Neural networks
Training neural networks - model selection
MACHINE LEARNING

Topics: training, validation and test sets, generalization

- Training set $D_{\text{train}}$ serves to train a model
- Validation set $D_{\text{valid}}$ serves to select hyper-parameters
- Test set $D_{\text{test}}$ serves to estimate the generalization performance (error)

- Generalization is the behavior of the model on unseen examples
  - this is what we care about in machine learning!
MODEL SELECTION

Topics: grid search

• To search for the best configuration of the hyper-parameters:
  ‣ you can perform a grid search
    - specify a set of values you want to test for each hyper-parameter
    - try all possible configurations of these values
  ‣ you can perform a random search
    - specify a distribution over the values of each hyper-parameters (e.g. uniform in some range)
    - sample independently each hyper-parameter to get a configuration, and repeat as many times as wanted

• Use a validation set performance to select the best configuration

• You can go back and refine the grid/distributions if needed
Topics: early stopping

- To select the number of epochs, stop training when validation set error increases (with some look ahead)