Optimization Algorithms

<u>:=</u>

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

Lecture Notes (Optional)

Programming Assignment

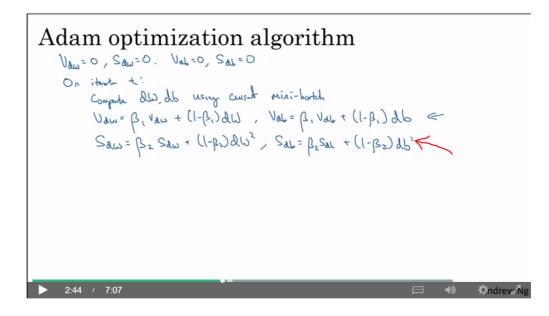
Heroes of Deep Learning (Optional)

Clarification about Upcoming Adam Optimization Video

Please note that in the next video at 2:44, the following picture is correct. However, later in the video, the db^2 lost the 2 .

The bottom right equation should be:

$$S_{db}=eta_2S_{db}+(1-eta_2)db^2$$



Mark as completed





