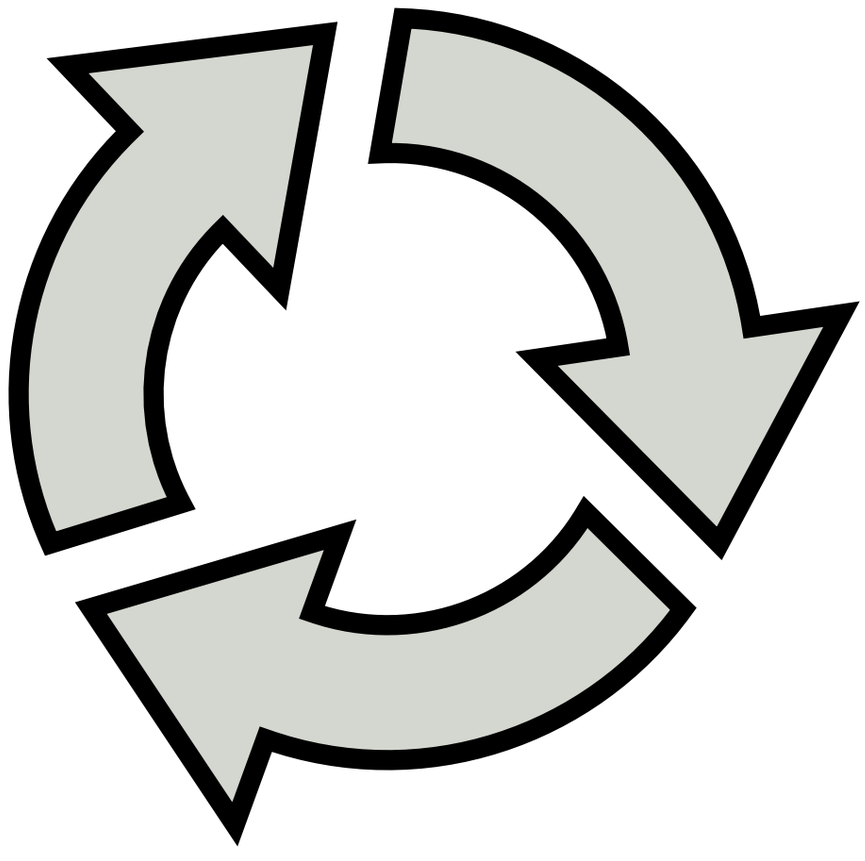
Summary, a practical guide to deep network optimization

Graduate student descent

semi-automated

Look at your data / model output



manual

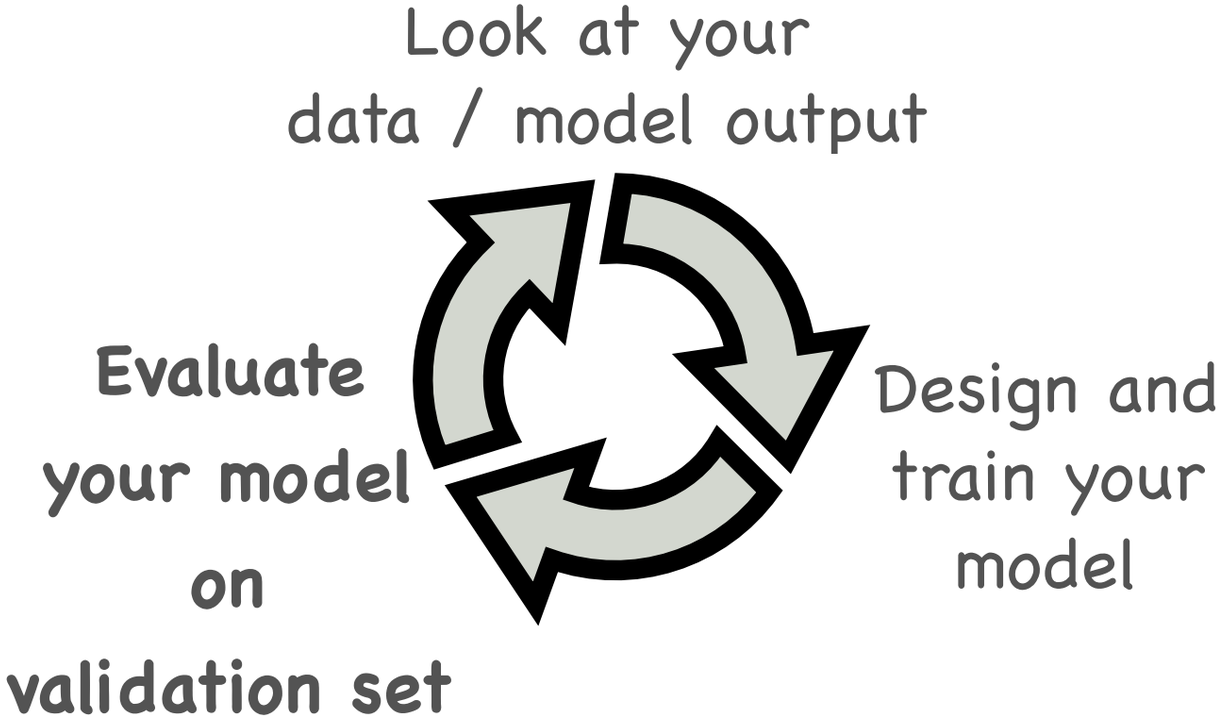
Design and train your model

Evaluate your model on validation set

automated

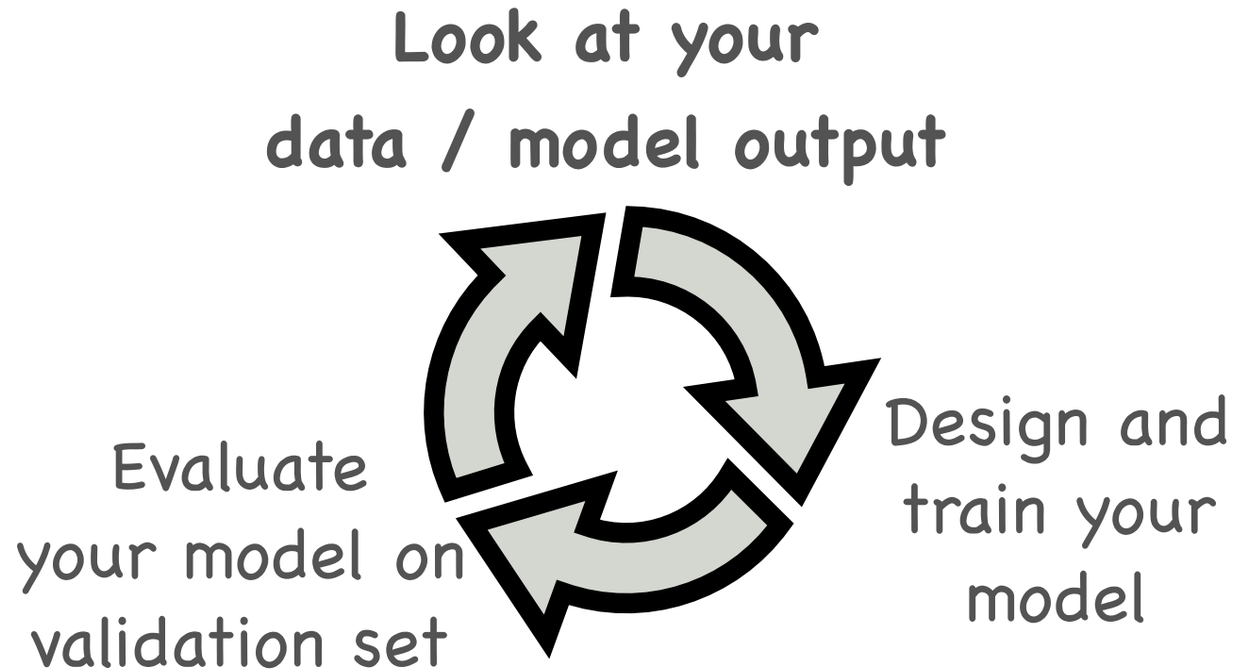
Evaluation on validation set

- Run during training
- Every epoch or n iterations
- Log in TensorBoard



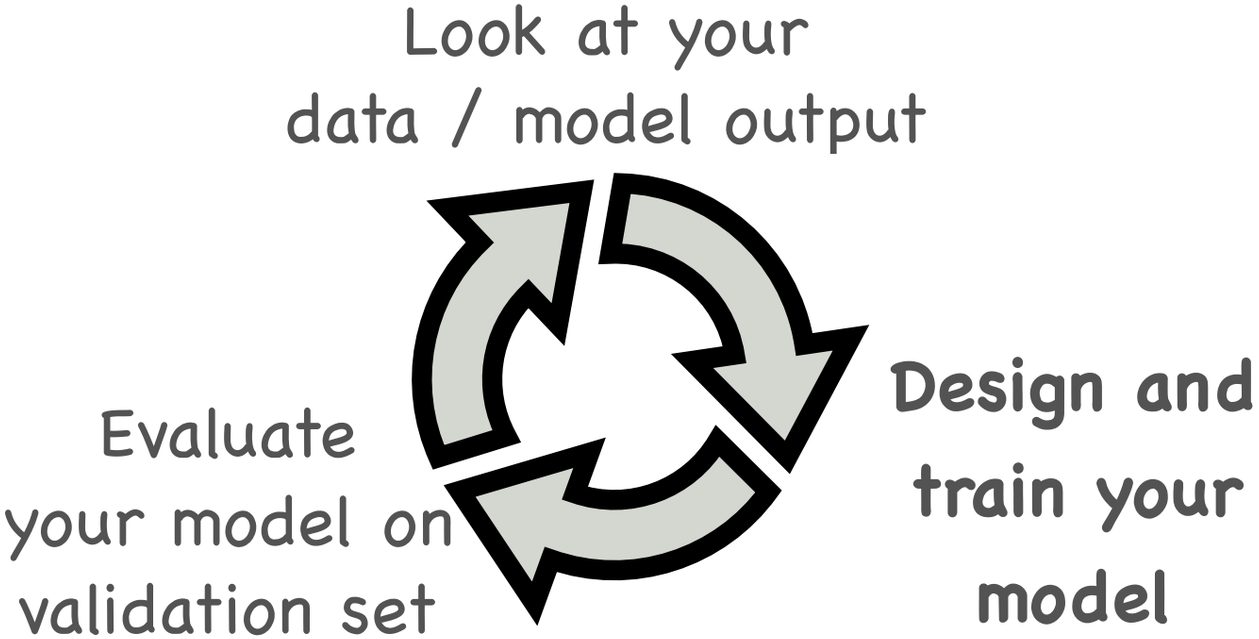
Look at your data / model output

- Run during training
  - Every epoch or n iterations
- Log in TensorBoard
- Select same training and validation images



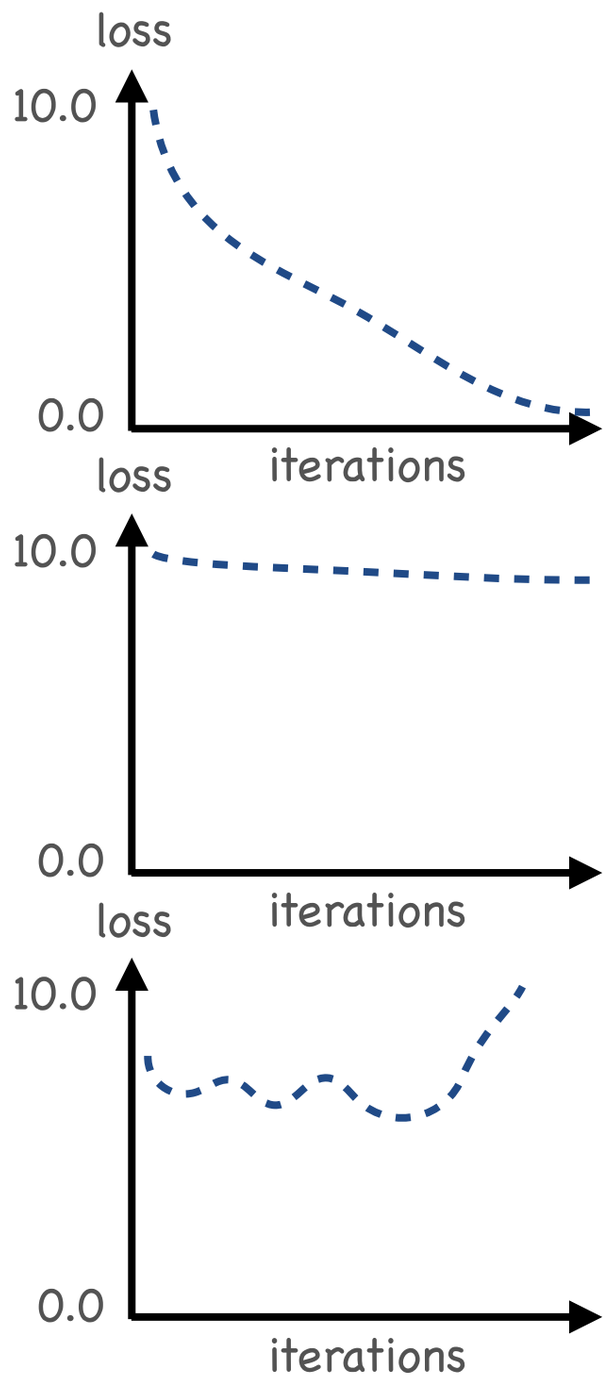
Design and train your model

- Mostly manual work



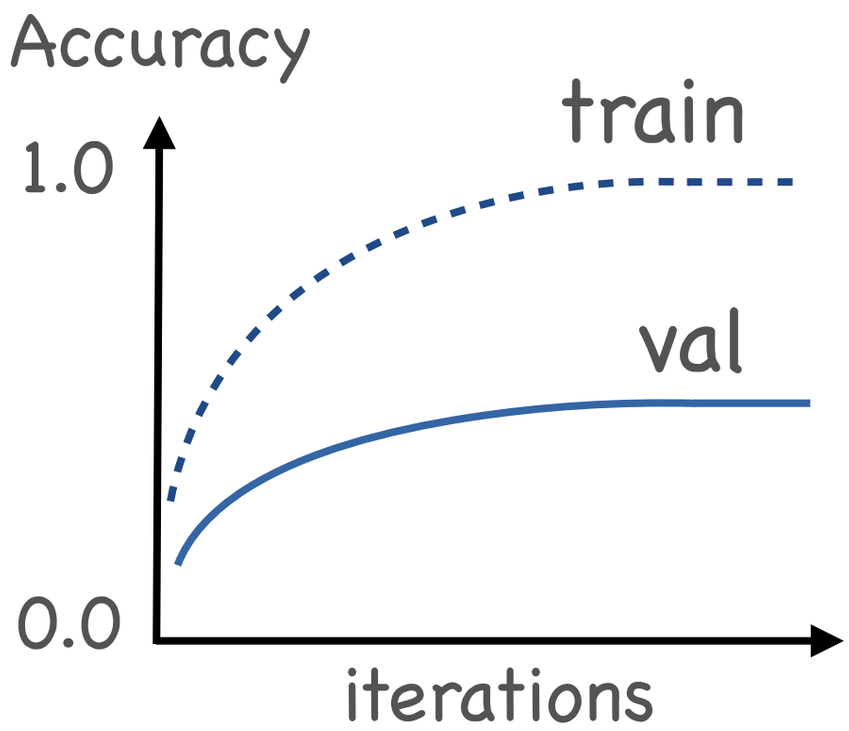
Design and train your model

- Network does not train
  - Vanishing or exploding gradients?
    - Fix initialization and learning rate
  - Slow training
    - Add normalization
    - Residual connections
- Iterate until model trains



Design and train your model

- Network overfits to training data
  - Add data augmentation
  - Early stopping
  - Try a pre-trained network
  - Collect more data
- Iterate until model generalizes well



Design and train your model

- Network fits training and validation data well
- Stop graduate student descent
- Take a break
- Evaluate on test set