

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II
DOTTORATO DI RICERCA / PhD PROGRAM IN
INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Seminar announcement

Wednesday 13 April 2022, Time: 14.30-16.00

Seminar Room, Floor I, Building 3/A, DIETI - Via Claudio, 21 - NAPOLI

MsTeams: <https://teams.microsoft.com/l/team/19%3agCflgCOe7qasFFjRFS2gz0E0bhvFk3cLfBB2N-UG4X41%40thread.tacv2/conversations?groupId=5775ad02-278d-441a-8852-5a560b1edca8&tenantId=2fcfe26a-bb62-46b0-b1e3-28f9da0c45fd>

Teams Code: **jnxqh3b**



Dr. Marco Valentino

University of Manchester, Manchester, United Kingdom,

Idiap Research Institute, Martigny, Switzerland

<https://www.marcovalentino.net>

Explainable Natural Language Inference

Abstract: The application of modern Deep Learning has led to remarkable results in complex Natural Language Inference (NLI) tasks. However, while architectures such as Transformers can achieve state-of-the-art performance on popular benchmarks, they tend to be inscrutable black-box models, posing serious concerns in terms of interpretability, trust, and evaluation. An appealing solution to overcome these limitations is to build intrinsically explainable models, capable of performing complex inference through the generation of natural language explanations. This talk will

review the main lines of research in explainable NLI, presenting benchmarks and evaluation methodologies along with state-of-the-art architectures. Specifically, we will investigate the problem of performing multi-hop and abductive natural language inference on large corpora, with an emphasis on hybrid models preserving robustness and scalability. The final part of the talk will discuss some of the challenges involved in evaluating explainability, presenting open research questions on quantitative metrics, faithfulness and inferential properties of natural language explanations.

Lecturer short bio: Marco Valentino was a master student at Federico II Computer Sciences degree and now is going to defend his thesis as PhD student in Computer Science at the Reasoning & Explainable AI group at the University of Manchester. He is also presently a Research Intern at the Idiap Research Institute, Martigny, Switzerland. His main research activity focuses on Explainability and Natural Language Inference, investigating the construction and evaluation of models capable of performing complex commonsense and scientific reasoning through the generation of natural language explanations.

For information: Prof. Francesco Cutugno (DIETI, UniNA) – francesco.cutugno@unina.it (organizer)