

Applied inductive learning/Introduction to statistical learning - Lecture 1

Pierre Geurts and Louis Wehenkel

Department of Electrical Engineering and Computer Science
University of Liège

Montefiore - Liège - September 17, 2015

Find slides: <http://montefiore.ulg.ac.be/~lwh/AIA/>

Problems addressed

Background and prerequisites

Teaching methodology, material, exams, contacts

Problems addressed in this course

*How can we design and/or make use of **algorithms** in order to extract from (possibly very large) datasets good decision strategies, predictive models, explanations and interpretations ?*

- ⇒ Batch-mode supervised learning (The main building block)
 - ⇒ Classical algorithms (Decision trees, nearest-neighbor, neural nets etc.)
 - ⇒ Theory (Sampling, likelihood, bias/variance, statistical learning theory)
 - ⇒ Advanced algorithms (Ensemble methods, kernel-based methods)
 - ⇒ Diverse learning protocols (SSL, RL, ActL, USL, on-line learning etc.)
-

Problems addressed in this course

*How can we design and/or make use of **algorithms** in order to extract from (possibly very large) datasets good decision strategies, predictive models, explanations and interpretations ?*

- ⇒ Batch-mode supervised learning (The main building block)
- ⇒ Classical algorithms (Decision trees, nearest-neighbor, neural nets etc.)
- ⇒ Theory (Sampling, likelihood, bias/variance, statistical learning theory)
- ⇒ Advanced algorithms (Ensemble methods, kernel-based methods)
- ⇒ Diverse learning protocols (SSL, RL, ActL, USL, on-line learning etc.)

-
- NB.
- ▶ Many practical problems are concerned
 - ▶ Most are related to complex and uncertain environments
 - ▶ Scalability considerations are important

Background and prerequisites

- ▶ Way of thinking:
 - ▶ probability theory, information theory, logic
- ▶ Mathematical analysis tools:
 - ▶ linear algebra, calculus, optimization theory, statistics
- ▶ Problem solving attitude:
 - ▶ algorithmics, complexity theory, intuition

Teaching methodology, material, exams

- ▶ Course web page: <http://www.montefiore.ulg.ac.be/~lwh/AIA/>
- ▶ Lectures: **mandatory** every Thursday PM (14h-17h)
- ▶ Personal work: readings, 3 projects by group of two students with **written report and oral defense for the last one**
- ▶ Material: slides of lectures and selected chapters of reference textbooks
- ▶ Evaluation: projects (30%); oral exam (70%) (January)
- ▶ Contact persons: P. Geurts, L. Wehenkel, A. Joly, A. Sutera
 - ▶ P.Geurts@ulg.ac.be and L.Wehenkel@ulg.ac.be: any question related to the course
 - ▶ A.Joly@ulg.ac.be, A.Sutera@ulg.ac.be: any question related to the projects

Lecture topics

- ▶ Introduction to machine learning Today
- ▶ Supervised learning
 - ▶ Decision trees
 - ▶ Linear models
 - ▶ Nearest-neighbor methods
 - ▶ Neural networks
 - ▶ Sampling, bias and variance, model assessment and selection
 - ▶ Ensemble methods and feature selection
 - ▶ Support vector machines and kernel-based methods
- ▶ Unsupervised learning
 - ▶ Clustering
 - ▶ Dimensionality reduction
- ▶ Plus a few special lectures related to the projects: introduction to scikit-learn, project explanations and corrections.