CORSO DI STUDIO: SCIENZE STATISTICHE (L41)

ANNO ACCADEMICO: 2023 - 2024

DENOMINAZIONE DELL'INSEGNAMENTO:

Fondamenti di Programmazione

General information	
Year of the course	Third year
Academic calendar (starting and ending date)	Second semester (19th february – 7th June 2024)
Credits (CFU/ETCS):	6
SSD	ING-INF/05
Language	Italian
Mode of attendance	Attendance is not mandatory but strongly recommended

Professor/ Lecturer	
Name and Surname	Rosalia Maglietta
E-mail	rosalia.maglietta@cnr.it
Telephone	080-5929454
Department and address	CNR, Via Amendola 122 D-O, Bari (8° floor)
Virtual room	Skype (username: magliettalia)
Office Hours (and modalities: e.g., by appointment, on line, etc.)	The reception of students will take place by appointment, to be arranged in advance via email

Work schedule			
Hours			
Total	Lectures	Hands-on (laboratory, workshops, working groups, seminars, field trips)	Out-of-class study hours/ Self-study hours
150	28	14	108
CFU/ETCS			
6	4	2	

Learning Objectives	The course focuses on acquiring skills in the main algorithms of machine learning and programming languages used in this field. Students learn to correctly define a data analysis methodology, critically leveraging machine learning strategies, interpret results, and accurately present the conclusions of the conducted analysis.
Course prerequisites	The course requires: understanding of fundamental mathematical concepts; in-depth knowledge of descriptive and inferential statistics; basic skills in at least one programming language, including: Python, MatLab, and/or R.

Teaching strategies	Lectures, Laboratories and exercises, Analysis and discussion of Case Studies, Study Groups, interactive sessions, and discussions

Expected learning outcomes	At the end of the course, the student demonstrates a solid understanding of the fundamentals of machine learning, is capable of successfully applying machine learning techniques using programming languages such as Python and/or Matlab to solve practical problems, and can critically evaluate the performance of models.

DD1 **Knowledge and understanding skills**	3In-depth understanding of the mathematical and statistical concepts underlying machine learning algorithms. Comprehensive understanding of the concepts of overfitting and underfitting. Knowledge of supervised and unsupervised learning algorithms. In-depth understanding of classification and regression concepts. Ability to preprocess data effectively, making them suitable for training machine learning models. Knowledge and ability to address issues related to the use of machine learning algorithms, including handling bias in data and models.
DD2 **Applied knowledge and understanding skills**	Ability to apply machine learning techniques to solve real-world problems, implementing models and interpreting results. Ability to select the most suitable machine learning algorithms to address the practical problem at hand. Ability to critically evaluate model performance using appropriate metrics.
DD3-5 **Soft skills**	Ability to clearly communicate results, conclusions, and applications of machine learning methods, both in written and oral form. Ability to independently learn new methodologies based on machine learning. Skills in working collaboratively in a team to solve complex machine learning problems, leveraging individual competencies collaboratively.
Teaching contents (Program)	Fundamental concepts and applications of machine learning. Introduction to Matlab and Python programming languages. Supervised and Unsupervised Learning. Regression and Classification Problems. Model Accuracy Evaluation: Overfitting, Underfitting, Bias, and Variance. Bayes Classifier. K-Nearest Neighbors Algorithm. Linear Regression Problem. Classification Problem. Resampling Methods: Cross-Validation and Bootstrap. Evaluation Metrics for Machine Learning Model Performance. Support Vector Machines Algorithm. Adaboost Algorithm. Random Forest Algorithm. Statistical Learning of Imbalanced Data. Rusboost Algorithm. Clustering Methods. Statistical Significance: p-value.

	Introduction to Deep Learning. Practical application of machine learning methods to real case studies.
Texts and readings	G. James, D. Witten, T. Hastie, R. Tibshirani, J. Taylor, An Introduction to Statistical Learning (2023) Springer Other books: E. Alpaydin, Introduction to Machine Learning. 3rd Edition, MIT Press C. Bishop, Pattern Recognition and Machine Learning, Springer
Notes on texts	
Additional materials	Scientific articles and technical reports used for in-depth analysis, laboratories, and study groups will be provided to students during the lectures.

Assessment	
Assessment methods	Oral examination with discussion of key concepts and theories of machine learning outlined in the program. In addition to assessing acquired knowledge and upon agreement with the instructor, the exam may involve the preparation and PowerPoint presentation of specific machine learning topics derived from reading and critically analyzing scientific articles on the chosen theme.
Assessment criteria	The student is evaluated based on the acquired knowledge of the main concepts of machine learning and programming covered during the course, as well as the ability to present them clearly and accurately. The student's final evaluation is influenced by their ability to form an independent critical judgment in the study and analysis of machine learning problems.
Final exam and grading criteria	The evaluation is expressed in thirtieths. The exam is considered passed if the student achieves a minimum score of 18/30.
Further information	