

Data collection was completed by applying a random torque on each of the links and getting the resulting data from the teacher dynamics model. The random torque range was chosen to lie within the range of ± 2 . The initial state of the robot arm was set based on the starting location of the arm at $-\pi/2, 0$ for position and $0, 0$ for angular velocities.

Once the starting state and the torque was acquired, the dynamics were run for the duration of the 5s (optionally for shorter or longer depending on the passed argument). The one link data was collected for 5s intervals and the two link data was collected for 0.01s (i.e. 1 time-step) since that seemed to work better for the two link model.

For each X, the state and action were concatenated as one data point. Each Y was the resultant state based on advancing the dynamics from action and state in X. Constant torque was applied per each step.

The network architecture of the 1-link, 2-link and 3-link all consisted of fully-connected linear layers with the ReLU activation function as shown below:

1-link	2-link	3-link
Layer 1 = Linear (3, 64)	Layer 1 = Linear (6, 256)	Layer 1 = Linear (9, 256)
Layer 2 = Linear (64, 128)	Layer 2 = Linear (256, 512)	Layer 2 = Linear (256, 128)
Layer 3 = Linear (128, 256)	Layer 3 = Linear (512, 128)	Layer 3 = Linear (128, 64)
Layer 4 = Linear (256, 128)	Layer 4 = Linear (128, 18)	Layer 4 = Linear (64, 32)
Layer 5 = Linear (128, 2)	Layer 5 = Linear (18, 4)	Layer 5 = Linear (32, 6)

Model Metadata:

- The best model for the 1link model was found at epoch # 1/2 with a training loss of 0.01097369 and testing loss of 0.01438949. Tested model as it was being trained and got a positive result so stopped training early.
- The best model for the 2link model was found at epoch # 4/7 with a training loss of 0.00001579 and testing loss of 0.00001318.
- The best model for the 3link model was found at epoch # 9/50 with a training loss of 0.00001573 and testing loss of 0.00001189.