Remarquable theorems, equations and methods that one should be able to prove and/or apply are indicated by ⋆

Part I: Classification variants

Multi-class learning

Rare-class learning

Ensemble methods

- Bucket of models
- Bagging, random forests
- Boosting, AdaBoost
- Stacking

Part II: Classification – Different paradigms

Semi-supervised learning

- Transductive SVM
- Clustering and classification
- Graph-based collective classification ⋆
- Self-training and co-training

Active learning

- Heterogeneity-based querying strategies, uncertainty sampling
- Performance-based querying strategies, expected error reduction
- Representativeness-based querying strategies

Part III: Mining Streams

Data stream paradigm

Synopsis data structures

- Sampling
  - Reservoir sampling ⋆
  - Bias-sensitive sampling
- Quality bounds
  - Markov's inequality ⋆
  - Chebychev's inequality ⋆
  - Chernoff bounds ⋆
  - Hoeffding's inequality ⋆
- Massive domain scenario
  - Approximate counting ⋆
  - Bloom filters ⋆
  - Count-min sketch ⋆
  - Flajolet-Martin algorithm ⋆
  - Alon-Matias-Szegedy sketch, mean-median trick ⋆
  - Frequent items, Lossy counting and exponentially decaying window ⋆

Classification

- Hoeffding's trees

Intro Temporal Data

Mining temporal data

- Temporal data characteristics and tasks
- Distance vs. similarity
  - Dynamic Time Warping (DTW), window constraint ⋆

Part IV: Mining Sequences

Distances

- Distances based on sequence alignment
  - Dynamic Time Warping (DTW) ⋆
  - Edit Distance (ED) ⋆
  - Longest Common Subsequence (LCS) ⋆
- Distances based on elements frequencies
  - Bag of Words and TF-IDF ⋆
  - n-grams, k-mers, and similarity kernels

Frequent pattern mining

- Subsequences and support ⋆
- Sequential pattern mining, GSP algorithm ⋆

Markov models

- Markov Chains
  - First and second order MC ⋆
  - Probabilistic suffix tree ⋆
- Hidden Markov Models
  - Evaluation: forward algorithm ⋆
  - Explanation: Viterbi algorithm ⋆
  - Training: Baum–Welch (aka forward-backward) algorithm ⋆

Part V: Mining Time-Series

Data preparation

- Interpolation ⋆
- Binning, smoothing, normalization and standardization ⋆
- Discretization, Symbolic Aggregate approXimation (SAX) ⋆

Transforms

- Discrete wavelet transform (DWT) ⋆
- Discrete Fourier Transform (DFT)

Models for time-series

- Stationarity, differencing
- Autocovariance and autocorrelation
- Periodicity

Part VI: Spatial Data

- Spatial and spatio-temporal data
- Distances and map projections
- Interpolation, density estimation, triangulation ⋆
- Contours and edges, shapes to time-series ⋆
- Discrete wavelet transform (DWT) ⋆
- Frequent trajectory patterns, tile transformation ⋆

Part VII: Outlier Analysis

Basics

- Applications
- Depth-based methods ⋆
- Deviation-based methods ⋆
- Density-based methods ⋆
- Statistical tests ⋆
- Mahalanobis distance ⋆
- Clustering models ⋆
- Distance-based models, k-NN distances, Local Outlier Factor (LOF) ⋆

High-dimensional data

- Angle-based method
- Subspace outlier detection
  - Grid-based sparsity coefficient, genetic algorithms
- Isolation-based methods, isolation trees

Temporal data

- Point outliers
- Combination and shape outliers (HOTSAX)