

Training loop

1. Draw a batch of training samples x and corresponding targets y
- 2. Run the network on x (this is called "forward pass"), obtain predictions y_{pred}**
- 3. Compute the "loss" of the network on the batch, a measure of the mismatch between y_{pred} and y**
4. Update all weights of the network in a way that slightly reduces the loss on this batch.

