Key concepts

- Learning from examples
- Non-parametric vs. parametric
- Optimization problem:
  \[
  \min_{params} \text{objective}(params, data) \\
  \text{s.t. constraint}(params)
  \]
- Convex optimization vs. non-convex
- Unsupervised vs. supervised
- Generative vs. discriminative
- Bias-variance tradeoff (or train-fit vs. generalization)
Unsupervised learning

- Clustering
  - K-means
  - Gaussian mixture model (GMM)
  - Expectation Maximization (EM)
- Dimensionality reduction
  - Principal component analysis (PCA)
Supervised learning

- Hypothesis: \( h: \mathcal{X} \rightarrow \mathcal{Y} \)
  - Classification: \( Y \) discrete (e.g. \( \{-, +\} \))
  - Regression: \( Y \) continuous (e.g. \([0,1]\))

```
function learning_algorithm(sample S) {
    ...
    return classifier h*
}
```
Support vector machine (SVM)

- Large margin
- Support vectors
- Dual optimization problem
- Kernel trick
Ensemble of classifiers

- Ensemble
- Boosting
  - Reduce bias (improve fit)
Random forest

- Decision tree
- Random forest
  - Reduce variance (improve generalization)
Gradient descent

$$\min_{\Theta} f(\Theta, X)$$
(artificial)
Neural networks

Multiple-layer-perceptron

\[
\Pr(\text{walking}|x) \quad \Pr(\text{shower}|x)
\]

Input features \(x\)

Gradient descent with Back-propagation
More neural networks

- Deep neural networks (DNN)
- Convolutional neural networks (CNN)
- Recurrent neural networks (RNN)
- Long short-term memory (LSTM)