



deeplearning.ai

Deep Neural Networks

Parameters vs Hyperparameters

What are hyperparameters?

Parameters: $W^{[1]}, b^{[1]}, W^{[2]}, b^{[2]}, W^{[3]}, b^{[3]} \dots$

Hyperparameters: learning rate α
#iterations

technically, alpha is parameter,
but is a parameter that
determines the real parameters

#hidden layers L

#hidden units $n^{[1]}, n^{[2]}, \dots$

choice of activation function

Because it is the hyper parameters that you know
somehow determine the final value of the
parameters W and b that you end up with.

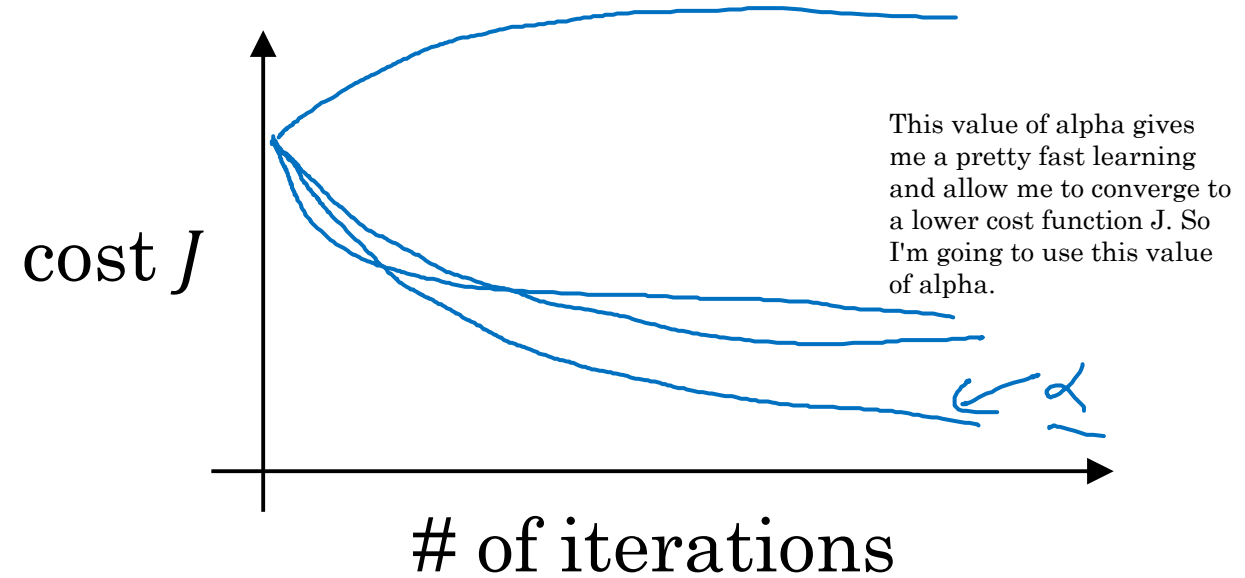
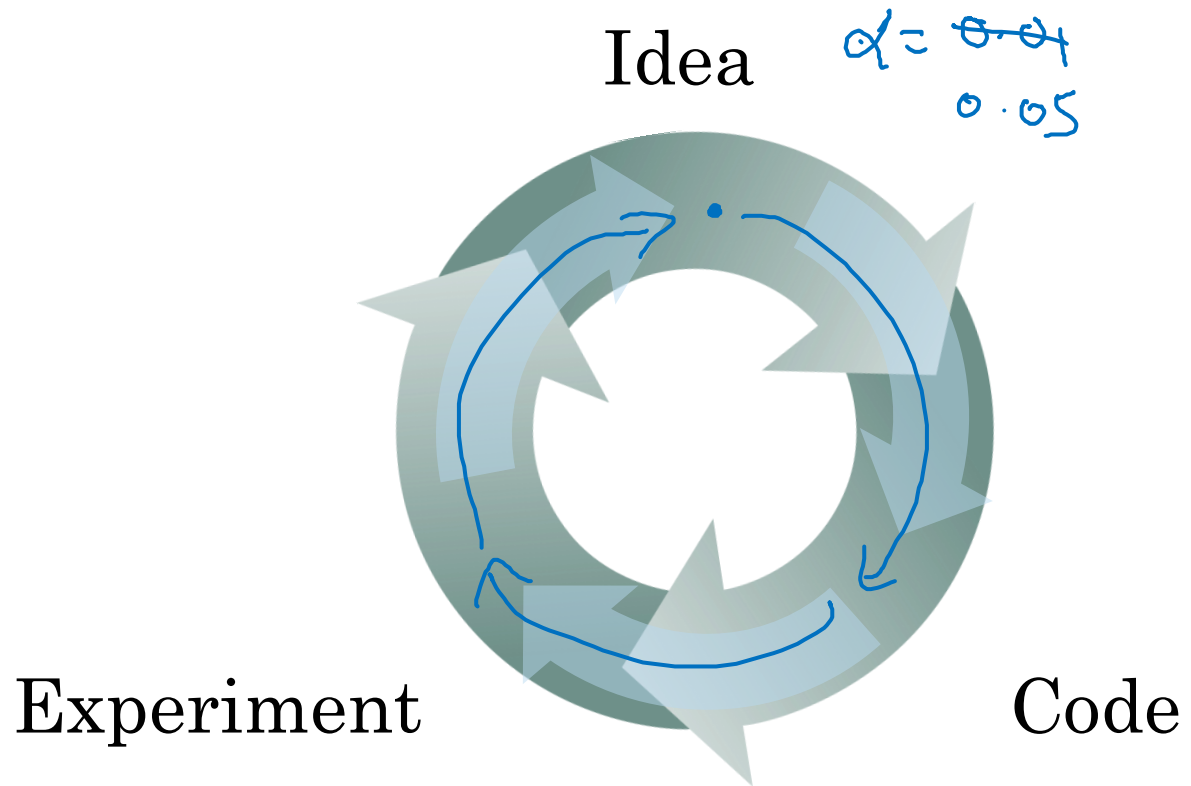
Later: Momentum, mini-batch size, regularizations, ...

momentum

various forms of
regularization parameters

So when you're training a deep net for your own application, you'll find that there maybe a lot of possible settings for the hyperparameters that you need to just try out.

Applied deep learning is a very empirical process



Vision, Speech, NLP, Ad, Search, Recommendation.

Just try a few values for the hyper parameters and double check if there's a better value for the hyper parameters and as you do so you slowly gain intuition as well about the hyper parameters that work for your problem.