



MITx 6.86x

Machine Learning with Python-From Linear Models to Deep Learning[Course](#)[Progress](#)[Dates](#)[Discussion](#)[Resources](#)[Course](#) / [Unit 3. Neural net...](#) / [Lecture 9. Feedforward Neural Networks, Back](#)[< Previous](#)

2. Back-propagation Algorithm

[Bookmark this page](#)

Exercises due Mar 29, 2023 08:59 -03 Completed

Back-propagation Algorithm



Video

[!\[\]\(de95854c7ee024cfadc48187bbb781b2_img.jpg\) Download video file](#)

Transcripts

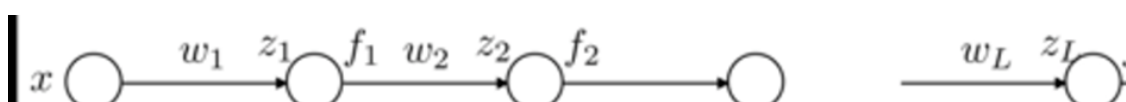
[!\[\]\(6059a5aa8b4ca7bb793408023d6c6e42_img.jpg\) Download SubRip \(.srt\) file](#)[!\[\]\(c50c8b7b2cc2cf9ff925edec0ee94c0d_img.jpg\) Download Text \(.txt\) file](#)

Once we set up the architecture of our (feedforward) neural network, our goal will be to find the weights that minimize our loss function. We will use the **stochastic gradient descent algorithm** (discussed in [Lecture 4](#) and revisited in [lecture 5](#)) to carry out the optimization.

This involves computing the gradient of the loss function with respect to the weight parameters.

Since the loss function is a long chain of compositions of activation functions with the weights entering at different stages, we will break down the computation of the gradient into different parts using the chain rule; this way of computing the gradient is called the back-propagation algorithm.

In the following problems, we will explore the main step in the stochastic gradient descent algorithm by training the following simple neural network from the video:



Gradient Descent Update

1/1 point (graded)

Let η be the learning rate for the stochastic gradient descent algorithm.

Recall that our goal is to tune the parameters of the neural network so as to minimize the loss. The following is the appropriate update rule for the parameter w_1 in the stochastic gradient descent algorithm.

☒ $w_1 \leftarrow w_1 - \eta \cdot \nabla_{w_1} \mathcal{L}(y, f_L)$

☐ $w_1 \leftarrow w_1 + \eta \cdot \nabla_{w_1} \mathcal{L}(y, f_L)$

☐ $w_1 \leftarrow \eta \cdot \nabla_{w_1} \mathcal{L}(y, f_L)$

☐ $w_1 \leftarrow -\eta \cdot \nabla_{w_1} \mathcal{L}(y, f_L)$

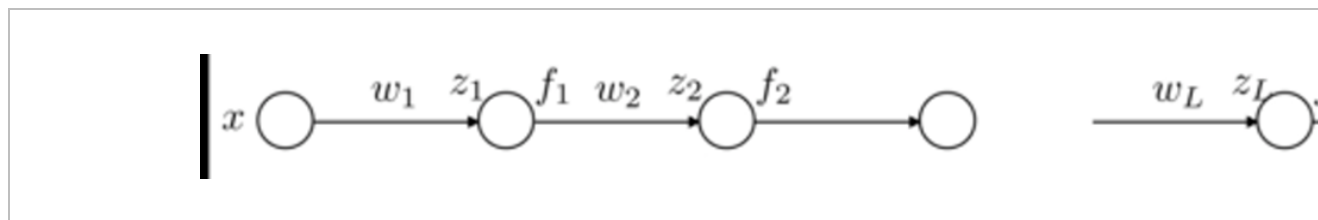


Submit

You have used 1 of 2 attempts

Recursive Expression - Part I

1/1 point (graded)



As above, let $\mathcal{L}(y, f_L)$ denote the loss function as a function of the predictions f_L and the target y .

$$z_1 = xw_1$$

$$\text{for } i = 2 \dots L: \quad z_i = f_{i-1}w_i \quad \text{where } f_{i-1} = f(z_{i-1}).$$

Let $\delta_i = \frac{\partial \mathcal{L}}{\partial z_i}$.

The first step to updating any weight w is to calculate $\frac{\partial \mathcal{L}}{\partial w}$.

**Submit**

You have used 2 of 2 attempts

Recursive Expression - Part II

1/1 point (graded)

As above, let $\mathcal{L}(y, f_L)$ denote the loss function as a function of the predictions f_L and the target y .

In this problem, we derive a recurrence relation between δ_i and δ_{i+1} .

Assume that f is the hyperbolic tangent function:

$$f(x) = \tanh(x)$$

$$f'(x) = (1 - \tanh^2(x)).$$

Which of the following option is the correct expression for δ_1 in terms of δ_2 ?

☒ $\delta_1 = (1 - f_1^2) \cdot w_2 \cdot \delta_2$

☐ $\delta_1 = (1 - f_1^2) \cdot w_1 \cdot \delta_2$

☐ $\delta_1 = (1 - f_2^2) \cdot w_2 \cdot \delta_2$

☐ $\delta_2 = (1 - f_1^2) \cdot w_2 \cdot \delta_1$

**Submit**

You have used 2 of 2 attempts

Final Expression of the Gradient

1/1 point (graded)

☒ $\frac{\partial \mathcal{L}}{\partial w_1} = x(1 - f_1^2)(1 - f_2^2) \cdots (1 - f_L^2)w_2w_3 \cdots w_L(2(f_L - y))$

☐ $\frac{\partial \mathcal{L}}{\partial w_1} = xw_2w_3 \cdots w_L(2(f_L - y))$



Submit

You have used 1 of 2 attempts

Discussion

< Previous

Next >

Topic: Unit 3. Neural networks (2.5 weeks):Lecture 9. Feedforward Neural Networks, Back Propagation, and Stochastic Gradient Descent (SGD) / 2. Back-propagation Algorithm



edX

[About](#)

[Affiliates](#)

[edX for Business](#)

[Open edX](#)

[Careers](#)

[News](#)

Legal

[Terms of Service & Honor Code](#)

[Privacy Policy](#)

[Accessibility Policy](#)

[Trademark Policy](#)

[Sitemap](#)

[Cookie Policy](#)

[Do Not Sell My Personal Information](#)

Connect



© 2023 edX LLC. All rights reserved.

深圳市恒宇博科技有限公司 [粤ICP备17044299号-2](#)