Year : 2018/19

24299 - Machine Learning

Syllabus Information

Academic Course: 2018/19
Academic Center: 337 - Polytechnic School
Study: 3377 - Bachelor's Degree in Computer Engineering
Subject: 24299 - Machine Learning
Credits: 5.0
Course: 2
Teaching languages: Theory: Grup 2: Catalan
Grup 1: English
Practice: Grup 201: Catalan
Grup 202: Catalan
Grup 203: Catalan
Grup 101: English
Grup 102: English
Seminar: Grup 201: Catalan
Grup 202: Catalan
Grup 203: Catalan
Grup 204: Catalan
Grup 205: Catalan
Grup 101: English
Grup 102: English
Grup 103: English

Teachers: Xavier Binefa Valls, Vicente Gomez Cerda
Teaching Period: Third Quarter
Schedule: ---

Presentation

The course is an introduction to the principles, methods and techniques most relevant to what is called Machine Learning (aka Pattern Recognition, or statistical learning). Machine learning is the data-driven approach to build intelligent systems. This approach consists in creating statistical models that by means of optimization algorithms conform to observed data to detect patterns or regularities are are able to predict satisfactorily new unobserved data. The large amounts of complex data that is nowadays being generated makes this course an indispensable tool in the fields of Big Data or Data Science.

In this course we introduce the main concepts and basic techniques of this discipline. It offers a tour of the most important paradigms, supervised learning and unsupervised learning. The course also includes an introduction to the most existing methods recently, such as deep learning. Basic concepts of calculation and basic linear algebra are assumed as well as notions of optimization. The course emphasizes the analytical, computational, statistical and data visualization aspects.

The course aims to be useful to students of the different degrees that follow it and uses theory classes, problems and laboratories to achieve practical skills and knowledge of the fundamental concepts.

Associated skills

Basic Competences

CB2. That the students can apply their knowledge to their work or vocation of a professional form in a professional way and possess the competences which are usually proved by means of the elaboration and defense of arguments and solving the solution of problems within their study area;

CB3. That the students have the ability of collecting and interpreting relevant data (normally within their study area) to issue judgements which include a reflection about relevant topics of social, scientific or ethical nature;

Transversal Competences

CT2. Applying leadership and coordination and proving initiative.

CT3. Applying with flexibility and creativity the acquired knowledge and adapting it to new contexts and situations.

Specific Competences

CE1. Solving the mathematical problems which can be set out in the engineering and apply the knowledge on: linear algebra; differential and integral calculus; numerical methods, numerical algorithms, statistics and optimization.

CE7. Managing and applying concepts of basic statistics and statistic interference to large data sets.

CE8. Using machine learning methods as a part of the process of analysis of large data sets within distributed computation environments.

Learning outcomes

RA.CE1.1 Applying linear algebra knowledge to solve problems which can be set out in the engineering

RA.CE1.5 Using knowledge of statistics to solve problems which can be set out in the engineering.

RA.CE7.3 Modelling real classifying problems by means of computational paradigms.

RA.CE8.1 Recognizing and applying optimization algorithms present in engineering tasks.

RA.CE8.4 Mastering advanced topics of machine learning using neuronal networks applied to massive data.

Prerequisites

It is advisable to have studied the subjects of Algebra and Calculus of 1st year as well as those of Probability and Statistics of 2nd year.

Contents

1. Introduction
2. Review of linear algebra and multivariate analysis
3. Bayesian decision theory
4. Linear models for regression
5. Linear models for classification
6. Kernel methods
7. Suppor vector machines
8. Clustering methods
9. Principal Component Analysis
10. Deep learning 1
11. Deep learning 2
12. Review

Teaching Methods

The theory lectures explain the theoretical contents of the subject and give guidelines to solve problems and practices. They are presented using slides and visualization programs. In some cases a short video is introduced.

The typical learning process begins with a theory session that presents certain theoretical-practical foundations. The student will have to complement this activity with a reading on his own notes and the additional material provided by the teacher. It is quite possible that the third day session of the next day starts with an electronic test about the contents explained in the previous session.

The Laboratory sessions are done both in the practical sessions as in the problems and will be done with a computer (the student must take them -in the library they make loans if necessary). Three evaluable projects will be solved in these sessions (one of the initials will not be evaluable). Each laboratory project will contain several theoretical questions and several programming objectives. Except the first, the same project will work in two or more sessions. The programming language to be used is Python.

Evaluation

Each of the three activities of the subject will be assessed: theory classes, practices and problems. In the continuous assessment, the final grade is obtained by doing the weighted average in the following way (it requires having at least 5 in theory and in the labs, otherwise the course would be suspended):

Final grade = 0.4 * Exam + 0.4 * Labs + 0.2 * Controls

The Labs score corresponds to two practice sessions and the control grade corresponds to two small controls in problem classes. The practices will be evaluated for the quality of the answers to theoretical questions, the quality of the code and the memory that is sent.

In July recovery the note will be

NoteJuliol = Maximum (NoteJuliol, 0.4 * NoteJuliol + 0.4 * Labs + 0.2 * Controls)

The controls and practices are not recovered either.

Bibliography and information resources

The Basic Bibliography will be (the books are accessible by web):

- C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer 2009
- M. Deisenroth A Aldo Faisal, and Cheng Soon Ong. *Mathematics For Machine Learning.* Cambridge University Press.
- D.MacKay. *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 2003
- I. Goodfellow, Y.Bengio, A. Courville *Deep Learning.*
- S. Theodoridis and K. Koutroumbas. *Pattern Recognition.* Elsevier
- Hastie, Tibshirani, Friedman, *The Elements of Statistical Learning*, Springer-Verlag, 2013

Material:

- Slides
- Some videos
- Collection of seminar exercises (solutions will be provided once they have been taught in class)
- Statements of Laboratories

There will be slides of the whole subject in the Aula Global