Learning from Data

- is a special case of
- is applied in

Unsupervised Learning

- Linear subspace methods
  - Principal Component Analysis (PCA)
    - dimension reduction
    - gaussian-like data
  - Independent Component Analysis (ICA)
    - non-gaussian data
  - Canonical Correlation Analysis (CCA)
    - multiview learning
- Non-linear subspace methods
  - kernel PCA
  - Laplacian eigen-map
    - dimension reduction
  - HGR-maximal correlation
    - nonlinear correlation
  - deep CCA.
    - multiview learning
  - non-parametric methods
- Clustering
  - k-means
  - floyd algorithm
  - Gaussian mixture model
  - spectral clustering
  - matrix factorization

Statistical & Mathematical Tools

- Model Estimation
  - Maximum likelihood estimation (MLE)
  - Maximum a-posteriori (MAP)
  - Variational Inference
  - Expectation Maximization (EM)
  - bayesian inference
- Optimization
  - Numerical optimization
    - gradient descent
    - newton's method
  - Convex optimization
    - Lagrange Duality
    - KKT Condition
- Matrix decomposition
  - Eigenvalue decomposition
  - Singular value decomposition
  - Rayleigh–Ritz Theorem
- Spectral graph theory
- Inequalities
  - Cauchy–Schwarz inequality
  - Jensen's inequality
  - union bound
  - Markov inequality

Reinforcement Learning

- value iteration
- policy iteration
- fitted value iteration
- Deep-Q learning
- model-based (MDP)
- model-free

Supervised Learning

- Linear models
  - linear regression
    - regression
  - logistic regression
    - binary classification
  - softmax regression
    - multi-label classific…
  - generalized linear models
  - perceptron algorithm
    - binary classification
  - Support Vector Machines (SVM)
    - binary classification
  - Gaussian discriminative analysis (GDA)
    - classification
  - Naive bayes
    - classification
    - discrete input
  - discriminative
  - generative
- Non-linear methods
  - kernel methods
    - kernel regression
    - kernel SVM
  - neural networks
    - MLP (multilayer perceptron)
    - CNN
  - Summary

Model selection

- Learning theory
  - VC dimension
  - Generalization bound (Uniform convergence theorem)
- General learning algorithm
  - Empirical Risk Estimation
  - Structured Risk Estimation
  - Instance-based learning
  - Rule-based learning
- Regularization
  - parameter norm penalty
  - MAP estimation
  - data augmentation
  - parameter sharing
  - drop out
- Bias-Variance tradeoff
- Cross-validation

Presented with xmind