

Optimization Algorithms

- ✓ **Video:** Mini-batch Gradient Descent
11 min
- ✓ **Video:** Understanding Mini-batch Gradient Descent
11 min
- ✓ **Video:** Exponentially Weighted Averages
5 min
- ✓ **Video:** Understanding Exponentially Weighted Averages
9 min
- ✓ **Video:** Bias Correction in Exponentially Weighted Averages
4 min
- ✓ **Video:** Gradient Descent with Momentum
9 min
- ✓ **Video:** RMSprop
7 min
- ✓ **Reading:** Clarification about Upcoming Adam Optimization Video
1 min
- ✓ **Video:** Adam Optimization Algorithm
7 min
- 📖 **Reading:** Clarification about Learning Rate Decay Video
1 min
- ▶ **Video:** Learning Rate Decay
6 min
- ▶ **Video:** The Problem of Local Optima
5 min

Lecture Notes (Optional)

Quiz

Programming Assignment

- 🔗 **Programming Assignment:** Optimization Methods
3h

Heroes of Deep Learning (Optional)

- ▶ **Video:** Yuanqing Lin Interview
13 min

Clarification about Learning Rate Decay Video

Please note that in the next video, at time 3:35, the values for alpha should be:

Epoch 1: alpha 0.1

Epoch 2: alpha 0.067

Epoch 3: alpha 0.05

Epoch 4: alpha 0.04

The formula for learning rate decay is:

$$\alpha = \frac{1}{1 + \text{decayRate} \times \text{epochNumber}} \alpha_0$$

Learning rate decay

1 epoch = 1 pass through data.

$$\alpha = \frac{1}{1 + \text{decay-rate} * \text{epoch-num}} \alpha_0$$

Epoch	α
1	0.1
2	0.067
3	0.05
4	0.04
⋮	⋮

$\alpha_0 = 0.2$
 $\text{decay-rate} = 1$

Mark as completed

