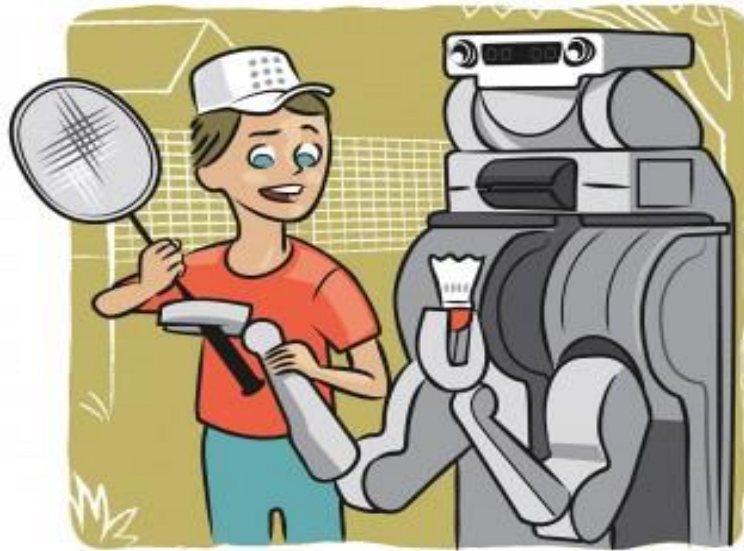
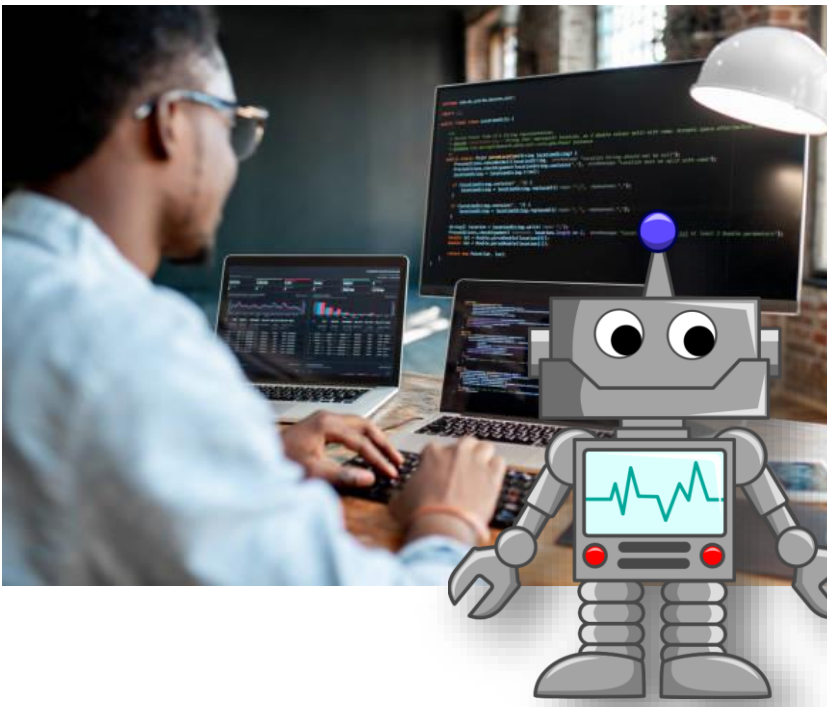


# Behavioral Cloning and Interactive Imitation Learning



Instructor: Daniel Brown

[Some slides adapted from Sergey Levine (CS 285) and Alina Vereshchaka (CSE4/510)]



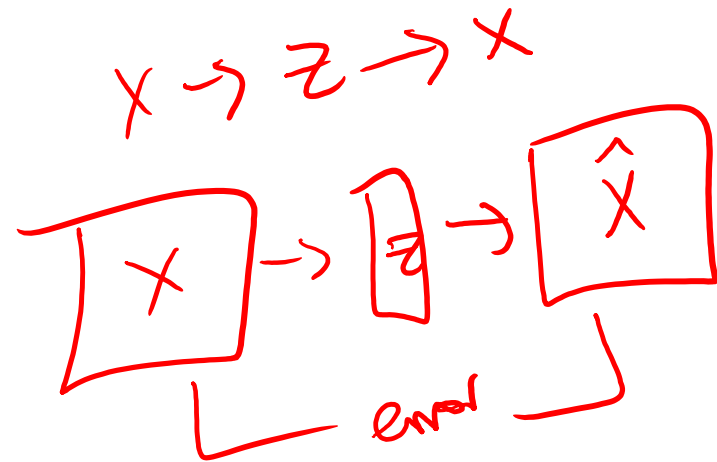


# Brief Machine Learning Refresher

There are roughly 3 main branches of machine learning

- Supervised Learning — labeled data  $(x, y)$
- Unsupervised Learning — structure representation learning
- Reinforcement Learning — trial & error learning, reward signal

$$f_{\theta}(x) \rightarrow y$$



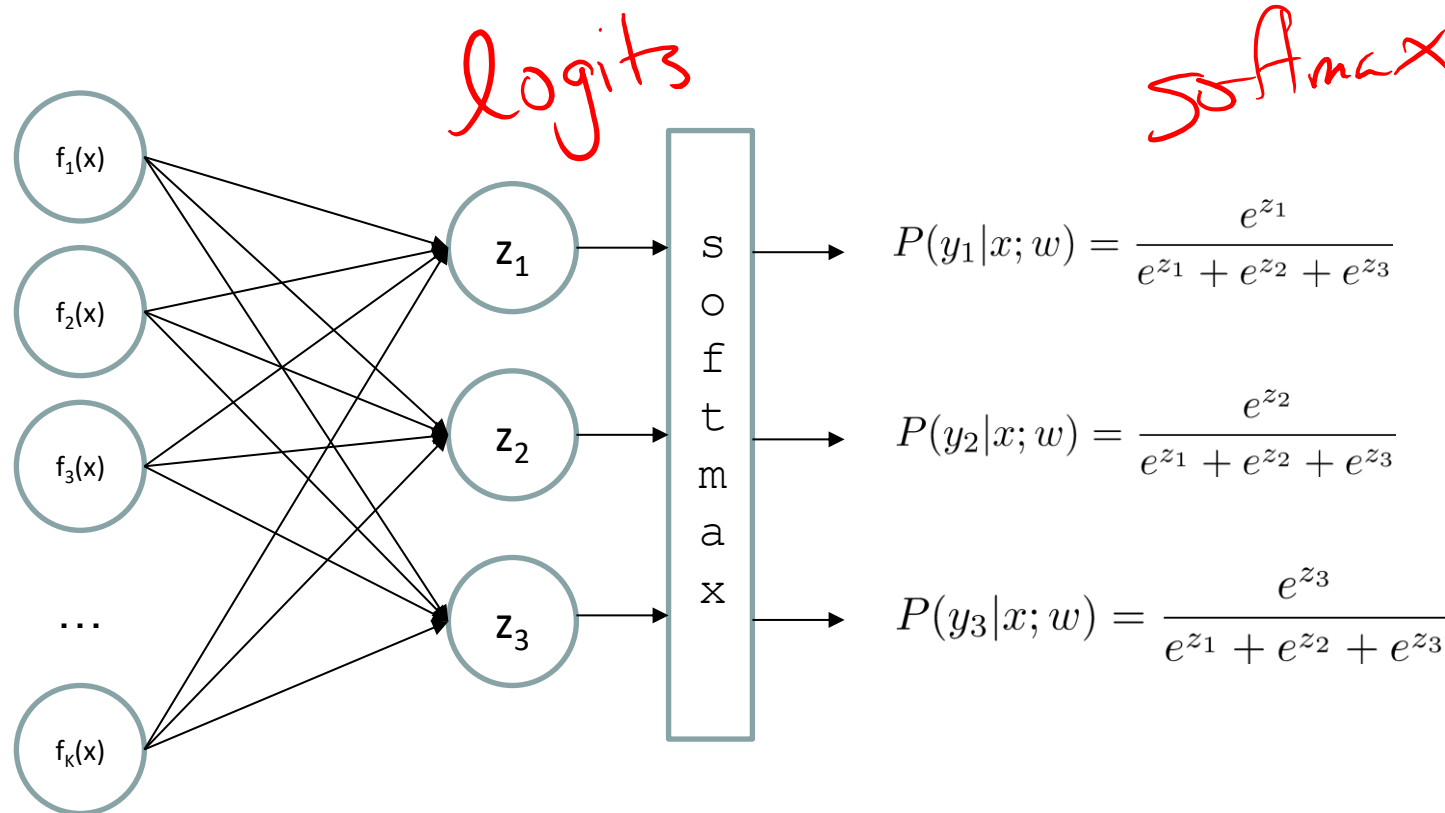
# Supervised Learning

- **Setting/Assumptions:** In supervised learning, the model is trained on labeled data, where the input data is paired with the correct output (i.e., the "ground truth").
- **Goal:** To learn a mapping from inputs to outputs so that the model can predict the output for new, unseen inputs.
- **Common Use Cases:**
  - Classification (e.g., spam email detection, image recognition).
  - Regression (e.g., predicting house prices, stock market trends).
- **Example models:**
  - Linear regression, decision trees, support vector machines, and neural networks.

# Multi-class Logistic Regression

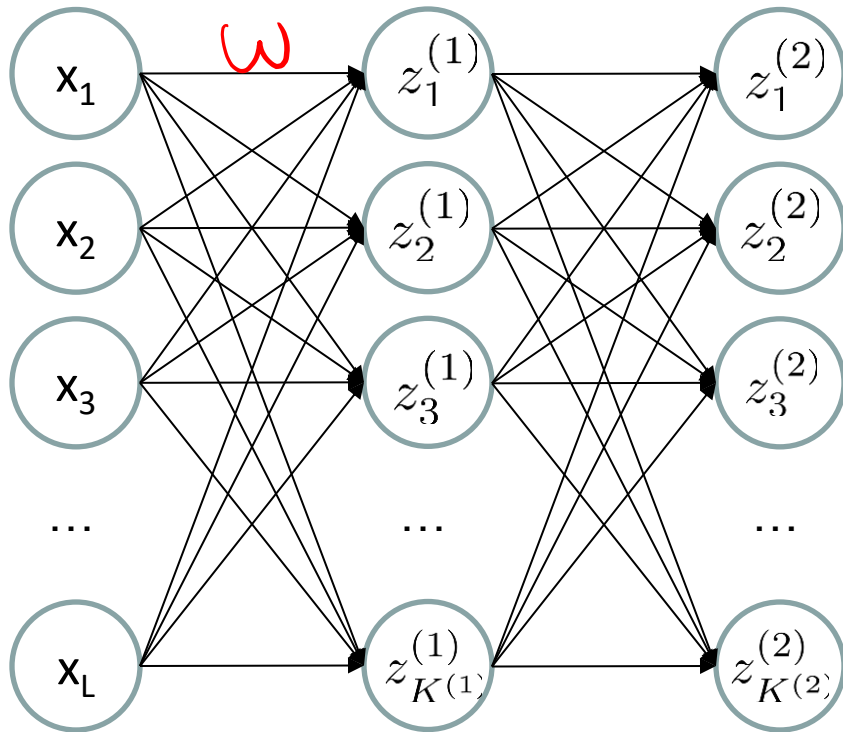
- = special case of neural network

# classes = 3

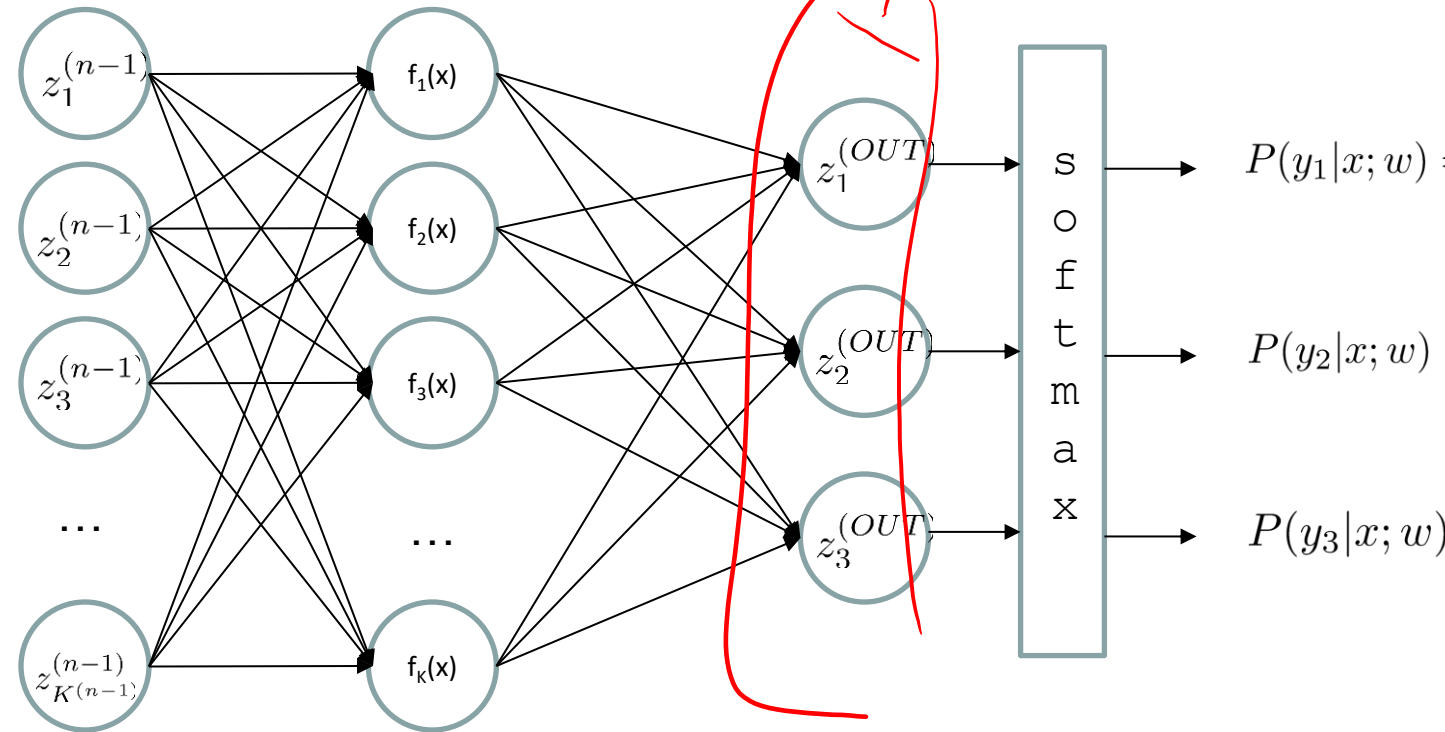


# Deep Neural Network = Also learn the features!

*g =*  *logits*



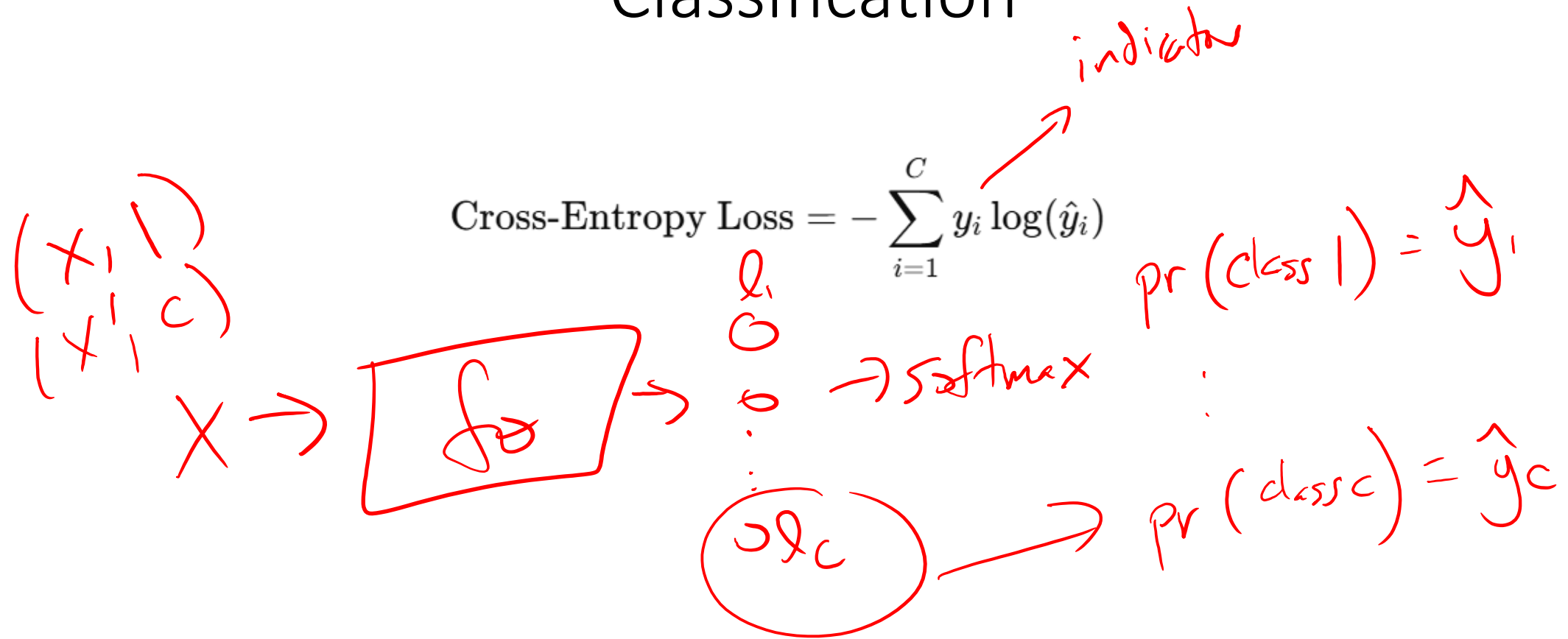
...



$$z_i^{(k)} = g\left(\sum_j W_{i,j}^{(k-1,k)} z_j^{(k-1)}\right)$$

**g = nonlinear activation function**

# Classification



# PyTorch Example

```
import torch.nn as nn
import torch.optim as optim
```

```
class ClassificationNetwork(nn.Module):
    def __init__(self, input_dim, num_classes):
        super(ClassificationNetwork, self).__init__()
        self.fc = nn.Linear(input_dim, num_classes)
```



fc1  
fc2

```
    def forward(self, x):
        return self.fc(x)
```

```
model = ClassificationNetwork(input_dim, num_classes)
criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(), lr=0.001)
```

```
for epoch in range(num_epochs):
    for inputs, labels in dataloader:
        optimizer.zero_grad()
        outputs = model(inputs)
        loss = criterion(outputs, labels)
        loss.backward()
        optimizer.step()
```



# PyTorch Example (MLP)

```
import torch.nn as nn
import torch.optim as optim
```

```
class ClassificationNetwork(nn.Module):
```

```
    def __init__(self, input_dim, num_classes):
```

```
        super(ClassificationNetwork, self).__init__()
```

```
        self.fc1 = nn.Linear(input_dim, num_hidden)
```

```
        self.relu = nn.ReLU()
```

```
        self.fc2 = nn.Linear(num_hidden, num_classes)
```

← layers?

```
    def forward(self, x):
```

```
        return self.fc2(self.relu(self.fc1(x)))
```

```
model = ClassificationNetwork(input_dim, num_classes)
```

```
criterion = nn.CrossEntropyLoss()
```

```
optimizer = optim.Adam(model.parameters(), lr=0.001)
```

```
for epoch in range(num_epochs):
```

```
    for inputs, labels in dataloader:
```

```
        optimizer.zero_grad()
```

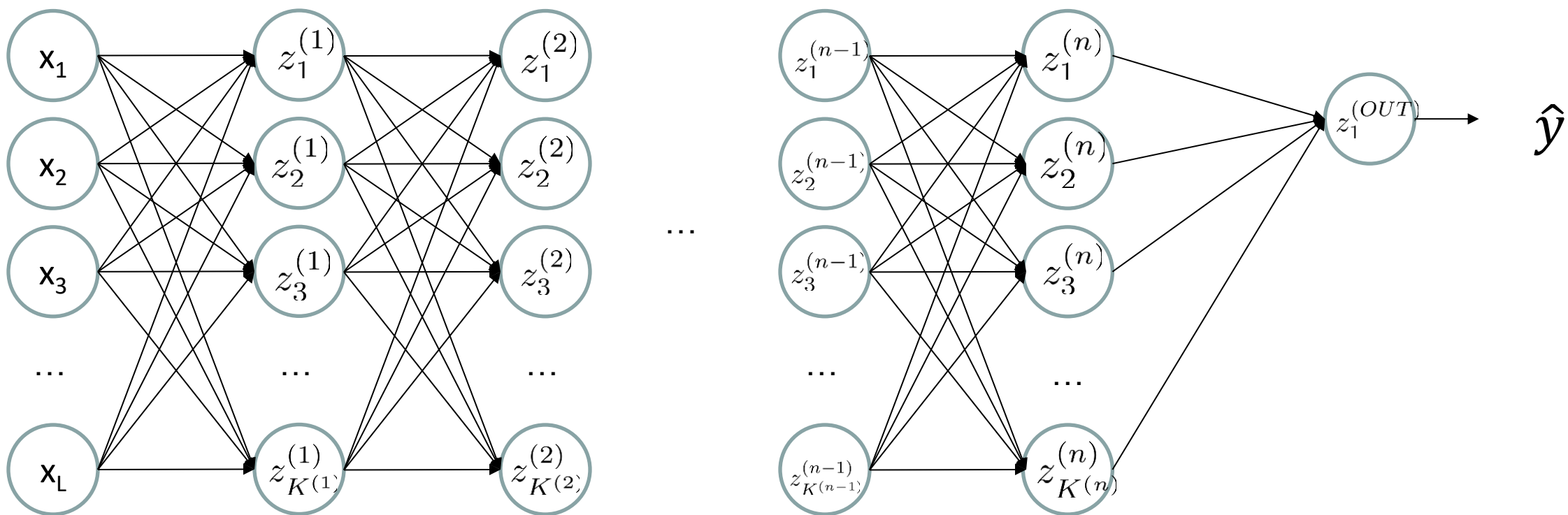
```
        outputs = model(inputs)
```

```
        loss = criterion(outputs, labels)
```

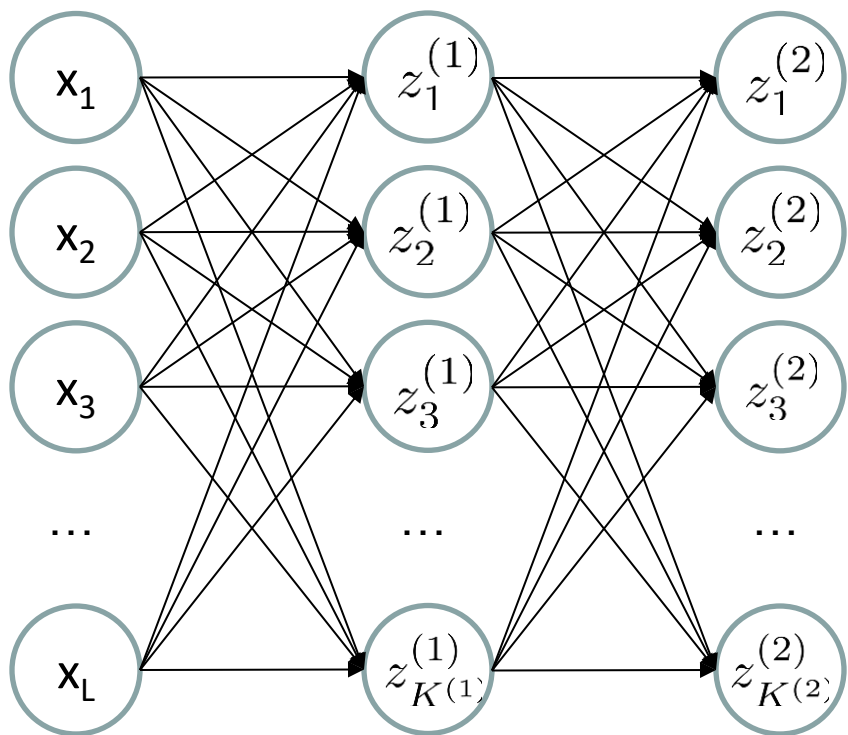
```
        loss.backward()
```

```
        optimizer.step()
```

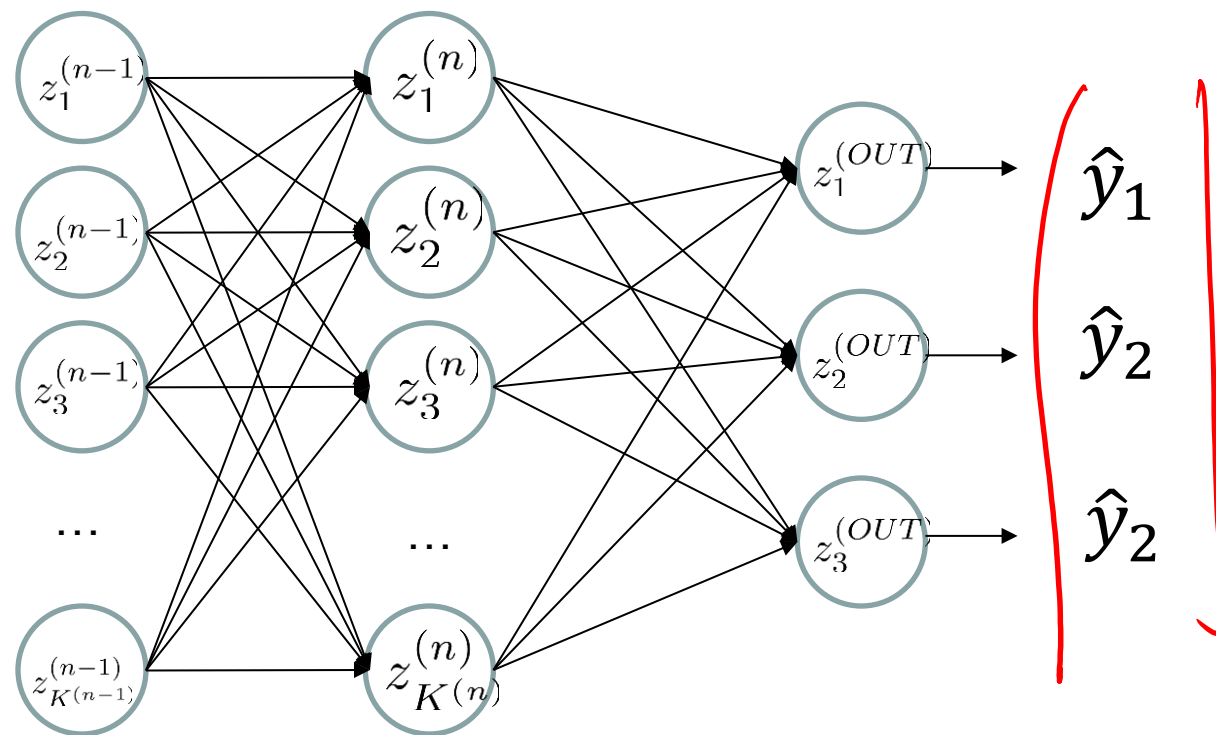
# Deep Neural Networks for Regression



# Deep Neural Networks for Regression



...



# Regression

$$\text{MSE Loss} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2$$

*pred* *true label*

↓ ↓

# PyTorch Example

```
import torch.nn as nn
import torch.optim as optim

class Classification Regression Network(nn.Module):
    def __init__(self, input_dim, num_classes):
        super(ClassificationNetwork, self).__init__()
        self.fc = nn.Linear(input_dim, num_classes)

    def forward(self, x):
        return self.fc(x)

model = ClassificationNetwork(input_dim, num_classes)
criterion = nn.CrossEntropyLoss() MSE
optimizer = optim.Adam(model.parameters(), lr=0.001)

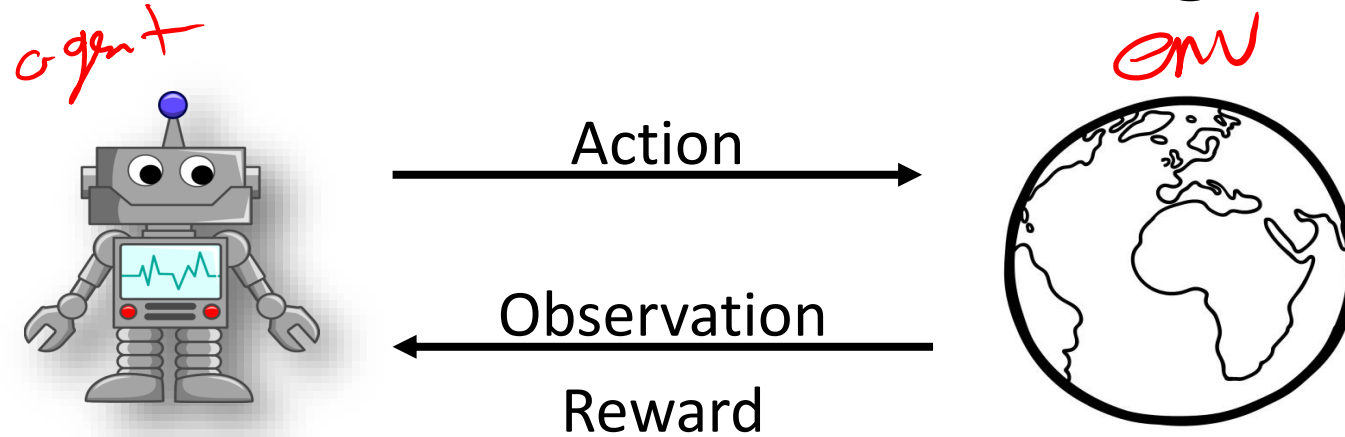
for epoch in range(num_epochs):
    for inputs, labels in dataloader:
        optimizer.zero_grad()
        outputs = model(inputs)
        loss = criterion(outputs, labels)
        loss.backward()
        optimizer.step()
```



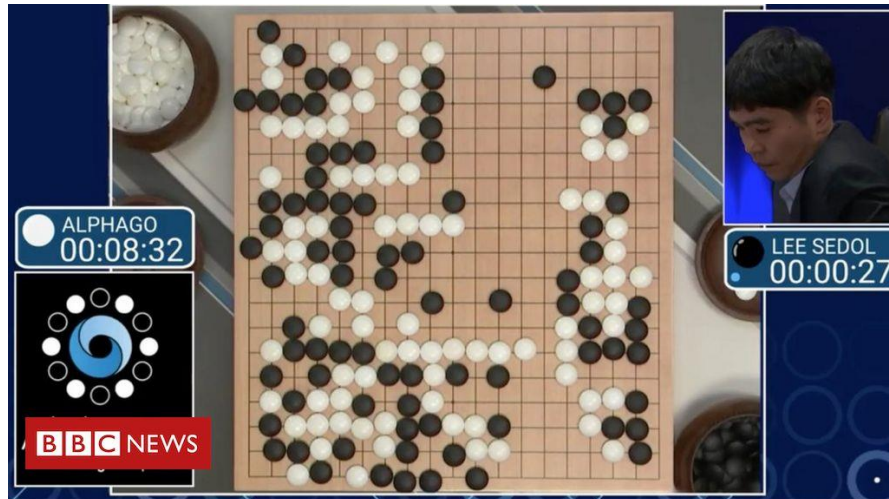
# Reinforcement Learning

- **Setting/Assumptions:** Reinforcement learning (RL) involves training an agent to make decisions by interacting with an environment. The agent learns through trial and error (receiving rewards and penalties), optimizing its behavior to maximize cumulative rewards.
- **Goal:** To learn a policy that maps states of the environment to actions that achieve the highest reward.
- **Common Use Cases:**
  - Game-playing AI (e.g., AlphaGo, chess-playing bots).
  - Robotics (e.g., autonomous navigation).
  - Dynamic resource allocation (e.g., in networking or traffic management).
- **Examples:**
  - Q-learning, Deep Q-Networks (DQN), and Proximal Policy Optimization (PPO).

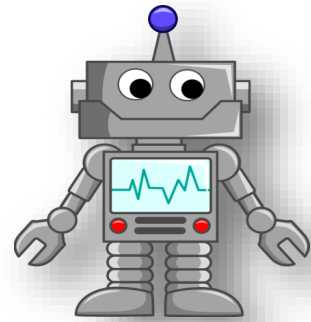
# Reinforcement Learning



$$r = +1 \text{ win} - 1 \text{ lose}$$



# Reinforcement Learning



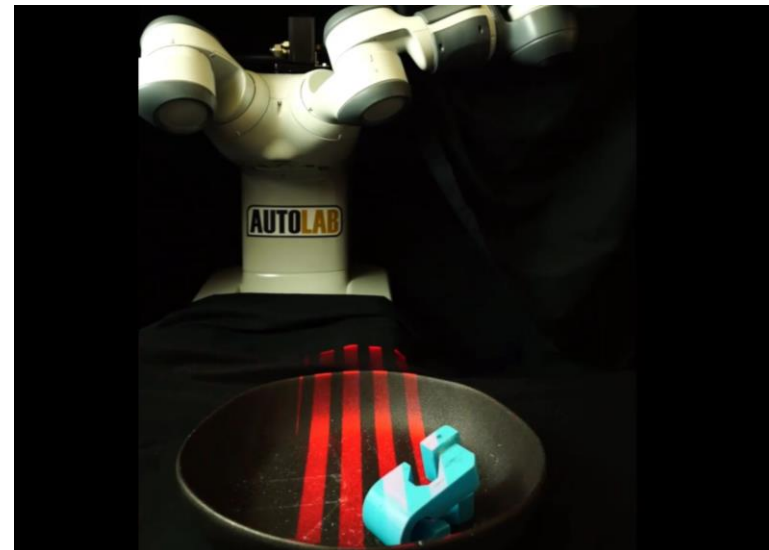
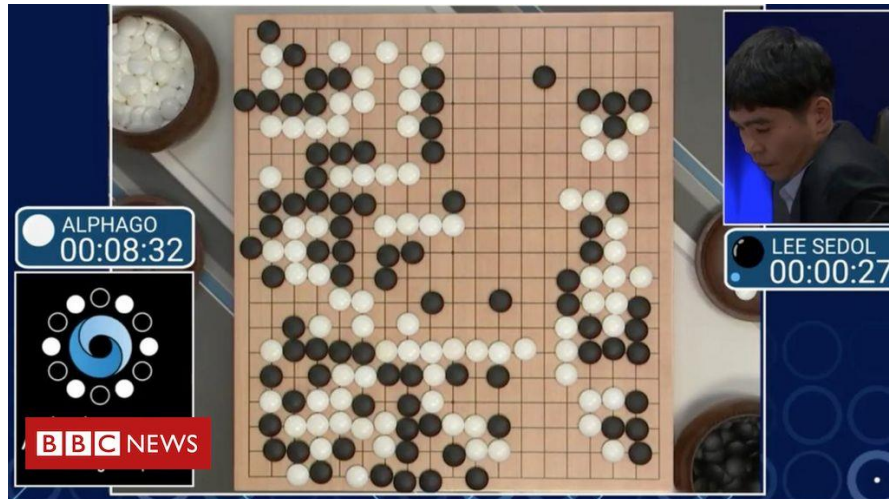
Action →

← Observation

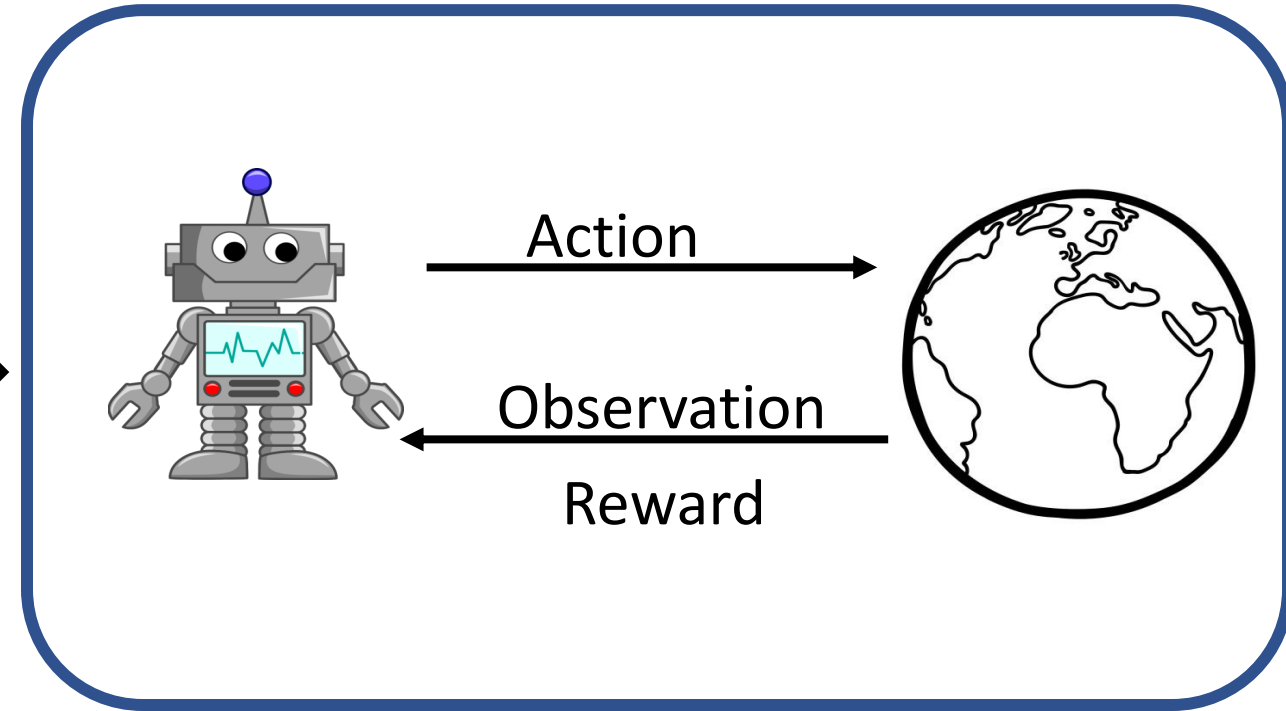
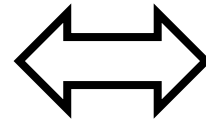
Reward



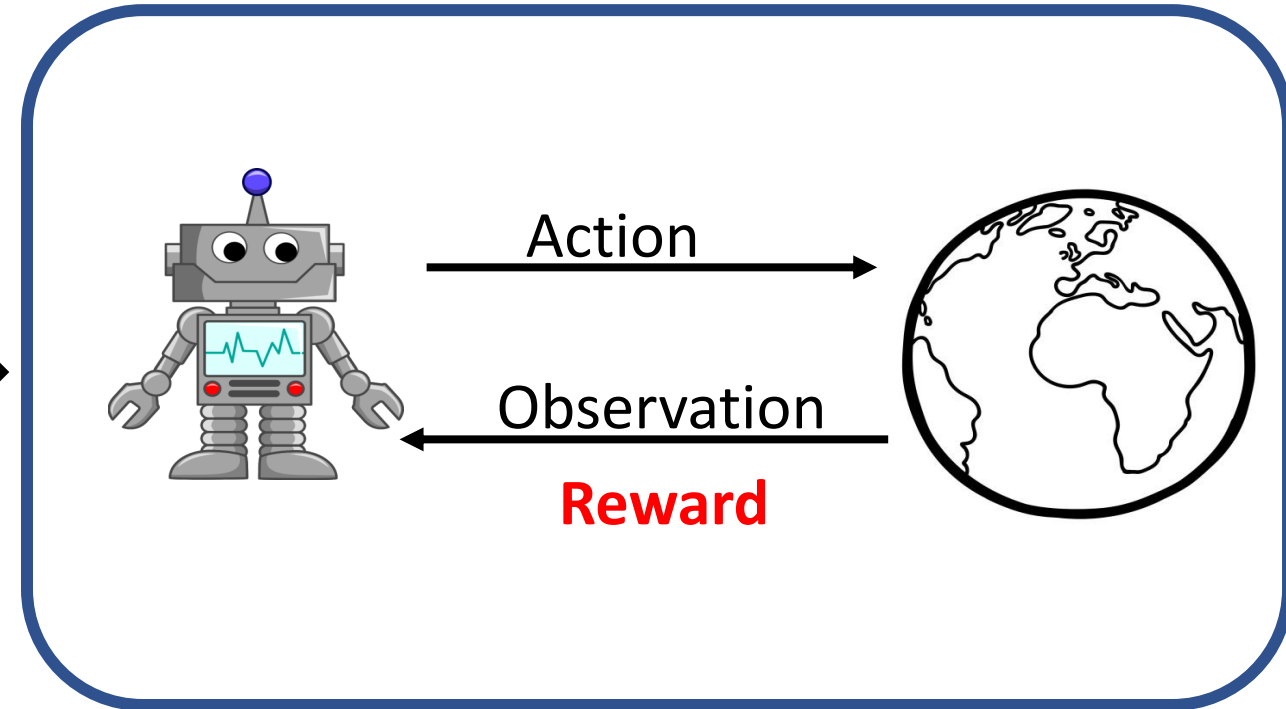
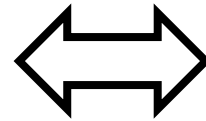
+ / 0



# Reward engineering is hard!

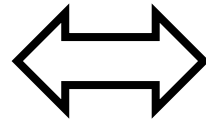


# Reward engineering is hard!

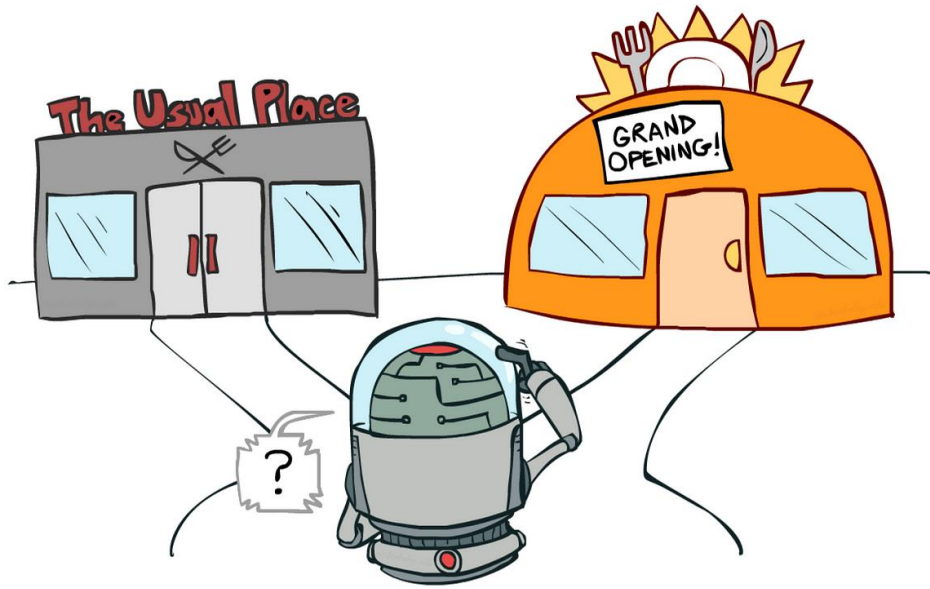




# Reward engineering is hard!



# Reinforcement learning is hard...even with a reward function!





# Imitation Learning (Learning from Demonstrations):

Learn a policy from examples of good behavior.

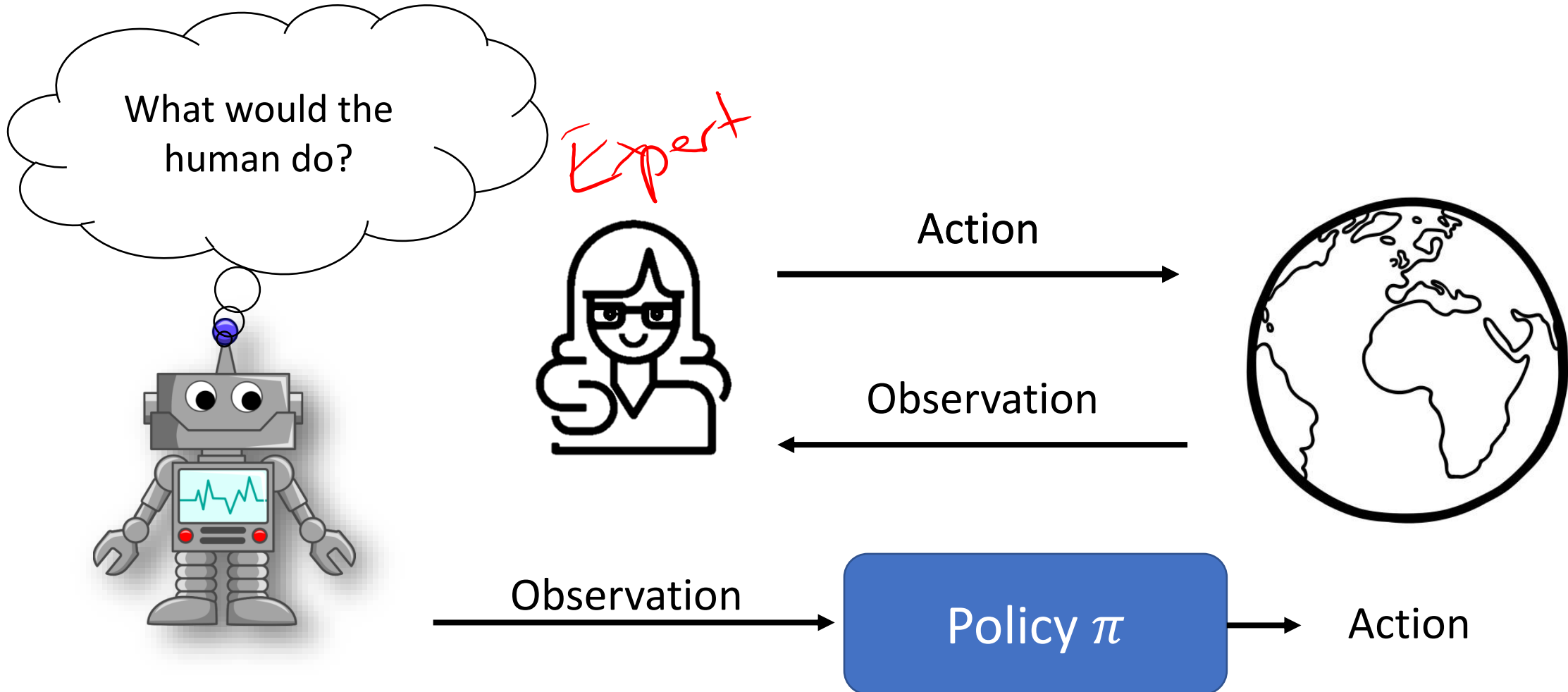
$$\pi(s) \rightarrow a$$



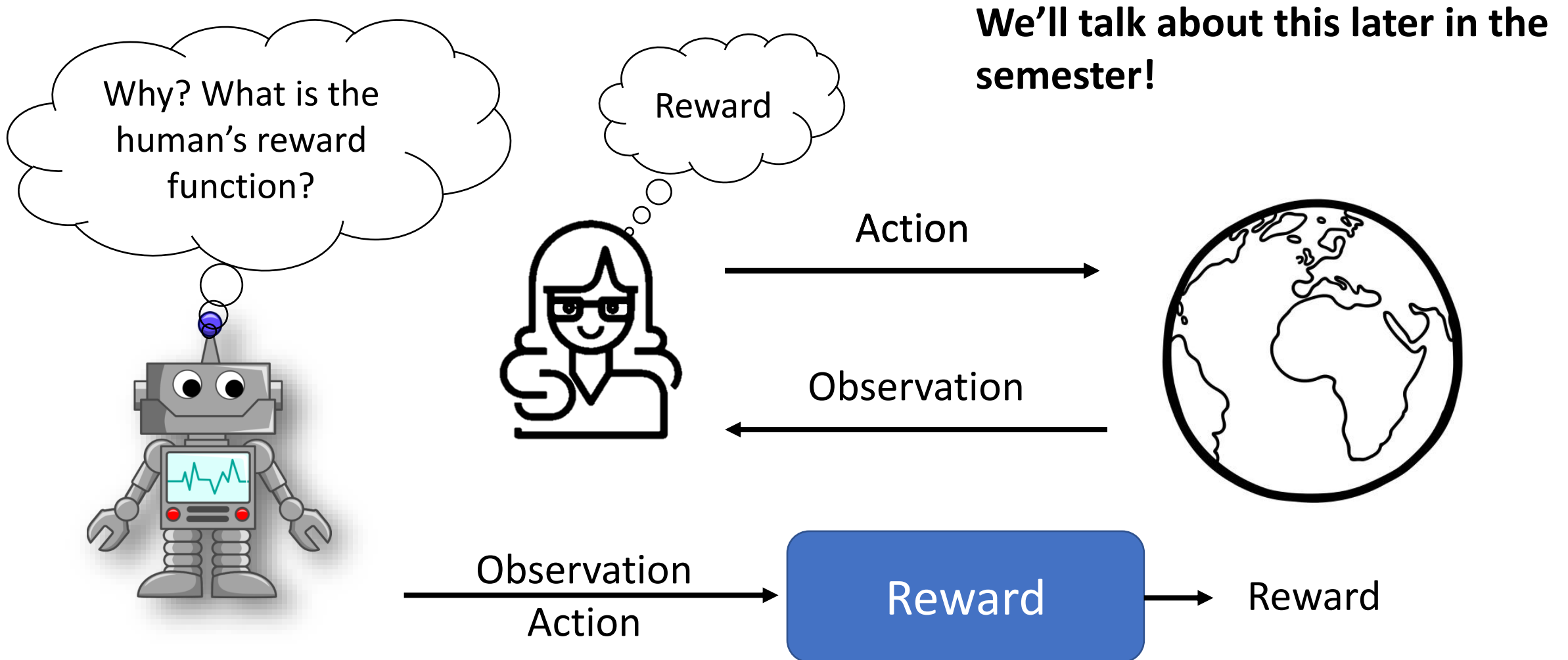
- Often showing is easier than telling.
- Alleviates problem of exploration.

$x$   $p(x)$

# Behavioral Cloning

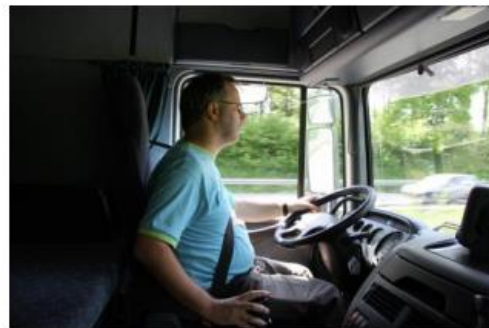
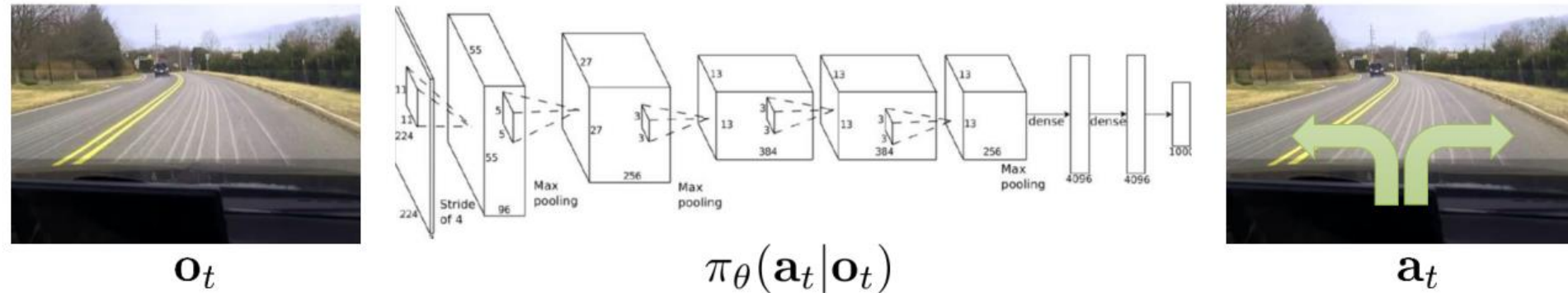


# Inverse Reinforcement Learning

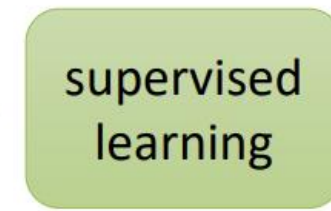




# Imitation Learning via Behavioral Cloning



$\mathbf{o}_t$   
 $\mathbf{a}_t$



*output input*  
 $\pi_{\theta}(\mathbf{a}_t | \mathbf{o}_t)$

# Live demo

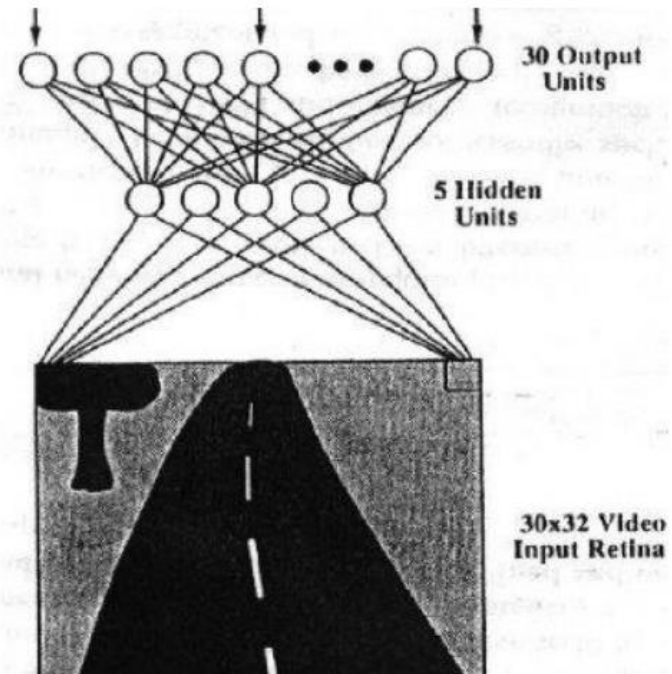
[https://github.com/dsbrown1331/imitation\\_learning/blob/main/README.md](https://github.com/dsbrown1331/imitation_learning/blob/main/README.md)

```
python test_gym.py
```

```
python mountain_car_bc.py --num_demos 1
```

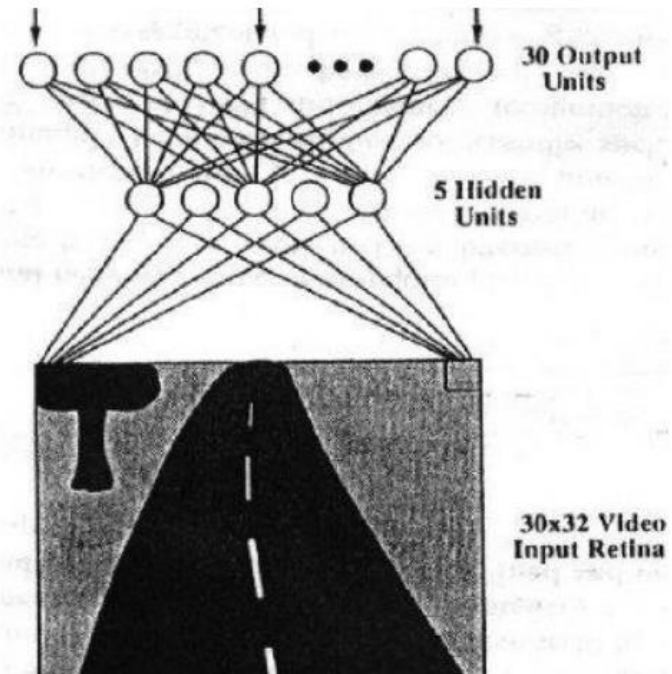
# ALVINN: One of the first imitation learning systems

ALVINN: **A**utonomous **L**and **V**ehicle **I**n a **N**eural **N**etwork  
1989



# ALVINN: One of the first imitation learning systems

ALVINN: **A**utonomous **L**and **V**ehicle **I**n a **N**eural **N**etwork  
1989





# What if you don't have actions?

The screenshot shows a YouTube video player in a Chrome browser window. The video is titled "How to Change a Tire | Change a flat car tire step by step" by the channel "Howdini". The video player shows a man in a white shirt kneeling next to a silver car, working on a tire. The video has 3.5M views and was uploaded 15 years ago. The description reads: "Nothing takes the joy out of a road trip like a flat tire. Do you know how to change it? We didn't, but we've learned from Allan Stanley of AAA. Download this video to your mobile phone just in case." The video player includes standard controls like play/pause, volume, and a progress bar. To the right of the video player is a sidebar with several recommended videos, including "How to change a tire | Dad, how do I?", "How to Replace your Flat Tire", "Steer Tire Change", "Winter Tire Swap", "How To Plug a Flat Tire (easily)", and "POV Tire Change with your Dad #shorts". The browser's address bar shows the URL "youtube.com/watch?v=joBmbh0AGSQ". The browser's top bar shows the date and time "Fri Aug 25 3:13:38 PM".

Chrome File Edit View History Bookmarks Profiles Tab Window Help

How to Change a Tire | Change a flat car tire step by step

youtube.com/watch?v=joBmbh0AGSQ

Incognito

YouTube

tire change

Sign in

AHSOKA Now streaming Disney+ STREAM NOW

Ahsoka Ad - www.disneyplus.com Watch

How to change a tire | Dad, how do I? Dad, how do I? 886K views · 3 years ago 13:24

How to Replace your Flat Tire Pushing Pistons 117K views · 2 years ago 0:57

Steer Tire Change The Tire Doctor 571K views · 1 year ago 1:01

Winter Tire Swap How To Change Your Tires Yourself - Winter/Summer Tire... Your Home Garage 48K views · 2 years ago 12:47

How to Plug a Flat Tire (easily) ChrisFix 922K views · 9 months ago 1:00

POV Tire Change with your Dad #shorts Charlie Berens 7.4M views · 6 months ago

How to Change a Tire | Change a flat car tire step by step

Howdini 714K subscribers Subscribe

36K 3:17 / 5:34

3.5M views 15 years ago

Nothing takes the joy out of a road trip like a flat tire. Do you know how to change it? We didn't, but we've learned from Allan Stanley of AAA. Download this video to your mobile phone just in case.



# Behavioral Cloning from Observation (Torabi et al. 2018)

$$D = \{ (s_a) \dots \}$$
$$D = \{ (s_0, s_1, s_2, \dots, s_T), (s'_0, s'_1, \dots, s'_T) \}$$