Review

Ziping Zhao

School of Information Science and Technology
ShanghaiTech University, Shanghai, China

CS182: Introduction to Machine Learning (Fall 2021)
http://cs182.sist.shanghaitech.edu.cn

Outline

Outline

Further Machine Learning Topics - I

- More kernel methods and kernel learning
  - All operations on data in terms of inner product of input vectors
  - Kernel trick: mapping instances from input space (not necessarily vector space) to an inner product space without explicit computation of mapping
  - Kernel PCA, kernel FDA, Gaussian process, kernel ridge regression, etc.
  - Learning kernel function or kernel matrix from data instead of fixing it in advance
- Semi-supervised learning
  - Training data contain both labeled and unlabeled data
  - In between supervised learning and unsupervised learning
  - Semi-supervised classification, semi-supervised clustering
- Active learning
  - A.k.a. query learning, optimal experimental design
  - Learner interactively chooses which data points to label
  - Substantially reduce the number of labels required
- Weakly supervised learning

Further Machine Learning Topics - II

- Distance metric learning
  - Learning meaningful distance metric from data
  - Related to kernel learning in kernel methods
  - Embedding of nonlinear data manifold
- Transfer learning and multi-task learning
  - Inspired by human learning
  - Transfer of knowledge learned from one task to a different but related task
  - Learning multiple tasks simultaneously
  - One/few-shot learning
- Statistical relational learning and network data
  - Data instances not i.i.d.
  - Relations (edges) between instances (vertices) represented as graph/network
  - Discovering network structure, e.g., community detection

Further Machine Learning Topics - III

- Approximate inference for Bayesian methods
  - As opposed to exact inference methods which often are intractable
  - Sampling-based methods: Gibbs sampling, Markov chain Monte Carlo (MCMC) algorithms
  - Variational approximation methods
- Nonparametric Bayesian methods
  - Number of model parameters increases with sample size
  - Gaussian process, relevance vector machine
  - Dirichlet process (Chinese restaurant process), beta process (Indian buffet process)
- Sparse coding and dictionary learning
  - Inspired by neural coding
  - Each object encoded by a small number of representative patterns
  - Representative patterns, forming a dictionary or basis set, learned from data automatically

Further Machine Learning Topics - IV

- Large-scale machine learning
  - Big Data
  - Online learning
  - Parallel and distributed algorithms
- Multi-modal machine learning and sensor fusion
- Self-Supervised Learning and contrastive learning
- Many more . . .

Outline

Review

- Basically, every topic is equally important.

Outline

Final Exam

- Jan. 8th, 2022, Saturday
- 9:00am - 11:00am
- Rm. 102, Teaching Center – allowed to enter 5-10 mins before exam starts
- Closed-book and closed-notes exam, but two A4-sized "cheat sheet" (both sides) allowed
- Calculator not needed and not allowed
- Bring ShanghaiTech student ID card
- All topics included
- Questions are mostly on topics covered in HW's

Hope this course can
get you on the way to the AI world