PyTorch and Neural Nets
Review Session

CS285
Instructor: Vitchyr Pong
Goal of this course

Train an agent to perform useful tasks
common training paradigm

- Collect data
- Train model

- Data
- Agent
How do we train a model?

$$\theta^* = \arg\min_{\theta} \sum_{(x,y) \in \mathcal{D}} ||f_\theta(x) - y||^2_2$$
How do we train a model?

\[ \theta^* = \arg \min_{\theta} \sum_{(x,y) \in \mathcal{D}} \| f_{\theta}(x) - y \|_2^2 \]
How do we train a model?

\[ \theta^* = \arg \min_{\theta} \sum_{(x,y) \in D} \| f_{\theta}(x) - y \|^2 \]
How do we train a model?

$$\theta^* = \arg \min_{\theta} \sum_{(x,y) \in D} \| f_{\theta}(x) - y \|^2_2$$

gradient descent  neural network
What is PyTorch?

Python library for…

● Defining neural networks
● Automatically computing gradients

\[ \theta^* = \arg \min_{\theta} \sum_{(x,y) \in \mathcal{D}} \| f_{\theta}(x) - y \|_2^2 \]
What is PyTorch?

Python library for...

- Defining neural networks
- Automatically computing gradients
- And more (GPU, optimizers, etc.)

\[
\theta^* = \arg \min_{\theta} \sum_{(x,y) \in D} \| f_\theta(x) - y \|_2^2
\]
PyTorch Alternatives

TensorFlow, JAX, Chainer, …

Basically all do the same
How does Pytorch work?
PyTorch: forward pass

\[
h_1 = \sigma(W_1x) \quad h_2 = \sigma(W_2h_1) \quad y = \sigma(W_3h_2)
\]
PyTorch: backward pass

\[ h_1 = \sigma(W_1 x) \quad h_2 = \sigma(W_2 h_1) \quad y = \sigma(W_3 h_2) \]

You define

\[
\frac{\partial y}{\partial W_1} = \frac{\partial y}{\partial h_2} \frac{\partial h_2}{\partial h_1} \frac{\partial h_1}{\partial W_1} \\
\frac{\partial y}{\partial W_2} = \frac{\partial y}{\partial h_2} \frac{\partial h_2}{\partial W_2} \\
\frac{\partial y}{\partial W_3}
\]

[picture from Stanford’s CS231n]
PyTorch Tutorial (Colab)

https://colab.research.google.com/drive/1r_-Ow0QYPN58cfuNjZDUy4O6HUvPDxyN?usp=sharing