



UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

PHD PROGRAM IN INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

PHD PROGRAM IN INFORMATION AND COMMUNICATION TECHNOLOGY FOR HEALTH

PhD Course announcement

Title: Machine Learning for Science and Engineering
Research

Lecturers: Proff. Anna Corazza, Roberto Prevete, Carlo Sansone

University of Naples "Federico II"

Email: anna.corazza@unina.it, roberto.prevete@unina.it, carlo.sansone@unina.it

Giorgio Satta

University of Padua

Credits: 5

Short bio notes:

Anna Corazza is associate professor of Computer Science at the Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione (DIETI) of the University of Naples Federico II. Since 1989 to 2000 she worked at the ITC-irst (today FBK) in Trento, in the Speech Processing group. Since 2000 to 2003 she has been an assistant professor at the University of Milan, Polo Didattico e di Ricerca di Crema, in the Information Technology Department. In 1993 she visited for four months the Speech Laboratory of McGill University, in Montreal, Canada, directed by Prof. Renato De Mori. In 1994 she visited for six months the Center of Language and Speech Processing, at the Johns Hopkins University, Baltimore, MD, USA, directed by prof. F. Jelinek. Her main research interests are in natural language processing and information retrieval, machine learning, privacy-preserving data mining, machine learning applications in software engineering.

Roberto Prevete (1994: MSc in Physics, 2000: PhD in Mathematics and Computer Science at University of Naples Federico II) is an Assistant Professor of Computer Science at the Dept. of Electrical Engineering and Information Technologies (DIETI), University of Naples Federico II, Italy. His main research interests concern computational models of brain mechanisms, theoretical aspects of machine learning, artificial neural networks, deep learning and their applications. His research has been published in international journals such as Neurocomputing, Neural Networks, Pattern Recognition Letters, Pattern Recognition, Knowledge-Based Systems, Scientific Reports, Biological Cybernetics, Experimental Brain Research, and Behavioral and Brain Sciences.

Carlo Sansone is currently Full Professor of Computer Engineering at the Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione of the University of Naples Federico II, where he coordinates the Master in Human-Centered Artificial Intelligence. His research interests cover the areas of image analysis and recognition, machine learning and deep learning. From an applicative point of view, his main contributions were in the fields of biomedical image analysis, biometrics, intrusion detection in computer networks and image forensics. He coordinated several projects in the areas of biomedical images



interpretation, network intrusion detection and cyberbullying detection by using artificial intelligence techniques.

Giorgio Satta is a Full Professor in the Department of Information Engineering at the University of Padua. He earned his Ph.D. in Computer Science from the University of Padua in 1990, followed by a postdoctoral fellowship at the Institute for Research in Cognitive Science at the University of Pennsylvania. Professor Satta's research program focuses on the field of Natural Language Processing (NLP). His scholarly contributions include one international monograph and over 160 peer-reviewed articles published in leading international journals, conference proceedings, and specialized workshops. Professor Satta has held important services within the scientific community, in the role of Chair of the European Chapter of the Association for Computational Linguistics (EACL) for the 2009–2010 term and in the role of Program Committee Chair of the Conference of the Association for Computational Linguistics in 2001.

Overview

The course introduces the main topics in machine learning for both supervised and unsupervised approaches. In addition to a general introduction to the field, we discuss a few topics that are widely considered very effective and promising. In particular, the concept of explainable AI will be discussed, with special attention to the case of neural networks.

There will be a final assessment.

Schedule

Lecture	Date	Time	Class	Topics	Lecturer
1	May, 28	8:30 - 10:30	T6	Supervised machine learning	Anna Corazza
2	May, 29	8:30 - 10:30	C2A	Unsupervised machine learning	Anna Corazza
3	June, 3	10:30 - 12:30	T6	From shallow to deep networks	Roberto Prevete
4	June, 4	8:30 - 10:30	T6	Deep Learning	Roberto Prevete
5	June, 5	8:30 - 10:30	C2A	Generative AI for images	Roberto Prevete
6	June, 8	8:30 - 10:30	T4	Ensemble methods	Carlo Sansone
7	June, 9	10:30 - 12:30	T4	Feature design	Anna Corazza
8	June, 10	10:30 - 12:30	T6	DL and GAI for sequences	Anna Corazza
9	June, 11	8:30 - 10:30	T6	Privacy for machine learning	Anna Corazza
10	June, 12	8:30 - 10:30	C2A	Foundations of Large Language Models	Giorgio Satta
Assessment test					

Content details

Lesson 1 – Introduction to the course. Supervised machine learning: definition of supervised machine learning with particular emphasis on classification, example of classification approaches in the vector space model (Rocchio, kNN), Support Vector Machines (linear versus non-linear classifiers, hard margin and soft margin support SVM, kernels). (Anna Corazza).

Lesson 2 – Unsupervised machine learning: introduction to clustering, flat clustering, K-means, clustering assessment, choice of the number of clusters. Hierarchical clustering: introduction, dendrograms, variants, discriminative cluster labelling, non discriminative cluster labelling. (Anna Corazza)

Lesson 3 – From shallow networks to deep networks: Structure and behaviour of Multi-layer Feed-Forward Neural Networks. Shallow networks as universal approximators. Error Functions and Optimization methods based on gradient descent. Back-propagation algorithm to compute error gradient. Neural Network



Evaluation. Introduction to data leakage and its prevention. Basic principles of Deep Learning. “ (Roberto Prevete)

Lesson 4 – Deep Learning: Deep Networks:: Rectified Linear Units (ReLU) and its variants. Trainable activation functions. Convolutional Neural Networks. (Roberto Prevete)

Lesson 5 -- Generative AI for images: Latent Variables in Machine Learning. GAN (Generative Adversarial Neural Network) and VAE (Variational AutoEncoder) approaches. (Roberto Prevete)

Lesson 6 – Ensemble methods: Combining Multiple Models. Bagging. Randomization: Random Subspace Ensemble, Random Forest, Rotation Forest. Boosting, Additive Regression. Stacking. Error Correcting Output Codes. (Carlo Sansone).

Lesson 7 – Feature design: Introduction to the problem of dimensionality reduction; geometrical introduction to Principal Component Analysis and its statistical interpretation; introduction to the feature selection problem. (Anna Corazza)

Lesson 8 – Deep learning and Generative AI for sequences. Recurrent NN, transformers and LLMs. (Anna Corazza)

Lesson 9 – Privacy for machine learning: general definitions and problems; differential privacy, information leaking for LLMs. (Anna Corazza)

Lecture 10 – Foundations of Large Language Models: Large Language Models (LLMs) represent a paradigm shift in the field of natural language processing, from task specialised architectures to a single universal model, leading to the generative artificial intelligence revolution that is having a large-scale societal impact today.

This seminar provides a high-level overview of the foundational aspects of LLMs, examining the evolution of high-dimensional word embeddings, the architectural innovation of the attention mechanism in capturing long-range dependencies within natural language, and the idea of autoregressive modelling as applied to the task of next-token prediction. Furthermore, we will detail the training pipeline of LLMs, from self-supervised pre-training on massive datasets using scaling laws to post-training refinement.

The talk will conclude by addressing open problems and debates, such as hallucination, interpretability, system biases, data exhaustion, and the stochastic parrot controversy, inviting a discussion on the future trajectories of the field. (Giorgio Satta)

By May 24th, 2026, participants are requested to join the following MS Teams group: [General | ML4SER 2026](#) (

<https://teams.microsoft.com/l/team/19%3ATQ4XRzE57tkccVZeEiQWfJj93BFW3eHDkZLNEdT4Ab01%40thread.tacv2/conversations?groupId=fa76bc18-486d-43ef-87f0-a32aa0dcadca&tenantId=2fcfe26a-bb62-46b0-b1e3-28f9da0c45fd>).

Once accepted in the Teams group, students must fill the following .xlsx file with their information:

<https://communitystudentiunina.sharepoint.com/:x/s/ML4SER2026/IQADbQkOixItRbKUShs8E-7FAUneIu6gB7moxgNOe0I0D4k?e=fbFUfN>

The course is in presence. However, students pursuing their PhD period abroad (for research purposes) have the option to request remote attendance for classes via MS Teams.

For information: Prof. Anna Corazza (DIETI, UniNA) – anna.corazza@unina.it (organizer)