UNIVERSITÀ DEGLI STUDI DI PALERMO

DEPARTMENT	Ingegneria
ACADEMIC YEAR	2019/2020
BACHELOR'S DEGREE (BSC)	ELECTRONIC ENGINEERING
SUBJECT	MACHINE LEARNING
TYPE OF EDUCATIONAL ACTIVITY	D
AMBIT	10432-A scelta dello studente
CODE	17878
SCIENTIFIC SECTOR(S)	ING-INF/03
HEAD PROFESSOR(S)	TINNIRELLO ILENIA Professore Ordinario Univ. di PALERMO
OTHER PROFESSOR(S)	
CREDITS	6
INDIVIDUAL STUDY (Hrs)	102
COURSE ACTIVITY (Hrs)	48
PROPAEDEUTICAL SUBJECTS	
MUTUALIZATION	MACHINE LEARNING - Corso: INGEGNERIA CIBERNETICA
YEAR	2
TERM (SEMESTER)	2° semester
ATTENDANCE	Not mandatory
EVALUATION	Out of 30
TEACHER OFFICE HOURS	**TINNIRELLO ILENIA** Monday 9:00 12:00 Ufficio del docente, presso il DEIM, secondo piano.

DOCENTE: Prof.ssa ILENIA TINNIRELLO

PREREQUISITES	The course is self-consistent. However, it is recommended to have some basics of probability theory, linear algebra and programming.
LEARNING OUTCOMES	Knowledge and understanding The student will be able to understand and analyze data classification problems and to apply concepts and methodologies for extracting useful information from the data. In particular, she/he will understand the behavior of Bayesian, linear and non-linear classifiers and Markov-based classifiers, as well as explore performance of different clustering and feature extraction algorithms. As special attention will be dedicated to case studies related to telecommunication systems. Applying Knowledge The student will be guided to implement some reference algorithms for applications on real datasets. She/he will be stimulated to extrapolate the concepts and the methods presented in the course for some specific case studies in order to apply them (and relevant design considerations) to different application scenarios and, in particular, to the analysis of audio and video signals. Judgements The student will be able to compare different approaches for data analysis according to the characteristics of the data sets and application scenarios. She/he will be also able to generalize the concepts and the methods acquired within the course and to related them to other statistical tools presented in other disciplines. Communication skills The student will learn the ability to rationally communicate her/his knowledge about the concepts and methods of the discipline, with a good level of clearness, fluency and correct use of technical language. In particular, she/he will be able to justify the design choices and the application of specific tools for solving the proposed analysis or synthesis problems. Learning skills The student will be able to read autonomously textbooks and scientific literature on machine learning, in order to study in depth approaches not discussed in the teacher-led lessons.
ASSESSMENT METHODS	EXAM ORGANIZATION The examination is based on a mandatory written test and an optional oral exam. The oral exam allows to improve the written test evaluation. To take the oral exam, it is required to have at least a sufficient evaluation of the written test. The grade of the written test is given in the range 0-30/30. The minimum grade to pass the test is 18/30. The oral test is evaluated in the range of 0-3/30 to be added to the grade of the written test. The final grade is given by the written test grade (in case the student does not take the oral exam) or by the sum of the written test and oral exam grades. DESCRIPTION OF THE TESTS The written test includes two open questions and three exercises, similar to the examples discussed in the course, in which the student has to apply the concepts and the methodologies presented during the lessons to simple problems of data analysis, system modeling and statistical inference. The written test lasts 2.5 hours. The test is devised to evaluate: - The knowledge and understanding levels of learning concepts and algorithms; - The ability of applying the acquired knowledge to solve autonomously learning problems and system optimizations; - The ability to communicate knowledge, analyses and conclusions, and justify the design choices. The oral exam lasts about 30 minutes. It is based on the presentation of a MATLAB project developed autonomously by the student on a case study. The exam allows to assess: - The ability of applying the learning schemes to real problems, by exploiting programming languages and libraries; - The ability to communicate knowledge, analyses and conclusions, with a good level of clearness, fluency and correct use of language; - The ability of reinterpretation of the concepts and interdisciplinary connections, showing evidence for autonomously undertaking further studies or professional activity. LEARNING OUTCOMES In order to provide the overall evaluation, we will estimate the results achieved in the following course objectives. Knowledge and understanding: Evaluation of knowledge, understanding and integration of principles, concepts, methods and techniques of the discipline.

	Applying knowledge: Evaluation of capabilities in applying theoretical and technical knowledge for tackling and solving problems; evaluation of the autonomy level and originality of proposed solutions. Making judgements: Evaluation of logical, analytical and critical abilities for reaching appropriate judgments and decisions, based on available information and data. Communication skills and learning skills: Evaluation of the ability to communicate knowledge, analysis and conclusions, with a good level of clearness, fluency and correct use of language. Evaluation of the capability of reinterpretation and interdisciplinary connection, showing evidence for autonomously undertaking further studies or professional activity. GRADES 30-30 and laude: Excellent. Full knowledge and understanding of concepts and methods of the discipline, excellent analytical skills even in solving original problems; excellent communication and learning skills. 27-29: Very good. Very good knowledge and understanding of concepts and methods of the discipline; very good communication skills; very good capability of concepts and methods applications. 24-26: Good. Good knowledge of main concepts and methods of the discipline; discrete communication skills; limited autonomy for applying concepts and methods for solving original problems. 21-23: Satisfying. Partial knowledge of main concepts and methods of the discipline; satisfying communication skills; scarce judgment autonomy. 18-20: Acceptable: Minimal knowledge of concepts and methods of the discipline; minimal communication skills; very poor or null judgement autonomy. 0-17: Non acceptable: Insufficient knowledge and understanding of concepts and methods of the discipline.
EDUCATIONAL OBJECTIVES	A first educational objective of the course is knowing and understanding some statistical approaches for data analysis and pattern recognition, which are especially useful for the analysis of audio and video signals and for the analysis of time series. A second education objective is applying the concepts and the methods presented in the course to real problems related to telecommunication systems, such as demodulation of coding signals and monitoring of internet traffic.
TEACHING METHODS	Teacher-led lessons and exercises; practical exercises in Python.
SUGGESTED BIBLIOGRAPHY	-SergiosTheodoridis. Pattern Recognition - Kevin P. Murphy. Machine Learning, a probabilistic perspective -Peter Harrington Machine Learning in Action -Lucidi del corso: http://www.tti.unipa.it/

SYLLABUS

Hrs	Frontal teaching
2	Introduction to machine learning.
4	Refreshing of probability and random variables.
8	Classification: hypothesis representation, decision regions, cost functions. Bayesian classifiers. Solutions for estimating the feature probability distributions: parametric and non-parametric estimators.
8	Linear classifiers and regressors: the perceptron and training techniques (gradient technique, cost functions, training for non-separable classes, etc.). Generalizations to multi-class classifiers. Classifiers with activation functions different from threshold functions: the logistic regressor. Introduction to support vector machines. Linear regressors and bias/variance dilemma.
8	Non linear classifiers: multi-stage neural networks. The XOR problem; three-level structures for multiple perceptrons and union of polyhedral regions. Backpropagation algorithm. Feature space transformation. Decision-trees and optimal threshold derivation. Combination of multiple non-linear classifiers.
8	Discrete-time Markov processes: transition matrix, equilibrium conditions, limit probabilities. Application example: the PageRank algorithm in Google. Viterbi algorithm for classifying a process with memory. Channel equalization. Approaches for training Hidden Markov Models.
2	Clustering: K-Means and fuzzy K-Means. Selection of the clusters number. Examples of clustering applications to user profiling in web applications.

Hrs	Practice
8	Introduction to Python and to the scikit learn library. Applications of all the proposed approaches to real data problems.