

Practical Considerations

What framework to use?

1. Keras
2. Tensorflow
3. Pytorch

What architecture to use?

1. Take whatever is the best on ILSVRC (latest ResNet)
2. Download a pre-trained model
3. Potentially add/delete some parts to it
4. Finetune it for your application

What hyperparameters to use?

1. Use whatever is reported to work best on ILSVRC
2. Play with regularization strength (dropout rates, Batch Norm, stochastic depth)

My First Convergence

- So now you want to become the deity of your deep architecture
- These appetites are completely natural

Follow your heart

1. Start with someone's architecture and modify from there
2. If nothing exists - follow the deep learning *mantra* - "*follow your heart*"
3. Add layers until you overfit and then add regularizers*

*If it doesn't immediately converge please don't go pouring salt water on your computer. You are brilliant, and the model you constructed is fantastic!

Learning Rate

1. Too small - takes days to converge
2. Too large - miss any local minimums
3. Start with static - if it's converging - use adaptive or decaying

Weight Initialization

1. Initialize with isotropic Gaussian and *follow your heart* with SD
2. Try values from 0.5 to 0.005
3. Typically larger the network, smaller the SD

Batch Normalization

Batch normalize the heck out of everything.

“For every complex problem there is an answer that is clear, simple, and wrong.”

- H. L. Mencken

Thanks