Announcements

- Please fill out course evaluations
- Last two lectures this week
- Unit 5 & 6 Test scheduled for next Tuesday
  - During class hours
  - Same rules as previous tests
  - Please make sure your laptop has enough power for the test.

Neural Network applications

Image Generation (DALL-E)

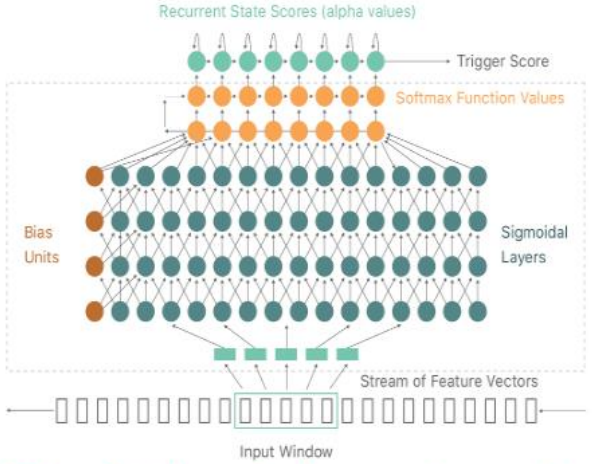
An astronaut lounging in a tropical resort. → Neural Net →

Teddy bears mixing chemicals. → Neural Net →

https://openai.com/dall-e-2/

Natural language & speech (Alexa/Siri)

"It's cold, bring your mittens!"



"Hey Siri, how's the weather today?"

Figure 2. The Deep Neural Network used to detect "Hey Siri." The hidden layers are actually fully connected. The top layer performs temporal integration. The actual DNN is indicated by the dashed box.

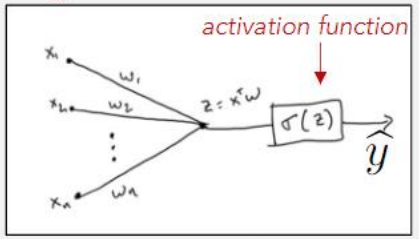
https://machinelearning.apple.com/2017/10/01/hey-siri.html

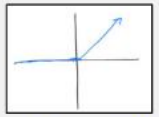
Activity: Neural Networks and SGD

Neurons

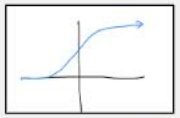
binary classification: $\sigma(z) = \text{sign}(z)$
regression: $\sigma(z) = z$

Single Neuron





ReLU: $\sigma(z) = \max(0, z)$

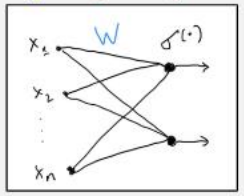


sigmoid/logistic: $\sigma(z) = \frac{1}{1+e^{-z}}$

Neural Networks

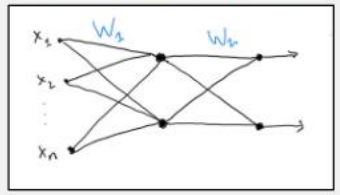
Single layer

$\hat{y} = \sigma(W^T x)$



Two layers

$\hat{y} = \sigma(W_2^T \sigma(W_1^T x))$



Training

n-layers

$$\min_{\boldsymbol{w}} \ell(\boldsymbol{w}) \rightarrow \min_{\boldsymbol{w}} \sum_i (\hat{y}_i - y_i)^2 \rightarrow \min \sum_i (\sigma(W_n^T \dots \sigma(W_2^T \sigma(W_1^T x_i)) \dots) - y_i)^2$$

- *Not convex, but we can still use gradient descent to find good W*
- *Derivatives for weights of each layer have simple expression*

Two ideas:

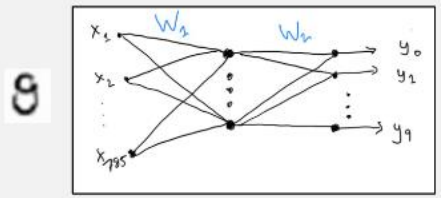
1) Stochastic Gradient Descent
2) Backpropagation

Need to compute the gradient!
Chain rule!

Example Application

MNIST database: recognizing handwritten digits

- Input: 28x28 image, stacked as vector $x \in \mathbb{R}^{785}$
- Output: likelihood of 0, 1, 2, … $\hat{y} \in [0,1]^{10}$



$$y = \begin{bmatrix} 0.05 \\ 0.01 \\ \vdots \\ 0.95 \\ 0.02 \end{bmatrix}$$

Other Applications

- natural language processing
- translation
- image processing
- Alexa/Siri (Hey Siri)
- self driving cars, …