edX

MITx 6.86x
Machine Learning with Python-From Linear Models to Deep Learning

Course   Progress   Dates   Discussion   Resources

‹ Previous   🎥 ✓   📖 ✓   ✎ ✓   ✎ ✓   ✎ ✓

6. Hidden Layer Models

🔖 Bookmark this page

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

Models with Hidden Layer

▶   **0:00 / 0:00**   ▸   **1.0x**

Video
⬇ Download video file

Transcripts
⬇ Download SubRip (.srt) file
⬇ Download Text (.txt) file

For the following set of problems, let's consider a simple 2-dimensional classification ta
made up of 4 points listed below:

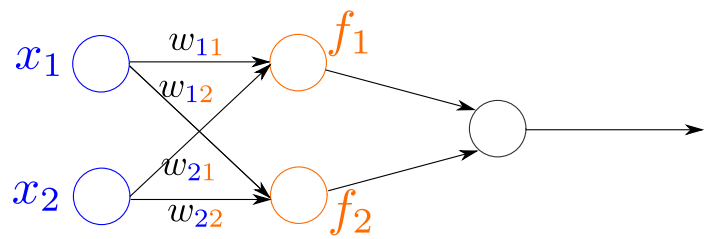$$x^{(1)} = (-1, -1) \quad , \qquad y^{(1)} = 1$$

$$x^{(2)} = (1, -1) \quad , \qquad y^{(2)} = -1$$

$$x^{(3)} = (-1, 1) \quad , \qquad y^{(3)} = -1$$

$$x^{(4)} = (1, 1) \quad , \qquad y^{(4)} = 1$$

The dataset is illustrated below (blue - positive, red - negative)

Let   denote the output of the two units in the hidden layer corresponding to th
respectively, i.e.

Consider the set   .

Assume that f is the linear activation function given by   .

For which of the following values of weights would the set   be linearly separable? (S

- [ ]
- [ ]
- [ ]
- [x] None of the above

✔

Submit   You have used 1 of 2 attempts

Non-linear Activation Functions

0/1 point (graded)
Again, let's focus on a network with one hidden layer with two units and use the same t
The weights of the network are given as follows:

✓

✗

Submit   You have used 2 of 2 attempts

Neural Network Learned parameters

1/1 point (graded)

Given a neural network with one hidden layer for classification, we can view the hidden
representation, and the output layer as a classifier using the learned feature representa

There're also other parameters that will affect the learning process and the performan
the learning rate and parameters that control the network architecture (e.g. number of
These are often called hyper-parameters.

Which of the following is/are optimized during a single training pass? (Note that cross-
before this point.) Check all that apply.

☐ The dimension of the feature representation

☑ The weights that control the feature representation

☐ The hyper-parameters

☑ The weights for the classifier

✓

Submit   You have used 1 of 2 attempts

Discussioı   ‹ Previous   Next ›

Topic: Unit 3. Neural networks (2.5 weeks):Lecture 8. Introduction to Feedforward
Neural Networks / 6. Hidden Layer Models

edX®

? **10 dim to 2 dim**

How do you plot a two dimensional version of a ten dimensional space?

? **Meaning of "Note that cross-validation is tuned before this point"**

Since we are doing a single training run, it seems to make no sense that any cross validation would have occ

💬 **Universal Approximation Theorem says that we can model any function with 3 layers,so why** networks require so many layer.Any comments to increase my understanding?

Universal Approximation Theorem says that we can model any function with 3 layers,so why do modern neur

💬 **Whats the professors point about redundancy,randomness is good? timestamp is 16:45**

Whats the professors point about redundancy,randomness is good?

? indices on w_ij are out of order?

💬 **Non-linear Activation Functions - no solution**

I have done the calculations for the second question and my answers show that none of the options work. I h

Legal

[Terms of Service & Honor Code](#)

[Privacy Policy](#)

[Accessibility Policy](#)

[Trademark Policy](#)

[Sitemap](#)

[Cookie Policy](#)

[Do Not Sell My Personal Information](#)

Connect

[Blog](#)

[Contact Us](#)

[Help Center](#)

[Security](#)

[Media Kit](#)