

Last Time

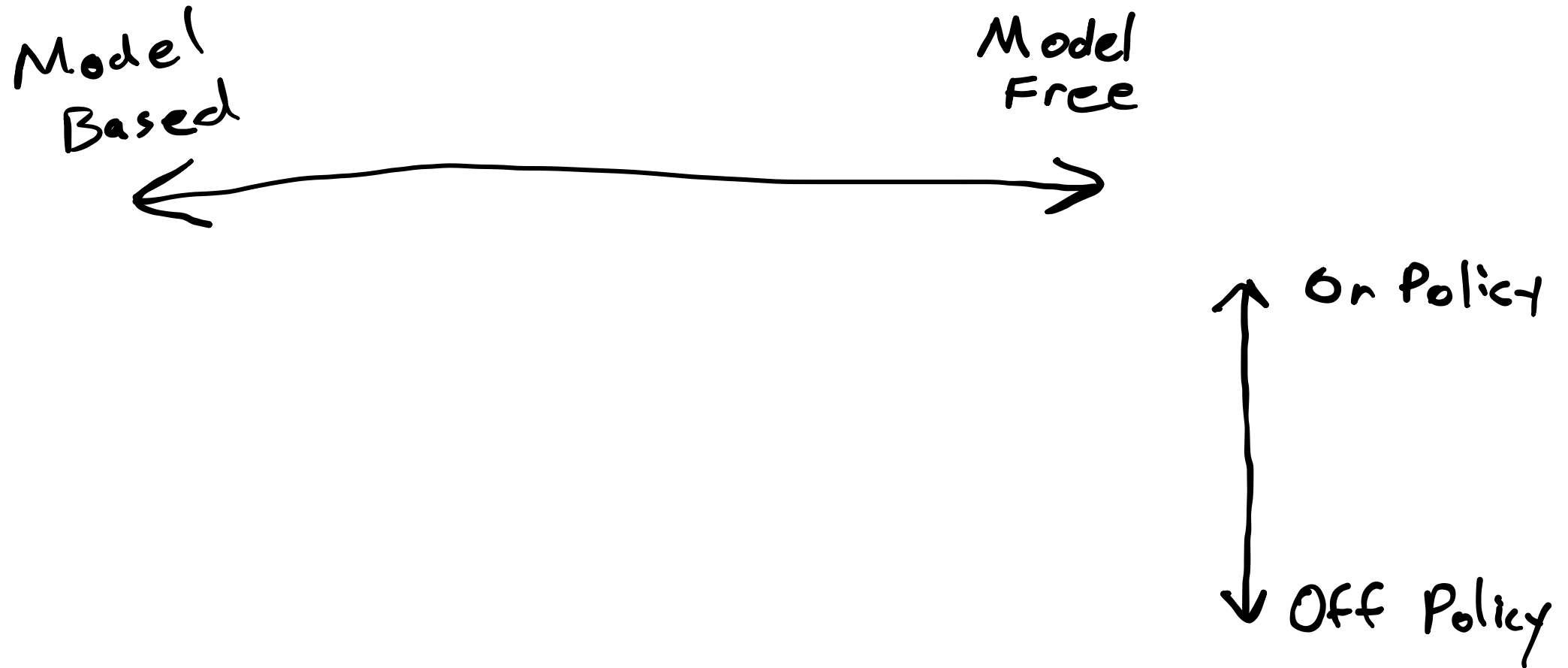
Last Time

Model
Based

Model
Free



Last Time



Last Time

Model
Based

Model
Free

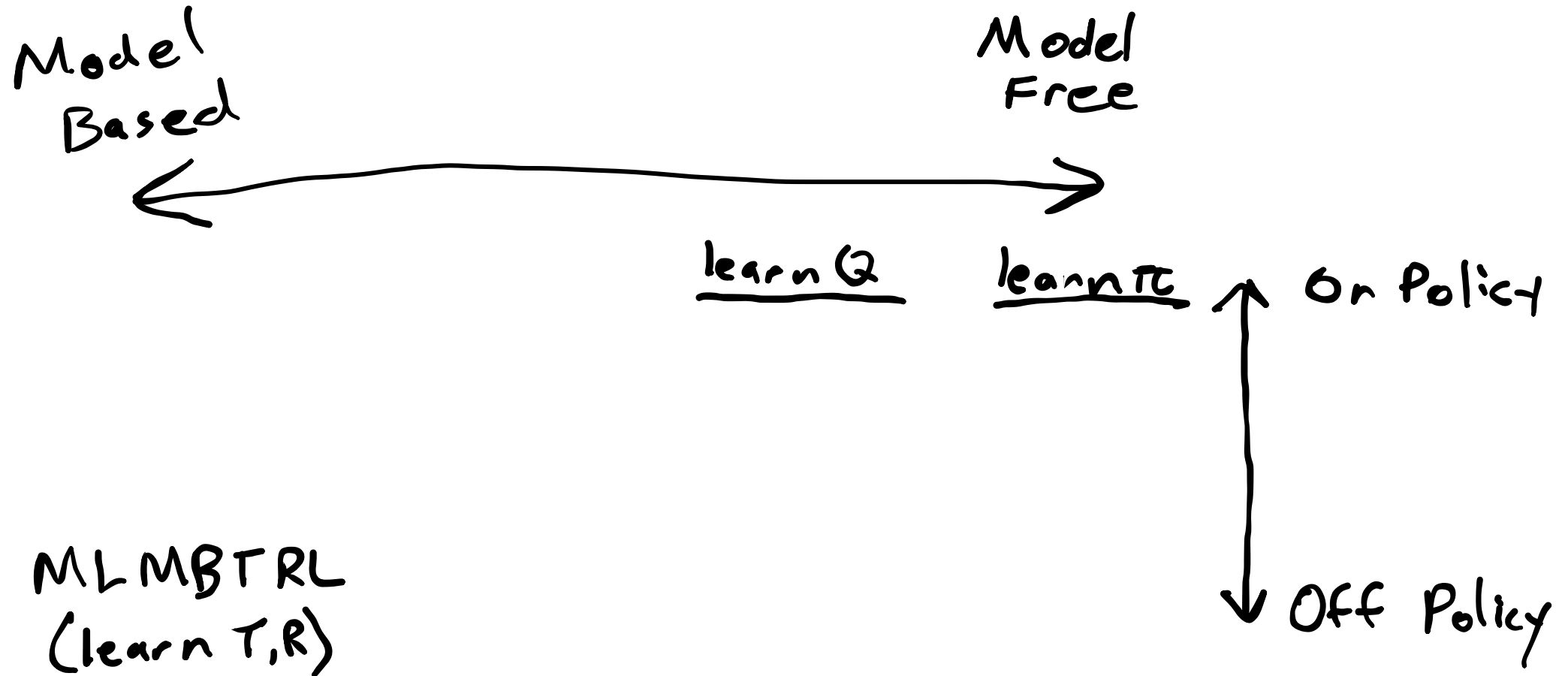


MLMBTRL
(learn T, R)

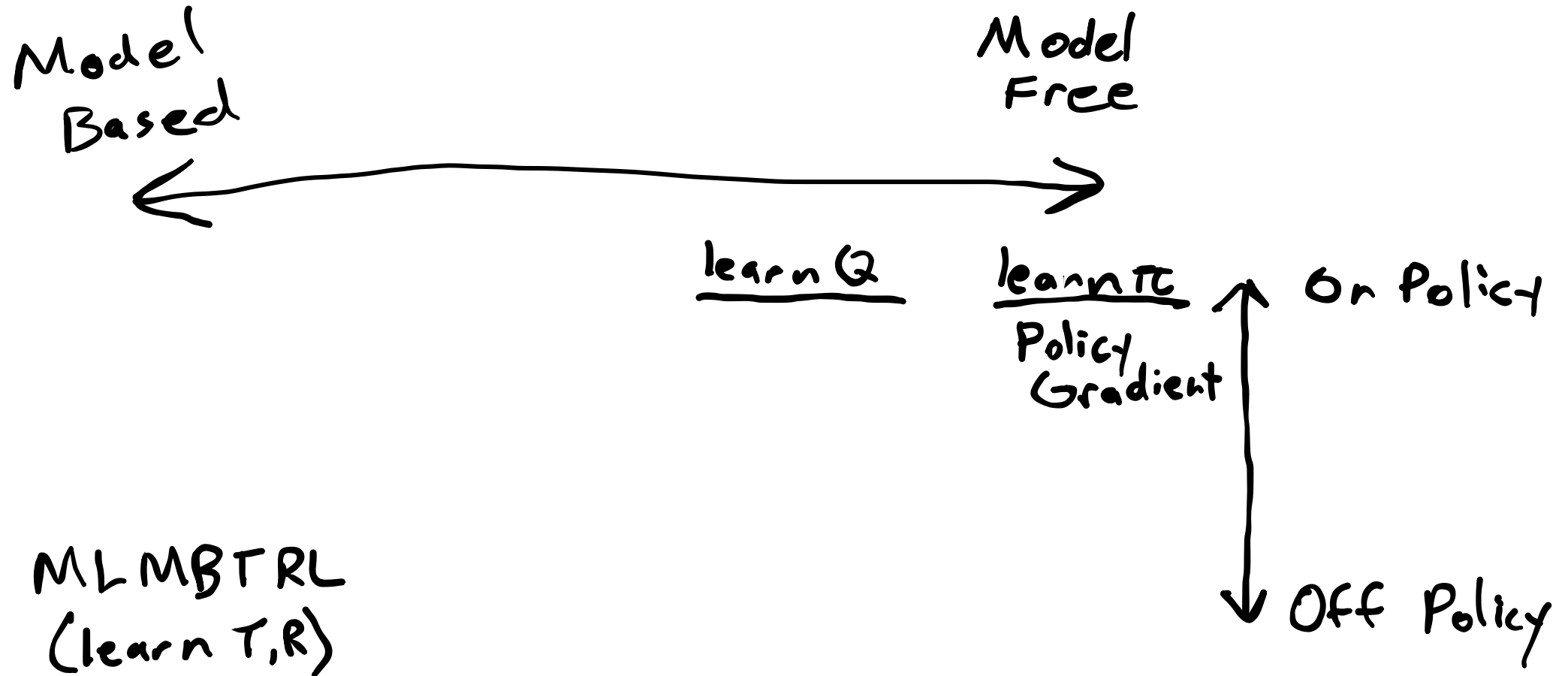
On Policy
Off Policy

A vertical double-headed arrow between 'On Policy' and 'Off Policy'.

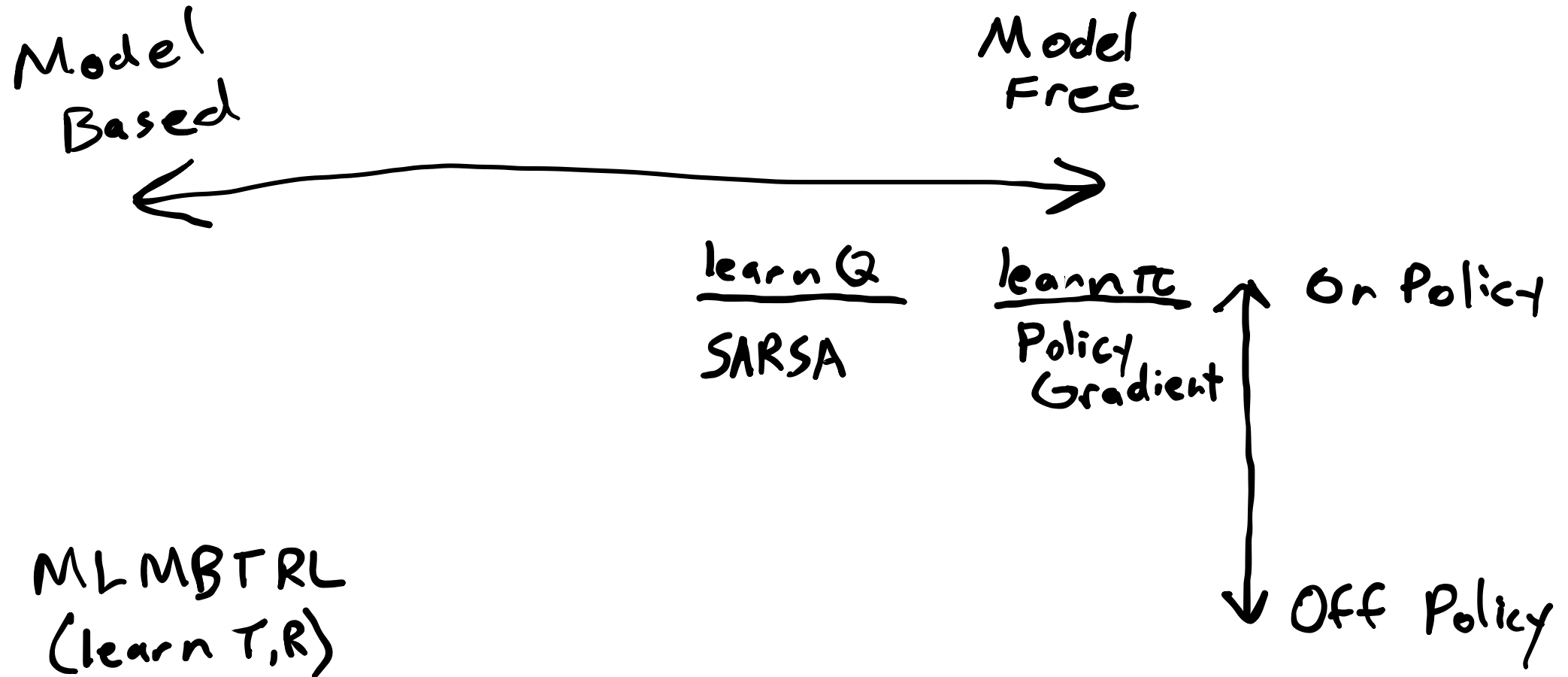
Last Time



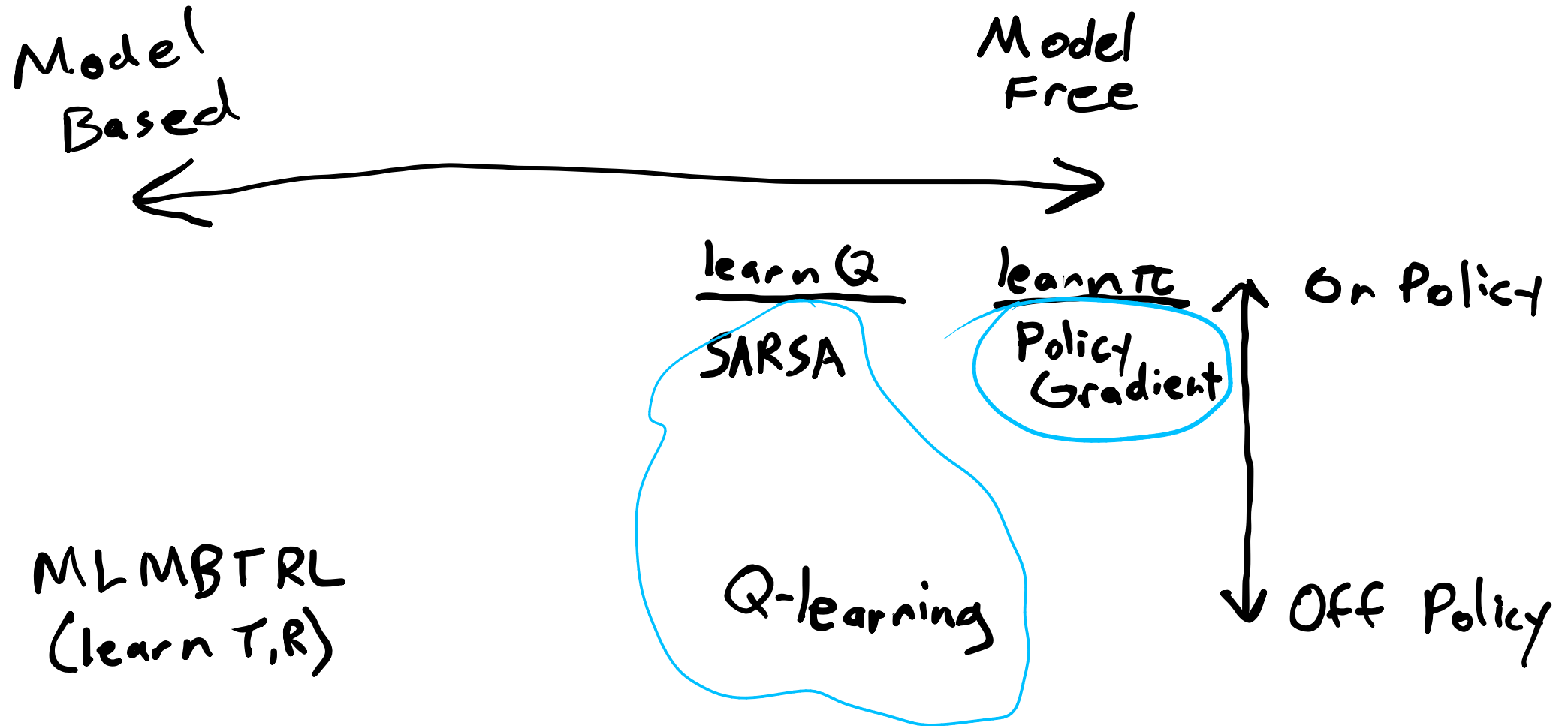
Last Time



Last Time




Last Time



This Time

Challenges in Reinforcement Learning:


- Exploration vs Exploitation
- Credit Assignment
- Generalization 

Function Approximation

Function Approximation

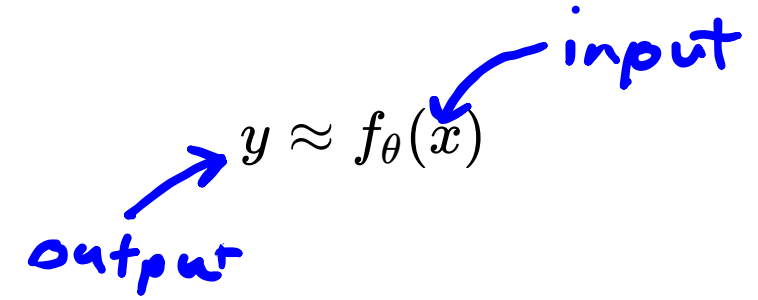
$$y \approx f_{\theta}(x)$$

Function Approximation

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


input

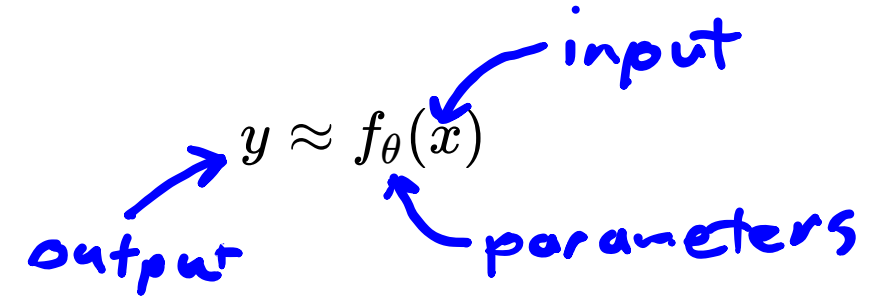
Function Approximation



A diagram illustrating the function approximation equation $y \approx f_{\theta}(x)$. The equation is centered, with a handwritten blue arrow pointing from the word "output" to the variable y on the left, and another handwritten blue arrow pointing from the word "input" to the variable x inside the function $f_{\theta}(x)$ on the right.

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

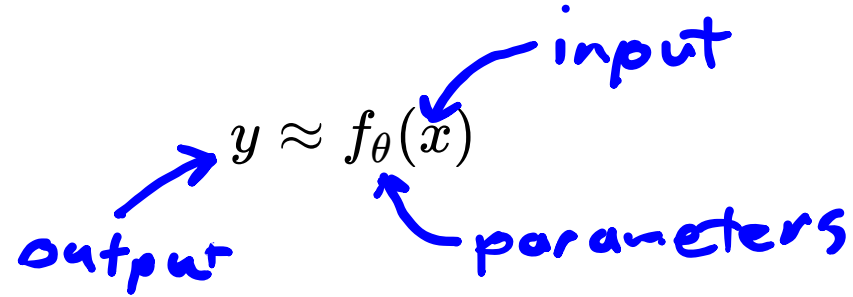
Function Approximation



A diagram illustrating the function approximation equation $y \approx f_{\theta}(x)$. The equation is centered, with handwritten blue arrows and labels identifying its components: an arrow points from the word "output" to y , an arrow points from the word "input" to x , and an arrow points from the word "parameters" to θ .

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

Function Approximation

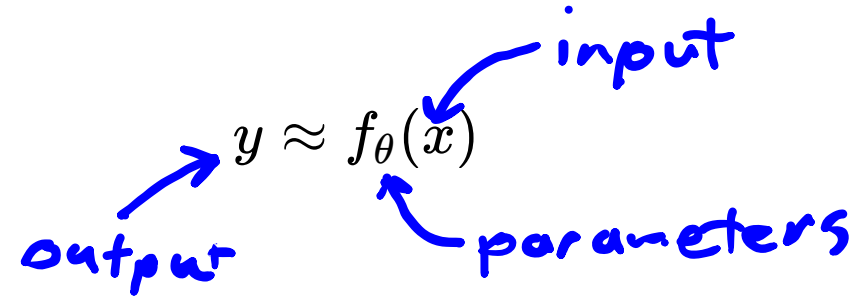


A diagram illustrating the function approximation equation $y \approx f_{\theta}(x)$. The equation is written in black. Three blue arrows point to different parts of the equation: one from the word "output" to y , one from the word "input" to x , and one from the word "parameters" to θ .

Previously, Linear:

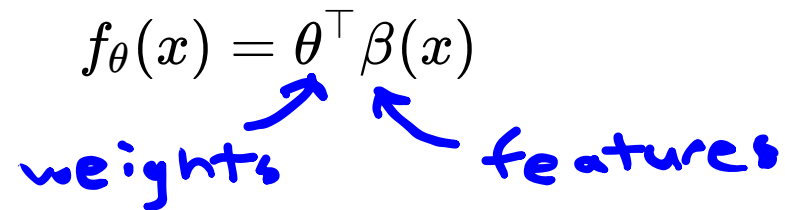
$$f_{\theta}(x) = \theta^{\top} \beta(x)$$

Function Approximation

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


A diagram illustrating the general function approximation equation $y \approx f_{\theta}(x)$. The equation is centered, with three handwritten blue arrows pointing to its components: an arrow from the word "output" points to y , an arrow from the word "input" points to x , and an arrow from the word "parameters" points to θ .

Previously, Linear:

$$f_{\theta}(x) = \theta^{\top} \beta(x)$$


A diagram illustrating the linear function approximation equation $f_{\theta}(x) = \theta^{\top} \beta(x)$. The equation is centered, with two handwritten blue arrows pointing to its components: an arrow from the word "weights" points to θ , and an arrow from the word "features" points to $\beta(x)$.

Function Approximation

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

output \swarrow \nwarrow input
parameters

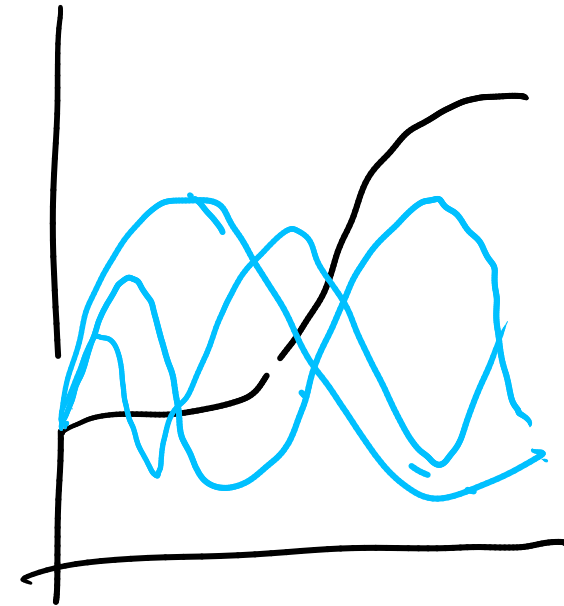
Previously, Linear:

$$f_{\theta}(x) = \theta^{\top} \beta(x)$$

weights \nwarrow \swarrow features

$\begin{bmatrix} \sin 7\pi x \\ \sin 2\pi x \\ \sin 3\pi x \end{bmatrix}$

e.g. $\beta_i(x) = \sin(i \pi x)$



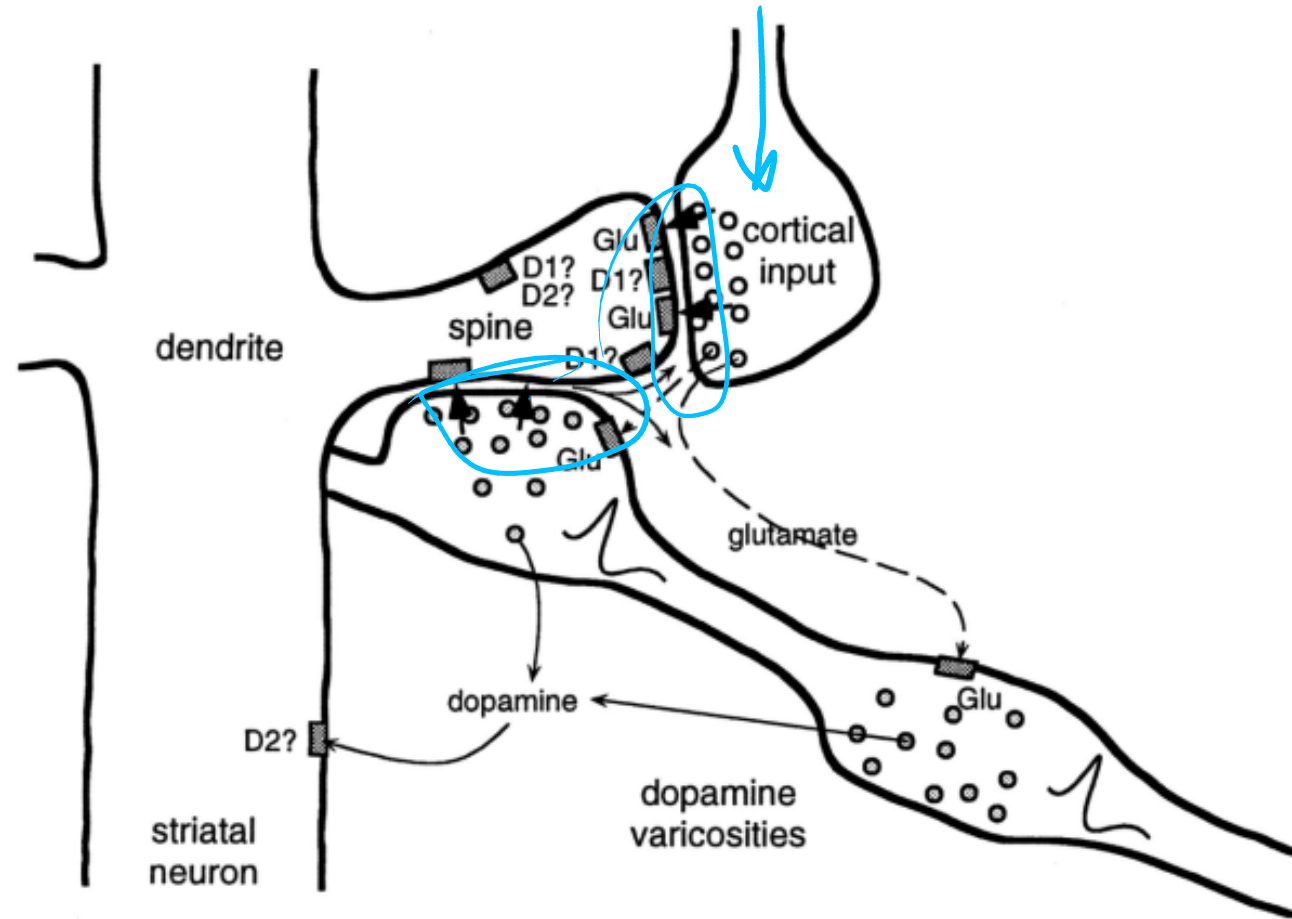
Neural Network

Neural Network

$$h(x) = \sigma(Wx + b)$$

Neural Network

$$h(x) = \sigma(Wx + b)$$



Neural Network

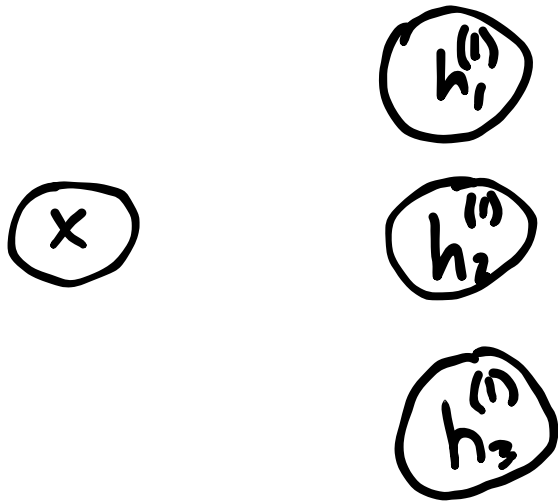
$$h(x) = \sigma(Wx + b)$$

$$f_{\theta}(x) = h^{(2)}(h^{(1)}(x))$$

Neural Network

$$h(x) = \sigma(Wx + b)$$

$$f_{\theta}(x) = h^{(2)}(h^{(1)}(x)) = \sigma^{(2)}(W^{(2)}\sigma^{(1)}(W^{(1)}x + b^{(1)}) + b^{(2)})$$

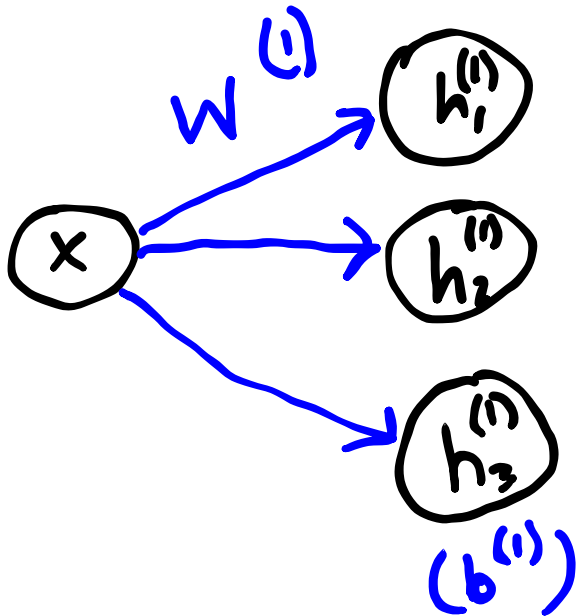


Neural Network

$$h(x) = \sigma(Wx + b)$$

$$f_{\theta}(x) = h^{(2)}(h^{(1)}(x)) = \sigma^{(2)}(W^{(2)}\sigma^{(1)}(W^{(1)}x + b^{(1)}) + b^{(2)})$$

$$\theta = (W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)})$$

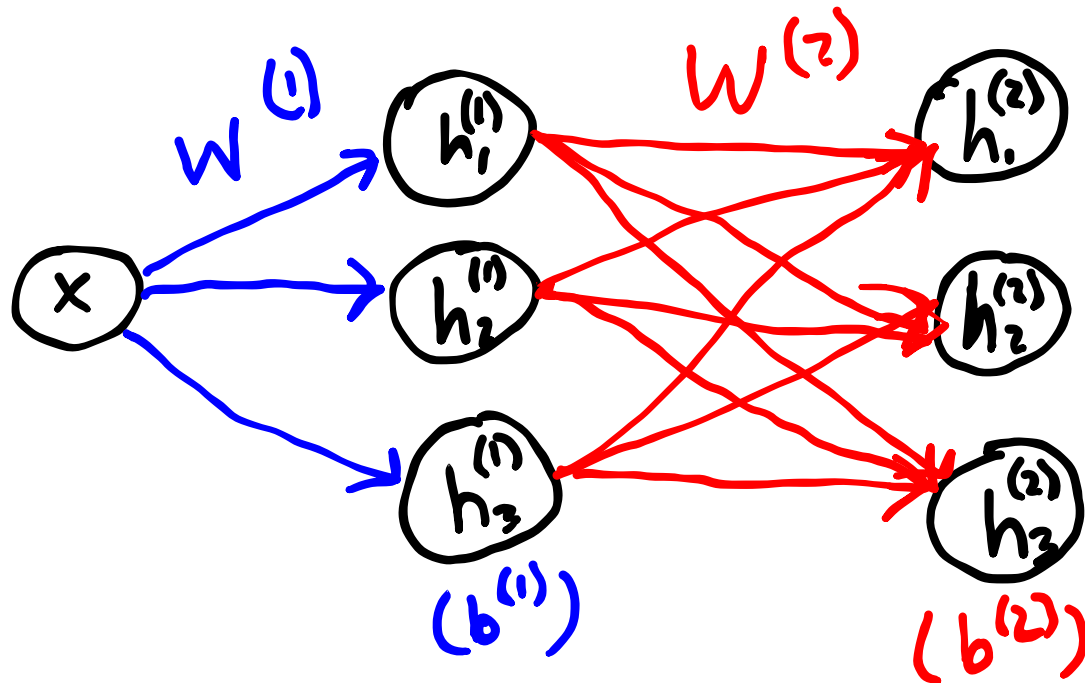


Neural Network

$$h(x) = \sigma(Wx + b)$$

$$f_{\theta}(x) = h^{(2)}(h^{(1)}(x)) = \sigma^{(2)}(W^{(2)}\sigma^{(1)}(W^{(1)}x + b^{(1)}) + b^{(2)})$$

$$\theta = (W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)})$$

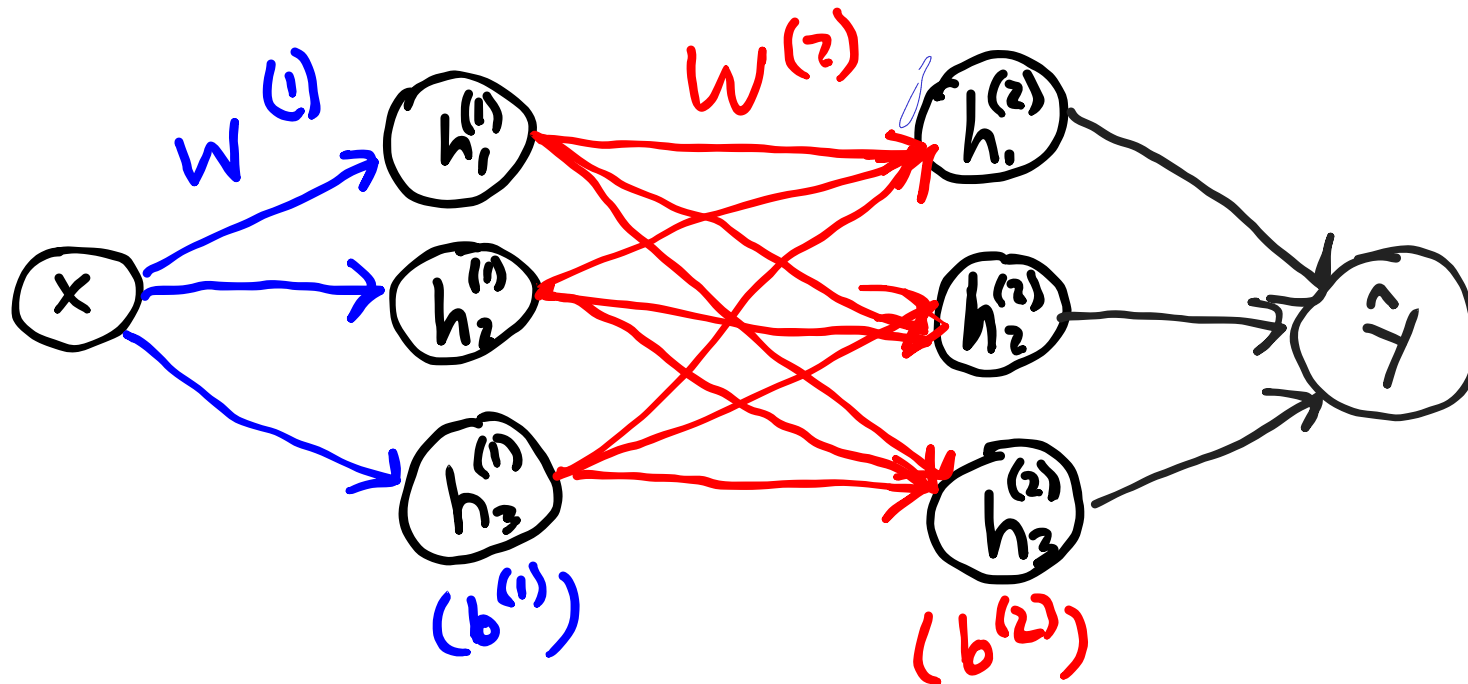


Neural Network

$$h(x) = \sigma(Wx + b)$$

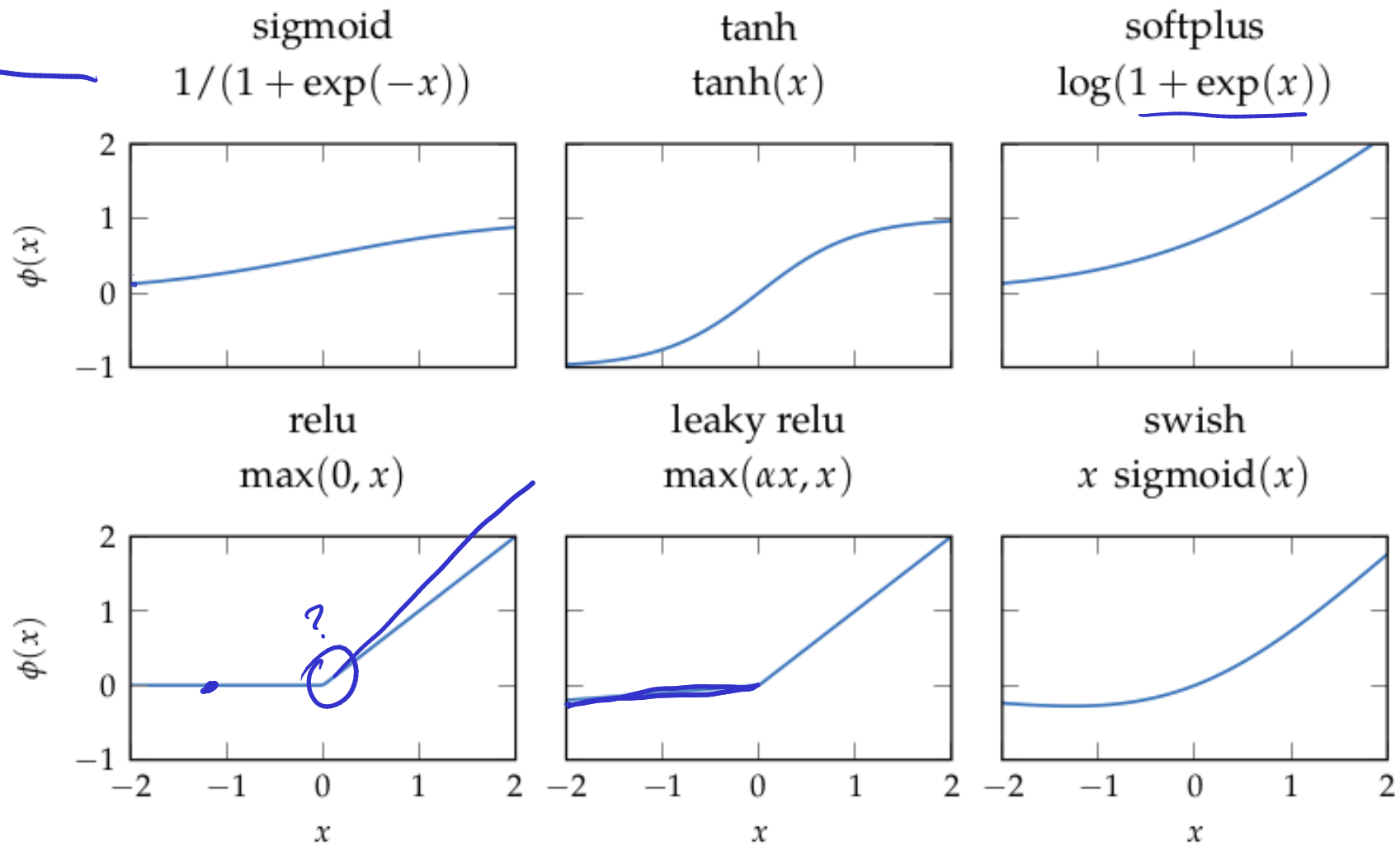
$$f_{\theta}(x) = h^{(2)}(h^{(1)}(x)) = \sigma^{(2)}\left(W^{(2)}\left(\sigma^{(1)}\left(\underline{W^{(1)}x + b^{(1)}}\right)\right) + b^{(2)}\right)$$

$$\theta = (W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)})$$



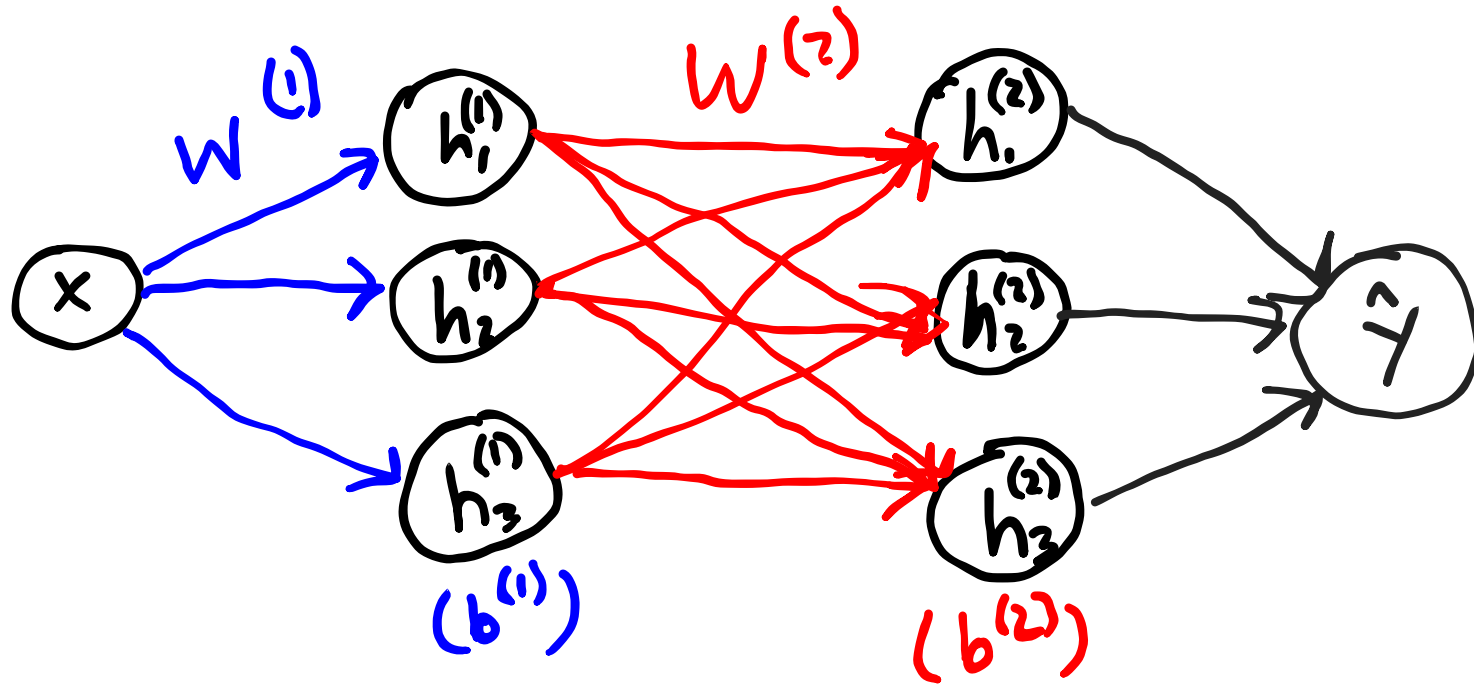
$$\begin{aligned} x &\in \mathbb{R}^2 \\ \hat{y} &\in \mathbb{R}^3 \\ W^{(1)} &\in \mathbb{R}^{3 \times 2} \\ b^{(1)} &\in \mathbb{R}^3 \end{aligned}$$

Nonlinearities

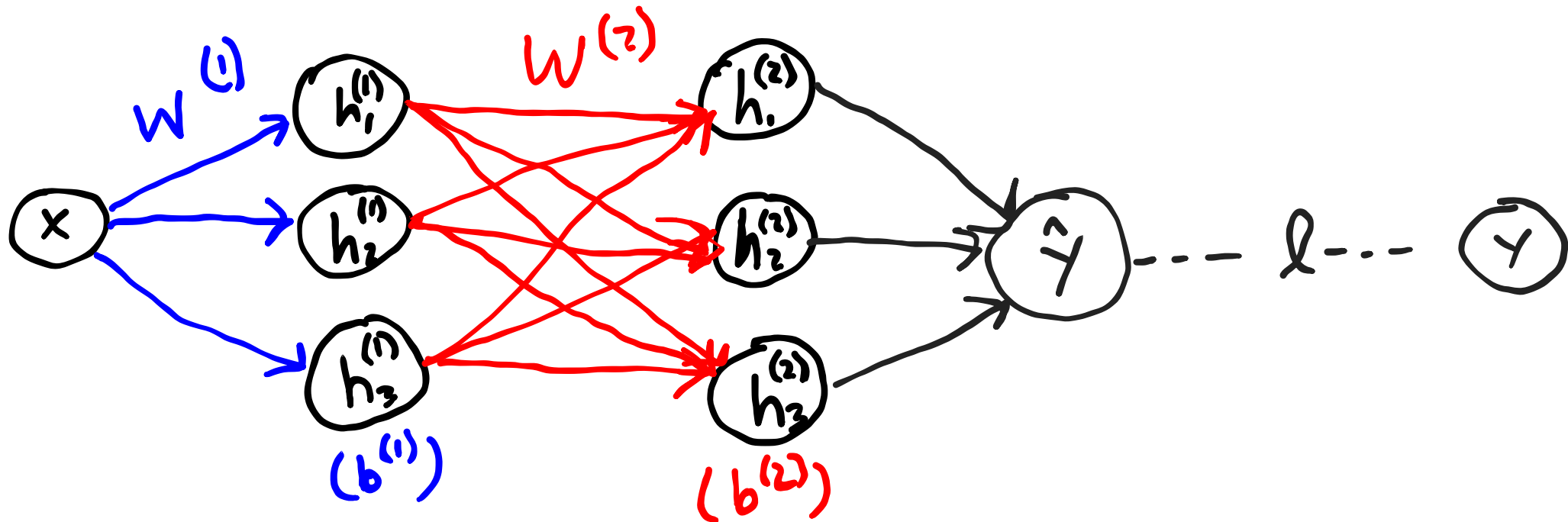


Training

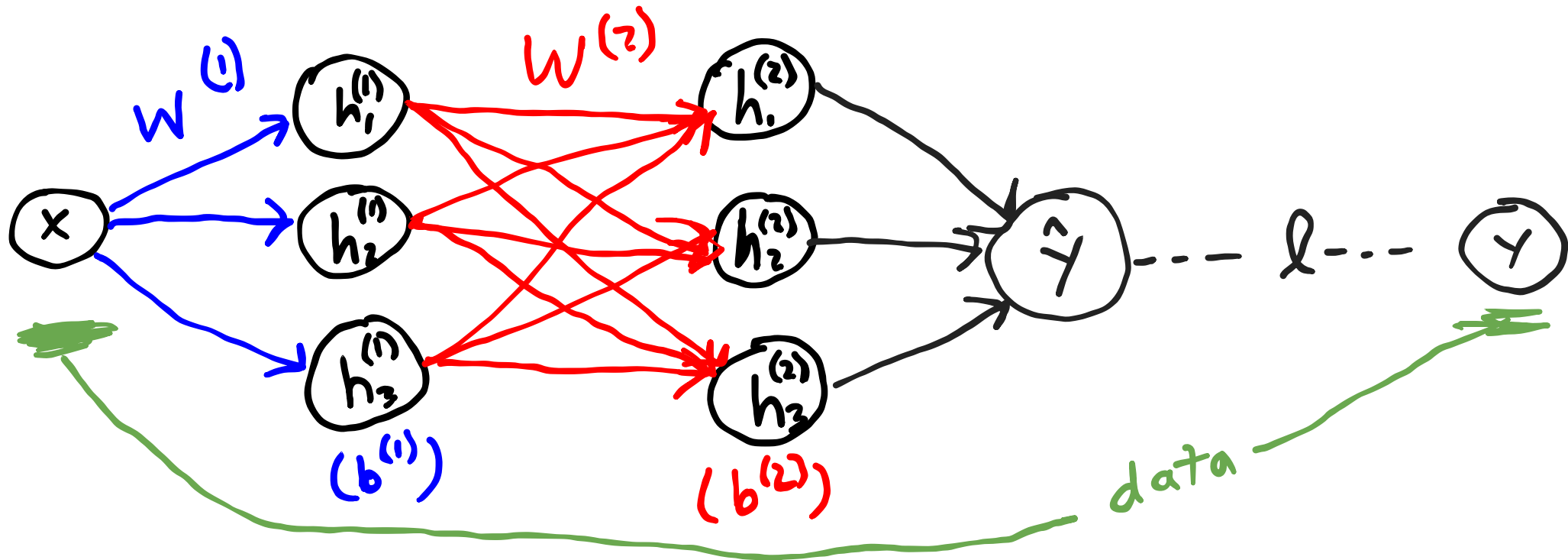
Training



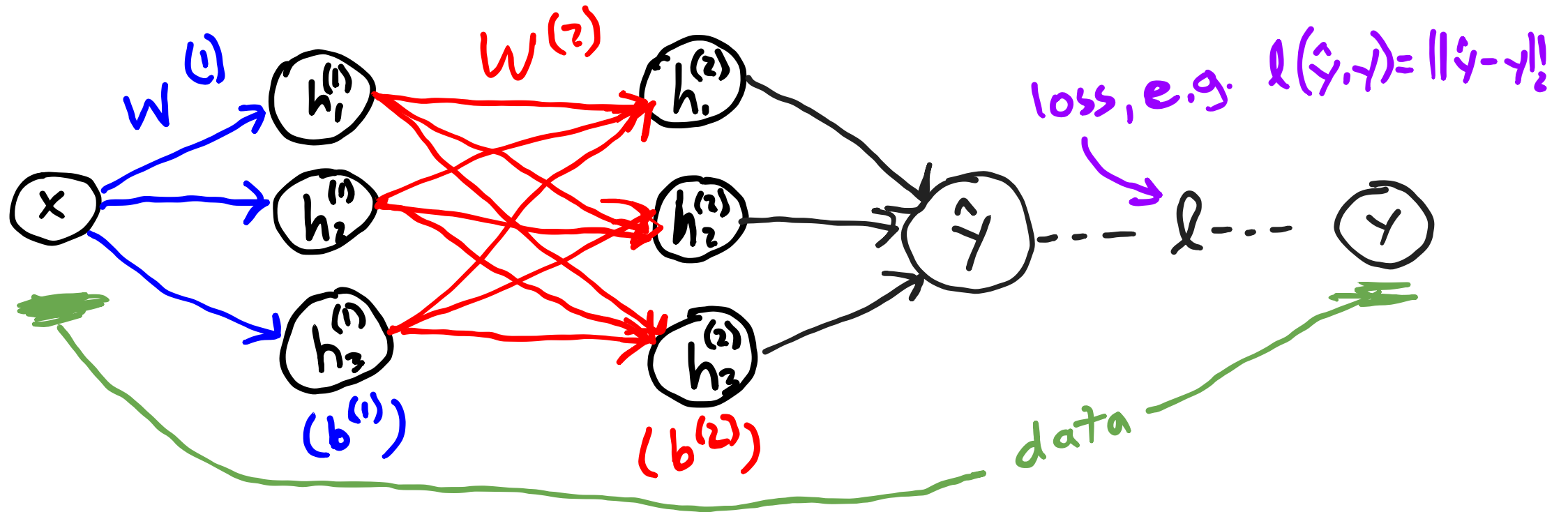
Training



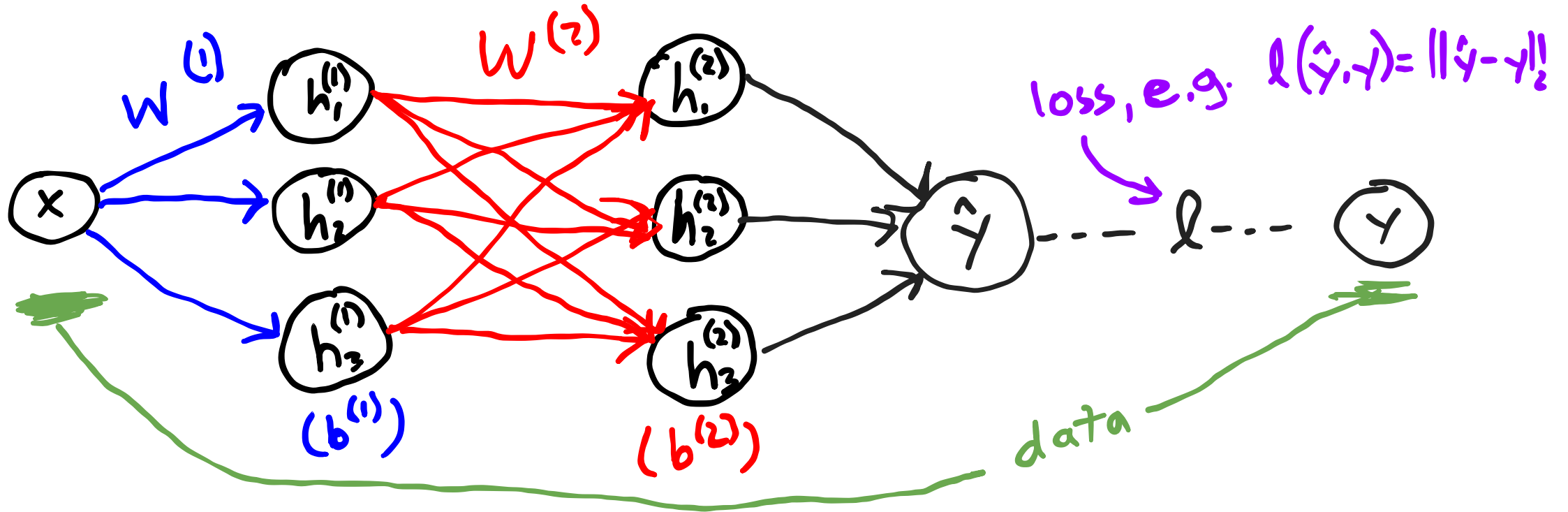
Training



Training

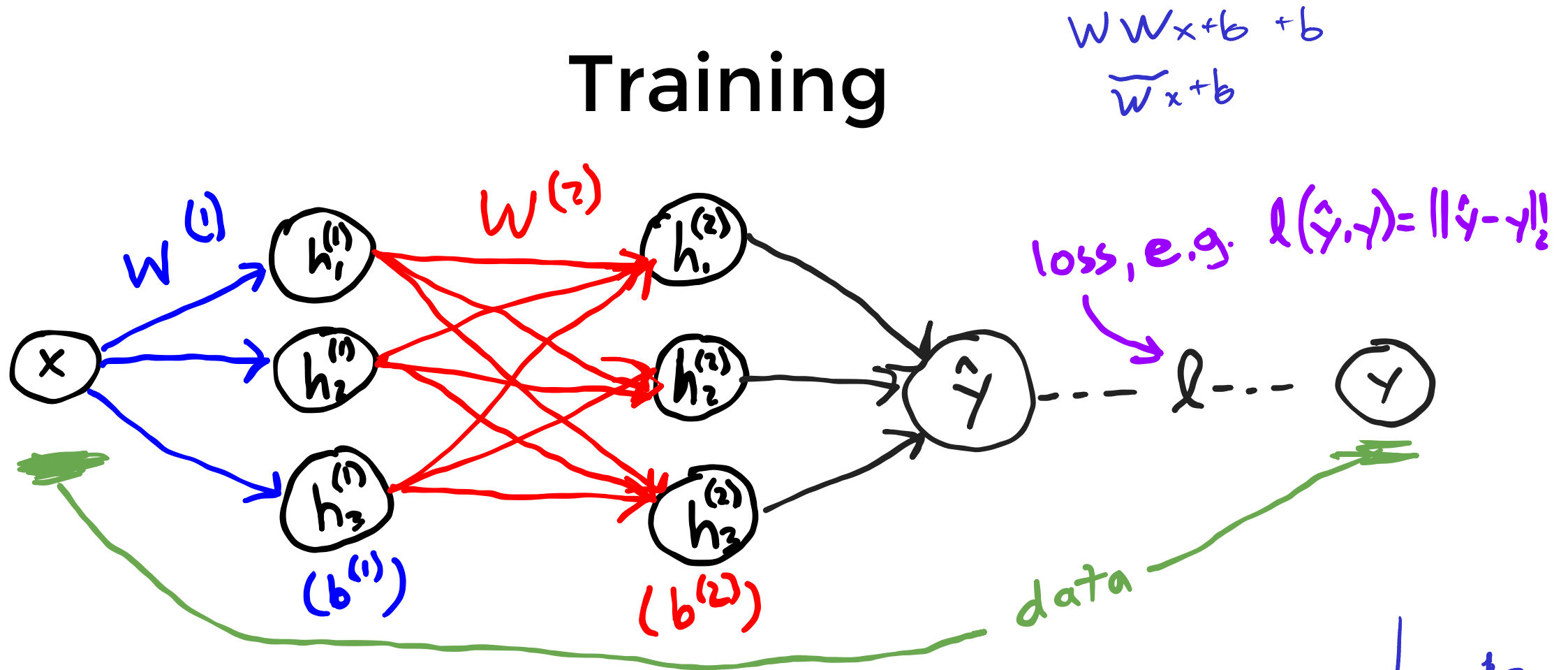


Training



$$\theta^* = \arg \min_{\theta} \sum_{(x,y) \in \mathcal{D}} l(f_{\theta}(x), y)$$

Training



$$\theta^* = \arg \min_{\theta} \sum_{(x,y) \in \mathcal{D}} l(f_{\theta}(x), y)$$

Stochastic Gradient Descent: $\theta \leftarrow \theta - \alpha \nabla_{\theta} l(f_{\theta}(x), y)$

Software

Chain Rule

Chain Rule

$$\frac{\partial f(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}}$$

Chain Rule

$$\frac{\partial f(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Chain Rule

$$\frac{\partial f(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Chain Rule

$$\frac{\partial f(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)}x + b^{(1)}) + b^{(2)}$$

Chain Rule

$$\frac{\partial f(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)}x + b^{(1)}) + b^{(2)}$$

$$\frac{\partial l}{\partial W^{(2)}} =$$

Chain Rule

$$\frac{\partial f(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)}x + b^{(1)}) + b^{(2)}$$

$$\frac{\partial l}{\partial W^{(2)}} = \frac{\partial l}{\partial \hat{y}} \left(\frac{\partial \hat{y}}{\partial W^{(2)}} \right)^\top$$

Chain Rule

$$\frac{\partial f(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)}x + b^{(1)}) + b^{(2)}$$

$$\frac{\partial l}{\partial W^{(2)}} = \frac{\partial l}{\partial \hat{y}} \left(\frac{\partial \hat{y}}{\partial W^{(2)}} \right)^\top = \frac{\partial l}{\partial \hat{y}} \sigma \left(W^{(1)}x + b^{(1)} \right)^\top$$

Chain Rule

$$\frac{\partial f(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)}x + b^{(1)}) + b^{(2)}$$

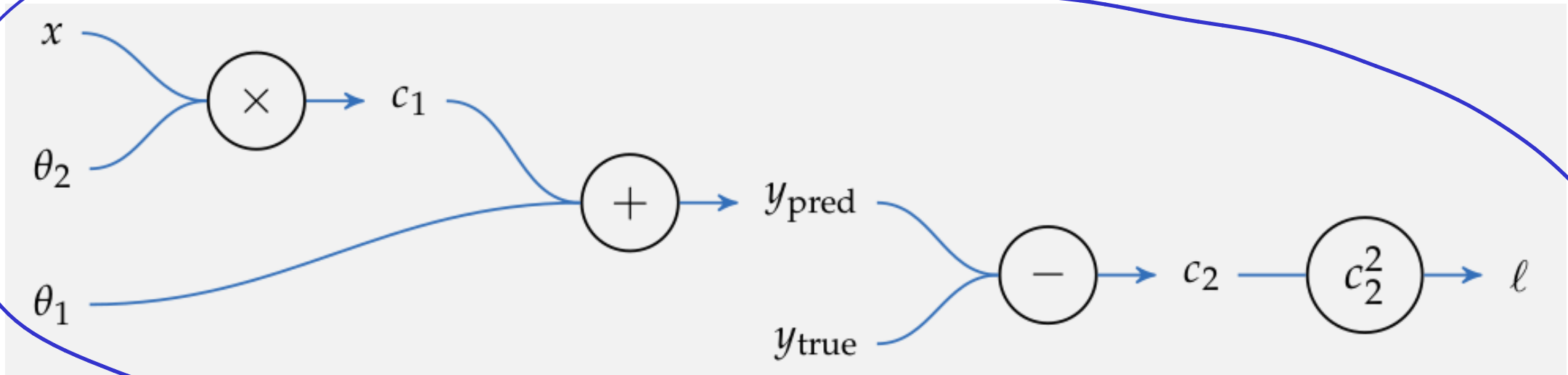
$$\frac{\partial l}{\partial W^{(2)}} = \frac{\partial l}{\partial \hat{y}} \left(\frac{\partial \hat{y}}{\partial W^{(2)}} \right)^\top = \frac{\partial l}{\partial \hat{y}} \underbrace{\sigma \left(W^{(1)}x + b^{(1)} \right)^\top}_{\text{output of first layer}}$$

$$W^{(2)} \leftarrow W^{(2)} - \alpha \frac{\partial l}{\partial W^{(2)}}$$

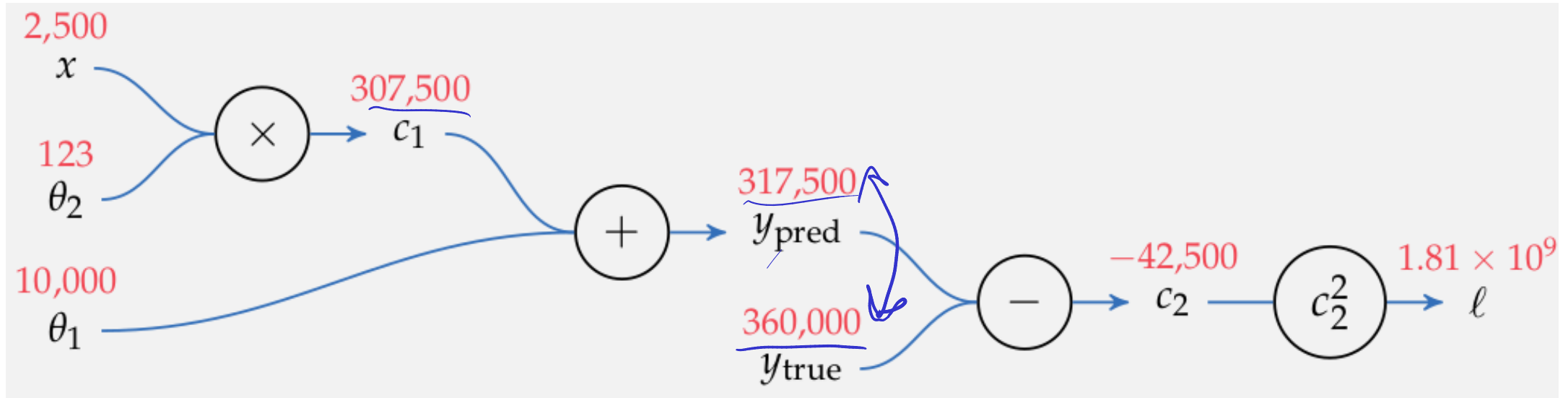
Backprop

Backprop

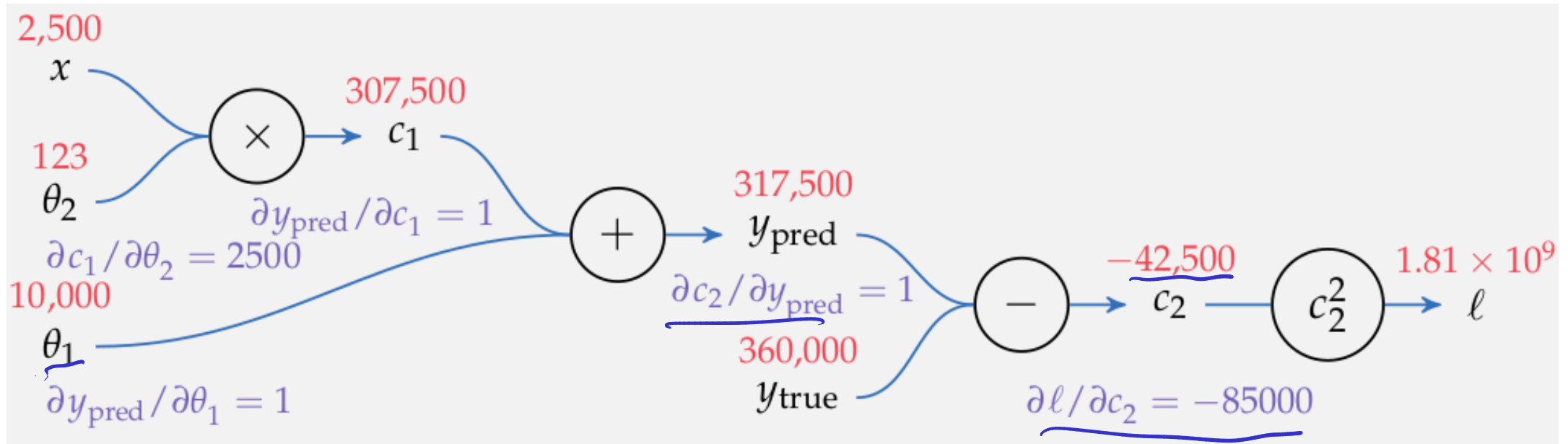
$$\hat{y} = \theta_2 x + \theta_1$$



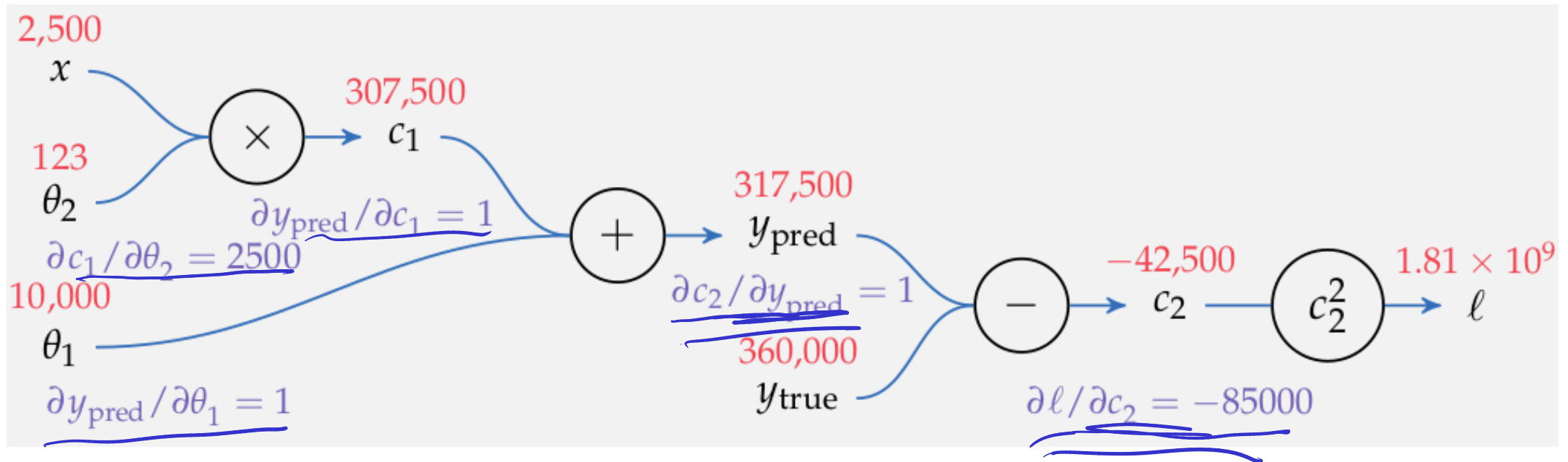
Backprop



Backprop



Backprop

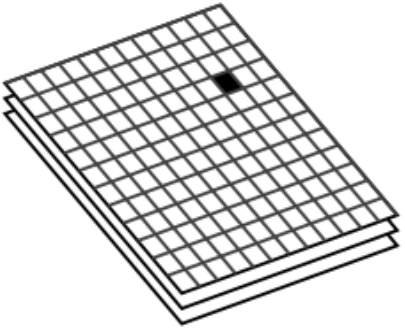


$$\frac{\partial \ell}{\partial \theta_1} = \frac{\partial \ell}{\partial c_2} \frac{\partial c_2}{\partial y_{\text{pred}}} \frac{\partial y_{\text{pred}}}{\partial \theta_1} = -85,000 \cdot 1 \cdot 1 = -85,000$$

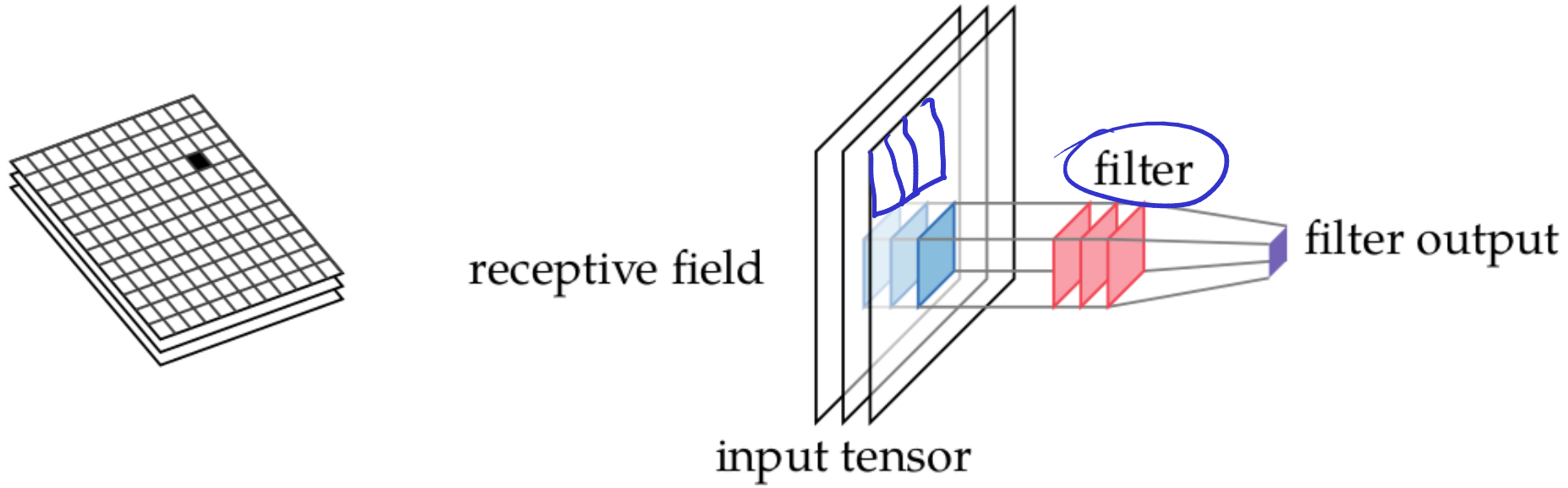
$$\frac{\partial \ell}{\partial \theta_2} = \frac{\partial \ell}{\partial c_2} \frac{\partial c_2}{\partial y_{\text{pred}}} \frac{\partial y_{\text{pred}}}{\partial c_1} \frac{\partial c_1}{\partial \theta_2} = -85,000 \cdot 1 \cdot 1 \cdot 2,500 = -2.125 \times 10^8$$

On Your Radar: ConvNets

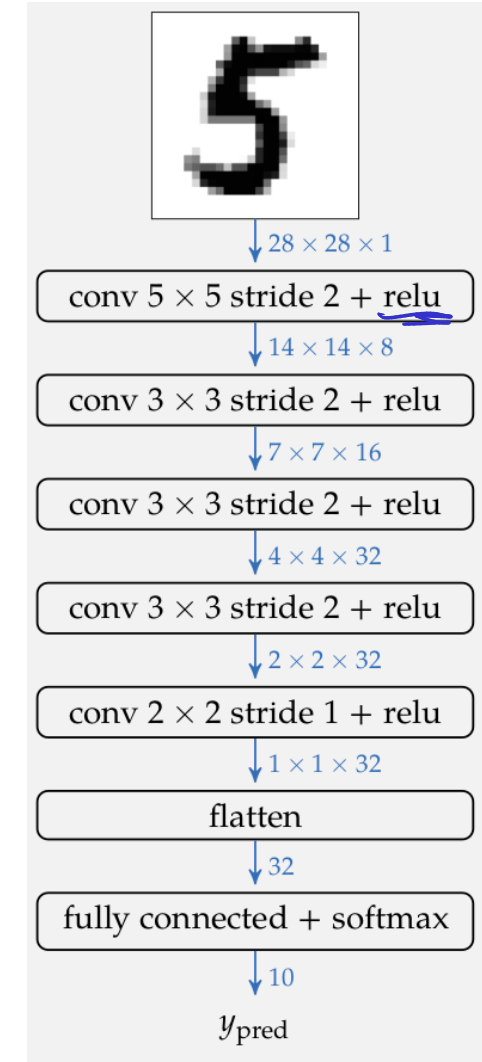
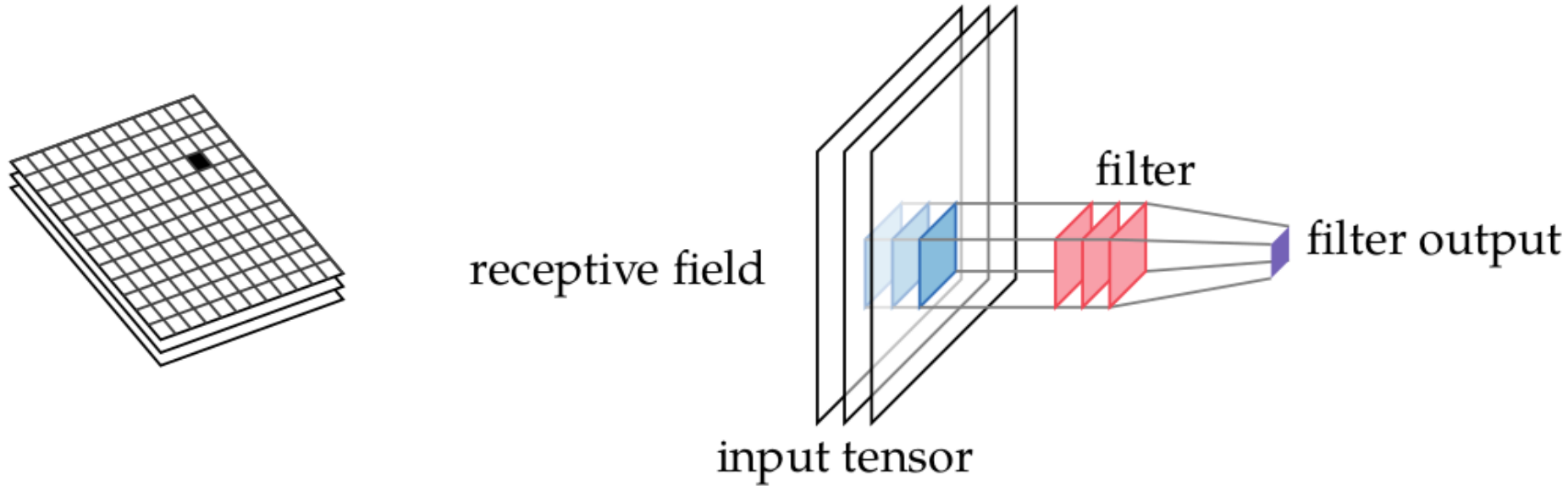
On Your Radar: ConvNets



On Your Radar: ConvNets



On Your Radar: ConvNets



On Your Radar: Regularization

On Your Radar: Regularization

$$\arg \min_{\boldsymbol{\theta}} \sum_{(x,y) \in \mathbf{D}} \ell(f_{\boldsymbol{\theta}}(x), y) - \beta \|\boldsymbol{\theta}\|^2$$

On Your Radar: Regularization

$$\arg \min_{\boldsymbol{\theta}} \sum_{(x,y) \in \mathbf{D}} \ell(f_{\boldsymbol{\theta}}(x), y) - \beta \|\boldsymbol{\theta}\|^2$$

e.g. Batch norm, dropout, resnets