Machine Learning: Algorithms and Applications

Floriano Zini

Free University of Bozen-Bolzano
Faculty of Computer Science

Academic Year 2011-2012

Contacts

- Floriano Zini
  - Room 2.18 (POS)
  - floriano.zini@unibz.it
- Office hours
  - Wednesday, 14:00 - 16:00
  - by prior arrangement via email
- Course web site
  - http://www.inf.unibz.it/~zini/ML
Course structure

- **Credits**: 4
- **Lectures**: 24 hours
- **Labs**: 12 hours
- **Assessment**
  - Project in small teams (max 2 people) [40 % of mark]
  - Final written exam [60 % of mark]

Timetable

- **Lectures**: **Mondays**, 10:30-12:30, E 411
- **Labs**: **Mondays**, 15:00-16:00, E 331
  - Starting on Monday, 05/03/2012
- Please check the course webpage and/or online timetable for variations
- No lecture/lab on
  - 09/04/2012
    - →10/4/2012, 17:00 - 20:00 E 531 ??
  - 30/4/2012
    - →2/5/2012, 16:00 - 19:00 E 531 ??
Goals of the course

- Introduce the fundamental principles and concepts of Machine Learning
- Present the principal algorithms and techniques focusing on
  - strengths and weaknesses
  - appropriateness for learning problems
- Teach how to design, implement, and evaluate ML algorithms
- Present practical applications of Machine Learning

How?

- Frontal lectures
- Labs
  - Exercises (fundamental for the written exam!)
    - Java will be used by the students to implement simple ML algorithms presented in the lectures
      - NetBeans (http://netbeans.org/) is the suggested IDE for development
      - Weka (http://www.cs.waikato.ac.nz/ml/weka/) will be used to explore the application of different ML algorithms
- Homework
  - Studying the slides is not enough
  - The students should also
    - Study on the suggested book
    - Do the proposed exercises
    - Read the proposed scientific articles
Syllabus

- Introduction to Machine Learning
  - What is ML, supervised, unsupervised, and reinforcement learning
- Classification and regression
  - Bayesian (probabilistic) learning, model probability estimation, dimensionality reduction, nonparametric methods, linear discrimination, artificial neural networks, genetic algorithms
- Design and analysis of ML experiments
  - Strategies, guidelines for good design, cross validation, performance measurement, hypothesis testing, algorithm comparison
- Advanced algorithms
  - Kernel Machines, Hidden Markov Models, Graphical models
- Reinforcement Learning

Course material

- Course slides
  - Available at [http://www.inf.unibz.it/~zini/ML](http://www.inf.unibz.it/~zini/ML) the day after the lecture
- Textbooks
    - Main text book, copies are available in the library
    - Most of the slides are taken from those provided by Alpaydin
    - A milestone book, copies are available in the library
- Selected scientific papers
  - Available at [http://www.inf.unibz.it/~zini/ML](http://www.inf.unibz.it/~zini/ML) the day after the lecture
- Software tools
  - Java NetBeans, Weka
- Datasets
Assessment

- Project (P) – maximum 12 points
  - In a team of maximum 2 students
  - Choose a ML technique introduced in the course to implement a learning system for an interesting task, and test it on some dataset(s)
  - Implementation in Java
- Written exam (E) – maximum 18 points
  - Questions and small exercises similar to those done in the lab
- **Positively assessed (i.e., at least 7 points) project work is required for taking the written exam!!**
- A positive mark of the project work counts for 3 regular exam sessions
- Final grade (G)
  - \( G = P + E \)
  - To pass: \( \geq 7 \) points of (P) and \( \geq 11 \) points of (E)

Project - proposal

- Freely propose the task, the ML algorithm, and the dataset(s) for your project work
- The project proposal should (be reasonably specific!)
  - be 1 - 2 pages long
  - describe the task you want to solve (i.e., application scenario)
  - indicate the learning algorithm you intend to apply
  - specify what are the input/output of the learning system, what kind of representation is used for the input/output
  - indicate the dataset(s) to be used
- Please send me by e-mail your proposal, including students’ names and ids
  - The deadline for proposal submission is ???(check course webpage)
  - I will revise the proposal and the students might be asked to improve it
Project - requirements

The deliverables of the project work include

- The source codes (in a zip file)
- The instruction (i.e., “readme”) file that describes in details how to install/compile/run the project program (and additional soft. pack. used)
- The project report (PDF file) that describes
  - The problem definition (i.e., which task the learning system solves)
  - The ML algorithm implemented and the datasets used
  - The system functions (and how to use these functions)
  - The experimental results of the system’s performance on the used datasets
  - The structure of the source code (and the role of each class and the main methods)
  - The major problems/issues occurred in the implementation, and how you solved them
  - Your discussions/findings/conclusions, and possible future improvements

Project - assessment

- The project should be delivered by email to me by ??? (check course webpage)
- The project work will be evaluated on
  - The complexity/difficulty of the task you have solved, and the quality of your solution
  - The evidence of the experimental results, and your discussions/arguments on these results
  - Quality of the report
  - The system implementation (e.g., functions, easy to use, etc.)
- If you reuse/extend/exploit existing codes/packages, you must explicitly and clearly describe in the report (and mention in the presentation)
  - what you have reused/extended/exploited; and
  - what are your own codes