UNIVERSITÀ DEGLI STUDI DI PALERMO

DEPARTMENT	Scienze Economiche, Aziendali e Statistiche
ACADEMIC YEAR	2023/2024
MASTER'S DEGREE (MSC)	STATISTICS AND DATA SCIENCE
INTEGRATED COURSE	NETWORK ANALYSIS AND OPTIMIZATION - INTEGRATED COURSE
CODE	21925
MODULES	Yes
NUMBER OF MODULES	2
SCIENTIFIC SECTOR(S)	SECS-S/06
HEAD PROFESSOR(S)	TUMMINELLO MICHELE Professore Ordinario Univ. di PALERMO
OTHER PROFESSOR(S)	TUMMINELLO MICHELE Professore Ordinario Univ. di PALERMO SIMONETTI ANDREA Ricercatore a tempo determinato Univ. di PALERMO
CREDITS	6
PROPAEDEUTICAL SUBJECTS	
MUTUALIZATION	
YEAR	1
TERM (SEMESTER)	2° semester
ATTENDANCE	Not mandatory
EVALUATION	Out of 30
TEACHER OFFICE HOURS	**SIMONETTI ANDREA** Tuesday 16:00 18:00 Edificio 13, I piano. Building 13, I floor, **TUMMINELLO MICHELE** Monday 14:00 16:00 Studio/Laboratorio: primo piano, ex DSSM Tuesday 14:00 16:00 Studio/Laboratorio: primo piano, ex DSSM

DOCENTE: Prof. MICHELE TUMMINELLO

PREREQUISITES	Vectors in R^n and their properties. Function of several variables. Matrix algebra. Differential and integral calculus. Gradient and Hessian of a function of several variables. Convexity of a function of several variables. First and second order condition of optimality. Elementary programming in R.
LEARNING OUTCOMES	1. Knowledge and understanding- definition and description of unconstrained and constrained optimization. Ability to Identify and discuss for convex, linear and quadratic optimization models. Definition and vector representation of discrete optimization problems. Ability to Identify and discuss the properties of a network. 2. Applying knowledge and understanding-Ability to Implement a classification model in Python by solving an optimization problem. Ability to analyze the properties of a network using Python or R. 3. Making judgements Ability to analyse a real optimization problem and choose the appropriate mathematical model to implement a classification model. Ability to analyse a real optimization problem and choice of the appropriate method to search for solutions. Ability to analyze e real network by choosing the appropriate indicators and metrics. 4. Communication skills-Present the results in professional way through pictures and spreadsheets. 5. Learning skills-Conduct research and analysis in the field of decision science using optimization and network models
ASSESSMENT METHODS	The exam consists of two parts related to the two modules taught. For the Optimization module, the test involves working on a project related to the study of a real dataset for classification models, preparing a brief report on the project, and giving an oral presentation describing its content. For the Networks module, the exam consists of working on a project related to the study of a real network, preparing a brief report on the project, and giving an oral presentation describing its content. Both projects are agreed upon with the instructor. For the Optimization module, a passing grade will be assigned to students who demonstrate their ability to describe the main properties of a classifier and the methods used for parameter estimation. For the Networks module, a passing grade will be assigned to students who demonstrate their ability to describe the main properties of the network, using the metrics considered during the course. The final grade is the arithmetic average of the grades obtained in both modules. The reports for both modules should be written in English. The final presentation for both modules will typically be conducted in English. However, the instructors may choose to discuss one or more topics in Italian if they deem it appropriate.
TEACHING METHODS	Lectures and practices

MODULE OPTIMIZATION *Prof. ANDREA SIMONETTI*	
SUGGESTED BIBLIOGRAPHY	
S. Boyd and L.Vandenberghe. Convex Optimization. Cambridge University Press. 2004. A. Ohri. Python for R users: a data science approach. John Wiley & Sons Inc. 2018 G. James, D. Witten, T. Hastie, R. Tibshirani. An introduction to statistical learning with applications in R. Springer. 2021	
AMBIT	50608-Matematico applicato
INDIVIDUAL STUDY (Hrs)	54
COURSE ACTIVITY (Hrs)	21
EDUCATIONAL OBJECTIVES OF THE MODULE	
At the end of the course, the student will be able to: 1) Define a constrained and unconstrained optimisation problem 2) Determine the maxima and minima of constrained and unconstrained optimisation problems 3) Use Python software to implement classification models 4) Understand the concept of overfitting and ross-validation	

SYLLABUS

Hrs	Frontal teaching
2	Description of course objectives. Constrained and unconstrained optimisation. Definition of convex functions and sets. First and second order conditions.
4	Constrained optimisation with equality constraints and inequality constraints. Convex and non-convex optimisation. Lagrange problems. Duality and Lagrangian duality.
4	Support Vector Machine (SVM). The kernel trick. Polynomial kernels and the radial basis function. Non-linear and soft margin kernels.
2	Machine Learning Approaches. Training a machine learning model, bias-variance trade-off, overfitting, cross-validation.

Hrs	Practice
2	Introduction to Python
3	Introduction to Numpy and Pandas modules for data manipulation
4	Implementation of Supervised Classification Models via Support Vector Machine

MODULE NETWORKS *Prof. MICHELE TUMMINELLO*	
SUGGESTED BIBLIOGRAPHY	
M. Newman, Networks: An Introduction, Oxford University Press. D. Pham, D. Karaboga, Intelligent Optimisation Techniques, Springer. D. Easley and J. Kleinberg, Networks, Crowds and Markets, Cambridge.	
AMBIT	50608-Matematico applicato
INDIVIDUAL STUDY (Hrs)	54
COURSE ACTIVITY (Hrs)	21
EDUCATIONAL OBJECTIVES OF THE MODULE	
OBJECTIVES OF THE UNIT are to: 1) construct a network model of a real world system and recognize its structure; 2) provide a vector representation of the space of solutions to the problem of modularity optimization and use heuristic stochastic optimization methods to identify sub-optimal solutions; 3) analyze the convergence of an iterative and stochastic algorithm that provides suboptimal solutions to the modularity optimization problem; 4) understand the difference between accuracy and precision of a solution; 5) describe the role of communities in a realization of the SIR model.	

SYLLABUS

Hrs	Frontal teaching
2	An introduction to networks. Descriptive analysis: degree, betweenness centrality, page rank, clustering coefficient.
2	Degree distribution, scale-free networks, Albert-Barabasi model
2	Stochastic processes on networks. Mean-field models. The SIR model.
2	Community detection through modularity optimization
2	Simulated annealing, genetic algorithms, taboo search, and extreme optimization to optimize modularity.
2	The infomap method

Hrs	Practice
4	Application of simulated annealing and genetic algorithms to real examples of optimization problems (e.g. the traveller salesman problem)
5	R and C tools for modularity optimization.