

MACHINE LEARNING AND FOR FINANCE AND INSURANCE

Course Description

The course aims to provide an up-to-date knowledge of the theoretical and empirical developments of the tools for modelling and understanding complex data sets. With the explosion of 'big data' problems, machine learning and statistical learning have become very hot areas in many scientific fields, including finance. The course covers both supervised learning problems, where the goal is to predict the value of an outcome measure based on a set of input measures, and unsupervised learning problems, where there is no outcome measure and the goal is to describe the associations and patterns between a set of input measures. Statistical properties of the methods will also be investigated. The statistical software R will be used.

Objectives

The students will be able to apply the most appropriate in their empirical analyses.

Main contents

1. Supervised Learning
 - 1.1. Linear Methods for Regression
 - 1.2. Classification
 - 1.3. Cross-Validation
 - 1.4. Linear Model Selection and Regularization
 - 1.5. Tree-Based Methods
2. Unsupervised Learning
 - 2.1. Principal Component Analysis
 - 2.2. Clustering

References

- Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685-725.
- Gareth, J., Daniela, W., Trevor, H., & Robert, T. (2013). *An introduction to statistical learning: with applications in R*. Springer.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer.