Neural Net Examples

CS 1678 Intro to Deep Learning
Feb. 4, 2021
First architecture

Input: $x_0$, $x_1$, $x_2$

Hidden: $z_0$, $z_1$

Output: $y_1$

Weights: $w^{(1)}_{10}$, $w^{(1)}_{11}$, $w^{(1)}_{12}$, $w^{(2)}_{10}$, $w^{(2)}_{11}$
Computing activations

• In all examples, $x = [x_0 \ x_1 \ x_2]$, where $x_0 = 1$
• Assume sigmoid activation function
• Initialize all weights to 0.1
• First example: $x = [1 \ 1 \ 0]$
• Second example: $x = [1 \ 0 \ 1]$
• Third example: $x = [1 \ 1 \ 1]$
Computing activations

• First example:
  • At hidden: $z_1 = ?$
  • At output: $y_1 = ? y_{\text{pred}} = ?$

• Second example:
  • At hidden: $z_1 = ?$
  • At output: $y_1 = ? y_{\text{pred}} = ?$

• Third example:
  • At hidden: $z_1 = ?$
  • At output: $y_1 = ? y_{\text{pred}} = ?$
Second architecture
Computing activations

• In all examples, \( x = [x_0 \ x_1 \ x_2] \), where \( x_0 = 1 \)
• Assume sigmoid activation function
• Initialize all weights to 0.05
• First example: \( x = [1 \ 1 \ 0] \)
• Second example: \( x = [1 \ 0 \ 1] \)
• Third example: \( x = [1 \ 1 \ 1] \)
Computing activations

• First, second, third example:
  • At hidden:
    • $z_1 = ?$
    • $z_2 = ?$
  • At output:
    • $y_1 = ?$
    • $y_2 = ?$
    • $y_{pred} = [1 \ 1]$
Training the first network

- Perform backpropagation using stochastic gradient descent (one sample at a time)
- Weights are initially all 0.1
- Learning rate is 0.3
- Sigmoid activation function at hidden and output
- $d s(x) / dx = s(x) (1 - s(x)) \, dx$
- Samples have the following labels:
  - First example: $x = [1 \ 1 \ 0]$, $y = 1$
  - Second example: $x = [1 \ 0 \ 1]$, $y = 0$
  - Third example: $x = [1 \ 1 \ 1]$, $y = 1$
- Preview: What do you expect final weights to be?
Learning from first example

• First example: \( x = [1 \ 1 \ 0], \ y = 1 \)
• Weights are \( w^{(1)}_{10} = w^{(1)}_{11} = w^{(1)}_{12} = w^{(2)}_{10} = w^{(2)}_{11} = 0.1 \)
• Activations are \( z_1 = 0.5498, \ y_1 = 0.5387 \)
• Compute errors:
  • \( \delta_{y_1} = ? \)
  • \( \delta_{z_1} = ? \)

• Update weights:
  • \( w^{(2)}_{10} = w^{(2)}_{10} - ? \)
  • \( w^{(2)}_{11} = w^{(2)}_{11} - ? \)
  • \( w^{(1)}_{10} = w^{(1)}_{10} - ? \)
  • \( w^{(1)}_{11} = w^{(1)}_{11} - ? \)
  • \( w^{(1)}_{12} = w^{(1)}_{12} - ? \)
Recap

• Do the $w^{(1)}$ weights we obtained make sense?