

Cours M2RI (second semestre)

Béatrice Laurent et Jean-Michel Loubes

Statistical Learning in high dimension

- **Introduction to supervised learning in regression and in classification**
 - o General Framework for classification : loss function, risk, optimal rules, Bayes classifier
 - o k- nearest neighbors algorithms, kernel algorithm, partition algorithms, curse of dimensionality
 - o empirical risk, penalised criterion, bias-variance trade-off, cross validation.
 - o Notions of Bias Detection in Machine Learning
- **Linear model in high dimension**
 - o Least square estimators with Ridge or Lasso regularisation. Risk bounds for the Lasso estimator
 - o Optimisation algorithms to compute the Lasso estimator
- **Classification an Regression trees (CART)**
 - o Definition of the algorithm
 - o Heterogeneity criteria
 - o Pruning algorithm
 - o Random forests
- **Support Vector Machine (SVM)**
 - o Linear SVM : geometric interpretation, convex optimisation problems (primal and dual problems)
 - o Nonlinear SVM and Reproducing Kernel Hilbert Spaces
- **Machine Learning for Anomaly Detection**
 - o One Class Modification for Classification Algorithm

- Variational Auto-Encoder
- **Mathematics Foundation of Transfert Learning**
 - Theory of Optimal Transport
 - Shift Assumption in Classification
 - Computational aspects and applications

References :

- **C.**
Giraud, *Introduction to high-dimensional statistics*, Monographs on Statistics and Applied Probability, CRC Press, Boca Raton, FL, 2015.
- **I.**
Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press
- **T.-**
Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning : data mining, inference, and prediction*, Springer 2009.
- **P.**
Massart, *Concentration inequalities and Model selection*, Springer, 2003.
- **B.**
Schölkopf and A. Smola, *Learning with Kernels Support Vector Machines, Regularization, Optimization and Beyond*, MIT Press, 2002.
- **V.N.**
Vapnik, *Statistical learning theory*, Wiley Inter science, 1999.
- **C.**
Molnar Interpretable Machine Learning, 2019. (available on line <https://christophm.github.io/interpretable-ml-book/>)