

Machine learning

Mathilde Mougeot

ENSIIE

Machine learning methods are today widely used in many applications as web page ranking, emails spam detection, energy models and forecasts...Over the past two decades, machine learning has become a key player for smart data analysis.

The purpose of this course is to provide an overview of the principal methods of machine learning to implement predictive models for a wide range of applications. The successive lessons will present the theoretical settings of machine learning in the regression and in the classification framework and also in the clustering framework and the implementation of these methods on real applications using the R software.

Program:

1. Introduction to machine learning, supervised learning and predictive models
2. Classification models. Performance criteria. Roc curves
3. Ensemble models. Classification and regression trees. Bagging. Random Forest.
4. Model aggregation. Boosting
5. Penalized regression. Performance criteria. Model selection
6. Clustering models

Bibliography:

- Hastié, T, Tibshirani, R, Friedman, J. The elements of statistical learning, 2009.
- James, G., Witten, D., Hastié, T. Introduction to statistical analysis with applications in R. 2013.
- Faraway, R. Extending the linear model with R, 2007.
- Hastié, T, Tibshirani, R, Wainwright, M. Statistical learning with sparsity, 2015.
- Goodfellow, I., Bengio, Y. and Courville, A. Deep learning, An MIT Press book, 2016.
-

Assessment:

Continuous monitoring & projects