Machine learning

Intro
A subfield of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence

Arthur Samuel (1959) defined machine learning as a "Field of study that gives computers the ability to learn without being explicitly programmed".

Algorithms learning from data and making predictions on data

Links to

- Computational statistics
- Optimization
Optimality

- Definition: the best, most favourable or desirable, under some restrictions

- Requires a formal definition!
  - E.g. Maximize separation between data

- Other examples
  - Minimize mean squared errors
  - Minimize jerk (motor control)
  - Maximize entropy
  - …
Curse of dimensionality

- Example: quadratic fit

\[ y(w, x) = w_0 + \sum_{i=1}^{D} w_i x_i + \sum_{i=1}^{D} \sum_{j=1}^{D} w_{ij} x_i x_j \]

- Dimension of data D

- Number of fit parameters grows with \( D^2 \)!
Curse of dimensionality

- The "curse of dimensionality" depends on the algorithm
  - Combinatorial explosion: $2^D$
  - Sampling data in $D$-dimensional space (suppose 100 samples / dimension): $10^{2xD}$
  - Machine learning: the higher $D$, the more sample data is needed $\rightarrow$ depends on algorithm
- …
- Always make sure there is enough data for the algorithm used!
- Beware of over-fitting!
Model assessment & selection

- **Selection**: chose the best model based on an estimate of each model’s performance
  - Requires criterion!
- **Assessment**: estimate the prediction (generalization) error of the chosen model on new data
Model assessment & selection

- Model complexity
  - Training (blue)
  - vs test (red)

- Bias-variance tradeoff
Inference & decisions