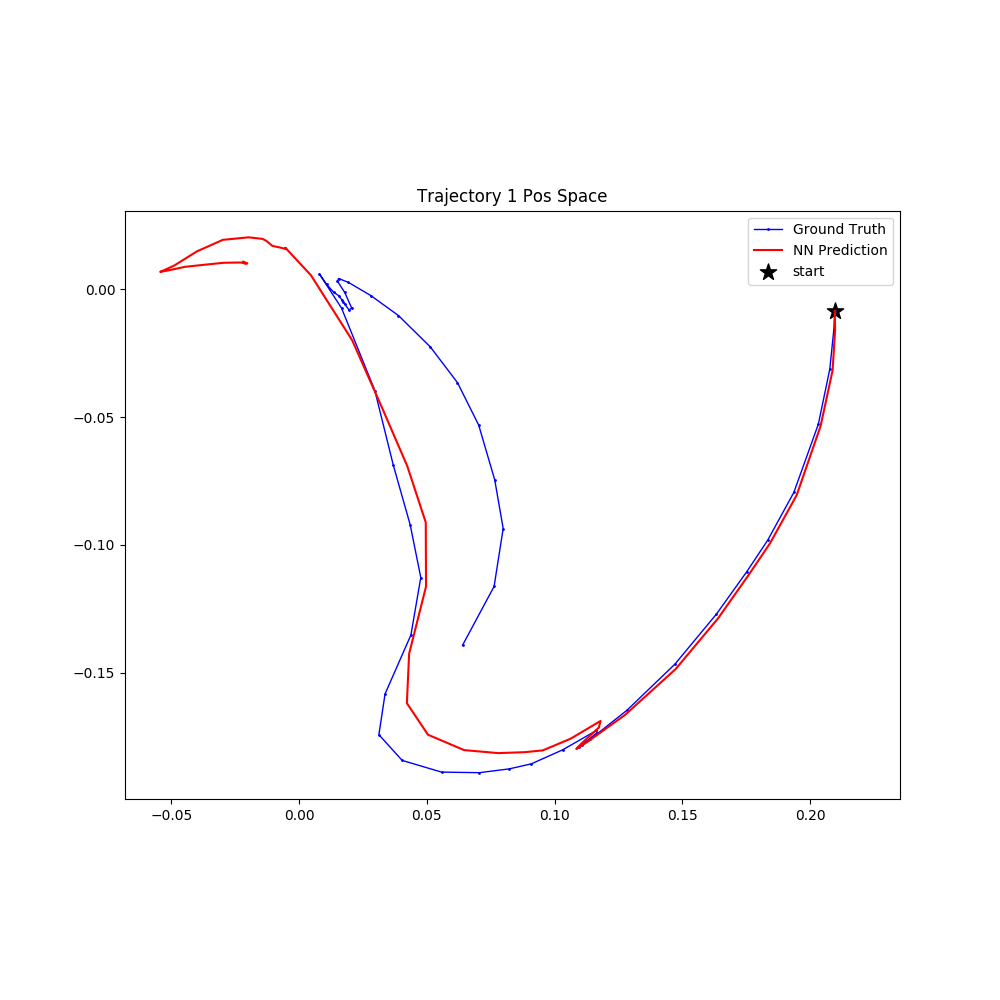
Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

* Reacher(100% Model):



* Reacher(100% Model):

A close up of a map

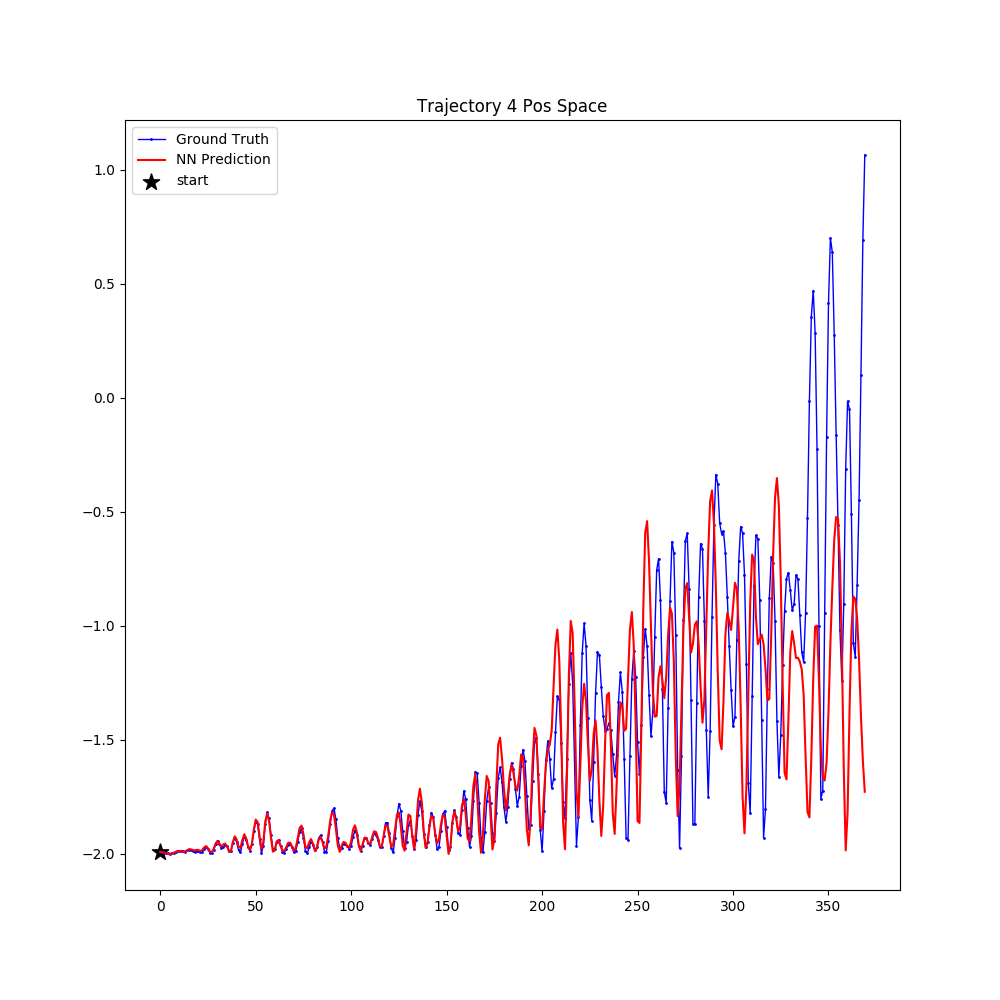
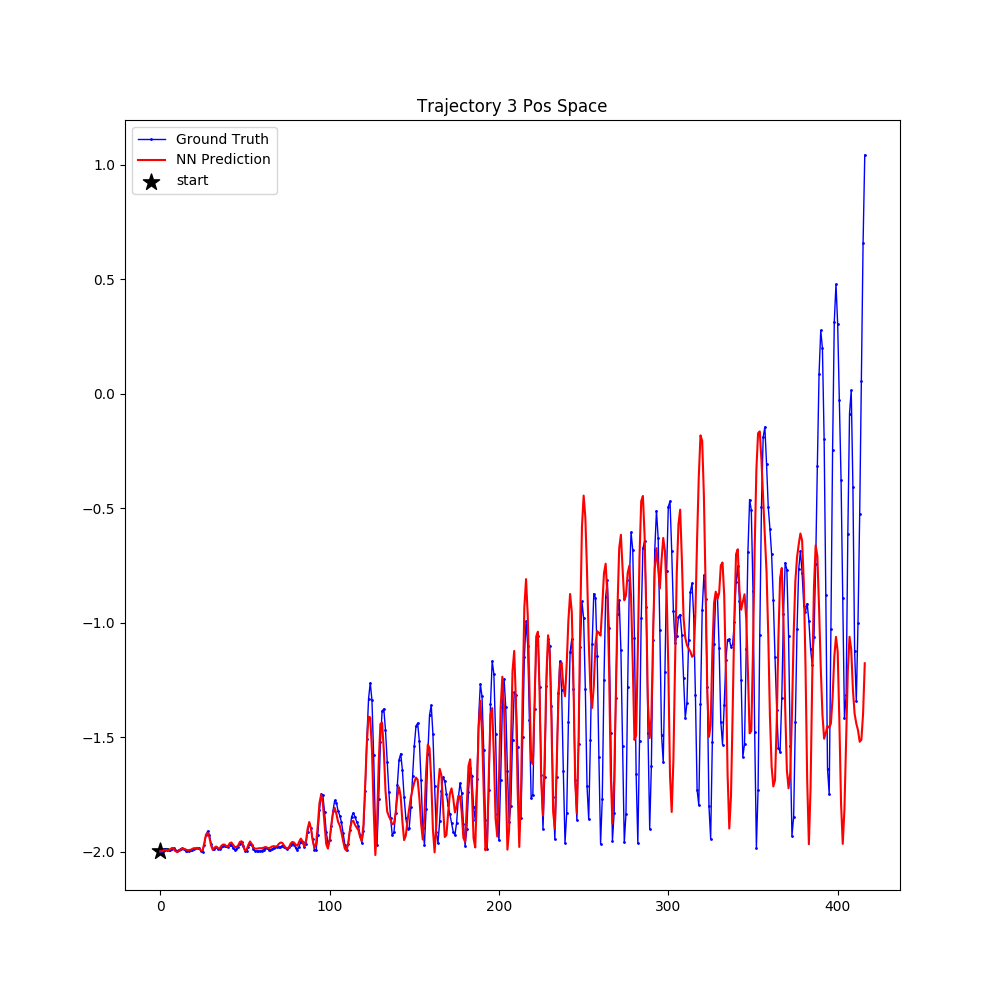
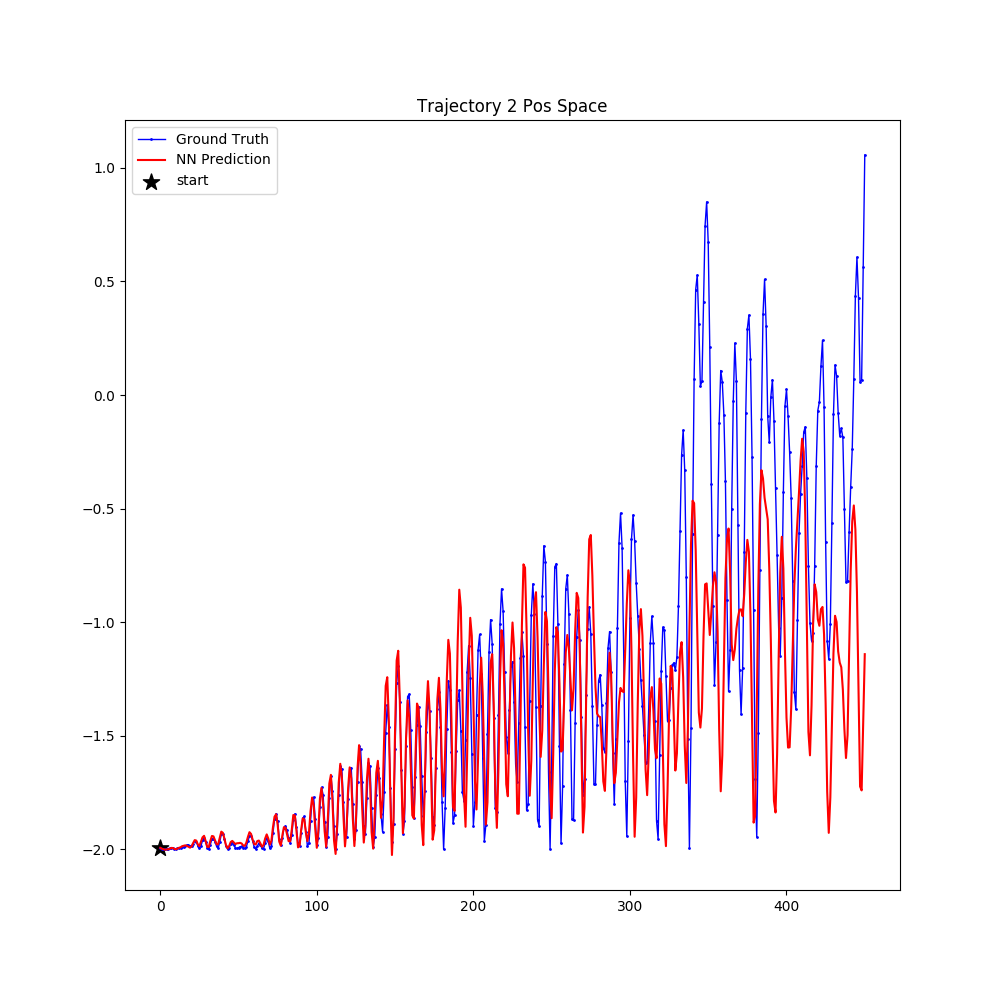
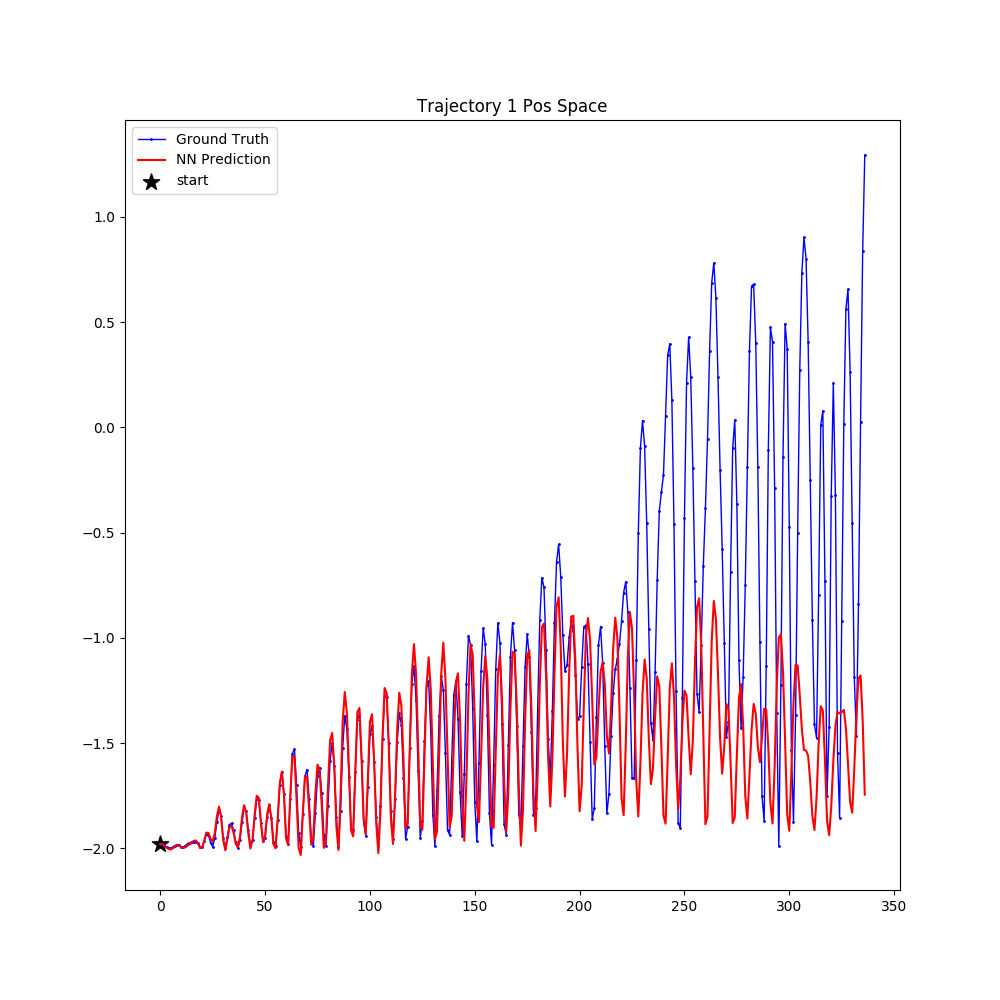
Description automatically generatedA close up of a map

Description automatically generatedA close up of text on a white background

Description automatically generatedA close up of a map

Description automatically generated

* Acrobot(100% Model):



**2. Gazebo Hand(100% Model):**

* Goal reach rate:

Table 4. Gazebo Hand (100% Model)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Goal Location | 0 | 2 | 7 | 8 | 15 | horseshoe |
| A\* | 0% | 100% | 100% | 100% | 100% | 90% |
| PPO | 100% | 100% | 60% | 40% | 20% | - |

* A\*:

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

* PPO (General Model: Goal location as a part of state; Continuous Reward):

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

**3. Reacher(100% Model):**

* Goal reach rate:

Table 2. Reacher (0.1% Model)

|  |  |  |  |
| --- | --- | --- | --- |
| Goal Location | 1 | 2 | 5 |
| A\* | 100% | 100% | 100% |
| PPO | 60% | 0% | 0% |

* A\*:

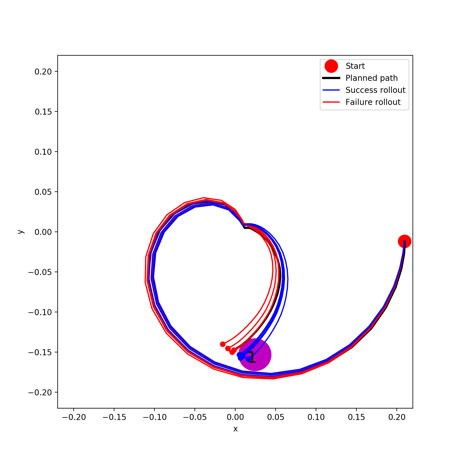
A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

* PPO (General Model: Goal location as a part of state; Continuous Reward):

A screenshot of a cell phone

Description automatically generatedA close up of a logo

Description automatically generated