Course guides
200900 - ML - Machine Learning

Last modified: 28/06/2020

Unit in charge: School of Mathematics and Statistics
Teaching unit: 723 - CS - Department of Computer Science.
715 - EIO - Department of Statistics and Operations Research.

Degree: MASTER'S DEGREE IN ADVANCED MATHEMATICS AND MATHEMATICAL ENGINEERING (Syllabus 2010). (Optional subject).
MASTER'S DEGREE IN STATISTICS AND OPERATIONS RESEARCH (Syllabus 2013). (Optional subject).

Academic year: 2020 **ECTS Credits:** 7.5 **Languages:** English

LECTURER

Coordinating lecturer: LUIS ANTONIO BELANCHE MUÑOZ

Others: Segon quadrimestre:
LUIS ANTONIO BELANCHE MUÑOZ - A
PEDRO FRANCISCO DELICADO USEROS - A

PRIOR SKILLS

The student should have knowledge of fundamental mathematical topics, such as linear algebra, calculus, probability distributions, optimization and basic (linear) statistical methods.

REQUIREMENTS

The student should have knowledge of basic machine learning concepts. These concepts can be acquired simultaneously, for example being enrolled in the "Statistical Learning" subject offered in the MESIO master.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
MAMME-CE2. MODELLING. Formulate, analyse and validate mathematical models of practical problems by using the appropriate mathematical tools.
MAMME-CE4. CRITICAL ASSESSMENT. Discuss the validity, scope and relevance of these solutions; present results and defend conclusions.

TEACHING METHODOLOGY

On-Site Learning: On-site learning will be organized into theoretical-practical sessions. All these sessions will be held in a standard classroom, although students should bring their own laptops. Lectures will normally combine a 75% of expository classes and another 25% of guided practical work. In the expository part of the sessions, the theoretical aspects are presented and discussed and accompanied by practical examples, using slides that will be previously supplied to the student. The fundamental work environment of the practical part of the sessions will be R, of which an intermediate knowledge is presumed (use of the environment and basic programming).

Off-Site Learning: Off-site learning will consist of the study and resolution of (mainly practical) problems that the student should turn in throughout the course. Some of these exercises will require completion of programming tasks in R and preparation of short reports using RMarkdown (or a similar tool).

LEARNING OBJECTIVES OF THE SUBJECT

Upon completion of the course, the student should have acquired advanced competences on the general topics of statistical machine learning and unsupervised topics, specially data visualization. In particular, the student should be able to produce machine learning solutions for many complex problems, including those in which a reduction of dimension is necessary, those where the data comes as variables of different mixed types, or those where the number of variables greatly exceeds the number of observations, such as problems typically found in genomics.

STUDY LOAD

Type	Hours	Percentage
Self study	127,5	68.00
Hours large group	60,0	32.00

Total learning time: 187.5 h

CONTENTS

Introduction to unsupervised learning

Description:
Definition and illustrative examples of unsupervised learning

Full-or-part-time: 2h
Theory classes: 2h

Nonlinear dimensionality reduction

Description:
a. Principal curves.
b. Local Multidimensional Scaling.
c. ISOMAP.
d. t-Stochastic Neighbor Embedding.
e. Applications

Full-or-part-time: 8h
Theory classes: 4h
Laboratory classes: 4h

Dimensionality reduction with sparsity

Description:
a. Matrix decompositions, approximations, and completion.
b. Sparse Principal Components and Canonical Correlation.
c. Applications

Full-or-part-time: 8h
Theory classes: 4h
Laboratory classes: 4h

General introduction to machine learning

Description:
Introduction to Bayesian thinking for machine learning. Learning by solving a regularized problem. Illustrative example.

Full-or-part-time: 5h
Theory classes: 2h
Practical classes: 3h

Learning in functional spaces

Description:
Reproducing kernel Hilbert spaces. The representer theorem. Example 1: Kernel ridge regression. Example 2: The Perceptron and the kernel Perceptron.

Full-or-part-time: 8h
Theory classes: 4h
Practical classes: 4h

Kernel functions in R^d

Description:
Description and demonstration of fundamental kernel functions in R^d. Polynomial and Gaussian kernels. General properties of kernel functions.

Full-or-part-time: 4h
Theory classes: 2h
Practical classes: 2h

The support vector machine for classification, regression and novelty detection

Description:
The support vector machine (SVM) is the flagship in kernel methods. Its versions for classification, regression and novelty detection are explained and demonstrated.

Full-or-part-time: 6h
Theory classes: 4h
Practical classes: 2h

Kernel functions for diferent data types

Description:
Some kernel functions for different data types are presented and demonstrated, such as text, trees, graphs, categorical variables, and others.

Full-or-part-time: 6h
Theory classes: 4h
Practical classes: 2h

Other kernel-based learning algorithms

Description:
Additional kernel-based learning methods are explained, such as kernel PCA and kernel FDA. These are illustrated in several application examples.

Full-or-part-time: 5h
Theory classes: 3h
Practical classes: 2h

Advanced ideas and techniques in kernel-based learning methods

Description:
Other advanced methods are briefly introduced, such as the RVM and GPs. Nyström acceleration and random Fourier features. Introduction to the idea of Deep Kernel Learning

Full-or-part-time: 2h
Theory classes: 2h

GRADING SYSTEM

The grading method will be based in two basic marks, as follows:

1) Pr done through the course: 50%
2) Final exam: 50%

The practical work will consist in a term project as well as several exercises, all of which can be done in groups (their fornat will be specified onsite), but the exam is completed as an individual task.

EXAMINATION RULES.

The precise format for the exam will be specified with sufficient advance. It may include restrictions on the allowed knowledge sources, such as written notes, books, internet connection, etc.

BIBLIOGRAPHY

Basic:
- Hastie, Trevor; Tibshirani, Robert; Wainwright, Martin. Statistical learning with sparsity : The Lasso and Generalizations [on line]. CRC raton, 2015 [Consultation: 30/06/2020]. Available on: https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=4087701. ISBN 78-1-4987-1216.
- Bishop, Christopher M. Pattern recognition and machine learning. New York: Springer, cop. 2006. ISBN 978-0387310732.
- Vapnik, Vladimir Naumovich. The Nature of statistical learning theory. 2nd ed. New York ; Barcelona [etc.]: Springer, cop. 2000. ISBN 978-0387987804.
- Kung, S. Y.. Kernel Methods and Machine Learning. Cambridge University Press, 2014.
- Schölkopf, Bernhard; Smola, Alexander J. Learning with Kernels : support vector machines, regularization, optimization, and beyond. Cambridge ; London: The MIT Press, cop. 2002. ISBN 9780262194754.
- Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome. The Elements of Statistical Learning : Data Mining, Inference, and Prediction [on line]. 2nd ed. New York, NY: Springer New York, 2009Available on: http://dx.doi.org/10.1007/978-0-387-84858-7. ISBN 978-0-387-84858-7.

Complementary:
- Smola, Alexander J. Advances in large margin classifiers. Cambridge, Mass.: MIT Press, 2000. ISBN 9780262194488.