Neural Networks: Prediction (forward pass)

Machine Learning
Fall 2023
Neural Networks

• What is a neural network?
• Predicting with a neural network
• Training neural networks
• Practical concerns
This lecture

• What is a neural network?

• Predicting with a neural network

• Training neural networks

• Practical concerns
Let us consider an example network

We will use this example network as to introduce the general principle of how to make predictions with a neural network.
Let us consider an example network

Naming conventions for this example
- Inputs: x
- Hidden: z
- Output: y
Let us consider an example network

Naming conventions for this example
- Inputs: \( x \)
- Hidden: \( z \)
- Output: \( y \)

Bias feature, always 1
Let us consider an example network

Naming conventions for this example
- Inputs: $x$
- Hidden: $z$
- Output: $y$

Sigmoid activations
Bias feature, always 1
Let us consider an example network

Naming conventions for this example
- Inputs: $x$
- Hidden: $z$
- Output: $y$

Bias feature, always 1

Sigmoid activations

Linear activation
Let us consider an example network

Naming Convention for Weights

$$w_{target\_layer\_from,to}$$
Let us consider an example network

Naming Convention for Weights

\( w_{target\_layer\ from,to} \)

\( w^o_{01} \)

From neuron #0 to neuron #1 in output layer
How to predict with a neural network: The forward pass

Given an input $x$, how is the output predicted?
The forward pass

Given an input $\mathbf{x}$, how is the output predicted

$$z_1 = \sigma(w_{01}^h + w_{11}^h x_1 + w_{21}^h x_2)$$
The forward pass

Given an input $x$, how is the output predicted

$$z_2 = \sigma(w_{02}^h + w_{12}^h x_1 + w_{22}^h x_2)$$

$$z_1 = \sigma(w_{01}^h + w_{11}^h x_1 + w_{21}^h x_2)$$
The forward pass

Given an input \( x \), how is the output predicted

\[
\text{output } y = w_{01}^o + w_{11}^o z_1 + w_{21}^o z_2
\]

\[
z_2 = \sigma(w_{02}^h + w_{12}^h x_1 + w_{22}^h x_2)
\]

\[
z_1 = \sigma(w_{01}^h + w_{11}^h x_1 + w_{21}^h x_2)
\]
The forward pass

Given an input $x$, how is the output predicted?

Output $y = w_{01}^o + w_{11}^o z_1 + w_{21}^o z_2$

$z_2 = \sigma(w_{02}^h + w_{12}^h x_1 + w_{22}^h x_2)$

$z_1 = \sigma(w_{01}^h + w_{11}^h x_1 + w_{21}^h x_2)$

Questions?