Deep Learning
Epochs / Batch Size / Iterations

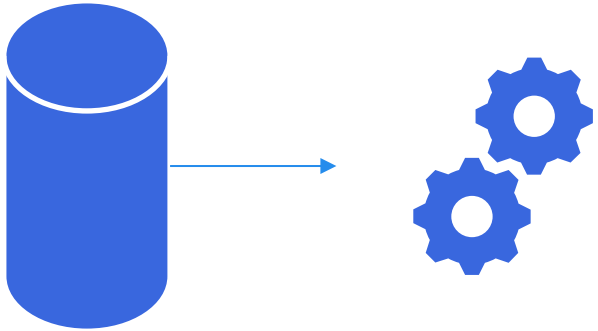
Bert Gollnick

Deep Learning Epoch/BatchSize/Iterations

Batch Size

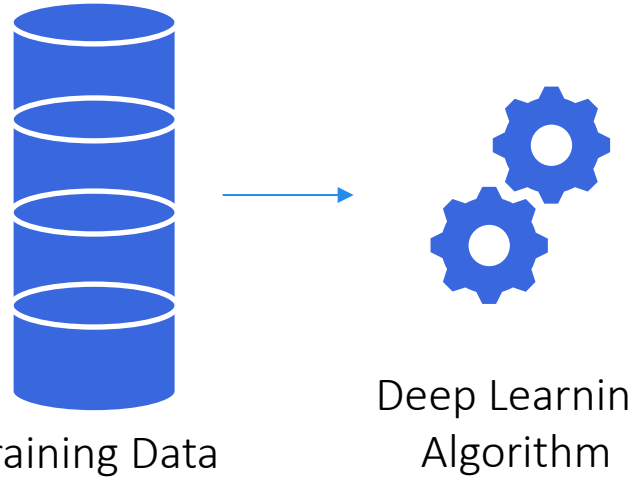
Problem

- Usually not possible to pass all data at once.



Solution

- Dataset divided into mini-batches.
- Batch Size represents training samples in mini-batch.



Deep Learning Epoch/BatchSize/Iterations

Epoch

- Epoch finished when complete dataset is passed through the neural network forward and backward.

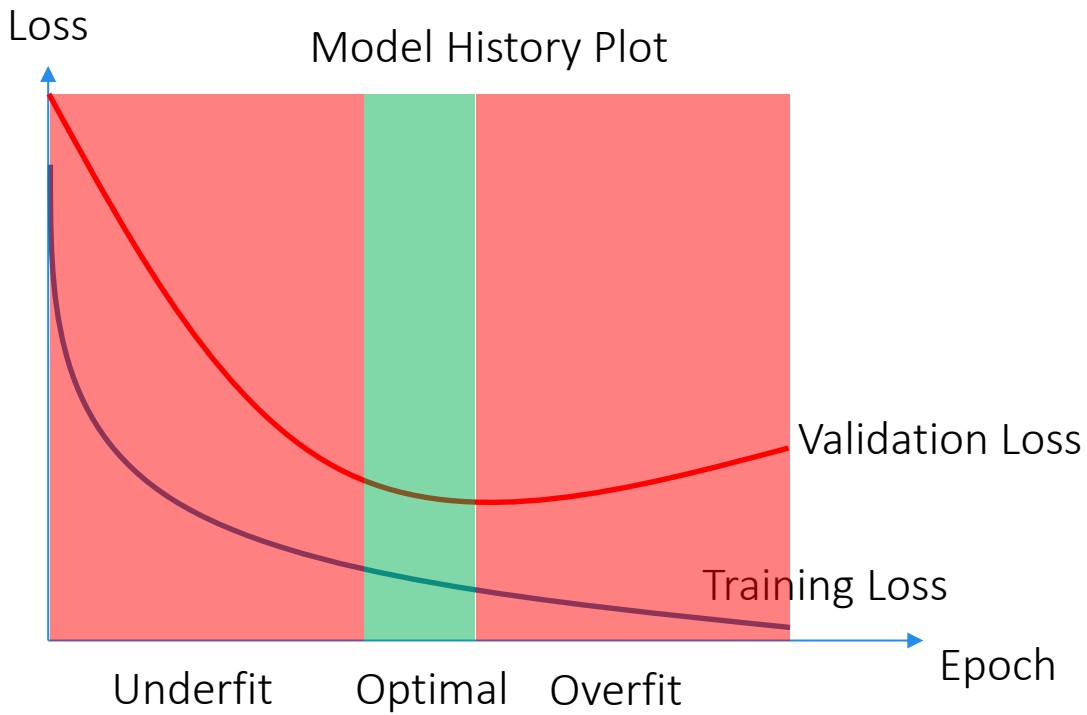
- An iteration is a single gradient update (update of model weights).

- Example:
  - Dataset of 500 images
  - Mini-batches of 50 images

  → 10 iterations required for finishing a single epoch

Deep Learning Epoch/BatchSize/Iterations

Best Practices: Number of Epochs



- Strategy: Early Stopping

Deep Learning Epoch/BatchSize/Iterations

Best Practices: Batch Size

- Batch size
  - Too small: high degree of variance within batches (small sample does not represent complete dataset)
  - Too large: possibly memory issues; overfitting
  - Optimum needs to be found
- Hyperparameter „batch size“ influences other hyperparameters
- Combination of hyperparameters relevant, e.g. batch size & learning rate

- Heuristic:
  - $BatchSize = \sqrt{Observations}$
  - Often: batch size defined as multiple of 2, e.g. 128, 256, …