

## Robot Learning Spring 2022

Sae Na Na  
DUV Reticle Stage MTD

April 26, 2024  
Wilton CT

# I took a Robot Learning class this spring

## An introduction to machine learning in robotics

- Supervised and unsupervised learning
- Deep learning
- Forward models
- Reinforcement learning

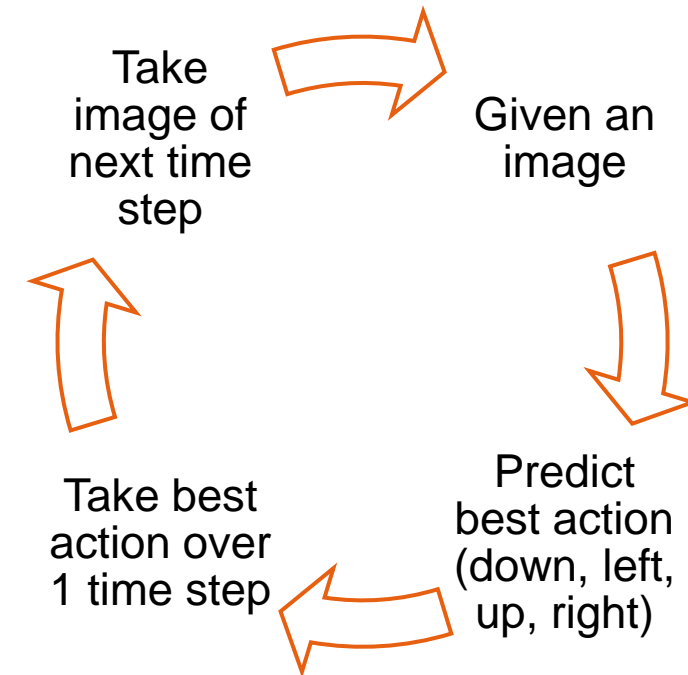
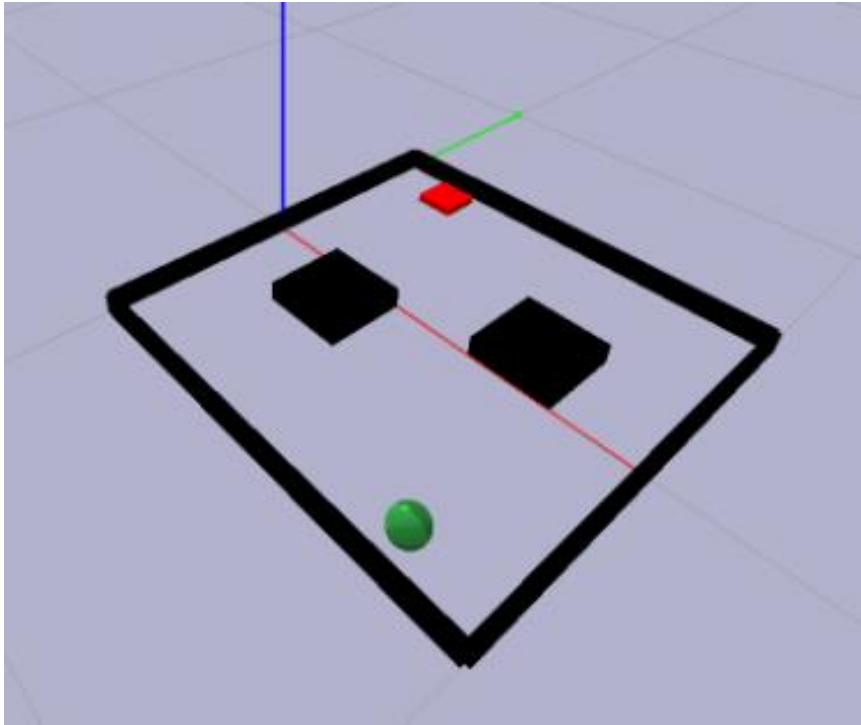
### *Structure of the class*

- lectures on theory + recent paper discussions
- 5 programming projects in Python

# Behavioral cloning

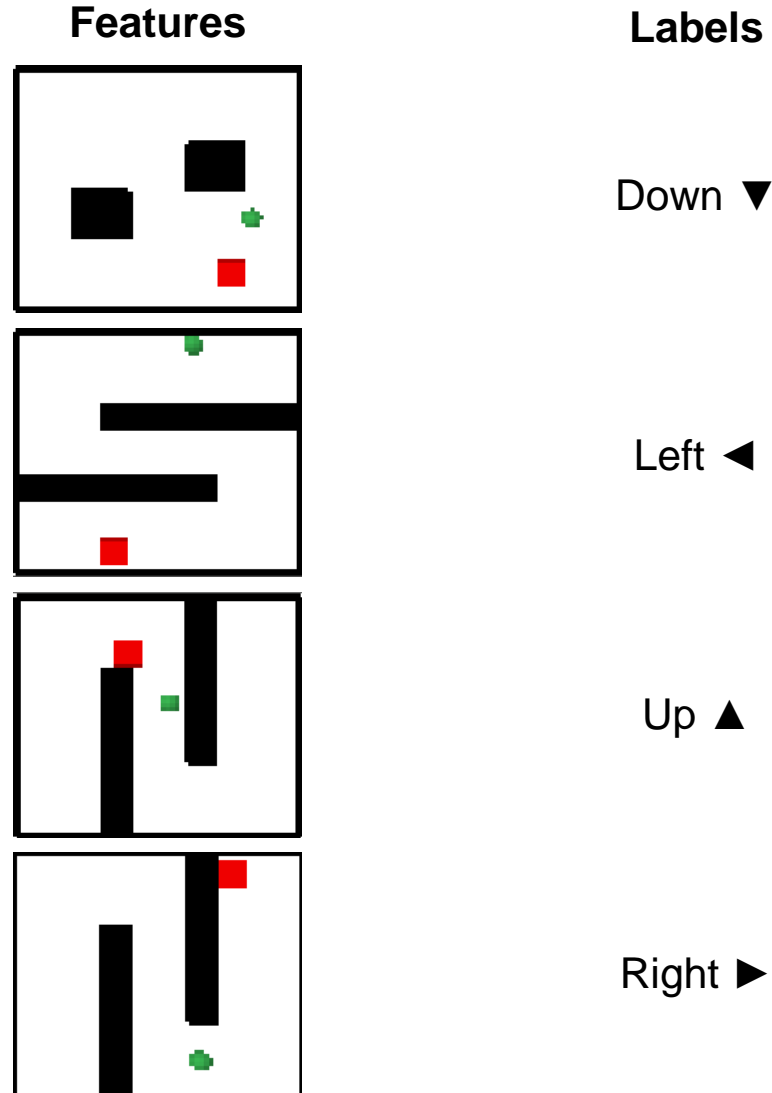
# Behavioral cloning

Given an RGB image of the world, choose the best action to reach the goal



# Behavioral cloning

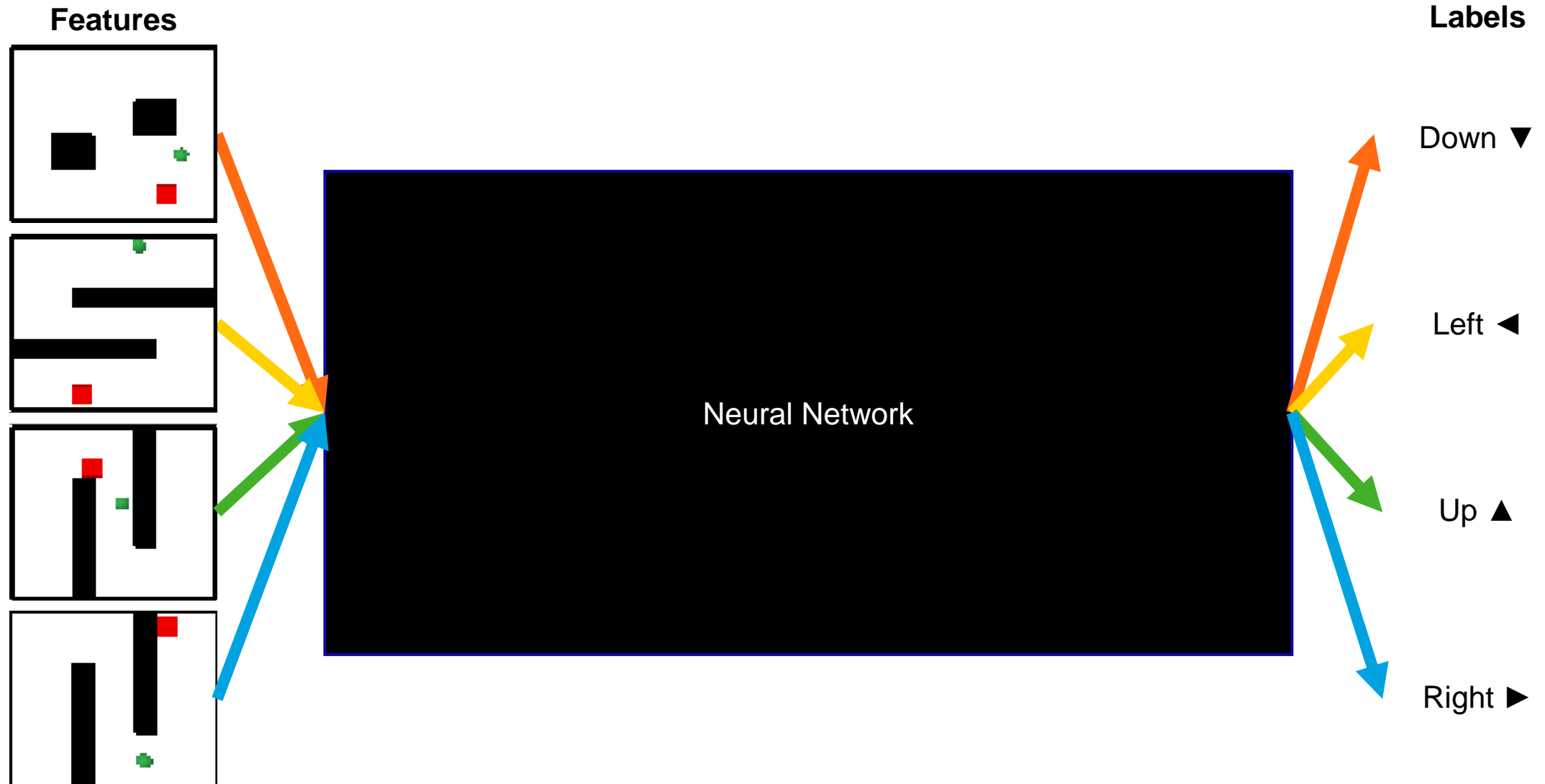
Given an RGB image of the world, choose the best action to reach the goal



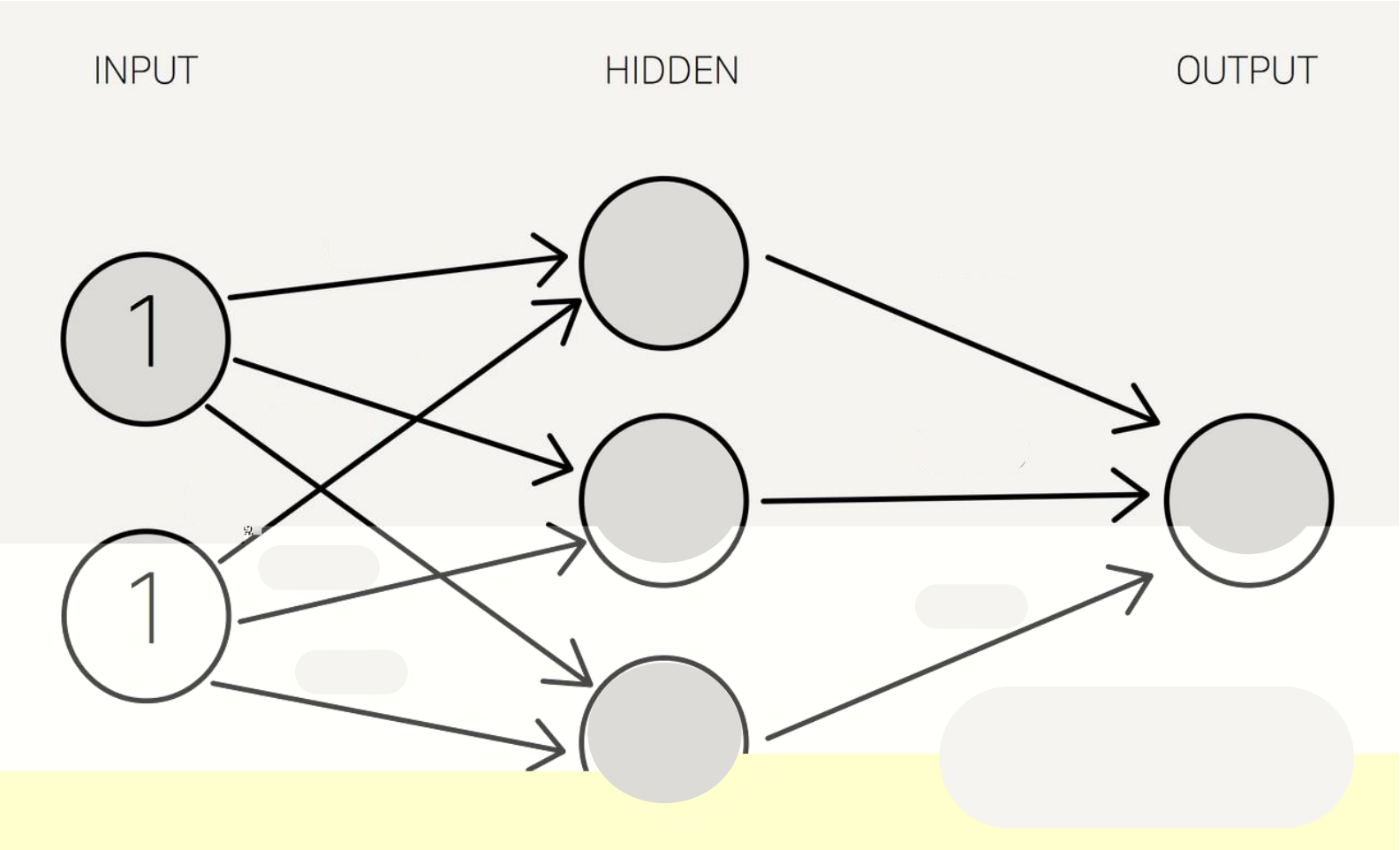
A dataset with 12000  
feature-label pairs is  
*provided to us*

# Behavioral cloning

Given an RGB image of the world, choose the best action to reach the goal

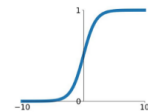


# Neural Networks

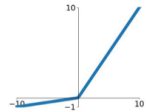


## Activation Functions

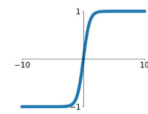
**Sigmoid**  
 $\sigma(x) = \frac{1}{1+e^{-x}}$



**Leaky ReLU**  
 $\max(0.1x, x)$



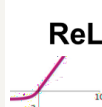
**tanh**  
 $\tanh(x)$



**Maxout**  
 $\max(w_1^T x + b_1, w_2^T x + b_2)$

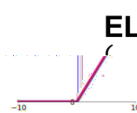
**ReLU**

$\max(0, x)$



**ELU**

$\begin{cases} x^n & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$

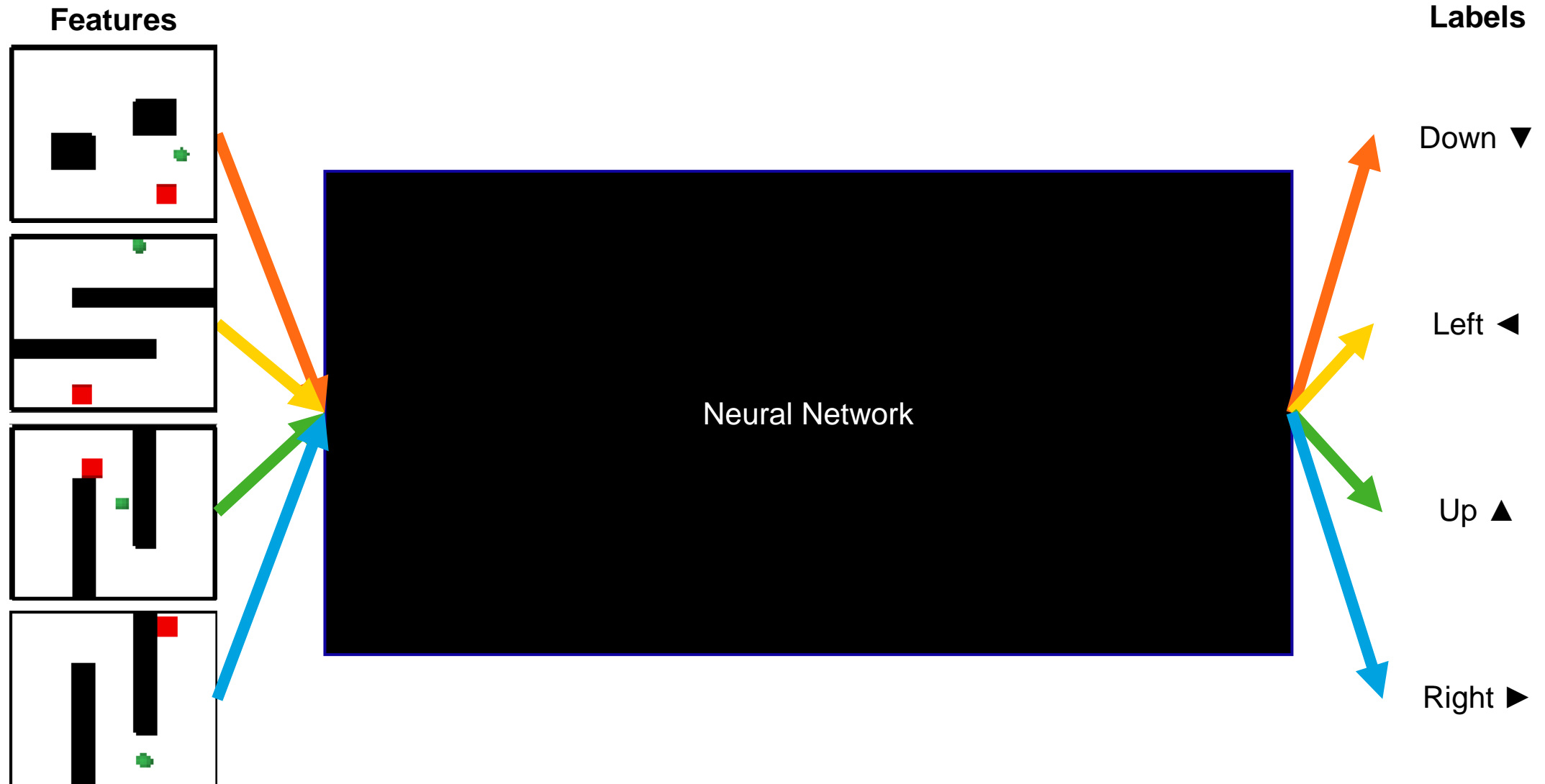


Forward propagation

Backpropagation

# Behavioral cloning

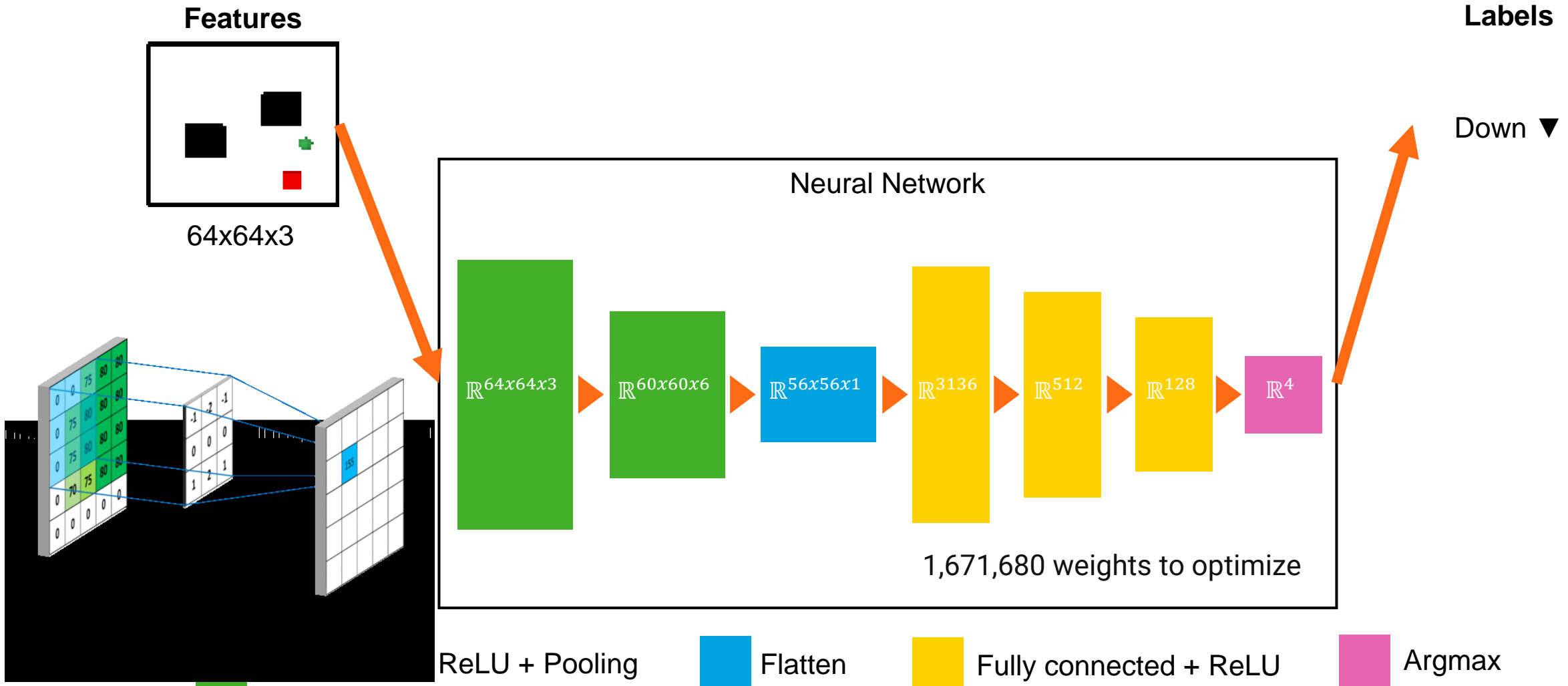
Given an RGB image of the world, choose the best action to reach the goal





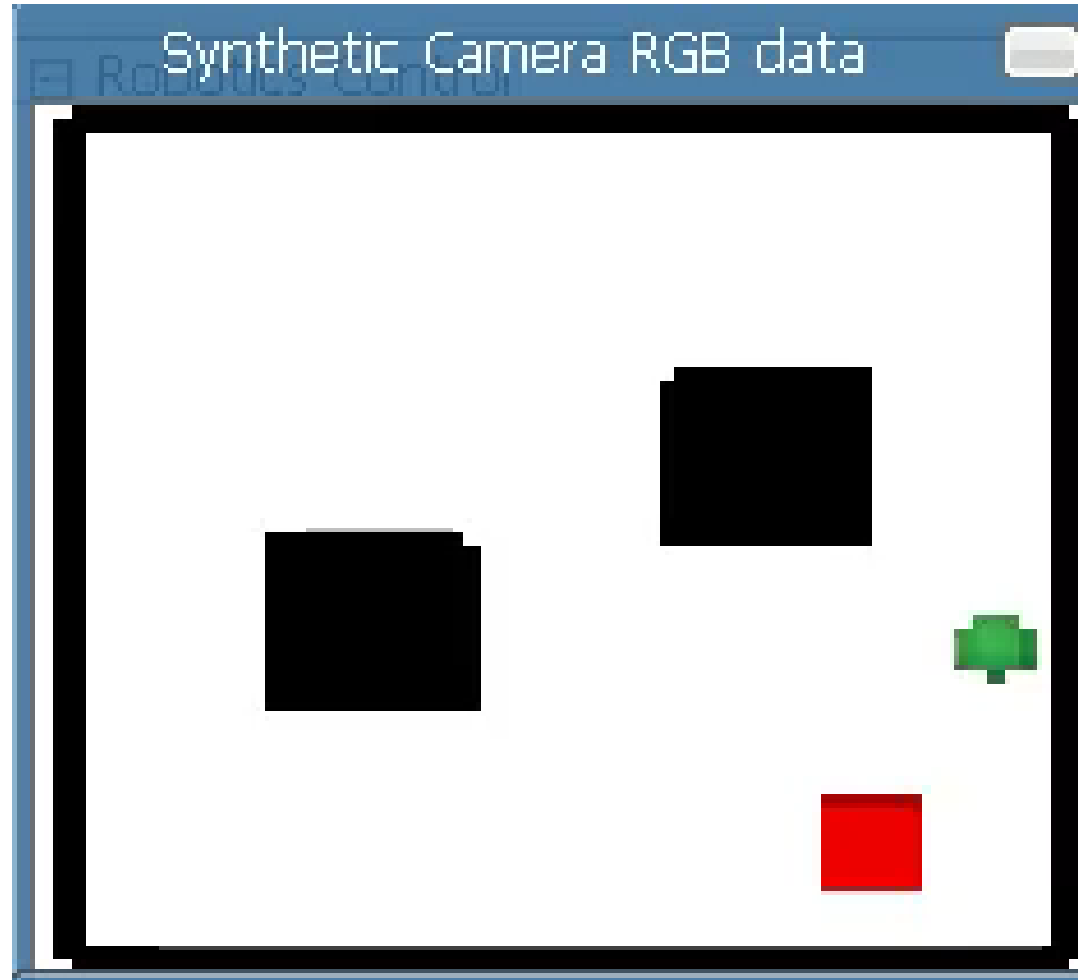
# Behavioral cloning

Given an RGB image of the world, choose the best action to reach the goal



# Behavioral cloning

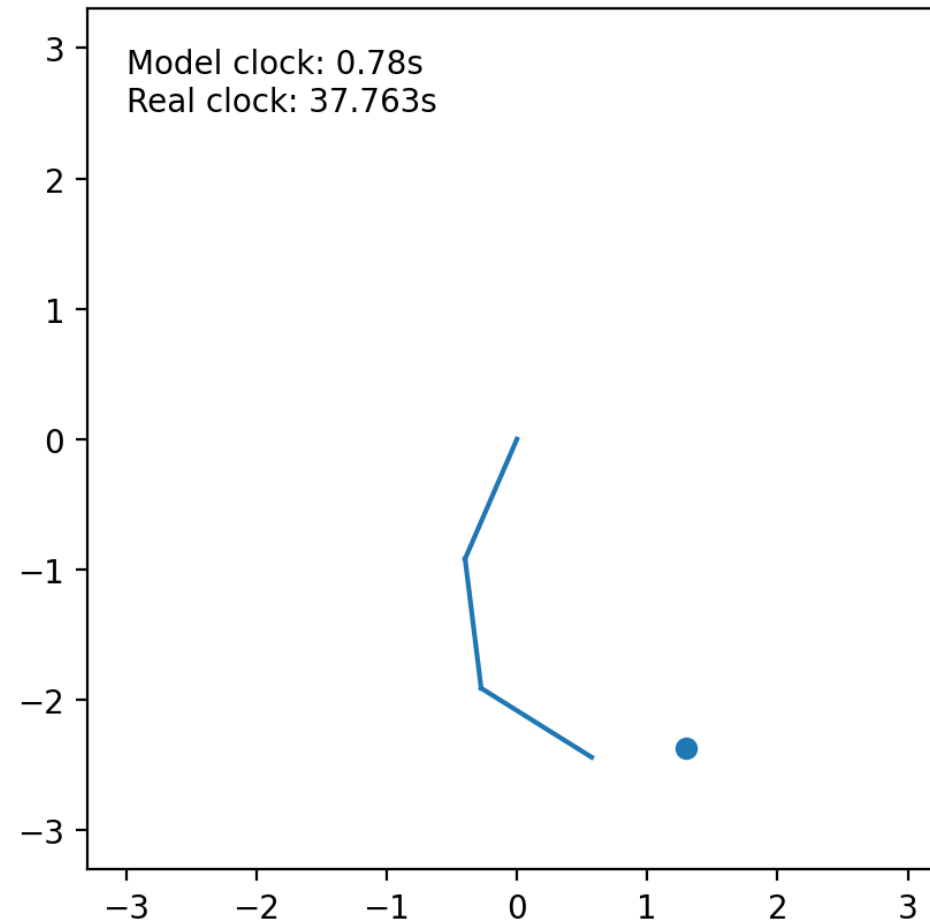
Given an RGB image of the world, choose the best action to reach the goal



# **Model Predictive Control, and learning a forward model**

# MPC and learning a forward model

Get the end-effector of the 1-, 2-, and 3-link pendulum to reach the goal within a threshold

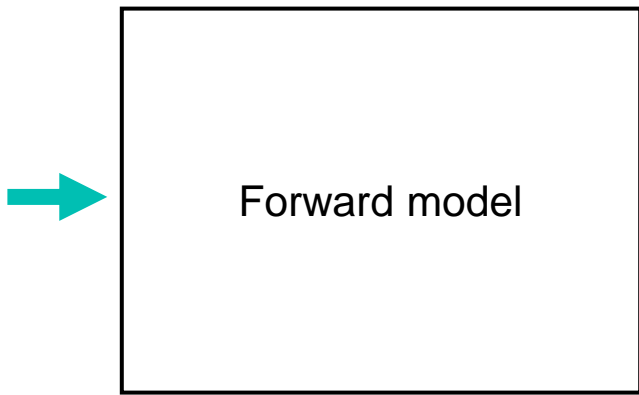


# MPC and learning a forward model

Get the end-effector of the 1-, 2-, and 3-link pendulum to reach the goal within a threshold

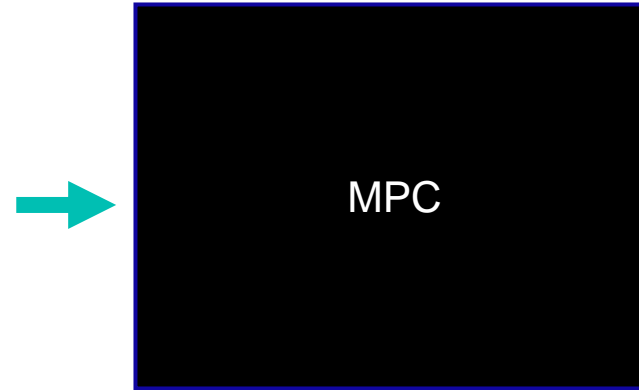
## State and action

$$x_k = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix}$$
$$u_k = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$



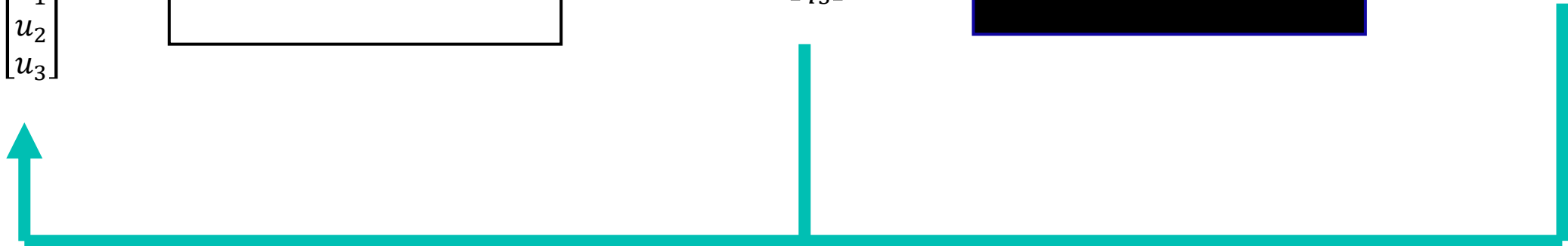
## Next state

$$x_{k+1} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix}$$



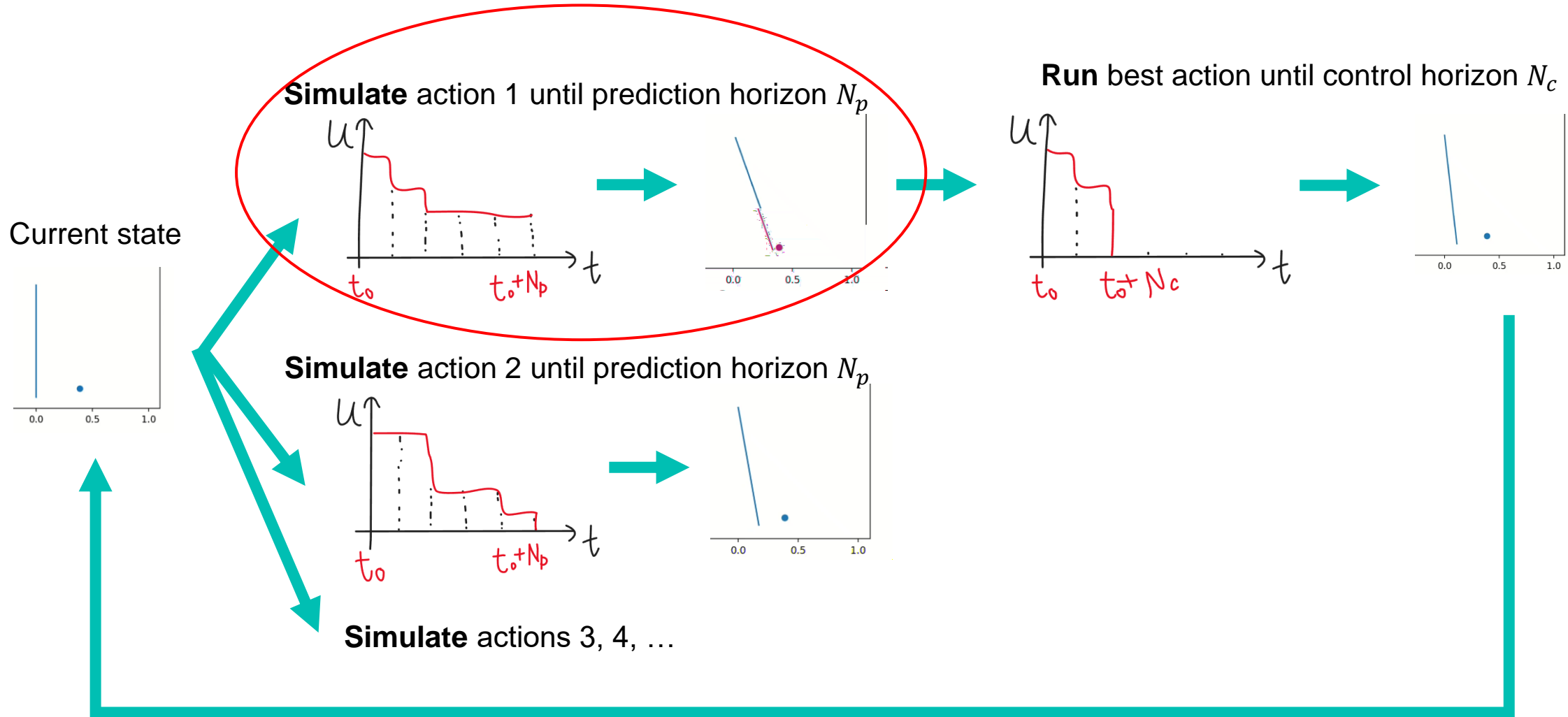
## Best action

$$u_{k+1} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$



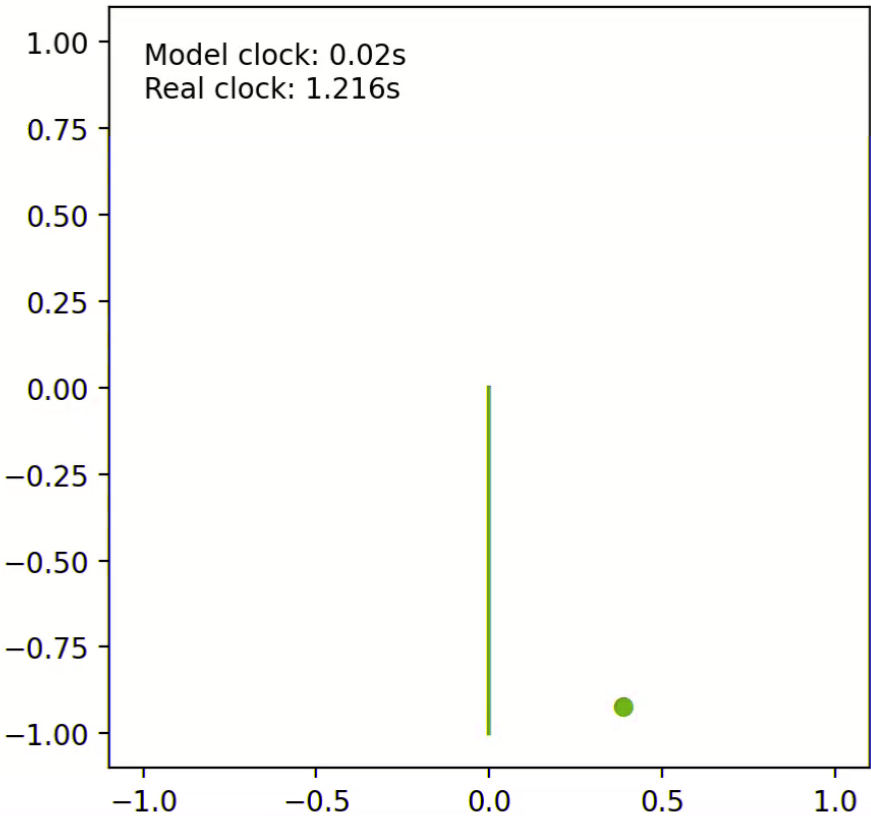
# Model Predictive Control (MPC)

Can be MIMO, considers future states, needs knowledge of forward model



# MPC and learning a forward model

Get the end-effector of the 1-, 2-, and 3-link pendulum to reach the goal within a threshold



# links	Prediction horizon	Control horizon	Delta u
1	50	10	0.1
2	40	2	0.1
3	90	10	0.2, 0.1

# MPC and learning a forward model

Get the end-effector of the 1-, 2-, and 3-link pendulum to reach the goal within a threshold

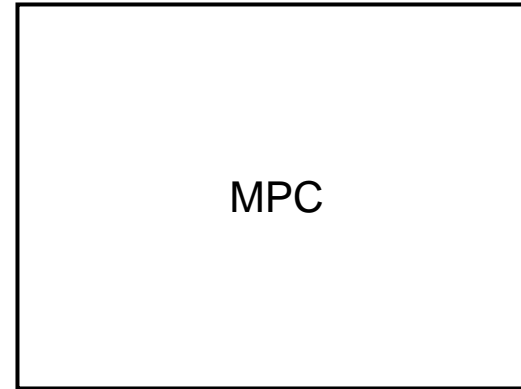
State and action

$$x_k = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix}$$
$$u_k = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$



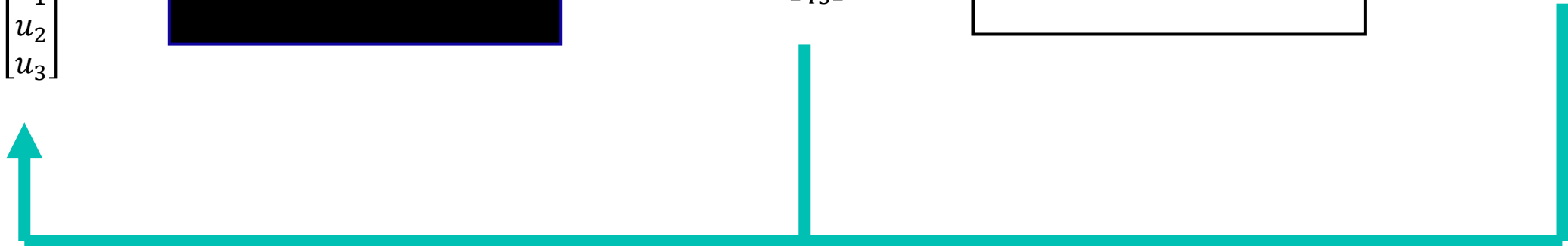
Next state

$$x_{k+1} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix}$$



Best action

$$u_{k+1} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$





# Forward model

Assume Markovian

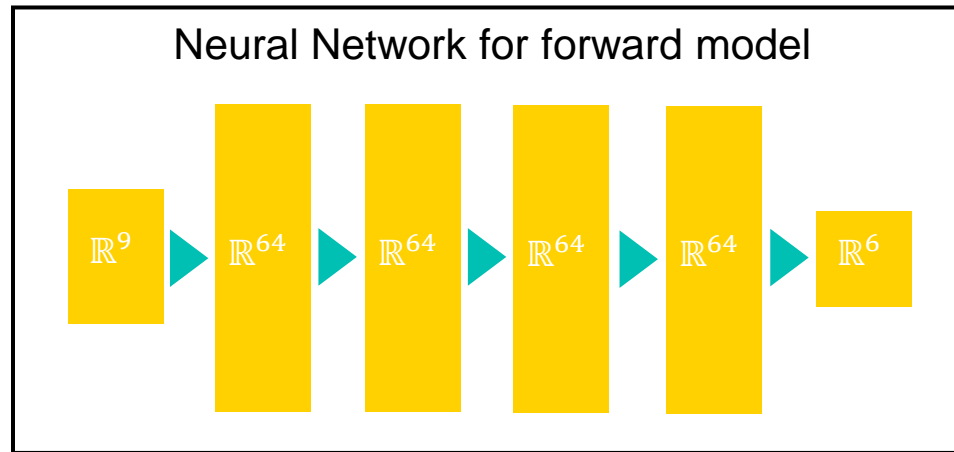
$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k)$$

- $\mathbf{x}_k$  - state of the system at time step  $k$
- $\mathbf{u}_k$  - action / command given to the robot at time step  $k$
- $\mathbf{x}_{k+1}$  - state of the system at the next time step,  $k+1$
- Using discretized time - size of time step  $\Delta t$  not explicitly part of forward model, but often very important

State and action

$$x_k = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix}$$

$$u_k = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$



Fully connected + ReLU

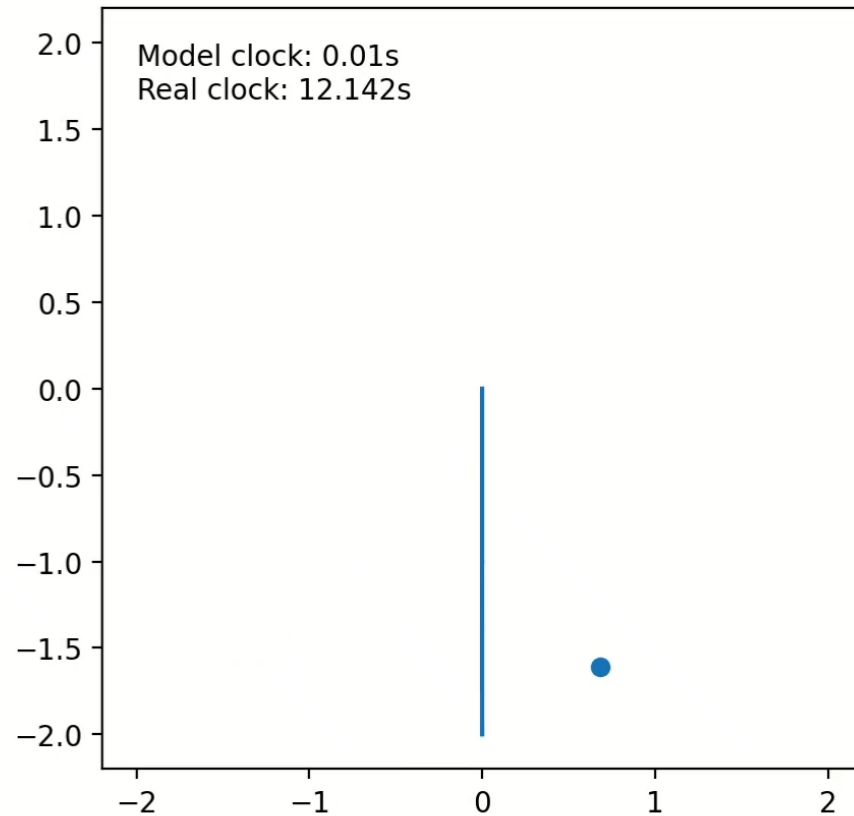
Next state

$$x_{k+1} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix}$$

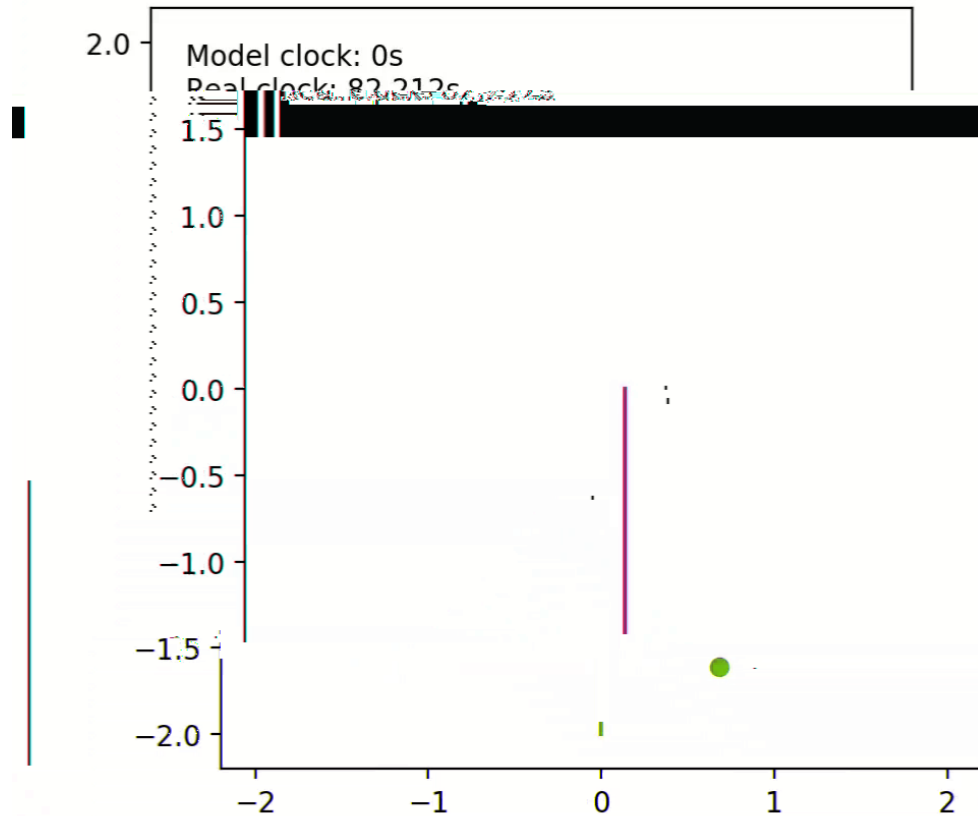
# MPC and learning a forward model

Get the end-effector of the 1-, 2-, and 3-link pendulum to reach the goal within a threshold

Analytical forward model



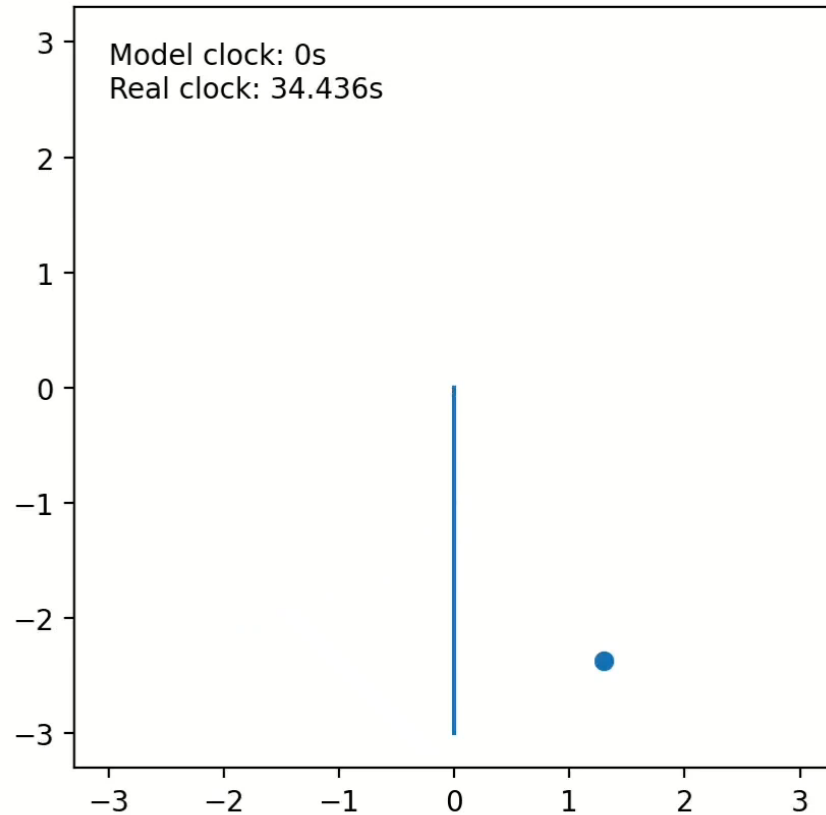
Neural network forward model



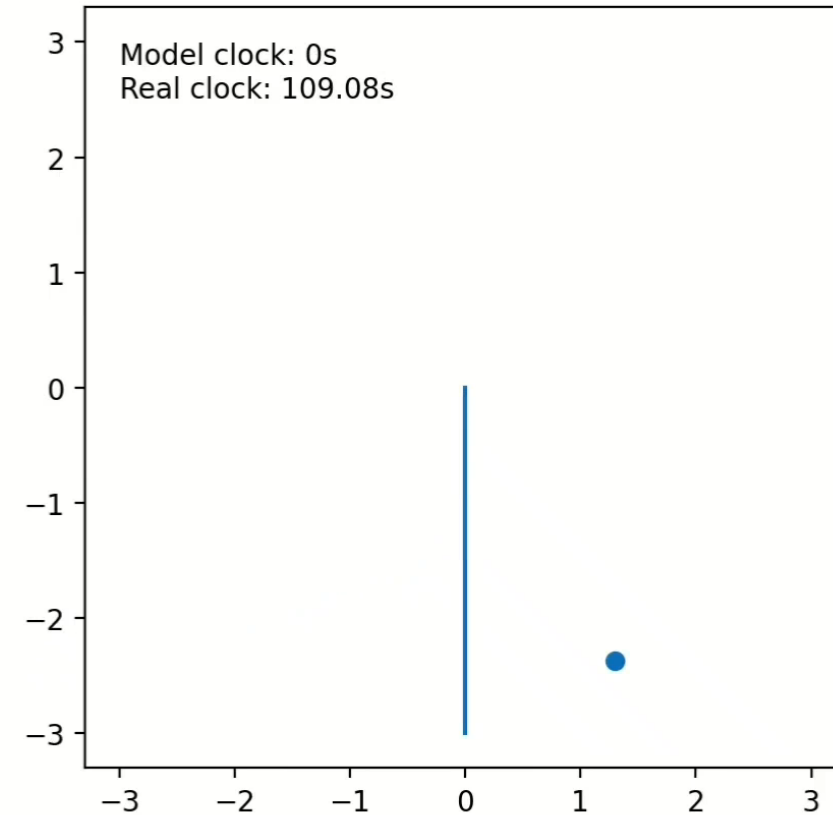
# MPC and learning a forward model

Get the end-effector of the 1-, 2-, and 3-link pendulum to reach the goal within a threshold

Analytical forward model



Neural network forward model



# Thanks



Sae Na Na  
DUV Reticle Stage MTD



Wilton CT