



deeplearning.ai

Hyperparameter tuning

Tuning process

Hyperparameters — How to organize & set hyperparams

P_0 α = Learning rate

P_1 β → momentum

P_3 ϵ → RMS/Adam [default: 10^{-8}]

P_2 #layers

P_1 # hidden units/layer

P_2 Learning Rate decay

P_1 mini batch size

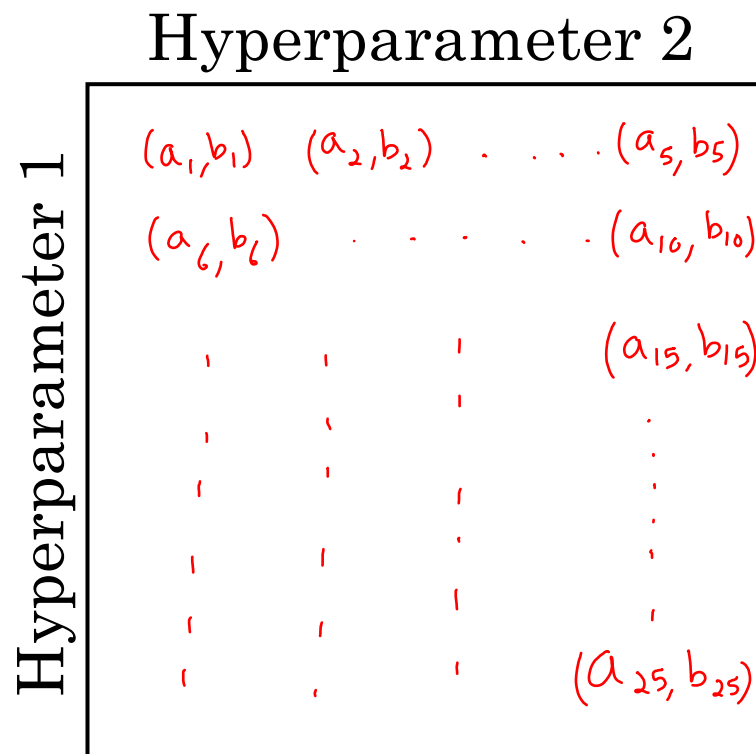
P_3 β_1/β_2 → Adam β values for momentum, RMS prop

[Generally default to $\beta_1 = 0.9$, $\beta_2 = 0.999$]

P_0 - set these first
 P_1 - set these second
 P_2 - set these 3rd

Try random values: Don't use a grid

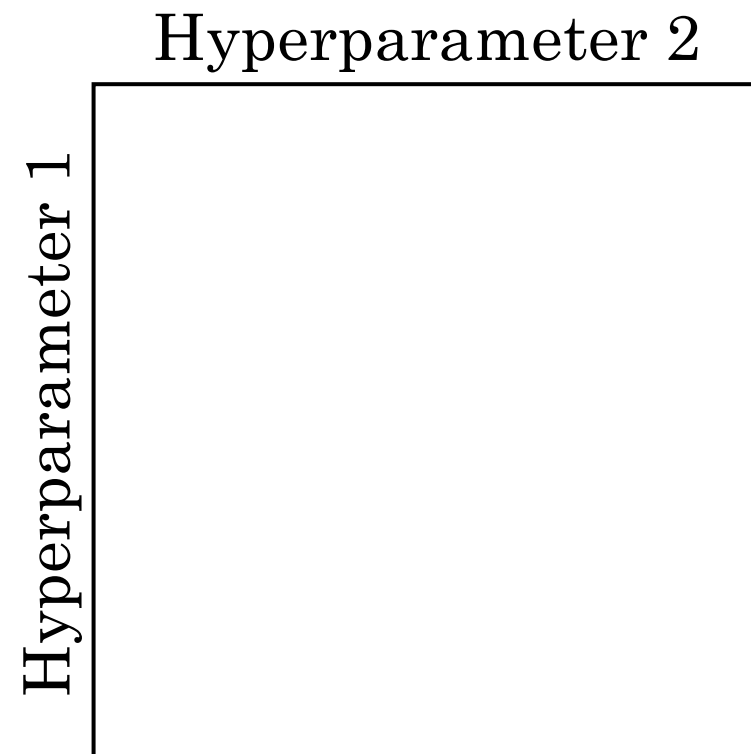
How to set values of param 1, 2 & see which combo works?



old common practice
(Don't do)

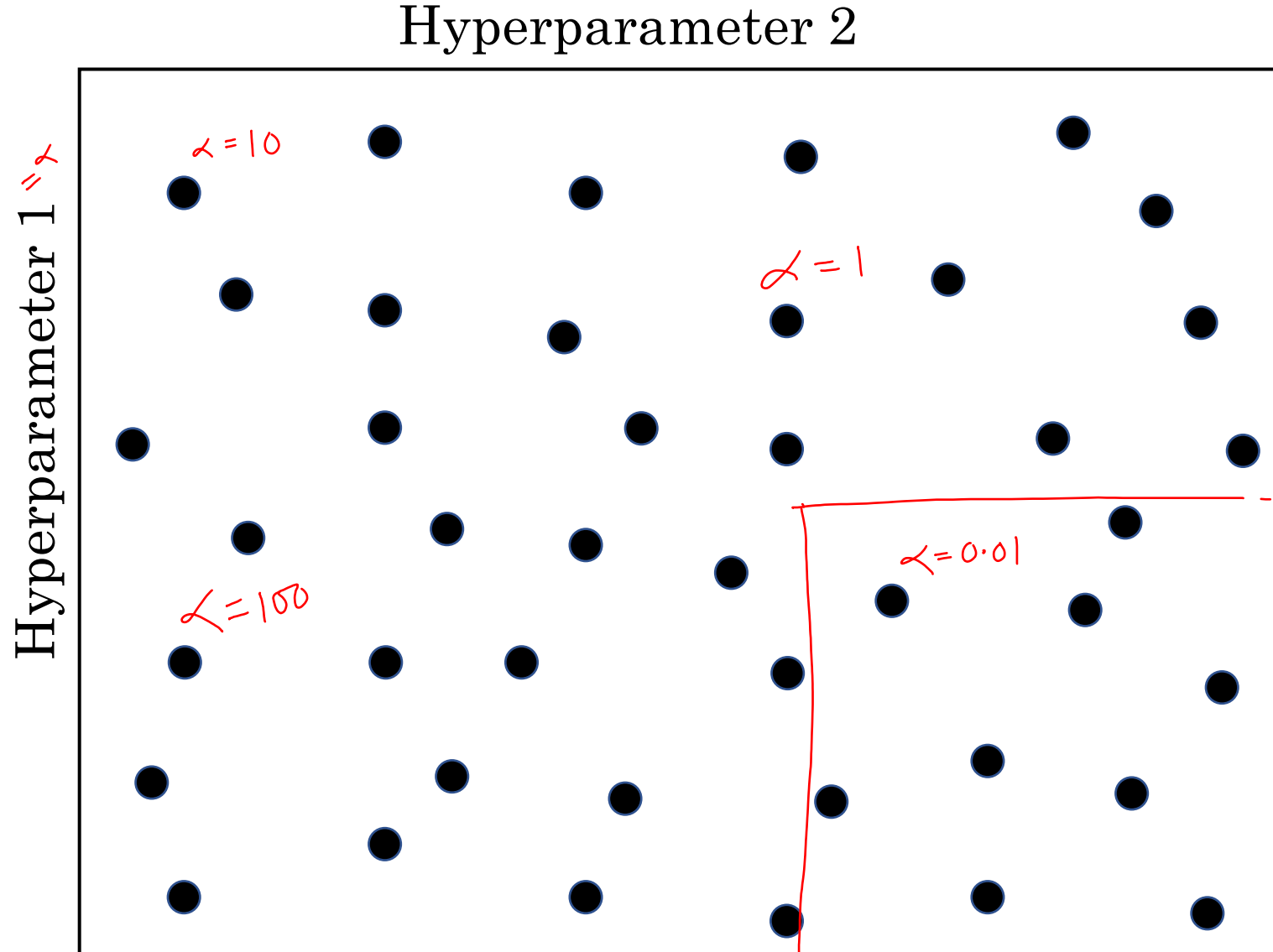
} works only
when # hyper
params are small

Try out a, b pairs & see which
values perform well together in the model



- Instead, Sample at Random
- Also give weight to the samples of values which give a good model perf. with the Imp params (α , #layers, #hidden units)

Coarse to fine



- First try, set α to values of diff order of magnitude $\alpha = 100, \alpha = 10, \alpha = 1, \alpha = 0.01, \alpha = 0.0001$ (COARSE)
- once you find it works well w/ say $\alpha = 0.01$, try $\alpha = 0.05$
 $\alpha = 0.09$
 $\alpha = 0.009$ etc. (FINE)